

# frontiers

## RESEARCH TOPICS

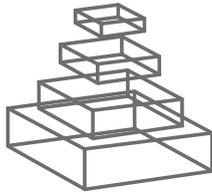
### NEURAL PROCESSING OF EMOTION IN MULTIMODAL SETTINGS

Topic Editors

Martin Klasen, Benjamin Kreifelts, Yu-Han Chen,  
Janina Seubert and Klaus Mathiak



frontiers in  
**HUMAN NEUROSCIENCE**



# frontiers

## **FRONTIERS COPYRIGHT STATEMENT**

© Copyright 2007-2015  
Frontiers Media SA.  
All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

Cover image provided by lbbl sarl, Lausanne CH

ISSN 1664-8714

ISBN 978-2-88919-414-8

DOI 10.3389/978-2-88919-414-8

## **ABOUT FRONTIERS**

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## **FRONTIERS JOURNAL SERIES**

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.

All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## **DEDICATION TO QUALITY**

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## **WHAT ARE FRONTIERS RESEARCH TOPICS?**

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: [researchtopics@frontiersin.org](mailto:researchtopics@frontiersin.org)

# NEURAL PROCESSING OF EMOTION IN MULTIMODAL SETTINGS

Topic Editors:

**Martin Klasen**, RWTH Aachen University, Germany

**Benjamin Kreifelts**, University of Tuebingen, Germany

**Yu-Han Chen**, The University of New Mexico School of Medicine, USA

**Janina Seubert**, Karolinska Institutet, Sweden

**Klaus Mathiak**, RWTH Aachen University, Germany

Our everyday life is characterized by a multitude of emotionally relevant cues that we perceive and communicate via various sensory channels. This does not only encompass the obvious cases of auditory and visual modalities, but also olfactory, gustatory, and even tactile stimuli. Any kind of emotional situation in a natural setting is usually a multimodal experience: A friend welcomes us with warm words, a smile, and a happy voice; the sight of our favourite food is accompanied by a seductive smell and a delicious taste; the thrill of watching an exciting movie scene is intensified by a gripping soundtrack. In these situations, the signals from various senses do not stand on their own; they interact and create a unified emotional experience. Recent neuroscientific research has begun to accommodate this inherent multimodality of emotions in natural situations by studying the interaction of affectively relevant information from more than one sensory channel. Fascinating new aspects emerge concerning the neurobiology of emotion processing, and there is evidence that integrating emotional cues from various sources invokes brain processes that go beyond the well-known patterns observed during unimodal stimulation.

The aim of this volume is to present novel and interesting studies dealing with the multimodality of emotions and their neural processing. This includes findings from novel paradigms beyond the classical stimulus-response pattern, fascinating new insights into the interaction of the chemical senses, new analysis methods, comprehensive reviews of selected topics, multimodality in social interactions, and clinical perspectives. Taken together, the studies of this volume thus help us to better understand the interplay of various senses in our daily emotional experiences.

# Table of Contents

- 05** ***Neural Processing of Emotion in Multimodal Settings***  
Martin Klasen, Benjamin Kreifelts, Yu-Han Chen, Janina Seubert and Klaus Mathiak
- 09** ***Dysregulated Left Inferior Parietal Activity in Schizophrenia and Depression: Functional Connectivity and Characterization***  
Veronika I. Müller, Edna C. Cieslik, Angela R. Laird, Peter T. Fox and Simon B. Eickhoff
- 20** ***Just Watching the Game Ain't Enough: Striatal fMRI Reward Responses to Successes and Failures in a Video Game During Active and Vicarious Playing***  
Jari Kätsyri, Riitta Hari, Niklas Ravaja and Lauri Nummenmaa
- 33** ***Context Counts! Social Anxiety Modulates the Processing of Fearful Faces in the Context of Chemosensory Anxiety Signals***  
Dirk Adolph, Lukas Meister and Bettina M. Pause
- 50** ***Perceiving Blocks of Emotional Pictures and Sounds: Effects on Physiological Variables***  
Anne-Marie Brouwer, Nelleke van Wouwe, Christian Mühl, Jan van Erp and Alexander Toet
- 60** ***On the Role of Crossmodal Prediction in Audiovisual Emotion Perception***  
Sarah Jessen and Sonja A. Kotz
- 67** ***Facial Reactions in Response to Dynamic Emotional Stimuli in Different Modalities in Patients Suffering From Schizophrenia: A Behavioral and EMG Study***  
Mariateresa Sestito, Maria Alessandra Umiltà, Giancarlo De Paola, Renata Fortunati, Andrea Raballo, Emanuela Leuci, Simone Maffei, Matteo Tonna, Mario Amore, Carlo Maggini and Vittorio Gallese
- 79** ***Experimental and Clinical Usefulness of Crossmodal Paradigms in Psychiatry: An Illustration From Emotional Processing in Alcohol-Dependence***  
Pierre Maurage and Salvatore Campanella
- 98** ***Food Related Processes in the Insular Cortex***  
Sabine Frank, Stephanie Kullmann and Ralf Veit
- 104** ***The Sensory Channel of Presentation Alters Subjective Ratings and Autonomic Responses Toward Disgusting Stimuli—Blood Pressure, Heart Rate and Skin Conductance in Response to Visual, Auditory, Haptic and Olfactory Presented Disgusting Stimuli***  
Ilona Croy, Kerstin Laqua, Frank Süß, Peter Joraschky, Tjalf Ziemssen and Thomas Hummel
- 114** ***Sex Differences in Chemosensation: Sensory or Emotional?***  
Kathrin Ohla and Johan N. Lundström

- 125** ***Avoidant Symptoms in PTSD Predict Fear Circuit Activation During Multimodal Fear Extinction***  
Rebecca K. Sripada, Sarah N. Garfinkel and Israel Liberzon
- 135** ***Non-Verbal Emotion Communication Training Induces Specific Changes in Brain Function and Structure***  
Benjamin Kreifelts, Heike Jacob, Carolin Brück, Michael Erb, Thomas Ethofer and Dirk Wildgruber
- 149** ***Emotional Sounds Modulate Early Neural Processing of Emotional Pictures***  
Antje B. M. Gerdes, Matthias J. Wieser, Florian Bublitzky, Anita Kusay, Michael M. Plichta and Georg W. Alpers
- 161** ***Attention and Multisensory Integration of Emotions in Schizophrenia***  
Mikhail Zvyagintsev, Carmen Parisi, Natalia Checkko, Andrey R. Nikolaev and Klaus Mathiak
- 168** ***Connecting Multimodality in Human Communication***  
Christina Regenbogen, Ute Habel and Thilo Kellermann
- 178** ***Dissociating Task Difficulty From Incongruence in Face-Voice Emotion Integration***  
Rebecca Watson, Marianne Latinus, Takao Noguchi, Oliver Garrod, Frances Crabbe and Pascal Belin
- 191** ***Multisensory Integration of Dynamic Emotional Faces and Voices: Method for Simultaneous EEG-fMRI Measurements***  
Patrick D. Schelenz, Martin Klasen, Barbara Reese, Christina Regenbogen, Dhana Wolf, Yutaka Kato and Klaus Mathiak
- 201** ***Situating Emotional Experience***  
Christine D. Wilson-Mendenhall, Lisa Feldman Barrett and Lawrence W. Barsalou
- 217** ***Neural Networks Underlying Affective States in a Multimodal Virtual Environment: Contributions to Boredom***  
Krystyna A. Mathiak, Martin Klasen, Mikhail Zvyagintsev, René Weber and Klaus Mathiak
- 224** ***Multisensory Integration Mechanisms During Aging***  
Jessica Freiherr, Johan N. Lundström, Ute Habel and Kathrin Reetz
- 230** ***Cross-Modal Integration of Emotions in the Chemical Senses***  
Moustafa Bensafi, Emilia Iannilli, Valentin A. Schriever, Johan Poncelet, Han-Seok Seo, Johannes Gerber, Catherine Rouby and Thomas Hummel
- 239** ***Subclinical Alexithymia Modulates Early Audio-Visual Perceptive and Attentional Event-Related Potentials***  
Dyna Delle-Vigne, Charles Kornreich, Paul Verbanck and Salvatore Campanella
- 253** ***Multimodal Emotion Perception After Anterior Temporal Lobectomy (ATL)***  
Valérie Milesi, Sezen Cekic, Julie Péron, Sascha Frühholz, Chiara Cristinzio, Margitta Seeck and Didier Grandjean
- 266** ***Emotional Valence and Spatial Congruency Differentially Modulate Crossmodal Processing: An fMRI Study***  
Dhana Wolf, Lisa Schock, Saurabh Bhavsar, Liliana R. Demenescu, Walter Sturm and Klaus Mathiak



# Neural processing of emotion in multimodal settings

Martin Klasen<sup>1,2\*</sup>, Benjamin Kreifelts<sup>3</sup>, Yu-Han Chen<sup>4</sup>, Janina Seubert<sup>5</sup> and Klaus Mathiak<sup>1,2</sup>

<sup>1</sup> Department of Psychiatry, Psychotherapy, and Psychosomatics, Medical School, RWTH Aachen University, Aachen, Germany

<sup>2</sup> Jülich Aachen Research Alliance–Translational Brain Medicine, RWTH Aachen University, Aachen, Germany

<sup>3</sup> Department of Psychiatry and Psychotherapy, University of Tuebingen, Tuebingen, Germany

<sup>4</sup> Department of Psychiatry, The University of New Mexico School of Medicine, Albuquerque, NM, USA

<sup>5</sup> Psychology Division, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

\*Correspondence: mklasen@ukaachen.de

## Edited and reviewed by:

John J. Foxe, Albert Einstein College of Medicine of Yeshiva University, USA

**Keywords: emotion, multisensory integration, social environment, EEG, fMRI**

Perhaps the most astonishing outcome of the Research Topic *Neural processing of emotion in multimodal settings* was the wide resonance. Not too long ago, emotions as well as multisensory integration both played outsider roles in neuroscience. However, nowadays the processing of emotional signals in the human brain has become an integrative part of basic neuroscience and clinical research. Considered a mere side effect of reasoning and thinking, the importance of emotions for human behavior has been underestimated for many years. The discovery of complex brain systems dedicated to the detection of harmful or positive situations, emotion recognition in others, and emotional experience have led to the conclusion that emotions are not at the periphery, but at the very core of human behavior. Among others, facial expressions, gestures, postures, and prosody express emotions. Thus, their integration is an essential part of face-to-face social interactions (De Gelder and Vroomen, 2000). Therefore, emotions have been described as inherently multimodal (Robins et al., 2009). This is also reflected on the psychological level, e.g., congruent bimodal emotions lead to shorter reaction times compared to faces alone (Massaro and Egan, 1996; Dolan et al., 2001).

Reflecting their evolutionary significance, emotional stimuli undergo preferred processing in the human brain (Klasen et al., 2011, 2012a). Emotion-relevant cues are delivered via multiple modalities: A picture of a beloved person evokes pleasant feelings; the furious barking of a dog signals danger; disgusting smell or taste helps to identify spoiled food. More than this, emotional cues mostly appear in combination: We recognize panic in another person by a fearful face and a frightened voice, but also by less obvious cues such as the perception of fear sweat. However, research has begun only recently to address behavioral and neural aspects of emotion integration (Klasen et al., 2012a). The aim of this volume is to fill in this gap. The studies reported here present a wide range of emotional stimuli—social and non-social—spanning the whole range of sensory modalities, from auditory and visual to touch and chemosensation.

Despite a considerable body of neuroimaging literature on emotion processing, the pathways of emotional information in the human brain are not fully understood. Considering multisensory emotions raises the additional question how these streams are integrated. Freiherr et al. (2013) provide an overview over sensory integration aspects and their development with healthy

aging. Recent neurobiological models propose multiple interactions between cortical and subcortical structures (Senkowski et al., 2008). Social emotion processing, however, is complex and involves bottom-up processes and top-down modulations. The full understanding of this complex interplay calls for methods that identify areas of emotional integration, but also show the time course and flow of information. Given the spatial proximity of unisensory and multisensory integration areas, there is a need for high resolution data in both time and space. The new technique of simultaneous EEG and fMRI recordings may adequately address this issue. Schelenz et al. (2013) present a novel source-localization driven analysis for EEG-informed fMRI. Applied to multisensory emotion paradigms, this method has the potential to map the exact cortical pathways of audiovisual signal integration.

Social emotion processing is disturbed in some clinical populations. In some psychiatric conditions this may even lie at the core of the affective symptomatology. Accordingly, a multitude of studies have addressed impairments in face processing in various psychiatric diseases such as major depression (Elliott et al., 2011), schizophrenia (Kohler et al., 2010), or alcoholism (Maurage et al., 2008). However, studies on auditory deficits are much less frequent, and multisensory emotion processing studies in clinical populations are largely missing, even though recent findings indicate that impairments in emotion integration may be equally important. This is nicely illustrated for the example of alcoholism in the review of Maurage and Campanella (2013). Complex emotional designs can also identify neural similarities between disorders. Using an audiovisual emotion paradigm, Müller et al. (2013) showed that both schizophrenia and depressive patients had a dysfunctional regulation in the same region of the angular gyrus. Even for subclinical deficits in emotion perception skills, multimodality may be the crucial factor. Delle-Vigne et al. (2014) investigated the processing of complex audiovisual stimuli in relation to alexithymia scores. Specifically for bimodal emotions, high alexithymic participants had higher amplitudes in the P100 and N100 components. This could not be observed in studies using unimodal stimulation.

A study by Zvyagintsev et al. (2013) addressed an aspect of integration which is particularly relevant for schizophrenia patients: the suppression of task-irrelevant information. Patient

ratings of visual stimuli were influenced by concurrent auditory information. This was the case for emotional and non-emotional material, indicating that modality-specific selective attention is disturbed in schizophrenia already at early sensory levels. Interestingly, healthy controls showed a similar effect solely for emotions, demonstrating an attentional capture effect across modalities. This is supported by the study of Adolph et al. (2013), showing that chemosensation interacts with visual perception. Here, the perception of sweat enhanced the allocation of attention to anxious faces. Moreover, sweat from social anxiety situations enhanced the processing of fearful facial stimuli only in socially anxious individuals—an impressive example of the integration of fear-relevant cues being influenced by personality traits. The interaction of visual emotion processing with irrelevant auditory cues was also subject of the study by Wolf et al. (2014). The authors demonstrated that visual emotion cues modulated tone processing in the auditory cortex. Thus, affective information in one sensory domain can influence even primary sensory cortex areas of another modality. Although there is overwhelming evidence for a functional specialization of sensory cortices, this contributes to the growing body of investigations suggesting that there is no cortex area which can be influenced solely by one sensory channel. Emotional content thus can trigger this cross-modal modulation. In a similar way, auditory emotional cues can enhance early cortical processing of visual stimuli. Gerdes et al. (2013) found an amplitude modulation of early visual P100 and P200 components when pictures were accompanied by emotional sounds. Emotional “crosstalk” between early auditory and visual areas thus seems to exist in both directions.

Emotional content can also modulate multisensory integration areas. Whereas matching affective information in different channels facilitates emotion recognition, non-matching information leads to emotional conflict. Watson et al. (2013) showed that audiovisual integration areas of the superior temporal cortex are sensitive to emotional congruency: Conflicting affective information enhanced activity in these sensory integration areas. Stronger cortical processing of incongruent emotional stimuli was also reported by Gerdes et al. (2013). They found enlarged P100 and P200 components for conflicting emotional information. Emotional sounds thus seem to modulate visual processing as early as 100 ms after stimulus presentation. These early interactions may be due not only to sensory integration, but also to crossmodal prediction. In real life, affective information from e.g., face and voice often do not arrive in perfect synchrony at the recipient’s eyes and ears; one modality often precedes the other one. Information from the earlier modality forms an expectation about the emotion in the other sense and modulates processing accordingly in a top-down fashion. Jessen and Kotz (2013) comprehensively review the literature on emotional crossmodal prediction and highlight its importance for stimulus integration.

Recent studies identify the amygdala and adjacent anterior temporal lobe structures as central for emotion evaluation and integration (Klasen et al., 2011; Mathiak et al., 2011). This is also highlighted in a lesion patients study by Milesi et al. (2014). Their findings confirm the role of the amygdala and anterior temporal lobe as parts of the visual system, but also show their importance for evaluating particularly positive emotional stimuli

across modalities. Moreover, these data show that a lacking ability to identify emotions in one domain can be compensated by cues from another. The same seems to be true in healthy controls when emotional information in one channel is missing. Regenbogen et al. (2013) investigated neural responses in various brain areas during video clips with emotional information in face, prosody, and speech content. If emotion from one channel was missing, input from the dorsomedial prefrontal cortex to the respective sensory cortex areas was increased, indicating a top-down modulation filling the sensory gap. The role of the amygdala for emotion processing was also highlighted in a multimodal fear conditioning study by Sripada et al. (2013). They investigated fear extinction processes in war veterans suffering from PTSD. Hyperactivation in fear-related brain circuits encompassing the amygdala during fear extinction was related to avoidance symptoms.

An important contribution to basic research with clinical perspectives is delivered by Kreifelts et al. (2013). They investigated the impact of emotion communication training on brain structure and function. Emotion-specific training modulated activity in cortical areas of face and voice processing, which shows their importance for emotion evaluation. Structural changes, however, were observed only in the fusiform face area (FFA). These findings support the notion that visual and auditory modalities support each other when emotions are categorized, but they also highlight the dominant role of vision. Visual dominance in emotion processing was also reported by Regenbogen et al. (2013). Here, the presence of facial emotions enhanced functional connectivity between the FFA and areas of the angular gyrus associated with audiovisual speech integration (Bernstein et al., 2008). Neural systems thus seem to prioritize emotional over neutral facial information. However, no such effect was observed for vocal emotions or auditory cortex. In a similar vein, Sestito et al. (2013) reported a prioritization of visual over auditory information for incongruent face-voice pairings. This was also reflected in autonomously triggered facial mimicry: Visual emotions led to stronger facial reactions than auditory ones. Peripheral physiological reactions triggered by affective signals play a decisive role in the genesis of emotional states (Brouwer et al., 2013). Accordingly, facial muscle reactions to emotional cues are reduced in schizophrenia patients who show emotion recognition impairments (Sestito et al., 2013). Taken together, visual information is more important than auditory for judging emotions; accordingly, bimodal emotional stimuli are primarily classified by their visual content (see also Klasen et al., 2011). This consistently reported prominence may in part be attributed to an unspecific visual dominance effect (Colavita, 1974); however, in the case of emotional cues, the fact that auditory signals are less reliable than visual ones may also add to the picture (Klasen et al., 2012a).

Recent evidence shows that interactions between emotional information are not limited to hearing and vision. Frank et al. (2013) discuss the multisensory integration of food-related cues in the insular cortex. Being a multimodal cortex region, the insula has been described as integrating interoceptive states with contextual information (Craig, 2009). Deviant stimulus processing in the insula has been discussed as the neural basis of various eating disorders; Frank et al. (2013) discuss the clinical implications of

this association. Being essential for the processing of food-related stimuli, the insula has been related to the processing of disgust-related stimuli from various modalities (Jabbi et al., 2008). The evolutionary significance of this function is obvious; checking if something is edible or spoiled relies on smell, taste, vision, and touch. The insula supports the integration of this information and thus seems to contribute essentially to the feeling of disgust. Accordingly, Croy et al. (2013) showed that disgust could be evoked via visual, auditory, tactile, and olfactory stimulation. Peripheral responses such as blood pressure, heart rate, or galvanic skin response, however, varied with modality.

An extraordinary, but important aspect of multisensory integration is investigated by Bensafi et al. (2013) and Ohla and Lundström (2013): the interaction between olfactory and trigeminal stimuli. These modalities are closely intertwined; in real life, there is almost no smell which does not trigger both systems. This sensory interplay is of high relevance for our perception of food and drinks. Bensafi et al. (2013) found shorter latencies of N1 and P2 responses and reduced N1 amplitudes to combined olfactory and trigeminal stimuli compared to both modalities in isolation. These findings suggest that trigeminal and olfactory cues support each other and reduce neural processing workload—in analogy to the findings from other modalities. Moreover, the authors identified the rostral anterior cingulate cortex as a binding region for olfactory and trigeminal stimuli. In a second study, Ohla and Lundström (2013) investigated gender effects in olfactory-trigeminal integration. The authors demonstrated that, despite comparable sensory sensitivity, women perceived trigeminal stimulation as more irritating than men. This was also reflected in enlarged late positive EEG components. These findings show a differential integration of olfactory and trigeminal stimulus aspects in men and women.

Recognizing that emotional experience in real life is a multisensory phenomenon leads to the conclusion that approaches using unimodal or static stimuli often lack external validity. This problem has been addressed by complex stimuli and innovative experimental designs. Another novel approach was applied by Wilson-Mendenhall et al. (2013). They employed the multisensory imagination of scenarios leading to negative emotion experience. This procedure creates actual emotional experience based on situational information and goes beyond reactive stimulus processing. Moreover, it takes into account that real-life emotional experience is not limited to some basic emotions and often goes far beyond a one-way stimulus-response pattern. Since the core function of emotions is guiding the individual's behavior via motivational processes, humans tend to actively search situations evoking positive emotions and to avoid situations associated with negative emotional outcomes. These degrees of freedom are difficult to realize in a traditional experiment. Virtual reality settings provide a promising tool for studying affective processes in multimodal environments. They are close to reality and allow the participants to individually select their actions based on rewarding values. Recent fMRI investigations have shown that video game paradigms are well suited to study the brain correlates of realistic behavior patterns using fMRI (e.g., Mathiak and Weber, 2006; Mathiak et al., 2011; Klasen et al., 2012b, 2013). Kätsyri et al. (2013) investigated responses of the brain reward system to

different types of events during free play of a multimodal violent video game using fMRI. They found that win and loss events differentially affected midbrain structures of the mesolimbic reward system; however, these effects did not predict subjective measures of emotional experience. Such insights into the neural processes underlying situational experience in video games come from the study by Mathiak et al. (2013). The authors used a combined approach integrating both game content and measures of game-induced affect. Their findings highlight the importance of cortex areas involved in self-referential emotion processing for the experience of more complex emotions in the virtual environment. Taken together, these findings indicate that reward-motivated behavior is strongly determined by striatal activity; the cognitive appraisal component which leads to perceived emotions, however, relies on cortex areas dedicated to the representation of inner states.

In summary, the investigations presented in this volume show that emotions from different senses interact at multiple levels, influence each other, and form holistic percepts, involving a variety of brain structures from unisensory cortices to high-level association areas. Importantly, they also clearly point out that emotional perception involves all human senses—not only hearing and seeing, but also touch, smell, taste, and even trigeminal signals. Moreover, they highlight the crucial necessity of taking into account the factor of multimodality when the neural processing of emotional situations is investigated.

## REFERENCES

- Adolph, D., Meister, L., and Pause, B. M. (2013). Context counts! social anxiety modulates the processing of fearful faces in the context of chemosensory anxiety signals. *Front. Hum. Neurosci.* 7:283. doi: 10.3389/fnhum.2013.00283
- Bensafi, M., Iannilli, E., Schriever, V. A., Poncelet, J., Seo, H. S., Gerber, J., et al. (2013). Cross-modal integration of emotions in the chemical senses. *Front. Hum. Neurosci.* 7:883. doi: 10.3389/fnhum.2013.00883
- Bernstein, L. E., Auer, E. T. Jr., Wagner, M., and Ponton, C. W. (2008). Spatiotemporal dynamics of audiovisual speech processing. *Neuroimage* 39, 423–435. doi: 10.1016/j.neuroimage.2007.08.035
- Brouwer, A. M., Van Wouwe, N., Muhl, C., Van Erp, J., and Toet, A. (2013). Perceiving blocks of emotional pictures and sounds: effects on physiological variables. *Front. Hum. Neurosci.* 7:295. doi: 10.3389/fnhum.2013.00295
- Colavita, F. B. (1974). Human sensory dominance. *Percept. Psychophys.* 16, 409–412. doi: 10.3758/BF03203962
- Craig, A. D. (2009). How do you feel—now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70. doi: 10.1038/nrn2555
- Croy, I., Laqua, K., Suss, F., Joraschky, P., Ziemssen, T., and Hummel, T. (2013). The sensory channel of presentation alters subjective ratings and autonomic responses toward disgusting stimuli—Blood pressure, heart rate and skin conductance in response to visual, auditory, haptic and olfactory presented disgusting stimuli. *Front. Hum. Neurosci.* 7:510. doi: 10.3389/fnhum.2013.00510
- De Gelder, B., and Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cogn. Emot.* 14, 289–311. doi: 10.1080/026999300378824
- Delle-Vigne, D., Kornreich, C., Verbanck, P., and Campanella, S. (2014). Subclinical alexithymia modulates early audio-visual perceptual and attentional event-related potentials. *Front. Hum. Neurosci.* 8:106. doi: 10.3389/fnhum.2014.00106
- Dolan, R. J., Morris, J. S., and De Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proc. Natl. Acad. Sci. U.S.A.* 98, 10006–10010. doi: 10.1073/pnas.171288598
- Elliott, R., Zahn, R., Deakin, J. F., and Anderson, I. M. (2011). Affective cognition and its disruption in mood disorders. *Neuropsychopharmacology* 36, 153–182. doi: 10.1038/npp.2010.77
- Frank, S., Kullmann, S., and Veit, R. (2013). Food related processes in the insular cortex. *Front. Hum. Neurosci.* 7:499. doi: 10.3389/fnhum.2013.00499

- Freiherr, J., Lundström, J. N., Habel, U., and Reetz, K. (2013). Multisensory integration mechanisms during aging. *Front. Hum. Neurosci.* 7:863. doi: 10.3389/fnhum.2013.00863
- Gerdes, A. B., Wieser, M. J., Bublatzky, F., Kusay, A., Plichta, M. M., and Alpers, G. W. (2013). Emotional sounds modulate early neural processing of emotional pictures. *Front. Psychol.* 4:741. doi: 10.3389/fpsyg.2013.00741
- Jabbi, M., Bastiaansen, J., and Keysers, C. (2008). A common anterior insula representation of disgust observation, experience and imagination shows divergent functional connectivity pathways. *PLoS ONE* 3:e2939. doi: 10.1371/journal.pone.0002939
- Jessen, S., and Kotz, S. A. (2013). On the role of crossmodal prediction in audiovisual emotion perception. *Front. Hum. Neurosci.* 7:369. doi: 10.3389/fnhum.2013.00369
- Kätsyri, J., Hari, R., Ravaja, N., and Nummenmaa, L. (2013). Just watching the game ain't enough: striatal fMRI reward responses to successes and failures in a video game during active and vicarious playing. *Front. Hum. Neurosci.* 7:278. doi: 10.3389/fnhum.2013.00278
- Klasen, M., Chen, Y. H., and Mathiak, K. (2012a). Multisensory emotions: perception, combination and underlying neural processes. *Rev. Neurosci.* 23, 381–392. doi: 10.1515/revneuro-2012-0040
- Klasen, M., Kenworthy, C. A., Mathiak, K. A., Kircher, T. T., and Mathiak, K. (2011). Supramodal representation of emotions. *J. Neurosci.* 31, 13635–13643. doi: 10.1523/JNEUROSCI.2833-11.2011
- Klasen, M., Weber, R., Kircher, T. T., Mathiak, K. A., and Mathiak, K. (2012b). Neural contributions to flow experience during video game playing. *Soc. Cogn. Affect. Neurosci.* 7, 485–495. doi: 10.1093/scan/nsr021
- Klasen, M., Zvyagintsev, M., Schwenzer, M., Mathiak, K. A., Sarkheil, P., Weber, R., et al. (2013). Quetiapine modulates functional connectivity in brain aggression networks. *Neuroimage* 75, 20–26. doi: 10.1016/j.neuroimage.2013.02.053
- Kohler, C. G., Walker, J. B., Martin, E. A., Healey, K. M., and Moberg, P. J. (2010). Facial emotion perception in schizophrenia: a meta-analytic review. *Schizophr. Bull.* 36, 1009–1019. doi: 10.1093/schbul/sbn192
- Kreifelts, B., Jacob, H., Bruck, C., Erb, M., Ethofer, T., and Wildgruber, D. (2013). Non-verbal emotion communication training induces specific changes in brain function and structure. *Front. Hum. Neurosci.* 7:648. doi: 10.3389/fnhum.2013.00648
- Massaro, D. W., and Egan, P. B. (1996). Perceiving affect from the voice and the face. *Psychon. Bull. Rev.* 3, 215–221. doi: 10.3758/BF03212421
- Mathiak, K. A., Klasen, M., Weber, R., Ackermann, H., Shergill, S. S., and Mathiak, K. (2011). Reward system and temporal pole contributions to affective evaluation during a first person shooter video game. *BMC Neurosci.* 12:66. doi: 10.1186/1471-2202-12-66
- Mathiak, K. A., Klasen, M., Zvyagintsev, M., Weber, R., and Mathiak, K. (2013). Neural networks underlying affective states in a multimodal virtual environment: contributions to boredom. *Front. Hum. Neurosci.* 7:820. doi: 10.3389/fnhum.2013.00820
- Mathiak, K., and Weber, R. (2006). Toward brain correlates of natural behavior: fMRI during violent video games. *Hum. Brain Mapp.* 27, 948–956. doi: 10.1002/hbm.20234
- Maurage, P., and Campanella, S. (2013). Experimental and clinical usefulness of crossmodal paradigms in psychiatry: an illustration from emotional processing in alcohol-dependence. *Front. Hum. Neurosci.* 7:394. doi: 10.3389/fnhum.2013.00394
- Maurage, P., Campanella, S., Philippot, P., Martin, S., and De Timary, P. (2008). Face processing in chronic alcoholism: a specific deficit for emotional features. *Alcohol. Clin. Exp. Res.* 32, 600–606. doi: 10.1111/j.1530-0277.2007.00611.x
- Milesi, V., Cekic, S., Peron, J., Fruhholz, S., Cristinzio, C., Seeck, M., et al. (2014). Multimodal emotion perception after anterior temporal lobectomy (ATL). *Front. Hum. Neurosci.* 8:275. doi: 10.3389/fnhum.2014.00275
- Müller, V. I., Cieslik, E. C., Laird, A. R., Fox, P. T., and Eickhoff, S. B. (2013). Dysregulated left inferior parietal activity in schizophrenia and depression: functional connectivity and characterization. *Front. Hum. Neurosci.* 7:268. doi: 10.3389/fnhum.2013.00268
- Ohla, K., and Lundström, J. N. (2013). Sex differences in chemosensation: sensory or emotional? *Front. Hum. Neurosci.* 7:607. doi: 10.3389/fnhum.2013.00607
- Regenbogen, C., Habel, U., and Kellermann, T. (2013). Connecting multimodality in human communication. *Front. Hum. Neurosci.* 7:754. doi: 10.3389/fnhum.2013.00754
- Robins, D. L., Hunyadi, E., and Schultz, R. T. (2009). Superior temporal activation in response to dynamic audio-visual emotional cues. *Brain Cogn.* 69, 269–278. doi: 10.1016/j.bandc.2008.08.007
- Schelenz, P. D., Klasen, M., Reese, B., Regenbogen, C., Wolf, D., Kato, Y., et al. (2013). Multisensory integration of dynamic emotional faces and voices: method for simultaneous EEG-fMRI measurements. *Front. Hum. Neurosci.* 7:729. doi: 10.3389/fnhum.2013.00729
- Senkowski, D., Schneider, T. R., Foxe, J. J., and Engel, A. K. (2008). Crossmodal binding through neural coherence: implications for multisensory processing. *Trends Neurosci.* 31, 401–409. doi: 10.1016/j.tins.2008.05.002
- Sestito, M., Umiltà, M. A., De Paola, G., Fortunati, R., Raballo, A., Leuci, E., et al. (2013). Facial reactions in response to dynamic emotional stimuli in different modalities in patients suffering from schizophrenia: a behavioral and EMG study. *Front. Hum. Neurosci.* 7:368. doi: 10.3389/fnhum.2013.00368
- Sripada, R. K., Garfinkel, S. N., and Liberzon, I. (2013). Avoidant symptoms in PTSD predict fear circuit activation during multimodal fear extinction. *Front. Hum. Neurosci.* 7:672. doi: 10.3389/fnhum.2013.00672
- Watson, R., Latinus, M., Noguchi, T., Garrod, O., Crabbe, F., and Belin, P. (2013). Dissociating task difficulty from incongruence in face-voice emotion integration. *Front. Hum. Neurosci.* 7:744. doi: 10.3389/fnhum.2013.00744
- Wilson-Mendenhall, C. D., Barrett, L. F., and Barsalou, L. W. (2013). Situating emotional experience. *Front. Hum. Neurosci.* 7:764. doi: 10.3389/fnhum.2013.00764
- Wolf, D., Schock, L., Bhavsar, S., Demenescu, L. R., Sturm, W., and Mathiak, K. (2014). Emotional valence and spatial congruency differentially modulate crossmodal processing: an fMRI study. *Front. Hum. Neurosci.* 8:659. doi: 10.3389/fnhum.2014.00659
- Zvyagintsev, M., Parisi, C., Chechko, N., Nikolaev, A. R., and Mathiak, K. (2013). Attention and multisensory integration of emotions in schizophrenia. *Front. Hum. Neurosci.* 7:674. doi: 10.3389/fnhum.2013.00674

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 September 2014; accepted: 26 September 2014; published online: 21 October 2014.

Citation: Klasen M, Kreifelts B, Chen Y-H, Seubert J and Mathiak K (2014) Neural processing of emotion in multimodal settings. *Front. Hum. Neurosci.* 8:822. doi: 10.3389/fnhum.2014.00822

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2014 Klasen, Kreifelts, Chen, Seubert and Mathiak. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Dysregulated left inferior parietal activity in schizophrenia and depression: functional connectivity and characterization

Veronika I. Müller<sup>1,2,3\*</sup>, Edna C. Cieslik<sup>1,2,3</sup>, Angela R. Laird<sup>4,5</sup>, Peter T. Fox<sup>4,6</sup> and Simon B. Eickhoff<sup>1,2,3</sup>

<sup>1</sup> Institute of Clinical Neuroscience and Medical Psychology, Heinrich Heine University, Düsseldorf, Germany

<sup>2</sup> Department of Neuroscience and Medicine, Research Center Jülich, INM-1, Jülich, Germany

<sup>3</sup> Department of Psychiatry, Psychotherapy, and Psychosomatics, Medical School, RWTH Aachen University, Aachen, Germany

<sup>4</sup> Research Imaging Institute, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA

<sup>5</sup> Department of Physics, Florida International University, Miami, FL, USA

<sup>6</sup> South Texas Veterans Administration Medical Center, San Antonio, TX, USA

## Edited by:

Martin Klasen, RWTH Aachen University, Germany

## Reviewed by:

Jennifer L. Robinson, Auburn University, USA  
Corrado Corradi-Dell'Acqua, University of Geneva, Switzerland  
Julia Sacher, University Leipzig, Germany

## \*Correspondence:

Veronika I. Müller, Department of Neuroscience and Medicine, Research Center Jülich, INM-1, Leo-Brandt-Straße, D-52428 Jülich, Germany  
e-mail: v.mueller@fz-juelich.de

The inferior parietal cortex (IPC) is a heterogeneous region that is known to be involved in a multitude of diverse different tasks and processes, though its contribution to these often-complex functions is yet poorly understood. In a previous study we demonstrated that patients with depression failed to deactivate the left IPC during processing of congruent audiovisual information. We now found the same dysregulation (same region and condition) in schizophrenia. By using task-independent (resting state) and task-dependent meta-analytic connectivity modeling (MACM) analyses we aimed at characterizing this particular region with regard to its connectivity and function. Across both approaches, results revealed functional connectivity of the left inferior parietal seed region with bilateral IPC, precuneus and posterior cingulate cortex (PrC/PCC), medial orbitofrontal cortex (mOFC), left middle frontal (MFG) as well as inferior frontal (IFG) gyrus. Network-level functional characterization further revealed that on the one hand, all interconnected regions are part of a network involved in memory processes. On the other hand, sub-networks are formed when emotion, language, social cognition and reasoning processes are required. Thus, the IPC-region that is dysregulated in both depression and schizophrenia is functionally connected to a network of regions which, depending on task demands may form sub-networks. These results therefore indicate that dysregulation of left IPC in depression and schizophrenia might not only be connected to deficits in audiovisual integration, but is possibly also associated to impaired memory and deficits in emotion processing in these patient groups.

**Keywords:** functional connectivity, depression, schizophrenia, inferior parietal cortex, resting-state

## INTRODUCTION

Depression and schizophrenia are both associated with social and affective dysfunctions as well as deficits in emotional processing (Bach et al., 2009; Bourke et al., 2010; Kohler et al., 2010; Comparelli et al., 2013). Research on affective deficits in psychiatric populations to date, however, has mainly focused on unimodal emotion processing, while in daily life emotion perception is generally based on the multimodal evaluation of information, such as hearing a laugh and seeing a smiling face. Importantly, in this context, emotional information from different sensory channels can be either congruent or incongruent, leading to faster responses when processing emotional congruent compared to incongruent information (De Gelder and Vroomen, 2000; Dolan et al., 2001; Collignon et al., 2008). Clinical studies have shown that patients with schizophrenia show aberrant audiovisual integration (De Gelder et al., 2005; De Jong et al., 2009; Van Den Stock et al., 2011). In contrast, Müller et al. (2012) and Müller et al. (2013) could not find any significant group difference in the behavioral rating of emotional faces while distracted by congruent and incongruent sounds, neither in a group of patients

with schizophrenia, nor in depression. Importantly, the neuronal correlates of crossmodal emotional processing in clinical populations are still sparse. In a recent study (Müller et al., 2013) investigating audiovisual (in) congruence processing in depression, we showed that compared to healthy controls, patients failed to deactivate the left inferior parietal cortex (IPC) and inferior frontal cortex when confronted with congruent happy audiovisual information, while there was no difference between groups for incongruent pairs. As we will show in this paper by using the same paradigm, a similar effect in the same region of the left IPC in patients with schizophrenia can be observed. In particular, schizophrenic patients also reveal decreased deactivation in congruent audiovisual conditions compared to controls. Thus, both schizophrenia and depression go along with IPC dysregulation during congruent audiovisual emotional processing, possibly indicating increased processing of unambiguous stimuli in these patient groups (Müller et al., 2013).

However, when interpreting these findings, one has to acknowledge that the IPC is a heterogeneous region, which is involved in a wide range of different functions ranging from

language and memory to action planning, higher social-cognition and other integrative processes (Glover, 2004; Wagner et al., 2005; Daselaar et al., 2006; Buckner et al., 2008; Binder et al., 2009; Spreng et al., 2009; Caspers et al., 2010; Arsalidou and Taylor, 2011; Bzdok et al., 2012; Schilbach et al., 2012; Seghier, 2013). Hence, it has been suggested that the IPC can be subdivided in different sub regions. Based on cytoarchitectonic mapping, Caspers et al. (2006, 2008) divided the IPC into seven different sub-regions. With regard to these divisions, the area that we found to be dysregulated in depression as well as in schizophrenia strongly overlaps with area PGp. Based on its anatomical and functional connections with temporal and lateral occipital (Caspers et al., 2011) as well as with frontal and parahippocampal areas (Uddin et al., 2010) it has been argued that PGp may mainly be involved in auditory-sensory integration and memory processes. Given the size and potential heterogeneity of this region in the posterior IPC, however, it remains open as to how the specific location disturbed in depression and schizophrenia relates to these roles. Furthermore, as the function of a specific brain area depends on those regions it interacts with, its role should not only be assessed in isolation but also together with regions it stands in interplay with (Stephan, 2004; Seghier, 2013). Therefore, the present study aims to investigate connectivity and function of the particular region, which has been found to be dysregulated in cross-modal affective integration across two different clinical groups, i.e., depression and schizophrenia. In particular, task-dependent and task-independent functional connectivity as well as behavioral characterization of the region of interest was carried out in healthy subjects in order to gain better insight of the role of this region from a system perspective.

## METHODS

### VOLUME OF INTEREST

The volume of interest used in the current study is based on two fMRI studies investigating the neural correlates of audiovisual incongruence processing in patients with depression as well as patients with schizophrenia. Before, describing the methods of the current study, we first describe the patient samples on which the VOI is based on as well as the experimental procedure of the audiovisual paradigm.

### Audiovisual paradigm and fMRI analysis

The stimuli and procedure used in the fMRI studies is the same as previously described (Müller et al., 2011, 2013). Thirty different pictures of faces from five males and five females, each showing a happy, neutral, and fearful expression (FEBA; Gur et al., 2002) were combined with 30 auditory stimuli, consisting of 10 yawns, 10 laughs, and 10 screams. Additionally 10 different blurred faces served as masks. This resulted in 180 stimulus pairs with 9 different conditions (fearful/scream, fearful/yawn, fearful/laugh, neutral/scream, neutral/yawn, neutral/laugh, happy/scream, happy/yawn, happy/laugh). Every trial started with the presentation of a sound concurrently with a mask. After 1000 ms the mask was displaced by an either neutral or emotional face and presented with the continuing sound for another 500 ms. Subjects had to ignore the sound and to just

rate the facial expression on an eight-point scale from extremely fearful to extremely happy.

fMRI data acquisition and statistical analysis was done as described in Müller et al. (2013). Images were acquired on a Siemens Trio 3T whole-body scanner (Erlangen, Germany) in the RWTH Aachen University hospital using blood-oxygen-level-dependent (BOLD) contrast [Gradient-echo EPI pulse sequence,  $TR = 2.2$  s, in plane resolution =  $3.1 \times 3.1$  mm, 36 axial slices (3.1 mm thickness)] covering the entire brain. Echo-planar imaging (EPI) images were corrected for head movement, normalized to the Montreal Neurological Institute (MNI) single subject template and spatially smoothed using an 8 mm FWHM Gaussian kernel. Data were then analysed using a General Linear Model as implemented in SPM8. For each subject, each experimental condition as well as the response event were separately modeled and simple main effects for each of the conditions computed by applying appropriate baseline contrasts. These individual first-level contrasts were then entered into a second-level group-analysis using an analysis of variance (ANOVA) employing a random-effects model. Based on these estimates of the second-level analysis, separate  $t$ -contrasts were calculated for the interaction congruence  $\times$  group by applying the respective contrast to the 2nd level parameter estimates. The resulting SPM (T) maps were then thresholded at a cluster level FWE rate of  $p < 0.05$  (cluster forming threshold:  $p < 0.001$  at voxel level).

### Subjects

The demographic and clinical characteristics of the patients with depression and the corresponding controls can be found in Müller et al. (2013).

Using the same audiovisual paradigm as in the previous study, we now tested a sample of 18 patients with schizophrenia and 18 healthy controls matched for age, gender, and education. Two patients and the corresponding controls were excluded from further analysis due to abnormal anatomy or an inability to understand the task. All participants were right handed, as confirmed by the Edinburgh Inventory (Oldfield, 1971) and reported normal or corrected-to-normal vision. **Table 1** presents the clinical profile of the schizophrenic patient group. Patients were recruited from the inpatient and outpatient units of the Department of Psychiatry, Psychotherapy and Psychosomatics, RWTH University Hospital. Of the 16 patients included in the final analysis, 14 met ICD-10 criteria for paranoid schizophrenia (F 20.0), whereas two were

**Table 1 | Demographic and clinical characteristics of the patient and control groups.**

	Patients	Controls
Gender	6 female/ 10 male	6 female/ 10 male
Age	38.37 $\pm$ 10.21	37.94 $\pm$ 10.38
Education in years	11.75 $\pm$ 2.29	12.93 $\pm$ 2.20
Age of onset	29.19 $\pm$ 8.65	
Duration of illness	9.25 $\pm$ 6.59	
PANSS positive	10.63 $\pm$ 3.83	
PANSS negative	12.75 $\pm$ 5.59	
PANSS general	22.31 $\pm$ 6.33	

diagnosed with the residual subtype (F 20.4). Only patients with no comorbid psychiatric or neurological illness and no substance addiction in the last 6 months were included in the study. All patients were medicated, in particular all of them were treated with atypical antipsychotics with one additionally taking typical antipsychotic medication. Furthermore, five patients were taking antidepressant agents and one was taking anticholinergic drugs.

Healthy controls had no history of neurological or psychiatric disorder and did not take any mood- or cognition-altering medication.

All subjects gave informed consent into the study which was approved by the ethics committee of the School of Medicine of the RWTH Aachen University.

### Definition of volume of interest

In the previous study in patients with depression (Müller et al., 2013), we found a significant incongruence  $\times$  group interaction in the left IPC. In particular, patients failed to deactivate this region in the congruent happy audiovisual condition. **Figure 1** (in blue) presents the dysregulated IPC region found in that study in depression (Müller et al., 2013). We now tested the same interaction in the schizophrenic sample, by using the same paradigm as well as applying the same methods and statistical analysis. When calculating the interaction contrast incongruence  $\times$  group, a significant effect was again found in the left IPC (**Figure 1**, red). Similar as in depression, this interaction was driven by a failure to deactivate this region in congruent conditions.

The current study now aims to investigate functional connectivity and characterization of that particular area which has been found to be dysregulated in audiovisual congruence processing in both depression and schizophrenia. For that, a conjunction of the interaction contrast of both groups was performed. The

resulting overlap then served as seed region (**Figure 1**, purple) for the calculation of functional connectivity.

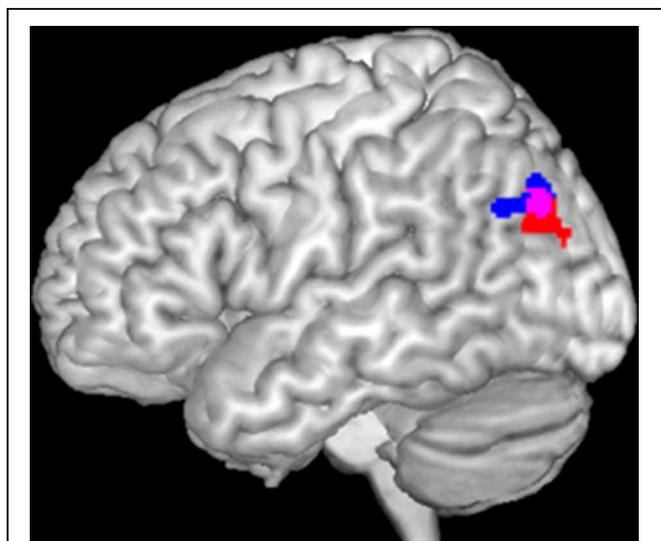
### TASK-DEPENDENT FUNCTIONAL CONNECTIVITY: META-ANALYTIC CONNECTIVITY MODELING (MACM)

To characterize the co-activation profile of the left inferior parietal seed region (**Figure 1**, purple), we used meta-analytic connectivity modeling (MACM). This approach to functional connectivity assesses which brain regions are co-activated above chance with a particular seed region across a large number of functional neuroimaging experiments. MACM thus takes advantage of the fact that functional imaging studies are normally presented in a highly standardized format using ubiquitously employed standard coordinate systems, and the emergence of large-scale databases that store this information, such as BrainMap (Laird et al., 2009a, 2011) or Neurosynth (Yarkoni et al., 2011). The first step in a MACM analysis is to identify all experiments in a database that activate the seed region. Subsequently, quantitative meta-analysis is employed to test for convergence across the foci reported in these experiments. Significant convergence of reported foci in other brain regions therefore, indicates consistent co-activation, i.e., functional connectivity with the seed (Laird et al., 2009b; Eickhoff et al., 2010; Robinson et al., 2010). Thus, we first identified all experiments in the BrainMap database ([www.brainmap.org](http://www.brainmap.org)) that featured at least one focus of activation in the seed region. Only studies reporting group analyses of functional mapping experiments of healthy subjects as well as activation only studies were included, while studies dealing with disease or drug effects were excluded. This resulted in inclusion of 160 experiments with a total of 2454 subjects and 2335 foci. Next, coordinate-based meta-analysis was performed in order to identify consistent co-activations across experiments by using the revised Activation Likelihood Estimation (ALE) algorithm (Eickhoff et al., 2009, 2012). This algorithm aims to identify areas showing a convergence of reported coordinates across experiments, which is higher than expected under a random spatial association. The results were thresholded at a cluster-level FWE corrected threshold of  $p < 0.05$  (cluster-forming threshold at voxel-level  $p < 0.001$ ).

### TASK-INDEPENDENT FUNCTIONAL CONNECTIVITY: RESTING-STATE

Resting state images were obtained from the Nathan Kline Institute “Rockland” sample, which are available online as part of the International Neuroimaging Datasharing Initiative ([http://fcon\\_1000.projects.nitrc.org/indi/pro/nki.html](http://fcon_1000.projects.nitrc.org/indi/pro/nki.html)). In total, the processed sample consisted of 132 healthy subjects between 18 and 85 years (mean age:  $42.3 \pm 18.08$  years; 78 male, 54 female) with 260 EPI images per subject. Images were acquired on a Siemens TrioTim 3T scanner using BOLD contrast [gradient-echo EPI pulse sequence, repetition time ( $TR$ ) = 2.5 s, echo time ( $TE$ ) = 30 ms, flip angle =  $80^\circ$ , in-plane resolution =  $3.0 \times 3.0$  mm, 38 axial slices (3.0 mm thickness), covering the entire brain].

Data was processed using SPM8 (Wellcome Trust Centre for Neuroimaging, London, <http://www.fil.ion.ucl.ac.uk/spm/software/spm8/>). Prior to further analyses, the first four scans were discarded and EPI images were then corrected for head



**FIGURE 1 | Significant interaction between incongruence  $\times$  group in left inferior parietal cortex in depression (blue) and schizophrenia (red). The overlap of both activations served as seed area (purple) for functional connectivity calculation.**

movement by affine registration using a two-pass procedure in which in a first step, images were aligned to the initial volumes and then subsequently to the mean of all volumes. Next, for every subject, the mean EPI image was spatially normalized to the MNI single-subject template (Holmes et al., 1998) by using the “unified segmentation” approach (Ashburner and Friston, 2005). Ensuing deformation was then applied to the individual EPI volumes and images smoothed by a 5-mm full width at half maximum Gaussian kernel to improve signal-to-noise ratio and to compensate for residual anatomical variations. Time-series of each voxel were processed as follows (Reetz et al., 2012; Sommer et al., 2012; Zu Eulenburg et al., 2012): Spurious correlations were reduced by excluding variance, which can be explained by the following nuisance variables: (1) the six motion parameters derived from image realignment; (2) their first derivatives; (3) mean GM, WM, and CBF intensity. All nuisance variables entered the model as first and also as second order terms (cf. Satterthwaite et al., 2013 for an evaluation of this framework). Finally, data was band-pass filtered with the cut-off frequencies of 0.01 and 0.08 Hz. Just as for the MACM analysis, the seed region was provided by the conjunction of the incongruence  $\times$  group interaction of the schizophrenia and depression study (**Figure 1**, purple). Time-courses of all voxels within that seed were then extracted and expressed as the first eigenvariate. To quantify resting-state functional connectivity linear (Pearson) correlation coefficients were computed between the ensuing characteristic time series of the seed region and the time-series of all other gray matter voxels of the brain. These voxel-wise correlation coefficients were then transformed into Fisher’s Z-scores and then fed into a second-level ANOVA including an appropriate non-sphericity correction as implemented in SPM8. Results were again thresholded at a cluster-level FWE corrected threshold of  $p < 0.05$  (cluster-forming threshold at voxel-level  $p < 0.001$ ). That is, the same criteria was used for both MACM and resting state analyses.

### CONJUNCTION BETWEEN MACM AND RESTING-STATE

We here aimed to identify regions that show functional connectivity with the seed across different mental states. Therefore, a conjunction analysis between MACM and resting state results was performed in order to detect areas showing both, task-dependent and task-independent functional connectivity with the seed region (extent threshold of 20 voxels). That is, by using the minimum statistics (Nichols et al., 2005; Jakobs et al., 2012) and computing the intersection of the thresholded connectivity maps derived from two different concepts of functional connectivity, we aimed to delineate consistent functional connectivity of the seed.

Finally, all results were anatomically labeled by using the SPM Anatomy Toolbox (Eickhoff et al., 2005, 2006, 2007).

### FUNCTIONAL CHARACTERIZATIONS OF DERIVED CO-ACTIVATIONS

Based on the conjunction of MACM and Resting-State, ensuing regions showing consensus of functional connectivity with the seed were further investigated. In particular, we were mainly interested in the functional role of the left IPC in co-activation with its connected regions. That is, we assessed functional properties (Laird et al., 2009b; Cieslik et al., 2012; Rottschy et al., 2012) for the IPC activation combined with all other regions

derived from the task-dependent and task-independent conjunction. Therefore, functional characterization of the derived network was performed by using the behavioral domain (BD) and paradigm class (PC) meta-data categories from the BrainMap database (Laird et al., 2009a, 2011; Turner and Laird, 2012), describing the classes of mental processes isolated by the archived experiments’ statistical contrasts. BDs include the main categories cognition, action, perception, emotion, interoception as well as their subcategories. In contrast, PCs classify the specific task employed (see <http://brainmap.org/subscribe/> for the complete list of BD and PC). For the functional characterization of the different networks, we proceeded as follows: First, we identified all experiments in the BrainMap database, which featured at least one focus of activation within the IPC and its connected region(s). That is, we identified all experiments activating the left IPC and simultaneously those regions it is connected with. Forward inference and reverse inference were calculated for each network in order to characterize the functional profiles of the respective networks. While forward inference on the functional characterization is based on the probability of observing activity in a brain region (or network) given knowledge of a psychological process, reverse inference tests the probability of a psychological process being present given knowledge of activation in a particular brain region (or network). In particular in the forward inference approach, we determined a network’s functional profile by identifying taxonomic labels for which the probability of finding activation in the respective network is significantly higher than the overall chance (across the entire database) of finding activation in that particular network. Significance was established using a binomial test ( $p < 0.001$ ). In the reverse inference approach, a network’s functional profile was determined by identifying the most likely BDs and paradigm classes given activation in a particular network. Significance was assessed by means of a chi-square test ( $p < 0.001$ ).

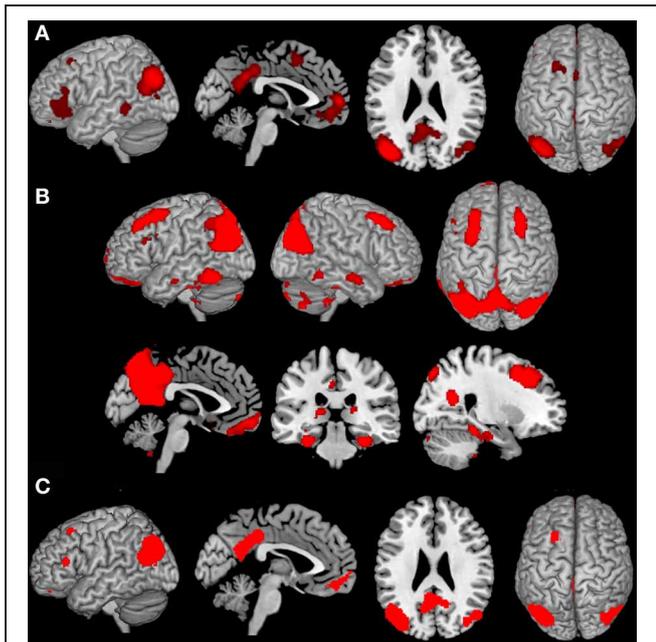
## RESULTS

### TASK-DEPENDENT FUNCTIONAL CONNECTIVITY

Analysis of task-based functional connectivity by MACM revealed significant co-activation of left IPC with its surrounding parietal areas (overlapping PGp and PGa, Caspers et al., 2006, 2008) extending into middle temporal and middle occipital gyri, with its right homologue (overlapping PGp, Caspers et al., 2006, 2008) also extending into middle temporal and occipital gyri, as well as with precuneus and posterior cingulate cortex (PrC/PCC), medial orbitofrontal cortex (mOFC), left inferior frontal gyrus (IFG) (overlapping area 45, Amunts et al., 1999), left middle/superior frontal gyrus and left middle temporal gyrus (**Figure 2A**).

### TASK-INDEPENDENT FUNCTIONAL CONNECTIVITY

Resting state connectivity of the seed region revealed significant connectivity with a broad network including bilateral inferior (overlapping PGp and PGa, Caspers et al., 2006, 2008) and superior parietal cortex (overlapping 7A, Scheperjans et al., 2008) extending into lateral occipital and temporal gyrus, PrC/PCC. Furthermore, functional connectivity was found with bilateral cerebellum (overlapping Lobule VI, IX, VIIa Crus I, VIIa Crus II, Diedrichsen et al., 2009), hippocampus (overlapping CA, SUB,



**FIGURE 2 | Results of the task-dependent (A), task-independent (B) analysis as well as the conjunction across both approaches (C). (A)** Task-dependent functional connectivity of the seed region with bilateral parietal cortex extending into middle temporal and middle occipital gyri, with posterior cingulate cortex and precuneus, medial orbitofrontal cortex, left inferior frontal gyrus, left middle/superior frontal gyrus and left middle temporal gyrus. **(B)** Task-independent functional connectivity of the seed with bilateral parietal cortex extending into lateral occipital and temporal gyrus, precuneus and posterior cingulate cortex, as well as with bilateral cerebellum, hippocampus, parahippocampal gyrus, thalamus, fusiform gyrus, inferior and middle temporal gyrus, middle frontal gyrus, left inferior frontal gyrus, and medial orbitofrontal cortex. **(C)** Conjunction of task-dependent and task-independent connectivity reveals significant functional connectivity of the seed area with bilateral inferior parietal cortex, precuneus and posterior cingulate cortex, left middle and inferior frontal gyrus and medial orbitofrontal cortex.

and EC, Amunts et al., 2005), parahippocampal gyrus, thalamus, fusiform gyrus, inferior and middle temporal gyrus, middle frontal gyrus (MFG), left IFG (overlapping area 44 and 45, Amunts et al., 1999), as well as mOFC (**Figure 2B**).

### CONJUNCTION OF TASK-DEPENDENT AND TASK-INDEPENDENT FUNCTIONAL CONNECTIVITY

The main interest of the present study was to characterize functional connectivity of the seed region that can be observed in both, task-dependent (MACM) and task-free (resting state), connectivity approaches. Areas consistently observed in both functional connectivity analyses consisted of all areas found in MACM except left middle temporal gyrus. That is, convergence between resting state and MACM connectivity of our left parietal seed could be found with its surrounding parietal (overlapping PGp and PGa, Caspers et al., 2006, 2008), middle occipital and temporal regions, its right homologue (overlapping PGp and PGa, Caspers et al., 2006, 2008), the PrC/PCC, left MFG extending into superior frontal sulcus, as well as IFG (overlapping area 45, Amunts et al., 1999) and mOFC (**Figure 2C**).

### FUNCTIONAL CHARACTERIZATION OF DERIVED NETWORKS

In a next step, forward and reverse inference were calculated for the whole network (IPC-PrC/PCC-mOFC-MFG-IFG) as well as each sub-network (left IPC co-activation with every combination of the regions of the derived network).

BDs that were overrepresented among experiments co-activating with the left IPC and the mOFC were explicit memory, social cognition and emotion. Furthermore, co-activation of those two regions was associated with paradigm classes related to episodic recall, subjective emotional picture discrimination, reward, imagination of objects/scenes and face monitoring. The same BDs and paradigm classes were revealed by reverse inference, except face monitoring, which revealed significance for forward inference only (**Figure 3A**).

Analysis of BDs overrepresented among experiments that co-activate left IPC and the PrC/PCC area revealed significant meta-data labels related to cognition, in particular with its subcategories explicit memory, social cognition and language. Moreover, there was a significant association with paradigm classes referring to theory of mind, explicit recognition, and subjective emotional picture discrimination. Reverse inference again revealed the same BDs and paradigm classes, except with imagined objects/scenes representing an additional paradigm class (**Figure 3B**).

In contrast, significant BDs overrepresented among experiments that co-activate left IPC and left IFG were also related to cognition, but with the subcategories reasoning, semantic language processing as well as explicit memory. Paradigm classes significantly associated with co-activation of IPC and IFG were overt and covert word generation and the Wisconsin card sorting test (**Figure 3C**).

Analysis of IPC and left MFG revealed above chance co-activation of those region in experiments related to explicit memory for both forward and reverse inference, but no significant specific paradigm class (**Figure 3D**).

Co-activation of IPC, mOFC, and PrC/PCC featured a significant overrepresentation of experiments related to explicit memory and social cognition. Analysis of the paradigm classes further revealed association with tasks involving episodic recall, imagined objects/scenes and face monitoring (**Figure 3E**).

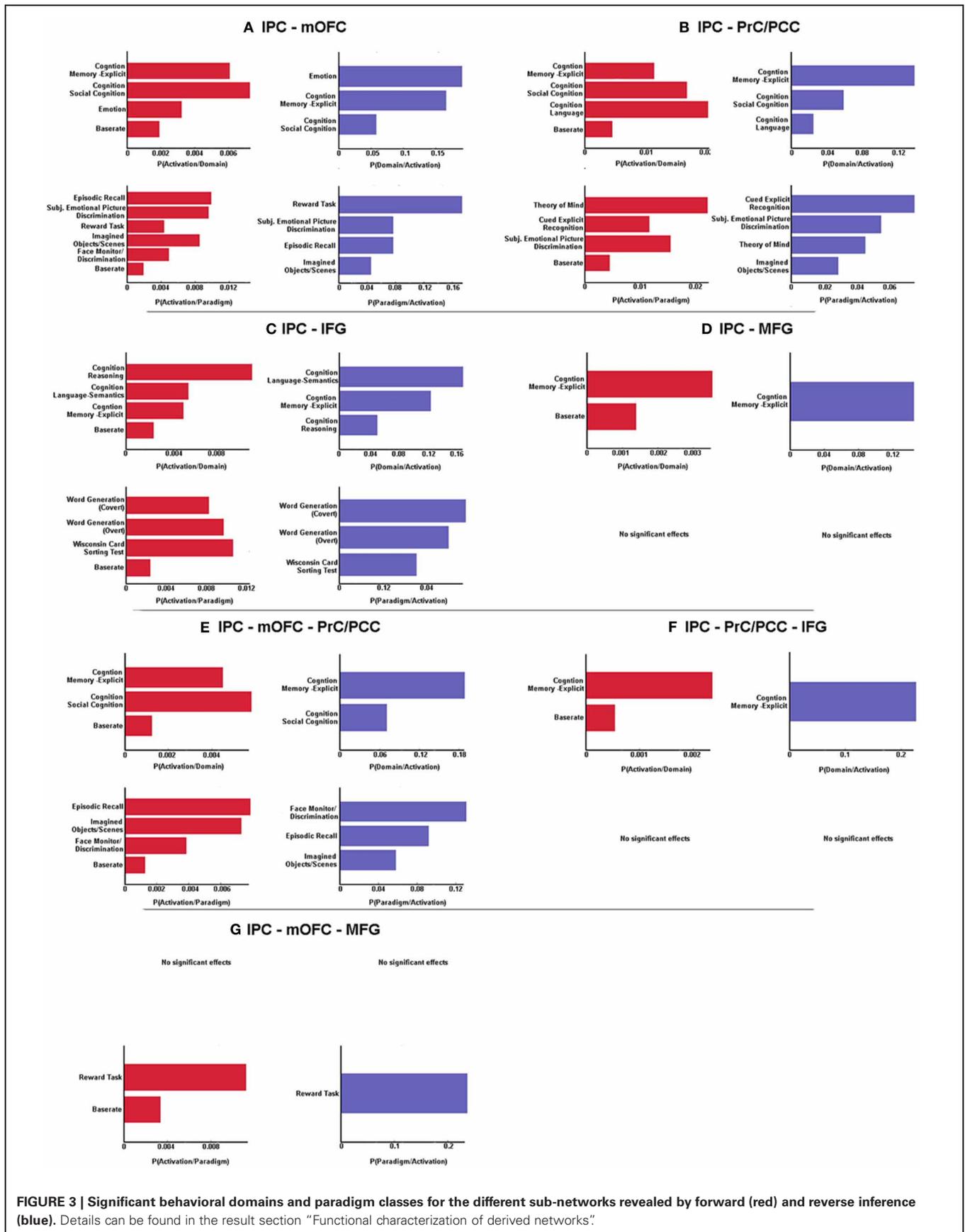
IPC, PrC/PCC, and IFG analysis revealed above chance co-activation of those three regions in experiments referring to explicit memory, but also no significant paradigm class (**Figure 3F**).

Additionally co-activation of IPC, mOFC, and MFG was only significantly associated with reward tasks, but with no specific BD (**Figure 3G**).

All other combinations (IPC-PrC/PCC-MFG, IPC-mOFC-IFG, IPC-MFG-IFG, IPC-PrC/PCC-mOFC-MFG, IPC-PrC/PCC-mOFC-IFG, IPC-PrC/PCC-MFG-IFG, IPC-mOFC-MFG-IFG and IPC-PrC/PCC-mOFC-MFG-IFG) didn't reveal any significant association to any BDs or paradigm classes.

### DISCUSSION

The current study investigated the connectivity and functional properties of a region in the left IPC, which has been found to be dysregulated in schizophrenia and depression during processing



**FIGURE 3 | Significant behavioral domains and paradigm classes for the different sub-networks revealed by forward (red) and reverse inference (blue).** Details can be found in the result section “Functional characterization of derived networks”.

of audiovisual emotional stimuli. That is, we investigated neuronal networks and their associated functions that center on an inferior parietal region showing aberrant responses during a multi-modal affective processing task in two major psychiatric disorders. First, results revealed that the seed region in the left posterior IPC functionally connects with PrC/PCC, mOFC as well as MFG and inferior (IFG) frontal gyrus. Quantitative functional characterization further indicates that on the one hand, all areas are engaged during experiments related to memory. On the other hand, sub-networks that relate to social cognition, reasoning, emotional, or language processing are also discernible.

### THE ROLE OF THE LEFT IPC

The IPC is a large and heterogeneous region that has been found to be involved in a wide range of different processes, ranging from action, memory, and language to mathematical problem solving and social cognition. Furthermore, it has been described as part of the default mode network (Glover, 2004; Wagner et al., 2005; Daselaar et al., 2006; Buckner et al., 2008; Binder et al., 2009; Spreng et al., 2009; Caspers et al., 2010; Arsalidou and Taylor, 2011; Bzdok et al., 2012; Schilbach et al., 2012; Seghier, 2013). Apart from this functional diversity, the IPC is also heterogeneous with regard to its macroanatomy. It has for instance been classified into BA39, the angular gyrus and BA40, the supramarginal gyrus. In addition, based on cyto-architectonic mapping, Caspers et al. (2006, 2008) subdivided the IPC into seven different regions. Thus, both the functional as well as architectural diversity strongly suggest that the IPC as a whole is too large and heterogeneous to allow any meaningful interpretations of IPC activity or IPC dysregulation.

Going more into detail, the IPC area which has been analyzed in the current study demonstrates correspondence with the angular gyrus and more precisely with area PGp (based on the subdivisions of Caspers et al., 2006, 2008). This part of the IPC is mainly involved in language processing (Hall et al., 2005; Binder et al., 2009; Price, 2010; Clos et al., 2012) and has in this context been suggested to act as a high-level supramodal conceptual integration area (Binder et al., 2009). In addition and in line with this view, investigation of audiovisual speech integration (Bernstein et al., 2008) point to angular gyrus involvement in crossmodal binding. Furthermore, Joassin et al. (2011) report greater left angular gyrus activity when presenting bimodal face-voice pairs compared to faces or voices alone, indicating a general and not only speech-specific role of this region in audiovisual binding. However, apart from crossmodal binding, the left angular region has also been associated with memory, the default mode network and social cognition (Seghier, 2013). Therefore, the common process involved in all these different tasks and domains associated with the angular gyrus may lie in the integration of information and concepts from different modalities and subsystems (Seghier, 2013). This indicates that it might be more informative to assess the functional role of the IPC from a network-based perspective. This idea is in line with the view that a specific role of a brain area cannot completely be determined by looking at it in isolation but should ideally be investigated together with regions it stands in interplay with (Stephan, 2004;

Seghier, 2013). Therefore, it may be argued that, in order to be able to interpret activity in such a complex multimodal region as the IPC as well as its dysregulation in schizophrenia and depression, it is crucial to investigate its functional role from a system perspective.

### CONNECTIVITY OF THE LEFT POSTERIOR IPC

By using two different approaches, investigation of functional connectivity revealed left IPC functional connectivity with numerous cortical and subcortical structures. In previous studies, area PGp, which overlaps with our seed, has been shown to structurally connect to temporal and lateral occipital areas (Caspers et al., 2011) but also with frontal (Caspers et al., 2011) and (para)hippocampal regions (Uddin et al., 2010). In addition, investigating resting-state functional connectivity, Uddin et al. (2010) also investigated functional connectivity of area PGp, demonstrating on the one hand similar but also different connections as found in anatomical studies. With a few subtle differences, possibly due to a larger seed in the former study, the results of the task-independent functional connectivity of the present study greatly resembles the results of Uddin et al. (2010), indicating a consistent network across different samples. In addition, as **Figure 2A** shows, task based co-activation of area PGp demonstrated co-activation with regions which have also been found to be functionally connected in the task-independent analysis, except left middle temporal gyrus. Therefore, left IPC, PrC/PCC, mOFC, MFG, and IFG may be regarded as a core network, as convergent connectivity patterns across both task dependent and task-independent approaches could be demonstrated and therefore they reflect regions that show coupling across two fundamentally different mental states (Jakobs et al., 2012). Now the question arises if this core network as a whole is involved in certain functions or if only the interplay of the left posterior IPC with specific regions of this core network is associated with particular functions. Therefore, we further analyzed the derived network with regard to the functional characterization of the different sub-networks. In this context, sub-networks are defined as association of specific functions to co-activation of the left IPC not with the whole but rather with only some (or only one) regions of the network.

### COMMON AND DIFFERENTIAL FUNCTIONAL ROLES OF DERIVED NETWORKS

#### *Common role: explicit memory*

Functional investigation of the sub-networks derived from the functional connectivity analyses reveals that all regions are involved in explicit memory processing. That is, memory processing was found to be above chance associated with co-activation of IPC-PrC/PCC, IPC-mOFC, IPC-IFG, IPC-MFG but also with co-activation of IPC-PrC/PCC-mOFC and IPC-PrC/PCC-IFG. All regions of this core network have already been associated with memory (Wagner et al., 2005; Brand and Markowitsch, 2008; Spreng et al., 2009). However, functional characterization of co-activation of all five regions together didn't reveal any significant result. This indicates that there is not only one specific memory process in which the whole network is involved, but rather, different sub-networks are associated with differential

memory processes. This may depend on the stimulus material used, for instance IPC-PrC/PCC-mOFC co-activation may play a role when social stimuli are processed, whereas conjoint activity of IPC, PrC/PCC, and IFG might be associated with recollection of verbal material. Furthermore, it should be noted that these results might also indicate that the label “explicit memory” in BrainMap is in some way too broad and therefore combines memory processes which are still heterogeneous.

In terms of processing of social stimuli it may be further speculated that co-activation of IPC and PrC/PCC mainly denotes recollection of the content of explicit memory. In contrast, IPC and mOFC interactions on the other side might be more involved in the emotional connotation of explicit memory. In line with this idea, Brand and Markowitsch (2006) suggested that OFC is always involved when a memory has a personal/emotional connotation. As the recall of explicit memory, especially autobiographical information, often involves both recollection of facts itself as well as the associated emotional hue, the result of elevated likelihood of IPC-mOFC-PrC/PCC activation during explicit memory experiments is not surprising.

In sum, current results suggest that even though the whole derived network is related to memory, different sub-networks are involved in differential memory processes.

### **Specific roles of sub-networks**

Besides the common process of memory, functional characterization also revealed specific roles of the different sub-networks. First of all, our results reveal that emotional processing as well as reasoning is sub-network specific. That is, emotion was associated with co-activation of IPC and mOFC only, while reasoning was exclusively overrepresented among experiments co-activating IPC and IFG. Thus, this results extent previous reports of involvement of mOFC in emotional processing (Phan et al., 2002; Kellermann et al., 2012) and IFG in reasoning (Prado et al., 2011), by showing both regions to be associated to these processes in co-activation with the left IPC.

Furthermore, with regard to our investigated networks, language was found to be associated with IPC and IFG as well as with IPC and PrC/PCC, but not with IPC, PrC/PCC, and IFG co-activation. This result is hence reasonable, considering that IPC and IFG co-activation was mainly associated to the subcategory semantic language processing while IPC-PrC/PCC co-activation was not related to any specific subcategory. On the one hand, IPC and IFG have already been reported to play a major role in semantic processing and resting state functional connectivity between those two regions have been demonstrated to correlate with reading comprehension (Hampson et al., 2006). On the other hand, IPC and PrC/PCC are important nodes of the mentalizing system (Mar, 2011). Therefore, it may be argued that IPC and IFG co-activation is more involved in general language comprehension, whereas conjoint activation of IPC, PrC/PCC mainly denotes language processes requiring theory of mind. In line with this view, Mar (2011) report IPC and precuneus activity only during story-based and non-story based theory of mind tasks but not just for narrative comprehension.

Moreover, in line with studies reporting a role of IPC, PrC/PCC, and mOFC in mentalizing and self-referential processing (Ochsner et al., 2005; Mar, 2011; Bzdok et al., 2012), functional characterization further reveal that IPC-PrC/PCC and IPC-mOFC, but also co-activation of all three regions together are related to social cognition. Thus, the two sub-networks (IPC-PrC/PCC and IPC-mOFC) may differently contribute to the overall process of social cognition. That is, IPC-mOFC co-activation may subserve the affective component of social cognition whereas IPC-PrC/PCC is more associated with introspection and self/other referential processing. As social cognition usually involves both, emotional but also perspective taking, co-activation of all three regions may be necessary for this process.

In sum, given the suggested role of the left posterior IPC as a higher-level integration area (Binder et al., 2009; Seghier, 2013), it may be speculated that depending on the specific regions the IPC co-activates with, the aspects which have to be integrated change, leading to association of different functions with differential sub-networks. These results therefore further highlight the importance of network-based investigations, indicating that the functional role of a specific brain area is highly dependent on those regions it interacts with.

### **POSTERIOR INFERIOR PARIETAL DYSFUNCTION IN PSYCHIATRIC DISORDERS**

Structural and functional deficits of the IPC have already been demonstrated, in schizophrenia as well as in depression (Canli et al., 2004; Torrey, 2007; Wang et al., 2008; Palaniyappan and Liddle, 2012; Zeng et al., 2012). Furthermore, in terms of schizophrenia, dysfunctions of this region have been associated with symptoms of thought disorder and depersonalization (Torrey, 2007). We investigated audiovisual emotional integration in schizophrenia and depression and found dysregulation of left posterior IPC in both patient groups. From a regional perspective, given the role of the IPC in combining information from different subsystems and in cross-sensory binding (Joassin et al., 2011; Seghier, 2013), dysregulation of this region in schizophrenia and depression suggests a deficit in audiovisual integration in both patient groups. In particular, deactivation possibly in order to inhibit binding of acoustic information with congruent visual target information was found to be impaired in schizophrenia and in depression. As especially deactivation of the congruent condition is impaired it may be suggested that patients show increased binding of congruent information compared to controls. This increased binding might in some way have a positive effect, possibly leading to increased salience of congruent pairs and as a result to normal face processing strategies. This assumption fits well with a previous EEG study in schizophrenia (Müller et al., 2012), demonstrating similar P1 amplitudes between patients and controls in emotional congruent audiovisual conditions, whereas in incongruent conditions patients showed a reduced P1 response.

However, as the current study now demonstrates that the left IPC is functionally connected with frontal cortices and PrC/PCC, which depending on task demands, form sub-networks, it may be speculated that deficits in this area might not only affect audiovisual processing of emotions but rather be associated to diverse areas of functioning. In particular, even though the seed was

defined by an area dysregulated specifically in an audiovisual task, this left posterior IPC deficit might rather reflect impaired integration in general. This might then further result in deficits also in those domains that are associated with the different sub-networks the IPC interacts with. That is, inferior parietal dysregulations might affect processing in the whole connected network and therefore also be associated with cognitive deficits and impairments in emotion processing. In line with this view, both disorders, schizophrenia as well as depression, go along with cognitive, social and emotional impairments (Bhalla et al., 2005; Lee et al., 2005; Bach et al., 2009; Bourke et al., 2010; Kohler et al., 2010; Wolkenstein et al., 2011; Young et al., 2011; Dimaggio et al., 2012; Fioravanti et al., 2012; Comparelli et al., 2013; Snyder, 2013), as well as with changes in connectivity within parts of the network (Karlsgodt et al., 2008; Zhou et al., 2010). In addition, illness severity in both, schizophrenia and depression correlates with the severity of impairments in cognition, social skills as well as emotional impairment (McDermott and Ebmeier, 2009; Gollan et al., 2010; Tanaka et al., 2012; Ventura et al., 2013). Interestingly, in schizophrenia associations of functions within these domains are more often found or larger for negative than for positive symptoms (Tanaka et al., 2012; Ventura et al., 2013). However, how these associations relate to impairments in posterior IPC and its related network remains an open question. Therefore, for future studies, it would be of interest to investigate the network derived in the current study by comparing functional

connectivity between patients and healthy controls, as well as correlate functional connectivity measures between these nodes with neuropsychological scores and symptomology.

## CONCLUSION

In summary, the present results demonstrate functional connectivity of left IPC with PrC/PCC, mOFC, left IFG, and MFG across task-dependent and task-independent approaches, which, depending on task demands, form sub-networks. While the whole network is associated with memory processes, specific sub-networks are involved when social cognition, reasoning, as well as emotional and language processes are required. Results therefore indicate that dysregulation of left IPC in depression and schizophrenia might not only reflect deficits in audiovisual integration, but is possibly also connected to impaired emotional and cognitive processing in these patient groups. Thus, the current study highlights the fact that in order to gain a better understanding of a region (and the meaning of its dysregulation), it is important to investigate its functional role from a system rather than a regional perspective.

## ACKNOWLEDGMENTS

This study was supported by the Deutsche Forschungsgemeinschaft (DFG, IRTG 1328), by the National Institute of Mental Health (R01-MH074457), and the Helmholtz Initiative on systems biology (The Human Brain Model).

## REFERENCES

- Amunts, K., Kedo, O., Kindler, M., Pieperhoff, P., Mohlberg, H., Shah, N. J., et al. (2005). Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: intersubject variability and probability maps. *Anat. Embryol. (Berl.)* 210, 343–352. doi: 10.1007/s00429-005-0025-5
- Amunts, K., Schleicher, A., Burgel, U., Mohlberg, H., Uylings, H. B., and Zilles, K. (1999). Broca's region revisited: cytoarchitecture and intersubject variability. *J. Comp. Neurol.* 412, 319–341. doi: 10.1002/(SICI)1096-9861(19990920)412:2<319::AID-CNE10>3.0.CO;2-7
- Arsalidou, M., and Taylor, M. J. (2011). Is 2+2=4? Meta-analyses of brain areas needed for numbers and calculations. *Neuroimage* 54, 2382–2393. doi: 10.1016/j.neuroimage.2010.10.009
- Ashburner, J., and Friston, K. J. (2005). Unified segmentation. *Neuroimage* 26, 839–851. doi: 10.1016/j.neuroimage.2005.02.018
- Bach, D. R., Buxtorf, K., Grandjean, D., and Strik, W. K. (2009). The influence of emotion clarity on emotional prosody identification in paranoid schizophrenia. *Psychol. Med.* 39, 927–938. doi: 10.1017/S0033291708004704
- Bernstein, L. E., Auer, E. T. Jr., Wagner, M., and Ponton, C. W. (2008). Spatiotemporal dynamics of audiovisual speech processing. *Neuroimage* 39, 423–435. doi: 10.1016/j.neuroimage.2007.08.035
- Bhalla, R. K., Butters, M. A., Zmuda, M. D., Seligman, K., Mulsant, B. H., Pollock, B. G., et al. (2005). Does education moderate neuropsychological impairment in late-life depression? *Int. J. Geriatr. Psychiatry* 20, 413–417. doi: 10.1002/gps.1296
- Binder, J. R., Desai, R. H., Graves, W. W., and Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* 19, 2767–2796. doi: 10.1093/cercor/bhp055
- Bourke, C., Douglas, K., and Porter, R. (2010). Processing of facial emotion expression in major depression: a review. *Aust. N.Z. J. Psychiatry* 44, 681–696. doi: 10.3109/00048674.2010.496359
- Brand, M., and Markowitsch, H. J. (2006). "Memory processes and the orbitofrontal cortex," in *The Orbitofrontal Cortex*, eds D. Zald and S. Rauch. (Oxford: Oxford University Press), 285–306. doi: 10.1093/acprof:oso/9780198565741.003.0011
- Brand, M., and Markowitsch, H. J. (2008). "The role of the prefrontal cortex in episodic memory," in *Handbook of Episodic Memory*, eds E. Dere, A. Easton, L. Nadel, and J. P. Huston (Amsterdam: Elsevier), 317–341.
- Buckner, R. L., Andrews-Hanna, J. R., and Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Ann. N.Y. Acad. Sci.* 1124, 1–38. doi: 10.1196/annals.1440.011
- Bzdok, D., Schilbach, L., Vogeley, K., Schneider, K., Laird, A. R., Langner, R., et al. (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Struct. Funct.* 217, 783–796. doi: 10.1007/s00429-012-0380-y
- Canli, T., Sivers, H., Thomason, M. E., Whitfield-Gabrieli, S., Gabrieli, J. D., and Gotlib, I. H. (2004). Brain activation to emotional words in depressed vs healthy subjects. *Neuroreport* 15, 2585–2588. doi: 10.1097/00001756-200412030-00005
- Caspers, S., Eickhoff, S. B., Geyer, S., Scheperjans, F., Mohlberg, H., Zilles, K., et al. (2008). The human inferior parietal lobule in stereotaxic space. *Brain Struct. Funct.* 212, 481–495. doi: 10.1007/s00429-008-0195-z
- Caspers, S., Eickhoff, S. B., Rick, T., Von Kapri, A., Kuhlén, T., Huang, R., et al. (2011). Probabilistic fibre tract analysis of cytoarchitecturally defined human inferior parietal lobule areas reveals similarities to macaques. *Neuroimage* 58, 362–380. doi: 10.1016/j.neuroimage.2011.06.027
- Caspers, S., Geyer, S., Schleicher, A., Mohlberg, H., Amunts, K., and Zilles, K. (2006). The human inferior parietal cortex: cytoarchitectonic parcellation and interindividual variability. *Neuroimage* 33, 430–448. doi: 10.1016/j.neuroimage.2006.06.054
- Caspers, S., Zilles, K., Laird, A. R., and Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage* 50, 1148–1167. doi: 10.1016/j.neuroimage.2009.12.112
- Cieslik, E. C., Zilles, K., Caspers, S., Roski, C., Kellermann, T. S., Jakobs, O., et al. (2012). Is there "One" DLPFC in cognitive action control? Evidence for heterogeneity from co-activation-based parcellation. *Cereb. Cortex*. doi: 10.1093/cercor/bhs256. [Epub ahead of print].

- Clos, M., Langner, R., Meyer, M., Oechslin, M. S., Zilles, K., and Eickhoff, S. B. (2012). Effects of prior information on decoding degraded speech: an fMRI study. *Hum. Brain Mapp.* doi: 10.1002/hbm.22151. [Epub ahead of print].
- Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., et al. (2008). Audio-visual integration of emotion expression. *Brain Res.* 1242, 126–135. doi: 10.1016/j.brainres.2008.04.023
- Comparelli, A., Corigliano, V., De Carolis, A., Mancinelli, I., Trovini, G., Ottavi, G., et al. (2013). Emotion recognition impairment is present early and is stable throughout the course of schizophrenia. *Schizophr. Res.* 143, 65–69. doi: 10.1016/j.schres.2012.11.005
- Daselaar, S. M., Fleck, M. S., and Cabeza, R. (2006). Triple dissociation in the medial temporal lobes: recollection, familiarity, and novelty. *J. Neurophysiol.* 96, 1902–1911. doi: 10.1152/jn.01029.2005
- De Gelder, B., and Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cogn. Emotion* 14, 289–311. doi: 10.1080/026999300378824
- De Gelder, B., Vroomen, J., De Jong, S. J., Masthoff, E. D., Trompenaars, F. J., and Hodiament, P. (2005). Multisensory integration of emotional faces and voices in schizophrenics. *Schizophr. Res.* 72, 195–203. doi: 10.1016/j.schres.2004.02.013
- De Jong, J. J., Hodiament, P. P., Van Den Stock, J., and De Gelder, B. (2009). Audiovisual emotion recognition in schizophrenia: reduced integration of facial and vocal affect. *Schizophr. Res.* 107, 286–293. doi: 10.1016/j.schres.2008.10.001
- Diedrichsen, J., Balsters, J. H., Flavell, J., Cussans, E., and Ramnani, N. (2009). A probabilistic MR atlas of the human cerebellum. *Neuroimage* 46, 39–46. doi: 10.1016/j.neuroimage.2009.01.045
- Dimaggio, G., Salvatore, G., Popolo, R., and Lysaker, P. H. (2012). Autobiographical memory and mentalizing impairment in personality disorders and schizophrenia: clinical and research implications. *Front. Psychol.* 3:529. doi: 10.3389/fpsyg.2012.00529
- Dolan, R. J., Morris, J. S., and De Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proc. Natl. Acad. Sci. U.S.A.* 98, 10006–10010. doi: 10.1073/pnas.171288598
- Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., and Fox, P. T. (2012). Activation likelihood estimation meta-analysis revisited. *Neuroimage* 59, 2349–2361. doi: 10.1016/j.neuroimage.2011.09.017
- Eickhoff, S. B., Heim, S., Zilles, K., and Amunts, K. (2006). Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. *Neuroimage* 32, 570–582. doi: 10.1016/j.neuroimage.2006.04.204
- Eickhoff, S. B., Jbabdi, S., Caspers, S., Laird, A. R., Fox, P. T., Zilles, K., et al. (2010). Anatomical and functional connectivity of cytoarchitectonic areas within the human parietal operculum. *J. Neurosci.* 30, 6409–6421. doi: 10.1523/JNEUROSCI.5664-09.2010
- Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., and Fox, P. T. (2009). Coordinate-based ALE meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. *Hum. Brain Mapp.* 30, 2907–2926. doi: 10.1002/hbm.20718
- Eickhoff, S. B., Paus, T., Caspers, S., Grosbras, M. H., Evans, A. C., Zilles, K., et al. (2007). Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. *Neuroimage* 36, 511–521. doi: 10.1016/j.neuroimage.2007.03.060
- Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25, 1325–1335. doi: 10.1016/j.neuroimage.2004.12.034
- Fioravanti, M., Bianchi, V., and Cinti, M. E. (2012). Cognitive deficits in schizophrenia: an updated metanalysis of the scientific evidence. *BMC Psychiatry* 12:64. doi: 10.1186/1471-244X-12-64
- Glover, S. (2004). Separate visual representations in the planning and control of action. *Behav. Brain Sci.* 27, 3–24. discussion: 24–78.
- Gollan, J. K., McCloskey, M., Hoxha, D., and Coccaro, E. F. (2010). How do depressed and healthy adults interpret nuanced facial expressions? *J. Abnorm. Psychol.* 119, 804–810. doi: 10.1037/a0020234
- Gur, R. C., Sara, R., Hagendoorn, M., Marom, O., Hughett, P., Macy, L., et al. (2002). A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. *J. Neurosci. Methods* 115, 137–143. doi: 10.1016/S0165-0270(02)00006-7
- Hall, D. A., Fussell, C., and Summerfield, A. Q. (2005). Reading fluent speech from talking faces: typical brain networks and individual differences. *J. Cogn. Neurosci.* 17, 939–953. doi: 10.1162/0898929054021175
- Hampson, M., Tokoglu, F., Sun, Z., Schafer, R. J., Skudlarski, P., Gore, J. C., et al. (2006). Connectivity-behavior analysis reveals that functional connectivity between left BA39 and Broca's area varies with reading ability. *Neuroimage* 31, 513–519. doi: 10.1016/j.neuroimage.2005.12.040
- Holmes, C. J., Hoge, R., Collins, L., Woods, R., Toga, A. W., and Evans, A. C. (1998). Enhancement of MR images using registration for signal averaging. *J. Comput. Assist. Tomogr.* 22, 324–333. doi: 10.1097/00004728-199803000-00032
- Jakobs, O., Langner, R., Caspers, S., Roski, C., Cieslik, E. C., Zilles, K., et al. (2012). Across-study and within-subject functional connectivity of a right temporo-parietal junction subregion involved in stimulus-context integration. *Neuroimage* 60, 2389–2398. doi: 10.1016/j.neuroimage.2012.02.037
- Joassin, F., Pesenti, M., Maurage, P., Verreckett, E., Bruyer, R., and Campanella, S. (2011). Cross-modal interactions between human faces and voices involved in person recognition. *Cortex* 47, 367–376. doi: 10.1016/j.cortex.2010.03.003
- Karlsgodt, K. H., Van Erp, T. G., Poldrack, R. A., Bearden, C. E., Nuechterlein, K. H., and Cannon, T. D. (2008). Diffusion tensor imaging of the superior longitudinal fasciculus and working memory in recent-onset schizophrenia. *Biol. Psychiatry* 63, 512–518. doi: 10.1016/j.biopsych.2007.06.017
- Kellermann, T. S., Sternkopf, M. A., Schneider, F., Habel, U., Turetsky, B. I., Zilles, K., et al. (2012). Modulating the processing of emotional stimuli by cognitive demand. *Soc. Cogn. Affect. Neurosci.* 7, 263–273. doi: 10.1093/scan/nsq104
- Kohler, C. G., Walker, J. B., Martin, E. A., Healey, K. M., and Moberg, P. J. (2010). Facial emotion perception in schizophrenia: a meta-analytic review. *Schizophr. Bull.* 36, 1009–1019. doi: 10.1093/schbul/sbn192
- Laird, A. R., Eickhoff, S. B., Fox, P. M., Uecker, A. M., Ray, K. L., Saenz, J. J., et al. (2011). The Brainmap strategy for standardization, sharing, and meta-analysis of neuroimaging data. *BMC Res. Notes* 4:349. doi: 10.1186/1756-0500-4-349
- Laird, A. R., Eickhoff, S. B., Kurth, F., Fox, P. M., Uecker, A. M., Turner, J. A., et al. (2009a). ALE meta-analysis workflows via the brainmap database: progress towards a probabilistic functional brain atlas. *Front. Neuroinform.* 3:23. doi: 10.3389/neuro.11.023.2009
- Laird, A. R., Eickhoff, S. B., Li, K., Robin, D. A., Glahn, D. C., and Fox, P. T. (2009b). Investigating the functional heterogeneity of the default mode network using coordinate-based meta-analytic modeling. *J. Neurosci.* 29, 14496–14505. doi: 10.1523/JNEUROSCI.4004-09.2009
- Lee, L., Harkness, K. L., Sabbagh, M. A., and Jacobson, J. A. (2005). Mental state decoding abilities in clinical depression. *J. Affect. Disord.* 86, 247–258. doi: 10.1016/j.jad.2005.02.007
- Mar, R. A. (2011). The neural bases of social cognition and story comprehension. *Annu. Rev. Psychol.* 62, 103–134. doi: 10.1146/annurev-psych-120709-145406
- McDermott, L. M., and Ebmeier, K. P. (2009). A meta-analysis of depression severity and cognitive function. *J. Affect. Disord.* 119, 1–8. doi: 10.1016/j.jad.2009.04.022
- Müller, V. I., Cieslik, E. C., Kellermann, T. S., and Eickhoff, S. B. (2013). Crossmodal emotional integration in major depression. *Soc. Cogn. Affect. Neurosci.* doi: 10.1093/scan/nst057. [Epub ahead of print].
- Müller, V. I., Habel, U., Derntl, B., Schneider, F., Zilles, K., Turetsky, B. I., et al. (2011). Incongruence effects in cross-modal emotional integration. *Neuroimage* 54, 2257–2266. doi: 10.1016/j.neuroimage.2010.10.047
- Müller, V. I., Kellermann, T. S., Seligman, S. C., Turetsky, B. I., and Eickhoff, S. B. (2012). Modulation of affective face processing deficits in schizophrenia by congruent emotional sounds. *Soc. Cogn. Affect. Neurosci.* doi: 10.1093/scan/nss107. [Epub ahead of print].
- Nichols, T., Brett, M., Andersson, J., Wager, T., and Poline, J. B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage* 25, 653–660. doi: 10.1016/j.neuroimage.2004.12.005
- Ochsner, K. N., Beer, J. S., Robertson, E. R., Cooper, J. C., Gabrieli, J. D., Kihlstrom, J. F., et al. (2005). The neural correlates of direct and reflected self-knowledge. *Neuroimage* 28, 797–814. doi: 10.1016/j.neuroimage.2005.06.069

- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- Palaniyappan, L., and Liddle, P. F. (2012). Dissociable morphometric differences of the inferior parietal lobule in schizophrenia. *Eur. Arch. Psychiatry Clin. Neurosci.* 262, 579–587. doi: 10.1007/s00406-012-0314-y
- Phan, K. L., Wager, T., Taylor, S. F., and Liberzon, I. (2002). Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage* 16, 331–348. doi: 10.1006/nimg.2002.1087
- Prado, J., Chadha, A., and Booth, J. R. (2011). The brain network for deductive reasoning: a quantitative meta-analysis of 28 neuroimaging studies. *J. Cogn. Neurosci.* 23, 3483–3497. doi: 10.1162/jocn\_a\_00063
- Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Ann. N.Y. Acad. Sci.* 1191, 62–88. doi: 10.1111/j.1749-6632.2010.05444.x
- Reetz, K., Dogan, I., Rolf, A., Binkofski, F., Schulz, J. B., Laird, A. R., et al. (2012). Investigating function and connectivity of morphometric findings—exemplified on cerebellar atrophy in spinocerebellar ataxia 17 (SCA17). *Neuroimage* 62, 1354–1366. doi: 10.1016/j.neuroimage.2012.05.058
- Robinson, J. L., Laird, A. R., Glahn, D. C., Lohval, W. R., and Fox, P. T. (2010). Metaanalytic connectivity modeling: delineating the functional connectivity of the human amygdala. *Hum. Brain Mapp.* 31, 173–184. doi: 10.1002/hbm.20854
- Rottschy, C., Caspers, S., Roski, C., Reetz, K., Dogan, I., Schulz, J. B., et al. (2012). Differentiated parietal connectivity of frontal regions for “what” and “where” memory. *Brain Struct. Funct.* doi: 10.1007/s00429-012-0476-4. [Epub ahead of print].
- Satterthwaite, T. D., Elliott, M. A., Gerraty, R. T., Ruparel, K., Loughhead, J., Calkins, M. E., et al. (2013). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *Neuroimage* 64, 240–256. doi: 10.1016/j.neuroimage.2012.08.052
- Scheperjans, F., Hermann, K., Eickhoff, S. B., Amunts, K., Schleicher, A., and Zilles, K. (2008). Observer-independent cytoarchitectonic mapping of the human superior parietal cortex. *Cereb. Cortex* 18, 846–867. doi: 10.1093/cercor/bhm116
- Schilbach, L., Bzdok, D., Timmermans, B., Fox, P. T., Laird, A. R., Vogeley, K., et al. (2012). Introspective minds: using ALE meta-analyses to study commonalities in the neural correlates of emotional processing, social and unconstrained cognition. *PLoS ONE* 7:e30920. doi: 10.1371/journal.pone.0030920
- Seghier, M. L. (2013). The angular gyrus: multiple functions and multiple subdivisions. *Neuroscientist* 19, 43–61. doi: 10.1177/1073858412440596
- Snyder, H. R. (2013). Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: a meta-analysis and review. *Psychol. Bull.* 139, 81–132. doi: 10.1037/a0028727
- Sommer, I. E., Clos, M., Meijering, A. L., Diederer, K. M., and Eickhoff, S. B. (2012). Resting state functional connectivity in patients with chronic hallucinations. *PLoS ONE* 7:e43516. doi: 10.1371/journal.pone.0043516
- Spreng, R. N., Mar, R. A., and Kim, A. S. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *J. Cogn. Neurosci.* 21, 489–510. doi: 10.1162/jocn.2008.21029
- Stephan, K. E. (2004). On the role of general system theory for functional neuroimaging. *J. Anat.* 205, 443–470. doi: 10.1111/j.0021-8782.2004.00359.x
- Tanaka, T., Tomotake, M., Ueoka, Y., Kaneda, Y., Taniguchi, K., Nakataki, M., et al. (2012). Clinical correlates associated with cognitive dysfunction in people with schizophrenia. *Psychiatry Clin. Neurosci.* 66, 491–498. doi: 10.1111/j.1440-1819.2012.02390.x
- Torrey, E. F. (2007). Schizophrenia and the inferior parietal lobule. *Schizophr. Res.* 97, 215–225. doi: 10.1016/j.schres.2007.08.023
- Turner, J. A., and Laird, A. R. (2012). The cognitive paradigm ontology: design and application. *Neuroinformatics* 10, 57–66. doi: 10.1007/s12021-011-9126-x
- Uddin, L. Q., Supekar, K., Amin, H., Yekhlevskaia, E., Nguyen, D. A., Greicius, M. D., et al. (2010). Dissociable connectivity within human angular gyrus and intraparietal sulcus: evidence from functional and structural connectivity. *Cereb. Cortex* 20, 2636–2646. doi: 10.1093/cercor/bhq011
- Van Den Stock, J., De Jong, S. J., Hodiament, P. P., and De Gelder, B. (2011). Perceiving emotions from bodily expressions and multisensory integration of emotion cues in schizophrenia. *Soc. Neurosci.* 6, 537–547. doi: 10.1080/17470919.2011.568790
- Ventura, J., Tom, S. R., Jetton, C., and Kern, R. S. (2013). Memory functioning and negative symptoms as differential predictors of social problem solving skills in schizophrenia. *Schizophr. Res.* 143, 307–311. doi: 10.1016/j.schres.2012.10.043
- Wagner, A. D., Shannon, B. J., Kahn, I., and Buckner, R. L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends Cogn. Sci.* 9, 445–453. doi: 10.1016/j.tics.2005.07.001
- Wang, L., Krishnan, K. R., Steffens, D. C., Potter, G. G., Dolcos, F., and McCarthy, G. (2008). Depressive state- and disease-related alterations in neural responses to affective and executive challenges in geriatric depression. *Am. J. Psychiatry* 165, 863–871. doi: 10.1176/appi.ajp.2008.07101590
- Wolkenstein, L., Schonenberg, M., Schirm, E., and Hautzinger, M. (2011). I can see what you feel, but I can't deal with it: impaired theory of mind in depression. *J. Affect. Disord.* 132, 104–111. doi: 10.1016/j.jad.2011.02.010
- Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., and Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* 8, 665–670. doi: 10.1038/nmeth.1635
- Young, K. D., Erickson, K., Nugent, A. C., Fromm, S. J., Mallinger, A. G., Furey, M. L., et al. (2011). Functional anatomy of autobiographical memory recall deficits in depression. *Psychol. Med.* 42, 345–357. doi: 10.1017/S0033291711001371
- Zeng, L. L., Liu, L., Liu, Y., Shen, H., Li, Y., and Hu, D. (2012). Antidepressant treatment normalizes white matter volume in patients with major depression. *PLoS ONE* 7:e44248. doi: 10.1371/journal.pone.0044248
- Zhou, Y., Yu, C., Zheng, H., Liu, Y., Song, M., Qin, W., et al. (2010). Increased neural resources recruitment in the intrinsic organization in major depression. *J. Affect. Disord.* 121, 220–230. doi: 10.1016/j.jad.2009.05.029
- Zu Eulenburg, P., Caspers, S., Roski, C., and Eickhoff, S. B. (2012). Meta-analytical definition and functional connectivity of the human vestibular cortex. *Neuroimage* 60, 162–169. doi: 10.1016/j.neuroimage.2011.12.032

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 21 March 2013; accepted: 24 May 2013; published online: 12 June 2013.

Citation: Müller VI, Cieslik EC, Laird AR, Fox PT and Eickhoff SB (2013) Dysregulated left inferior parietal activity in schizophrenia and depression: functional connectivity and characterization. *Front. Hum. Neurosci.* 7:268. doi: 10.3389/fnhum.2013.00268

Copyright © 2013 Müller, Cieslik, Laird, Fox and Eickhoff. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



# Just watching the game ain't enough: striatal fMRI reward responses to successes and failures in a video game during active and vicarious playing

Jari Kätsyri<sup>1,2\*</sup>, Riitta Hari<sup>3,4</sup>, Niklas Ravaja<sup>2,5,6</sup> and Lauri Nummenmaa<sup>3,7,8</sup>

<sup>1</sup> Department of Media Technology, Aalto University School of Science, Espoo, Finland

<sup>2</sup> School of Business, Aalto University, Helsinki, Finland

<sup>3</sup> Brain Research Unit, O.V. Lounasmaa Laboratory, Aalto University School of Science, Espoo, Finland

<sup>4</sup> Advanced Magnetic Imaging Centre, Aalto Neuroimaging, Aalto University, Espoo, Finland

<sup>5</sup> Helsinki Institute for Information Technology, Aalto University, Espoo, Finland

<sup>6</sup> Department of Social Research, University of Helsinki, Helsinki, Finland

<sup>7</sup> Department of Biomedical Engineering and Computational Science, Aalto University School of Science, Espoo, Finland

<sup>8</sup> Turku PET Centre, University of Turku, Turku, Finland

## Edited by:

Martin Klasen, RWTH Aachen University, Germany

## Reviewed by:

Lutz Jäncke, University of Zurich, Switzerland

Krystyna A. Mathiak, RWTH Aachen University, Germany

## \*Correspondence:

Jari Kätsyri, Department of Media Technology, Aalto University School of Science, PO Box 15500, FI-00076 Aalto, Finland  
e-mail: jari.katsyri@aalto.fi

Although the multimodal stimulation provided by modern audiovisual video games is pleasing by itself, the rewarding nature of video game playing depends critically also on the players' active engagement in the gameplay. The extent to which active engagement influences dopaminergic brain reward circuit responses remains unsettled. Here we show that striatal reward circuit responses elicited by successes (wins) and failures (losses) in a video game are stronger during active than vicarious gameplay. Eleven healthy males both played a competitive first-person tank shooter game (active playing) and watched a pre-recorded gameplay video (vicarious playing) while their hemodynamic brain activation was measured with 3-tesla functional magnetic resonance imaging (fMRI). Wins and losses were paired with symmetrical monetary rewards and punishments during active and vicarious playing so that the external reward context remained identical during both conditions. Brain activation was stronger in the orbitomedial prefrontal cortex (omPFC) during winning than losing, both during active and vicarious playing. In contrast, both wins and losses suppressed activations in the midbrain and striatum during active playing; however, the striatal suppression, particularly in the anterior putamen, was more pronounced during loss than win events. Sensorimotor confounds related to joystick movements did not account for the results. Self-ratings indicated losing to be more unpleasant during active than vicarious playing. Our findings demonstrate striatum to be selectively sensitive to self-acquired rewards, in contrast to frontal components of the reward circuit that process both self-acquired and passively received rewards. We propose that the striatal responses to repeated acquisition of rewards that are contingent on game related successes contribute to the motivational pull of video-game playing.

**Keywords:** emotion, motivation, natural stimulation, reward system, striatum, video-game playing

## INTRODUCTION

Video game playing is intrinsically motivating (cf. Ryan and Deci, 2000): most people play video games because they are inherently interesting and enjoyable rather than because they provide financial rewards or other external outcomes (Ryan et al., 2006; Przybylski et al., 2009, 2010). Accordingly, brain imaging studies have demonstrated that video game playing engages key motivational systems of the brain, as evidenced by increases in dopamine release (Koepp et al., 1998) and hemodynamic activations (Hoeft et al., 2008) in the striatum (see also Kätsyri et al., 2012). Major motivational events during the gameplay consist of successes and failures to achieve specific game goals, such as managing to eliminate one's opponents or avoiding getting eliminated oneself. Successes and failures are

among the most potent triggers for pleasant and unpleasant emotions (Nummenmaa and Niemi, 2004), and their affective salience is amplified when they can be attributed to internal (as during active gameplay) rather than external causes (Weiner, 1985). In line with this, brain imaging studies have shown that self-acquired rewards—such as those contingent on correct motor responses—rather than those delivered at random evoke stronger neural responses in the striatum (e.g., Zink et al., 2004). Consequently, it is possible that the motivational pull of video games could be explained by the amplified reward responses triggered by actively obtaining rewards during gameplay. Here we tested this hypothesis by contrasting reward circuit responses to success- and failure-related gameplay events during *active* and *vicarious* video-game playing—that is, situations

in which players have complete versus no control over their game character.

Success- and failure-related gameplay events fulfil the three characteristics of rewards and punishments considered in animal learning (Schultz, 2004, 2006; Berridge and Kringelbach, 2008). First, they contribute to learning by providing direct feedback on the players' performance. Second, they are associated with approach and withdrawal behaviors, given that players strive to succeed and to avoid failing in the game (see Clarke and Duimering, 2006). Third, successes are generally associated with pleasant and failures with unpleasant emotional responses—even though in some games this mapping may be more complex (Ravaja et al., 2006, 2008). Dopaminergic pathways extending from the midbrain (ventral tegmental area and substantia nigra, VTA/SN) to the ventral and dorsal striatum (nucleus accumbens, caudate nucleus, and putamen) and frontal cortex (orbitomedial and medial prefrontal cortex; omPFC and vmPFC) are involved in processing rewards and punishments (Kelley, 2004; O'Doherty, 2004; Bressan and Crippa, 2005; Knutson and Cooper, 2005; Schultz, 2006; Berridge and Kringelbach, 2008; Hikosaka et al., 2008; Haber and Knutson, 2010; Koob and Volkow, 2010). This dopaminergic circuitry also likely contributes to encoding successes and failures during video-game playing. For example, neurons in monkey lateral PFC are differentially activated by successes and failures in a competitive shooting game (Hosokawa and Watanabe, 2012). Moreover, functional magnetic resonance imaging (fMRI) studies in humans have shown that successes in a video game evoke stronger activations than do failures in nucleus accumbens, caudate, and anterior putamen, as well as mPFC (Mathiak et al., 2011; Kätsyri et al., 2012; Klasen et al., 2012), and that the most anteroventral striatal activations correlate with the players' self-rated hedonic experiences during these events (Kätsyri et al., 2012).

The striatum is extensively connected to associative, motor, and limbic circuits, thereby being in an ideal anatomical position to combine both motor and affective information (Haber and Knutson, 2010). Both animal and human studies have consistently indicated that striatal reward responses are contingent on the rewards themselves as well as the actions performed to acquire them (cf. Delgado, 2007). Monkey caudate neurons fire more frequently during motor actions leading to expected rewards than during non-rewarded actions (Kawagoe et al., 1998; Schultz et al., 2000). Human fMRI studies have similarly demonstrated contingency between action and reward in the striatum. For example, dorsal caudate responds differentially to rewards and punishments only when they are perceived to be contingent on the participants' button presses (Tricomi et al., 2004). Similarly, reward activations in putamen are elevated only when the rewards are contingent on button presses (Elliott et al., 2004). Furthermore, activations in the whole striatum have been found for button presses that were executed to obtain rewards or to avoid punishments (Guitart-Masip et al., 2011). Particularly the ventral striatum shows increased activation after verbal feedback following successful motor performance, both in the absence and presence of monetary rewards (Lutz et al., 2012). Unlike the striatum,

the omPFC processes reward independently of motor actions both in monkeys (Schultz et al., 2000) and humans (Elliott et al., 2004). Following these findings, successes should evoke stronger activations in the striatum than failures only during active video game playing, whereas the omPFC should show stronger activations to successes both during active and vicarious playing.

Up to date, few brain imaging studies have compared neural activations during active and vicarious video game playing. One study using electroencephalography demonstrated that active versus vicarious playing evokes increased fronto-parietal cortical activations, along with higher self-reported spatial presence in the game (Havranek et al., 2012). Haemodynamic responses to active and vicarious playing have been explicitly compared in only one study (Cole et al., 2012): the onset of video game activated the striatum (nucleus accumbens, caudate, and putamen) and frontal nodes adjacent to mPFC (i.e., anterior cingulate cortex), with stronger activation during active than vicarious gameplay. The striatal activations decreased following the offset of playing. However, the fMRI responses to success events in the game did not differ between active and vicarious playing; furthermore, failure events were not included in the game. It is possible that the applied between-subjects design (i.e., comparison between participant groups playing and watching a video game) was not powerful enough to reveal success-related differences between active and vicarious playing. Furthermore, the study naturally begs the question of whether failure events would evoke differential activations during active and vicarious gameplay.

Here we investigated whether the reward-system activations elicited by successes and failures in a competitive video game would differ between active and vicarious video-game playing in a fully within-subjects design. We used a simplified tank shooter game that was customized for the fMRI setting (cf. Kätsyri et al., 2012). The major success and failure events in the game consisted of wins (eliminating the opponent) and losses (getting eliminated oneself) against one's opponent. We reanalyzed parts of our previously published data on active gameplay (Kätsyri et al., 2012), and compared them with novel data from watching the same game. Unlike in our previous analysis of active gameplay data, we now contrasted win and loss events separately, given that recent evidence suggests that striatal activations decrease both during wins and losses during active gameplay (Mathiak et al., 2011). We paired win and loss events with symmetric monetary rewards and punishments during both active and vicarious playing, so that the external reward for wins and losses remained identical in both conditions. Based on the previous literature, we predicted that the striatum (particularly, nucleus accumbens, ventral caudate, and anterior putamen) would show a stronger difference between wins and losses during active than vicarious gameplay, and that these effects would be associated with corresponding amplified experiences of pleasant and unpleasant emotions. We also predicted that wins would evoke greater responses in the mPFC (in particular, omPFC) than losses both during active and vicarious playing, and that these differential activations would correlate with self-rated pleasantness and unpleasantness evaluations.

## MATERIALS AND METHODS

### PARTICIPANTS

The participants were eleven right-handed male volunteers with a mean age of 25.6 years (range 22–33 years) and with abundant experience in gaming (mean 7.8 h/week, range 1–20 h/week). Additional six participants were scanned but excluded from the analysis due to technical problems (one participant), deviant playing strategies (extensive button pressing; one participant), or excessive head movements (four participants). The total playing time reported by all participants was below 30 h/week, which is an often-used criterion for addictive video game playing (Ko et al., 2009; Han et al., 2010). None of the participants had prior experience of the game played in the present study. All but one participant reported playing first-person shooter games on a regular basis with modest weekly play times (mean 3.2 h/week, range 0.5–10 h/week). All participants were Finnish under- or post-graduate university students. Only male participants were recruited because men typically have more experience of video games, are generally more motivated by such games, and show higher preference than women for competitive video games (Lucas, 2004). Participants with self-reported history of neurological or psychiatric disorders were excluded. All participants provided written informed consent as part of a protocol approved by the Ethics Committee of the Helsinki and Uusimaa University District and received monetary compensation for their lost working hours.

### STATISTICAL POWER

For statistical power calculations, we used previous active gameplay data ( $N = 43$  participants) from Cole et al. (2012). Given that their experiment did not include explicit comparison between wins and losses during active versus vicarious gameplay, we instead adopted their reported statistics on NAcc responses to active gameplay onsets ( $M = 0.234$  and  $SD = 0.2015$ ). Next, using G\*Power software (Faul et al., 2007), we estimated the *a priori* statistical power of the present experiment to detect similar effect sizes ( $\gamma = 0.234/0.2015 = 1.16$ ). The estimated power

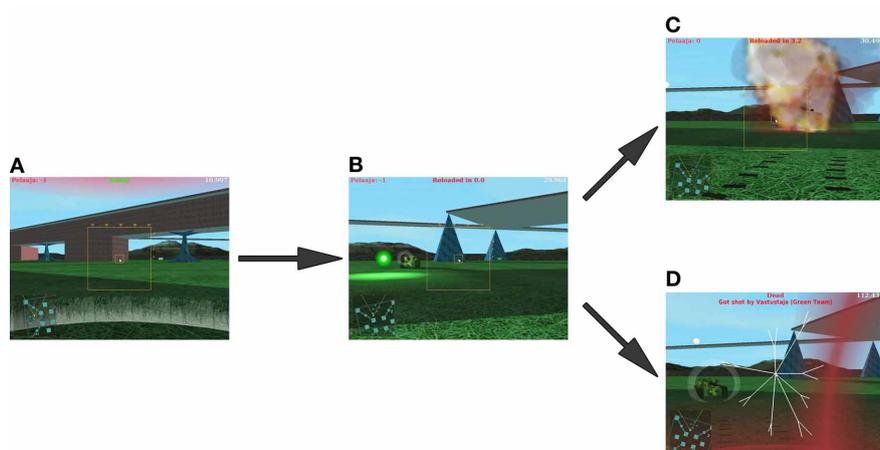
was 93%, which was considered satisfactory for the present purposes.

### EXPERIMENTAL PROCEDURE

Our experimental setting has been described in detail previously (Kättsyri et al., 2012). Briefly, during scanning the participants played two sessions of a first-person tank-shooter game “BZFlag” (an in-house modified version of 2.0.14; <http://bzflag.org>) against alleged human and computer opponents, respectively, and watched one pre-recorded gameplay video. Sessions lasted 10 min each and were presented in a counterbalanced order. However, to avoid possible reward-response biases due to competing against another human (cf. Kättsyri et al., 2012), we here analyzed only the computer-opponent session. One participant whose video-watching data were missing was replaced with a new participant; otherwise, the computer-opponent data were identical to our previous data (Kättsyri et al., 2012).

Effects of active versus vicarious gameplay on win- and loss-related activations were evaluated in a 2 (win versus lose)  $\times$  2 (play vs. watch) within-subjects design. During active playing, the participant’s task was to seek and destroy the opponent’s tank from the battlefield without getting destroyed himself (Figure 1). The corresponding win and loss events, as well as joystick movements, were time-referenced to fMRI scans and logged automatically for statistical analyses. During vicarious gameplay, the participant’s task was to follow a gameplay video recorded with the video capture software FRAPS (<http://www.fraps.com>) from one player, who did not take part in the actual study. Frequencies of wins and losses in the gameplay video were similar to those in the gameplay sessions (cf. Table 1). The final video had similar resolution (video: 1024  $\times$  768 pixels sampled at 30 fps, audio sampled at 48 kHz) and visual quality (15 Mbit/s after video compression with XVID codec; <http://www.xvid.com>) as the actual video game. The gameplay video was presented using Presentation software (<http://www.neurobs.com>).

To control for the external reward context for active and vicarious playing conditions, we introduced symmetric monetary



**FIGURE 1 |** A sample sequence of gameplay events: player spawns in a random location on the battlefield and starts searching for the opponent (A), the player and opponent engage each other (B), until the player either manages to eliminate the opponent (C) or gets eliminated himself (D).

rewards and punishments to wins and losses, respectively. Participants were told that in addition to a fixed compensation (20 Euros), they would gain money (+0.33 Euros) when winning and lose money (−0.33 Euros) when losing during the gameplay or when watching the player winning or losing on the video. In reality, all participants received an equal monetary compensation (30 Euros), which exceeded the sum any of them would otherwise have gained.

### SELF-REPORTS

Before the experiment, participants filled a 20-item self-evaluation questionnaire related to their dispositional behavioral inhibition and activation system (BIS/BAS) sensitivities (Carver and White, 1994). The BIS and BAS regulate aversive and appetitive motivation, modulating behavioral and affective responses towards punishments and rewards, respectively (Carver and White, 1994). The BIS scale is comprised of seven items (e.g., “I feel pretty worried or upset when I think or know somebody is angry at me”). The BAS scale is comprised of three subscales: drive (4 items; e.g., “I go out of my way to get things I want”), reward responsiveness (5 items; e.g., “When I get something I want, I feel excited and energized”), and fun seeking (4 items; e.g., “I crave excitement and new sensations”). Each of the items was rated on a 4-point scale, ranging from 1 (very false with me) to 4 (very true for me). The psychometric properties of the instrument have been shown to be acceptable (Carver and White, 1994).

To assess participants’ subjective experiences during active and vicarious playing, we asked them to complete a series of self-reports after both gameplay sessions. The order of questions was randomized, and the responses were given by moving the joystick. We used the Game Experience Questionnaire (Ijsselstein et al., 2008) to quantify the following facets of gaming experience: challenge, competence, flow, positive affect, negative affect, immersion, and tension (two items per scale). Spatial presence—the experience of being physically present in the game environment (Lombard and Ditton, 1997)—was measured with the Spatial Presence scale of the ITC Sense of Presence Inventory (Lessiter et al., 2001). The Spatial Presence scale is comprised of 19 items (e.g., “I had a sense of being in the game scenes”). To measure the participants’ experience of taking part in a social interaction with their opponent, we used the Social Presence in Gaming Questionnaire (de Kort et al., 2007), which consists of 17 items related to empathy (e.g., “I empathized with the other”), involvement with the other player’s actions (e.g., “My actions depended on the other’s actions”), and negative feelings towards him (e.g., “I felt revengeful”). Two additional questions were used for evaluating overall pleasantness of all win and all loss events during a session, on a scale ranging from 1 (extremely unpleasant) to 5 (neither pleasant nor unpleasant) to 9 (extremely pleasant).

### JOYSTICK REGRESSORS

Horizontal and vertical joystick coordinates were digitized at 200 Hz and collapsed into Euclidean distances from the joystick’s central position. Resulting position and velocity (i.e., the first derivative of position) tracks were low-pass filtered at 5 Hz using a first-order smoothing filter (Savitzky and Golay, 1964). Mean

joystick position and velocity values were extracted separately for each fMRI scan of each participant. Finally, to remove any overlap between these time courses, the joystick velocity time course was orthogonalized with respect to the joystick position track. Consequently, the joystick position regressor measured the overall tank movement, whereas the joystick velocity regressor measured how much the player exerted control over the tank’s movement during each fMRI scan. Similar regressors were extracted for the watching condition from the game logs of the player whose gameplay session was shown on the video. These variables were subsequently used as nuisance covariates in the fMRI data analysis.

## ACQUISITION AND ANALYSIS OF fMRI DATA

### Data acquisition and preprocessing

Functional and anatomical volumes were collected with a General Electric Signa 3.0 T MRI scanner at the Advanced Magnetic Imaging Centre of Aalto University. Whole-brain functional images were acquired using weighted gradient-echo planar imaging, sensitive to BOLD signal contrast (35 oblique slices without gaps, slice thickness = 4 mm, TR = 2070 ms, TE = 32 ms, FOV = 220 mm, flip angle = 75°, interleaved slice acquisition, 293 volumes per session with a resolution of  $3.4 \times 3.4 \text{ mm}^2$ ). The first three volumes were discarded to allow for equilibration effects. T1-weighted structural images were acquired at a resolution of  $1 \times 1 \times 1 \text{ mm}^3$  using a sequence with ASSET calibration.

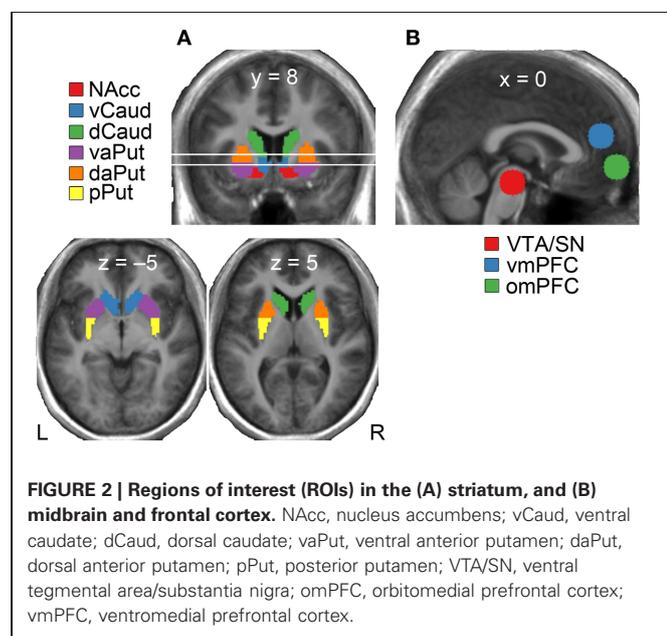
Preprocessing and analysis of fMRI data were performed using SPM8 software package (Wellcome Department of Imaging Neuroscience, London) in Matlab (version 7.11). The EPI images were sinc interpolated in time to correct for slice timing differences and realigned to the first scan by rigid-body transformations to correct for head movements. ArtRepair toolbox (version 4; <http://spnl.stanford.edu/tools/ArtRepair>; Mazaika et al., 2009) was used to correct for movement artifacts. Realigned functional volumes were first motion-adjusted and outlier volumes (head position change exceeding 0.5 mm or global mean BOLD signal change exceeding 1.3%) were then replaced by linear interpolation between the closest non-outlier volumes. Four participants with more than 10% outlier volumes were removed from further analysis. On average, 2.5% of volumes during the video game playing session and 1.5% of volumes during the video game watching were classified as outliers—the number of outliers was not significantly different between these conditions (Wilcoxon’s  $T_{(10)} = 0.77$ ,  $p = \text{n.s.}$ ). EPI and structural images were coregistered and normalized to the ICBM152 standard template in Montreal Neurological Institute (MNI) space (resolution  $2 \times 2 \times 2 \text{ mm}^3$ ) using linear and nonlinear transformations and smoothed spatially with a Gaussian isotropic kernel of 6-mm full width half maximum. The functional data were filtered temporally using an autoregressive model (AR-1) and a high-pass filter with 171.5 s cut-off (corresponding to the duration of the longest game rounds).

### Statistical analyses

We analysed our unconstrained video game playing data using event-related fMRI by focusing the analyses on win and loss events, whose timings were annotated automatically for every

participant. Specifically, a random-effects model was implemented using a two-stage process. At the first level, each participant's hemodynamic responses to wins and losses during active and vicarious playing were modeled as delta (stick) functions, which were convolved with the hemodynamic response function (HRF). Joystick position and velocity time courses were included as nuisance regressors—head motion regressors were not included given that the motion adjustment procedure of ArtRepair toolbox (Mazaika et al., 2009) had already accounted for these. Individual contrast images for the conditions “winning while playing,” “winning while watching,” “losing while playing,” and “losing while watching” were then generated. At the second level, the first-level contrast images were subjected to a 2 (win vs. loss)  $\times$  2 (play vs. watch) factorial analysis, assuming dependency and unequal variances between the levels of both variables. With balanced designs at the first level (i.e., similar events for each subject, in similar numbers), this second-level analysis closely approximated a true mixed-effects design, with both within- and between-subject variance. At the second-level, we tested the main effects of contrasts “win > loss,” “loss > win,” “play > watch,” and “watch > play” with t-tests. To identify brain regions showing differential sensitivities to wins and losses during active and vicarious playing, we specified additional interaction contrasts “(play: win > loss) > (watch: win > loss)” and “(play: loss > win) > (watch: loss > win).” Statistical threshold in these analyses was set to family-wise error (FWE) corrected  $P < 0.05$ .

We defined *a priori* regions of interest (ROI) for testing activations within mesial, striatal, and frontal parts of the reward circuit (Figure 2). Given that the striatum encompasses several anatomically and functionally segregated regions (cf. Haber and Knutson, 2010), we divided it into the following six subregions using the same classification as in our previous study (Kättsyri et al., 2012): nucleus accumbens (NAcc), ventral caudate (vCaud), dorsal caudate (dCaud), ventral anterior putamen (vaPut), dorsal anterior putamen (daPut), posterior putamen (pPut).



anterior putamen (daPut) and posterior putamen (pPut). A spherical 10-mm ROI was defined for the VTA/SN (MNI coordinates 0, -22, -18) based on a previous study (O’Doherty et al., 2002). A spherical 10-mm ROI was derived for the vmPFC (MNI 0, 46, 18) based on a previous meta-analysis (Steele and Lawrie, 2004). Given that some fMRI studies on reward processing have reported more inferior reward-sensitive activation clusters, an additional 10-mm spherical ROI was extracted for the omPFC (MNI 0, 58, -6) from a previous study (Xue et al., 2009).

Correlations between self-ratings and mean beta responses in our predefined ROIs, during active versus vicarious playing, were tested with non-parametric Spearman’s rank correlation tests. Correlations between pleasantness evaluations for specific game events and overall game experience evaluations were tested similarly. For these analyses, difference scores were first calculated between the playing and watching conditions for the pairs of variables in question, and correlations between them ( $R_{\text{Play-Watch}}$ ) were then tested. Given that such difference scores may produce spurious correlations (Cohen et al., 1983), we additionally calculated separate correlation coefficients for the variables constituting the difference scores ( $R_{\text{Play}}$  and  $R_{\text{Watch}}$ ) and set a criterion that their relative magnitudes should follow those of the difference scores (i.e.,  $R_{\text{Play}} > R_{\text{Watch}}$  when  $R_{\text{Play-Watch}} > 0$ ; and  $R_{\text{Play}} < R_{\text{Watch}}$  when  $R_{\text{Play-Watch}} < 0$ ) for a difference score correlation to be considered as significant. Significance level thresholds for difference score correlations, when they were unplanned, were adjusted using false discovery rate (FDR) correction (Benjamini and Hochberg, 1995) at  $P < 0.05$ .

## RESULTS

### BEHAVIORAL EVALUATIONS

Table 1 shows results from self-reports for active and vicarious playing conditions. Pleasantness ratings for wins and losses were significantly different from the scale’s middle point (neutral emotional state) both during active (Wilcoxon signed rank tests:  $Z = 3.0$  and  $-2.8$ ,  $P = 0.003$  and  $0.004$ ; effect sizes: *Pearson’s*  $R = 0.64$  and  $-0.61$ ) and vicarious playing ( $Z = 2.9$  and  $-2.9$ ,  $P = 0.004$  and  $0.004$ ,  $R = 0.62$  and  $-0.61$ ). Active versus vicarious playing did not differ with regard to number of wins ( $R = -0.31$ ), number of losses ( $R = -0.35$ ), or game end scores (number of wins minus losses;  $R = 0.01$ ). These manipulation checks confirmed that players associated wins and losses with rewards and punishments, respectively, and that the numbers of wins and losses did not differ between active and vicarious playing conditions.

In contrast to the measures above, participants’ experiences were clearly different during active and vicarious gameplay, with higher flow experience ( $R = 0.57$ ), lower negative affect ( $R = -0.52$ ), higher immersion ( $R = 0.49$ ), and higher spatial presence ( $R = 0.57$ ) during active playing. Similarly, players rated loss events as more unpleasant during active than vicarious playing ( $R = 0.57$ ). Following the general guidelines of Cohen (1992), these results represent medium ( $R > 0.3$ ) to large effect sizes ( $R > 0.5$ ). Additionally, we observed borderline effects ( $P < 0.10$ ) for higher challenge ( $R = 0.41$ ) and higher positive affect ( $R = 0.38$ ) during active playing with medium effect sizes. In contrast, players did not report significantly different social presence between active and vicarious playing ( $R = 0.08$  for empathy,

0.32 for involvement, and 0.18 for negative feelings)—apparently, watching the game and playing it against an alleged computer-controlled opponent were associated with similar low levels of social presence.

We also tested whether participants' pleasantness evaluations for specific game events during active versus vicarious gameplay conditions were associated with their overall game experiences or BIS/BAS scores. The results showed that pleasantness difference

scores for win events (active minus vicarious playing) were correlated positively with BAS fun seeking scores ( $R_{\text{Play-Watch}} = 0.79$ ,  $P = 0.004$ , FDR-corrected  $P_{\text{thr}} = 0.010$ ), and that the correlations for the constituent scores were meaningful ( $R_{\text{Play}} = 0.74 > R_{\text{Watch}} = -0.21$ ). Significant difference score correlations were observed also between pleasantness evaluations for win events and competence, negative affect, and positive affect; however, these findings were rejected as spurious given that their constituent scores showed correlations whose relative magnitudes were opposite to expected.

**Table 1 | Mean  $\pm$  SEM behavioral and self-report measures from active (playing) and vicarious (watching) video-game playing sessions.**

Behavioral measure	Playing	Watching	<i>P</i>
Game score	3.1 $\pm$ 1.5	3.0 <sup>a</sup>	0.964
Wins	18.6 $\pm$ 1.2	20.0 <sup>a</sup>	0.309
Losses	15.6 $\pm$ 1.2	17.0 <sup>a</sup>	0.247
Game experience			
Challenge	3.0 $\pm$ 0.2	2.2 $\pm$ 0.3	0.057†
Competence	3.0 $\pm$ 0.3	2.6 $\pm$ 0.3	0.237
Flow	3.2 $\pm$ 0.3	2.3 $\pm$ 0.3	0.007**
Negative affect	1.4 $\pm$ 0.2	2.4 $\pm$ 0.3	0.014*
Positive affect	3.3 $\pm$ 0.3	2.6 $\pm$ 0.3	0.074†
Immersion	2.9 $\pm$ 0.2	2.2 $\pm$ 0.2	0.022*
Tension	1.9 $\pm$ 0.3	2.1 $\pm$ 0.3	0.469
Spatial presence	3.1 $\pm$ 0.2	2.4 $\pm$ 0.3	0.008**
Social presence			
Empathy	1.8 $\pm$ 0.2	1.7 $\pm$ 0.2	0.725
Involvement	3.8 $\pm$ 0.2	3.4 $\pm$ 0.3	0.139
Negative feelings	2.4 $\pm$ 0.3	2.2 $\pm$ 0.2	0.403
Pleasantness			
Win events	7.5 $\pm$ 0.2	7.0 $\pm$ 0.4	0.238
Loss events	2.8 $\pm$ 0.5	3.4 $\pm$ 0.3	0.034*

The self-rating measures were scored on a scale ranging from 1 to 5 except for pleasantness ratings, which ranged from 1 (most unpleasant) to 9 (most pleasant). Game scores refer to numbers of wins minus numbers of losses. Significance values are from Wilcoxon matched-pairs tests.

<sup>a</sup>Identical for all participants.

† $P < 0.10$ .

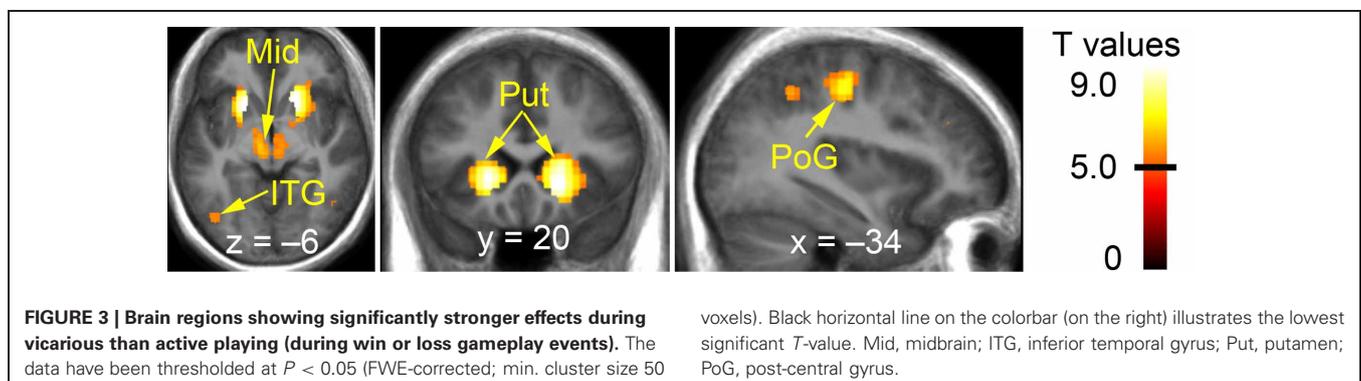
\* $P < 0.05$ .

\*\* $P < 0.01$ .

#### FULL VOLUME ANALYSIS OF fMRI DATA

Contrasting vicarious with active playing revealed activation clusters in the bilateral striatum, midbrain (including VTA/SN), sensorimotor cortices (pre- and postcentral gyri), and ventral visual stream (e.g., inferior temporal gyrus; **Figure 3** and **Table 2**). To test whether these clusters reflected *activations during active playing* or *deactivations during vicarious playing* (or both), we defined contrasts for these effects (i.e., “watch+” and “play–”) and used them as implicit masks ( $P < 0.001$ ) for the contrast between vicarious and active playing. All of the identified clusters in **Table 2** survived implicit masking by deactivations during active playing (“play–”), whereas none of them survived implicit masking by activations during vicarious playing (“watch+”), confirming that the findings reflect systematic deactivations during active gameplay events.

No significant activation or deactivation clusters were observed using the *a priori* significance threshold for the main effects of winning versus losing or vice versa, or for the interaction effects between wins versus losses and active versus vicarious playing or vice versa. However, using a small-volume correction for our *a priori* regions of interest (FWE-corrected threshold  $P < 0.05$  at cluster-level) and a slightly more lenient threshold  $P < 0.001$  (uncorrected) at voxel-level, we found stronger activations for wins versus losses in omPFC and bilateral ventral striatum. Furthermore, ventral striatal activations for wins versus losses were stronger during active than vicarious playing (**Table 3**). Using similar masking procedure as above, we found that even though wins evoked relatively stronger responses than losses during active gameplay, both events evoked BOLD signal decreases relative to the active gameplay baseline. Next we used detailed region-of-interest analyses as described below to decompose these effects.



**Table 2 | Brain regions responding to vicarious versus active playing (pooled over both win and loss events).**

Region	Hemisphere	Voxels	Peak T	MNI coordinates		
				x	y	z
Striatum (Put, Caud)	R	1553	10.99	18	20	-4
Striatum (Put, Caud)	L	1067	10.70	-18	14	-4
Pre- and postcentral gyri (BA3/4)	L	361	7.92	-34	-20	50
Midbrain (incl. VTA/SN)	L/R	560	7.10	2	-18	-10
Inferior parietal lobule (BA40)	R	93	6.96	54	-26	40
Inferior temporal gyrus (BA37)	L	125	6.95	-48	-70	0
Inferior parietal lobule (BA40)	L	54	6.32	-34	-48	46
Precuneus	R	66	6.03	28	-62	42

Activation clusters were thresholded at FWE-corrected  $P < 0.05$  (minimum cluster size 50 contiguous voxels).

**Table 3 | Brain regions showing statistically significant activation clusters after small-volume correction for all regions of interest.**

Region	Hemisphere	Voxels	Peak T	MNI coordinates		
				x	y	z
<b>WIN &gt; LOSS</b>						
omPFC (medial frontal gyrus)	L/R	109	6.96	-2	56	-4
Striatum (Put, Caud)	R	362	5.28	24	10	-6
Striatum (Put, Caud)	L/R	246	5.08	-24	8	-4
<b>WIN &gt; LOSS × PLAY &gt; WATCH</b>						
Striatum (Put, Caud)	R	164	4.30	20	22	2
Striatum (Put, Caud)	L	72	3.81	-22	10	-4

Data were thresholded at FWE-corrected  $P < 0.05$  ( $P < 0.001$  at voxel-level).

### REGION-OF INTEREST ANALYSIS IN THE REWARD CIRCUIT

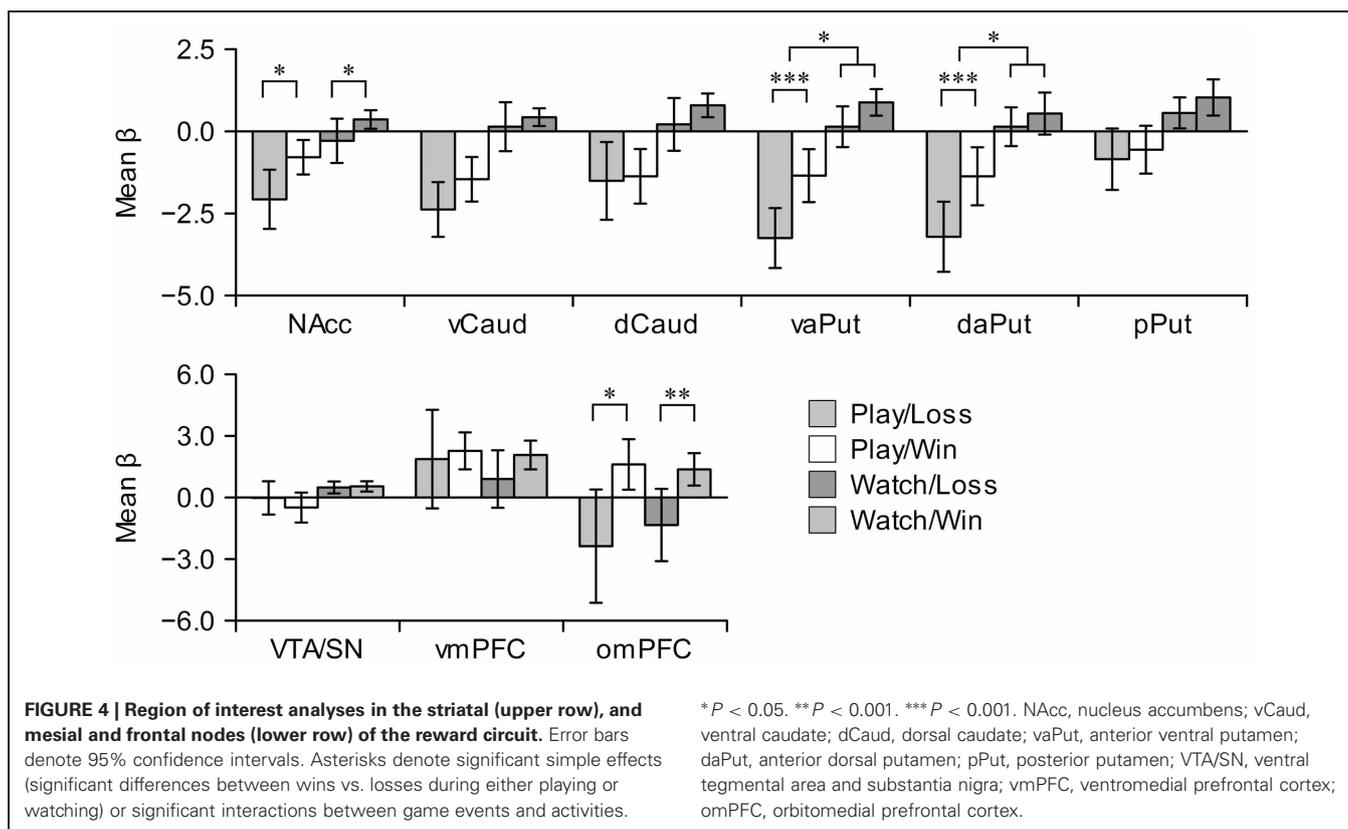
We calculated mean beta values in our *a priori* ROIs and subjected them to analyses of variance (ANOVAs). First, we used omnibus analysis with 9 (Region: all striatal, frontal, and mesial ROIs) × 2 (Activity: playing, watching) × 2 (Event: win, loss) repeated-measures ANOVA to confirm that the following interactions with region were statistically significant: region × activity ( $F_{(8, 80)} = 12.08$ ,  $P < 0.001$ ,  $\eta^2 = 0.07$ ), region × event [ $F_{(8, 80)} = 8.89$ ,  $P < 0.001$ ,  $\eta^2 = 0.06$ ], and region × activity × event [ $F_{(8, 80)} = 3.45$ ,  $P = 0.002$ ,  $\eta^2 = 0.01$ ]. To break down these regional interactions, we conducted 2 (Activity) × 2 (Event) repeated-measures ANOVAs separately in all regions.

**Figure 4** shows mean beta responses for win and loss events during active and vicarious playing conditions in all ROIs. Individual bar plots illustrate the activation directions (i.e., activations or deactivations) during win and loss events, and asterisks highlight significant differences between wins and losses. Wins versus losses evoked significantly greater effects, regardless of activity, in NAcc [playing:  $F_{(1, 10)} = 6.01$ ,  $P = 0.03$ ,  $\eta^2 = 0.34$ ; watching:  $F_{(1, 10)} = 6.14$ ,  $P = 0.03$ ,  $\eta^2 = 0.31$ ] and omPFC [playing:  $F_{(1, 10)} = 8.85$ ,  $P = 0.014$ ,  $\eta^2 = 0.69$ ; watching:  $F_{(1, 10)} = 24.77$ ,  $P = 0.001$ ,  $\eta^2 = 0.71$ ]. In contrast, wins versus losses evoked significantly greater effects only during active playing in vaPut [ $F_{(1, 10)} = 44.22$ ,  $P < 0.001$ ,  $\eta^2 = 0.77$ ] and daPut [ $F_{(1, 10)} = 70.08$ ,  $P < 0.001$ ,  $\eta^2 = 0.81$ ]. Interaction between action and event reached statistical significance in

vaPut [ $F_{(1, 10)} = 8.09$ ,  $P = 0.02$ ,  $\eta^2 = 0.03$ ] and daPut [ $F_{(1, 10)} = 13.35$ ,  $P = 0.004$ ,  $\eta^2 = 0.04$ ]. The main effect of activity was significant in all striatal regions ( $F_s > 11.45$ ,  $P_s < 0.007$ ,  $\eta^2 > 0.38$ ) and, as can be seen in **Figure 4**, this clearly resulted from deactivations during active playing. Similar trend was evident also in VTA/SN [ $F_{(1, 10)} = 5.08$ ,  $P = 0.048$ ,  $\eta^2 = 0.23$ ]. Taken together, these results demonstrate that although both win and loss events elicited deactivations in the striatum during active playing, activations in NAcc and aPut (both vaPut and daPut) returned closer to baseline levels during win events; furthermore, wins versus losses evoked greater activation changes in aPut during active than vicarious playing. Inspection of individual mean beta responses demonstrated that the latter result was robust; that is, mean beta responses to wins versus losses in aPut were greater during active than vicarious playing with nine out of eleven participants. In contrast to the aPut region, omPFC showed greater activations during win events regardless of active and vicarious playing.

### CORRELATIONS BETWEEN BEHAVIORAL AND fMRI RESPONSES

We predicted that the players' self-evaluations for pleasantness of wins versus losses during active versus vicarious playing would be associated with the corresponding BOLD signal changes in the striatum. To test this hypothesis, we calculated differences between wins and losses during active versus vicarious playing [i.e., contrast "(play: win > loss) > (watch: win > loss)"] both



for pleasantness ratings and mean beta values. Contrary to our predictions, no statistically significant correlations between these variables were found in any striatal region ( $R_s < 0.51$ ;  $P_s > 0.11$ ). Similarly, we failed to find statistically significant correlations between pleasantness ratings and mean beta values for wins versus losses, pooled over active and vicarious playing, in either of the frontal ROI ( $R_s < 0.18$ ,  $P_s > 0.59$ ).

One possibility for explaining the systematic deactivations in midbrain and striatum during active versus vicarious playing (cf. **Figure 4**) is that their activations remained elevated throughout the active gameplay due to anticipatory or hedonic reward processing but returned closer to baseline levels during both win and loss events. To test this, we calculated difference scores between active and vicarious playing for positive and negative affect measures, and compared them against mean beta values for active versus vicarious playing (pooled over wins and losses) in our predefined ROIs. Consistently, positive affect difference scores showed a significant correlation with deactivation strengths in VTA/SN, dCaud, and vaPut, and a marginally significant correlation with deactivations in vCaud (**Table 4**); all of these effects were large ( $R > 0.50$ ). When calculated separately, correlation coefficients in these ROIs were more negative during active than vicarious playing (i.e.,  $R_{\text{play}} < R_{\text{watch}}$ ). Bivariate scatter plots for these correlations are shown in **Figure 5**. In other words, the greater deactivations these regions exhibited during the win and loss gameplay events, the higher positive affects the players reported after active than vicarious playing.

## DISCUSSION

In the present investigation we studied fMRI responses to win and loss gameplay events (relative to activation levels during generic video game playing) during active and vicarious gameplay. Our results revealed two main effects in the striatum. First, replicating similar previous findings (Mathiak et al., 2011), both win and loss events evoked *deactivations* with respect to generic gameplay levels during active but not during vicarious playing. Second, in addition to this main effect of gameplay activity, win events evoked higher activation levels (i.e., weaker deactivations during active playing and stronger activations during vicarious playing) than loss events. Furthermore, our results showed an interaction between these two effects; that is, activation changes due to wins versus losses in the striatum, particularly in the anterior putamen, were larger during active than vicarious playing. This interaction effect demonstrates for the first time that winning versus losing in a complex video game evokes stronger effects in the striatum during active than vicarious gameplay. This finding is consistent with both animal electrophysiology (Kawagoe et al., 1998; Schultz et al., 2000) and human neuroimaging (Elliott et al., 2004; O'Doherty et al., 2004; Tricomi et al., 2004; Zink et al., 2004; Guitart-Masip et al., 2011), showing that striatal reward responses depend critically on the recipients' own actions. These previous studies have employed simple tasks where rewards were associated with specific motor actions (e.g., pressing one of two buttons), whereas the present study extends these findings by demonstrating action-reward contingency in the striatum during a complex, ecologically valid task (video

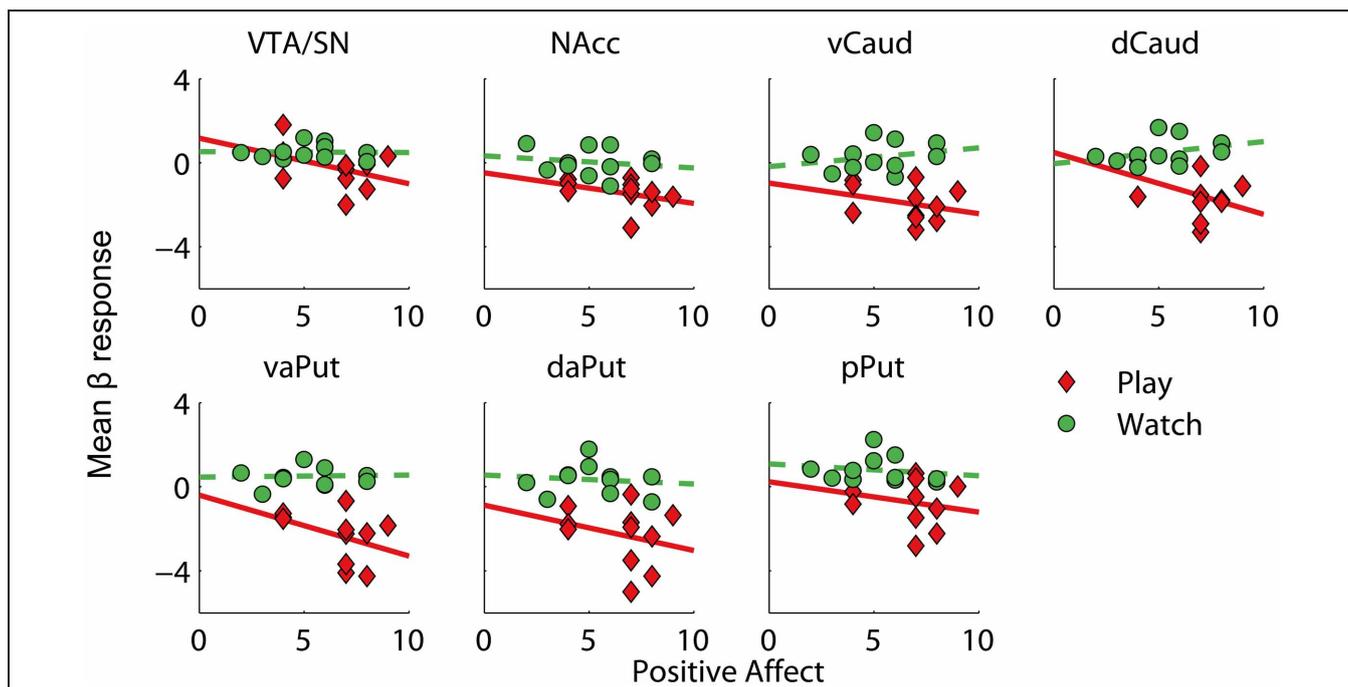
**Table 4 | Correlations between difference scores (active minus vicarious playing) for positive and negative affect measures, and mean beta values (active versus vicarious playing) in mesial and striatal regions ( $R_{\text{Play-Watch}}$ ).**

Region	NA				PA			
	$R_{\text{Play-Watch}}$	$R_{\text{Play}}$	$R_{\text{Watch}}$	$P_{\text{Play-Watch}}$	$R_{\text{Play-Watch}}$	$R_{\text{Play}}$	$R_{\text{Watch}}$	$P_{\text{Play-Watch}}$
NAcc	-0.06	0.05	-0.27	0.872	-0.60	-0.58	-0.09	0.054
vCaud	-0.04	-0.08	-0.26	0.914	-0.63	-0.27	0.18	0.037†
dCaud	0.12	0.21	-0.17	0.726	-0.81	-0.35	0.40	0.002*
vaPut	0.12	-0.19	0.27	0.715	-0.74	-0.50	-0.04	0.010*
daPut	0.18	-0.31	0.52	0.607	-0.45	-0.21	-0.18	0.163
pPut	0.15	-0.35	0.48	0.665	-0.48	-0.22	-0.34	0.132
VTA/SN	0.53	0.17	0.78	0.094	-0.70	-0.15	0.01	0.016*

Separate correlation coefficients for active and vicarious conditions ( $R_{\text{Play}}$  and  $R_{\text{Watch}}$ ) are also displayed to facilitate the interpretation of difference score correlations. Significance values are from Spearman's rank correlation tests.

† $P < 0.05$  (uncorrected).

\* $P < 0.05$  (FDR-corrected).



**FIGURE 5 | Bivariate scatter plots for Positive Affect evaluations versus mean beta responses in regions of interest, for active and vicarious playing conditions.** The solid and dashed lines depict best linear fits to the

data. VTA/SN, ventral tegmental area and substantia nigra; NAcc, nucleus accumbens; vCaud, ventral caudate; dCaud, dorsal caudate; vaPut, anterior ventral putamen; daPut, anterior dorsal putamen; pPut, posterior putamen.

game playing) that simulates free-ranging human motivated behavior.

We were also able to dissociate the coding of actively versus passively obtained rewards in the striatum and frontal cortex: whereas the anterior putamen was more sensitive to wins than losses only during active gameplay, the omPFC showed stronger activation to winning than losing during both active and vicarious playing. Action-independent reward activations in the omPFC have been observed previously in both animal (Schultz et al., 2000) and human neuroimaging studies (Elliott et al., 2004).

Given that the win and loss events were associated with external monetary rewards and punishments, the omPFC activations are also consistent with the known role of omPFC in processing monetary gains and other secondary rewards (Xue et al., 2009). However, nucleus accumbens in the striatum also showed greater activations to wins than losses during both active and vicarious gameplay. It is possible that, unlike the anterior putamen, nucleus accumbens was generally sensitive to receiving rewards similarly as omPFC. The dissociable response patterns of anterior putamen and nucleus accumbens results could stem

from the different connectivity patterns of ventromedial striatum (including nucleus accumbens) and dorsolateral striatum (including putamen): whereas the ventromedial striatum receives visceral afferents, the more dorsolateral regions are connected predominantly with higher-order associative and sensorimotor regions (Voorn et al., 2004).

The observed striatal deactivations during both wins and losses during active playing significantly extended our previous findings (Kätsyri et al., 2012). Although similar striatal deactivations have been observed previously (Mathiak et al., 2011), reward circuit deactivations associated with rewarding gameplay events nevertheless warrant consideration. One possible explanation is that the striatum showed tonic activations when the player was actively competing against his opponent, and that these activations returned closer to baseline levels whenever a break in the game restrained him from pursuing this goal; that is, both after he became incapacitated (loss events) and after he managed to eliminate his opponent (win events). Unfortunately, we were not able to test this hypothesis directly: as the strength of a raw BOLD signal is arbitrary, comparing the intercepts of independently scanned active and vicarious playing sessions would have been nonsensical. However, previous fMRI and PET studies have already demonstrated that active gameplay evokes tonic increases in striatal activations (Koepp et al., 1998; Hoefft et al., 2008), and one previous study has shown that active gameplay onsets and offsets evoke striatal fMRI activations and deactivations, respectively (Cole et al., 2012).

Above we suggested that striatal deactivations taking place at the times of wins and losses could be caused by tonic activation levels during generic gameplay, which returned closer to baseline levels when the gameplay activity was interrupted. Although this suggestion is speculative, there are at least two potential explanations for why video game playing would evoke tonic activations in the striatum. First, such activations, particularly during active playing, could reflect the inherently rewarding nature of playing *per se* (cf. Przybylski et al., 2010; see also Koepp et al., 1998). Our results tentatively support this view, given that the striatal and mesial deactivations caused by gameplay events (wins and losses) during active versus vicarious gameplay were correlated with the players' positive affect self-ratings for the corresponding whole sessions. Second, it is possible that the tonic striatal activations would reflect sustained anticipatory rather than hedonic reward processes—that is, 'wanting' rather than 'liking' components of reward (see Berridge, 2007; Diekhof et al., 2012). This is a plausible explanation, given that in our fast-paced video game (with 20–30 s. mean round durations; see **Table 1**), all activities following the onset of a new game round (i.e., finding and engaging the opponent) were ultimately associated with reward seeking. It is, however, uncertain why such anticipatory responses should be greater during active than vicarious playing. Furthermore, anterior putamen (during active playing) and nucleus accumbens (during both active and vicarious playing) were sensitive also to reward outcomes, as their responses were greater for wins than losses. The suggestions on striatal responses to anticipated and obtained rewards are not mutually exclusive. In fact, a recent meta-analysis of brain imaging studies demonstrated that the ventral striatum, unlike mPFC, is

sensitive to both anticipated and received rewards (Diekhof et al., 2012).

In addition to affective evaluations, active and vicarious playing conditions evoked differential spatial presence and flow experiences. Spatial presence has been associated with activations in a wide network including the ventral visual stream, the parietal cortex, the premotor cortex, and the brainstem (Jäncke et al., 2009). Interestingly, our results demonstrated that in addition to the striatum, also these regions showed strong deactivations during win and loss events during active playing (cf. **Table 2**). Hence, it is possible that also the network contributing to the experience of spatial presence showed tonic activations during active playing, which returned to baseline levels after win and loss events. Nevertheless, it is clear that future studies are needed for disseminating the tonic and phasic fMRI activations and their behavioral correlates (e.g., spatial presence) during video game playing.

In line with the attribution theory (Weiner, 1985), players' self-ratings confirmed that losses were experienced as more unpleasant during active than vicarious playing, even though the external monetary rewards and punishments for wins and losses were identical during active and vicarious playing conditions. The perceived pleasantness of win events during active playing was also linked to individual differences in appetitive motivation (i.e., tendency for fun seeking). Our results nevertheless did not provide evidence for associations between players' pleasantness self-ratings and their fMRI responses to win and loss events in general, or between the active and vicarious playing activities. However, it should be noted that players made only two evaluations with respect to all the win and loss events of a game, respectively, and it is possible that such overall evaluations may not have been as accurate as *post-hoc* evaluations for all game events would have been. In future, this problem could be solved by showing participants video recording of their gameplay sessions, and asking them to continuously rate their emotional feelings during the gameplay; this technique has been proven successful for example when studying the brain basis of emotions elicited by movies (Nummenmaa et al., 2012) and already utilized in previous fMRI game studies (Klasen et al., 2008).

Our subjects used precision hand actions to manipulate the joystick, and thus it is critical to control for sensorimotor processes related to the acquisition of rewards, especially because the striatum is also involved in sensorimotor control over corrective hand movements (Siebner et al., 2001; Turner et al., 2003). This issue is particularly important for win and loss events, given that these events are typically followed by different changes in movement demands (e.g., continuation of gameplay vs. total immobility). To our knowledge, however, previous brain imaging studies have not explicitly tried to control for joystick movement confounds. Even after we included continuous confound regressors both for overall movements and for movement direction changes, our results clearly demonstrated similar striatal effects for video game playing events as previously reported (Klasen et al., 2012), implying that such results cannot be accounted for by sensorimotor effects. Nevertheless, the effects of varying movement demands following win and loss events could be studied more

explicitly in the future; for example, by manipulating whether the player is able to move after specific game events or not. Future studies with explicit focus on testing the role of reward anticipation versus reward reception in striatal responses should also be conducted. Such studies should utilize slower-paced video games with sufficiently long durations between critical actions (such as shooting) and their outcomes.

Although the present sample size was comparable to those of several recent fMRI studies utilizing video game stimuli (Mathiak and Weber, 2006; Weber et al., 2006; Mobbs et al., 2007; Ko et al., 2009; Mathiak et al., 2011; Klasen et al., 2012), future studies should consider using larger sample sizes to detect potentially more fine-grained differences between active and vicarious gameplay. We conducted retrospective power analysis for our data using G\*Power software (Faul et al., 2007) to estimate the minimum sample sizes that should be used in future within-subjects studies to detect similar effects with 80% statistical power (at 5% significance level). These calculations showed that five participants would be sufficient for detecting similar win versus loss responses in the ventral anterior putamen ( $M = 1.90$ ,  $SD = 0.95$ , and  $\gamma = 2.0$ ). However, to replicate the differential win versus loss responses during active versus vicarious playing in the same region, a larger sample of at least thirteen participants should be used ( $M = 1.16$ ,  $SD = 1.35$ ,  $\gamma = 0.86$ ). As the present study has demonstrated, automatic annotation of gameplay events allows easy acquisition of large datasets from naturalistic video game playing tasks.

## REFERENCES

- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57, 289–300.
- Berridge, K. C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology* 191, 391–431. doi: 10.1007/s00213-006-0578-x
- Berridge, K. C., and Kringelbach, M. L. (2008). Affective neuroscience of pleasure: reward in humans and animals. *Psychopharmacology* 199, 457–480. doi: 10.1007/s00213-008-1099-6
- Bressan, R., and Crippa, J. (2005). The role of dopamine in reward and pleasure behaviour: review of data from preclinical research. *Acta Psychiatr. Scand.* 111, 14–21. doi: 10.1111/j.1600-0447.2005.00540.x
- Carver, C. S., and White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS Scales. *J. Pers. Soc. Psychol.* 67, 319–333. doi: 10.1037/0022-3514.67.2.319
- Clarke, D., and Duimering, P. (2006). How computer gamers experience the game situation: a behavioral study. *Comput. Entertain.* 4:6. doi: 10.1145/1146816.1146827
- Cohen, J. (1992). A power primer. *Psychol. Bull.* 112, 155–159. doi: 10.1037/0033-2909.112.1.155
- Cohen, J., Cohen, P., West, S. G., and Aiken, L. S. (1983). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc. doi: 10.1037//0033-2909.112.1.155
- Cole, S. W., Yoo, D. J., and Knutson, B. (2012). Interactivity and reward-related neural activation during a serious videogame. *PLoS ONE* 7:e33909. doi: 10.1371/journal.pone.0033909
- Delgado, M. R. (2007). Reward-related responses in the human striatum. *Annal. N.Y. Acad. Sci.* 1104, 70–88. doi: 10.1196/annals.1390.002
- Diekhof, E. K., Kaps, L., Falkai, P., and Gruber, O. (2012). The role of the human ventral striatum and the medial orbitofrontal cortex in the representation of reward magnitude – an activation likelihood estimation meta-analysis of neuroimaging studies of passive reward expectancy and outcome processing. *Neuropsychologia* 50, 1252–1266. doi: 10.1016/j.neuropsychologia.2012.02.007
- Elliott, R., Newman, J. L., Longe, O., and Deakin, J. F. W. (2004). Instrumental responding for rewards is associated with enhanced neuronal response in subcortical reward systems. *Neuroimage* 21, 984–990. doi: 10.1016/j.neuroimage.2003.10.010
- Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G\*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Meth.* 39, 175–191. doi: 10.3758/BF03193146
- Guitart-Masip, M., Fuentemilla, L., Bach, D. R., Huys, Q. J. M., Dayan, P., Dolan, R. J., et al. (2011). Action dominates valence in anticipatory representations in the human striatum and dopaminergic midbrain. *J. Neurosci.* 31, 7867–7875. doi: 10.1523/JNEUROSCI.6376-10.2011
- Haber, S. N., and Knutson, B. (2010). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology* 35, 4–26. doi: 10.1038/npp.2009.129
- Han, D. H., Bolo, N., Daniels, M. A., Arenella, L., Lyoo, I. K., and Renshaw, P. F. (2010). Brain activity and desire for Internet video game play. *Compr. Psychiatry* 52, 88–95. doi: 10.1016/j.comppsy.2010.04.004
- Havranek, M., Langer, N., Cheetham, M., and Jäncke, L. (2012). Perspective and agency during video gaming influences spatial presence experience and brain activation patterns. *Behav. Brain Funct.* 8:34. doi: 10.1186/1744-9081-8-34
- Hikosaka, O., Bromberg-Martin, E., Hong, S., and Matsumoto, M. (2008). New insights on the subcortical representation of reward. *Curr. Opin. Neurobiol.* 18, 203–208. doi: 10.1016/j.conb.2008.07.002
- Hoefl, F., Watson, C. L., Kesler, S. R., Bettinger, K. E., and Reiss, A. L. (2008). Gender differences in the mesocorticolimbic system during computer game-play. *J. Psychiatr. Res.* 42, 253–258. doi: 10.1016/j.jpsychires.2007.11.010
- Hosokawa, T., and Watanabe, M. (2012). Prefrontal neurons represent winning and losing during competitive video shooting games between monkeys. *J. Neurosci.* 32, 7662–7671. doi: 10.1523/JNEUROSCI.6479-11.2012
- Ijsselstein, W., van den Hoogen, W., Klimmt, C., de Kort, Y., Lindley,

- C., Mathiak, K., et al. (2008). "Measuring the experience of digital game enjoyment," in *Proceedings of Measuring Behavior*, eds A. J. Spink, M. R. Ballintijn, N. D. Bogers, F. Grieco, L. W. S. Loijens, L. P. J. J. Noldus, et al. (Netherlands: Maastricht), 7–8.
- Jäncke, L., Cheetham, M., and Baumgartner, T. (2009). Virtual reality and the role of the prefrontal cortex in adults and children. *Front. Neurosci.* 3, 52–59. doi: 10.3389/neuro.01.006.2009
- Kätsyri, J., Hari, R., Ravaja, N., and Nummenmaa, L. (2012). The opponent matters: elevated fMRI reward responses to winning against a human versus a computer opponent during interactive video game playing. *Cereb. Cortex*. doi: 10.1093/cercor/bhs259
- Kawagoe, R., Takikawa, Y., and Hikosaka, O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. *Nat. Neurosci.* 1, 411–416. doi: 10.1038/1625
- Kelley, A. E. (2004). Ventral striatal control of appetitive motivation: role in ingestive behavior and reward-related learning. *Neurosci. Biobehav. Rev.* 27, 765–776. doi: 10.1016/j.neubiorev.2003.11.015
- Klasen, M., Weber, R., Kircher, T. T. J., Mathiak, K. A., and Mathiak, K. (2012). Neural contributions to flow experience during video game playing. *Soc. Cogn. Affect. Neurosci.* 7, 485–495. doi: 10.1093/scan/nsr021
- Klasen, M., Zvyagintsev, M., Weber, R., Mathiak, K. A., and Mathiak, K. (2008). Think aloud during fMRI: neuronal correlates of subjective experience in video games. *Lect. Notes Comput. Sci.* 5294, 132–138. doi: 10.1007/978-3-540-88322-7\_13
- Knutson, B., and Cooper, J. C. (2005). Functional magnetic resonance imaging of reward prediction. *Curr. Opin. Neurobiol.* 18, 411–417. doi: 10.1097/01.wco.0000173463.24758.f6
- Ko, C.-H., Liu, G.-C., Hsiao, S., Yen, J.-Y., Yang, M.-J., Lin, W.-C., et al. (2009). Brain activities associated with gaming urge of online gaming addiction. *J. Psychiatric Res.* 43, 739–747. doi: 10.1016/j.jpsychires.2008.09.012
- Koepf, M. J., Gunn, R. N., Lawrence, A. D., Cunningham, V. J., Dagher, A., Jones, T., et al. (1998). Evidence for striatal dopamine release during a video game. *Nature* 393, 266–268. doi: 10.1038/30498
- Koob, G. F., and Volkow, N. D. (2010). Neurocircuitry of addiction. *Neuropsychopharmacology* 35, 217–238. doi: 10.1038/npp.2009.110
- de Kort, Y. A. W., Ijsselstein, W. A., and Poels, K. (2007). "Digital games as social presence technology: development of the social presence in gaming questionnaire (SPGQ)," in *Proceedings of Presence* (Barcelona), 195–203.
- Lessiter, J., Freeman, J., Davidoff, J. B., and Keogh, E. (2001). A cross-media presence questionnaire: the ITC-sense of presence inventory. *Pre. Tele. Virt. Environ* 10, 282–298. doi: 10.1162/105474601300343612
- Lombard, M., and Ditton, T. (1997). At the heart of it all: the concept of presence. *J. Comput. Med. Commun.* 3. <http://jcmc.indiana.edu/vol3/issue2/lombard.html>. doi: 10.1111/j.1083-6101.1997.tb00072.x
- Lucas, K. (2004). Sex differences in video game play: a communication-based explanation. *Commun. Res.* 31, 499–523. doi: 10.1177/0093650204267930
- Lutz, K., Pedroni, A., Nadig, K., Luechinger, R., and Jäncke, L. (2012). The rewarding value of good motor performance in the context of monetary incentives. *Neuropsychologia* 50, 1739–1747. doi: 10.1016/j.neuropsychologia.2012.03.030
- Mathiak, K. A., Klasen, M., Weber, R., Ackermann, H., Shergill, S. S., and Mathiak, K. (2011). Reward system and temporal pole contributions to affective evaluation during a first person shooter video game. *BMC Neurosci.* 12:66. doi: 10.1186/1471-2202-12-66
- Mathiak, K., and Weber, R. (2006). Toward brain correlates of natural behavior: fMRI during violent video games. *Hum. Brain Mapp.* 27, 948–956. doi: 10.1002/hbm.20234
- Mazaika, P. K., Hoeft, F., Glover, G. H., and Reiss, A. L. (2009). "Methods and software for fMRI analysis for clinical subjects," in *Poster Session Presented at the Meeting of Human Brain Mapping 2009* (San Francisco, CA). doi: 10.1016/S1053-8119(09)70238-1
- Mobbs, D., Petrovic, P., Marchant, J. L., Hassabis, D., Weiskopf, N., Seymour, B., et al. (2007). When fear is near: threat imminence elicits prefrontal-periaqueductal gray shifts in humans. *Science* 317, 1079–1083. doi: 10.1126/science.1144298
- Nummenmaa, L., and Niemi, P. (2004). Inducing affective states with success–failure manipulations: a meta-analysis. *Emotion* 4, 207–214. doi: 10.1037/1528-3542.4.2.207
- Nummenmaa, L., Glerean, E., Viinikainen, M., Jääskeläinen, I., Hari, R., and Sams, M. (2012). Emotions promote social interaction by synchronizing brain activity across individuals. *PNAS* 109, 9599–9604. doi: 10.1073/pnas.1206095109
- O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr. Opin. Neurobiol.* 14, 769–776. doi: 10.1016/j.conb.2004.10.016
- O'Doherty, J. P., Dayan, P., Schultz, J., Deichmann, R., Friston, K., and Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304, 452–454. doi: 10.1126/science.1094285
- O'Doherty, J. P., Deichmann, R., Critchley, H. D., and Dolan, R. J. (2002). Neural responses during anticipation of a primary taste reward. *Neuron* 33, 815–826. doi: 10.1016/S0896-6273(02)00603-7
- Przybylski, A. K., Rigby, C. S., and Ryan, R. M. (2010). A motivational model of video game engagement. *Rev. Gen. Psychol.* 14, 154–166. doi: 10.1037/a0019440
- Przybylski, A. K., Ryan, R. M., and Rigby, C. S. (2009). The motivating role of violence in video games. *Pers. Soc. Psychol. Bull.* 35, 243–259. doi: 10.1177/0146167208327216
- Ravaja, N., Saari, T., Salminen, M., Laarni, J., and Kallinen, K. (2006). Phasic emotional reactions to video game events: a psychophysiological investigation. *Media Psychol.* 8, 343–367. doi: 10.1207/s1532785xmep0804\_2
- Ravaja, N., Turpeinen, M., Saari, T., Puttonen, S., and Keltikangas-Järvinen, L. (2008). The psychophysiology of James Bond: phasic emotional responses to violent video game events. *Emotion* 8, 114–120. doi: 10.1037/1528-3542.8.1.114
- Ryan, R. M., and Deci, E. L. (2000). Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemp. Edu. Psychol.* 25, 54–67. doi: 10.1006/ceps.1999.1020
- Ryan, R. M., Rigby, C. S., and Przybylski, A. (2006). The motivational pull of video games: a self-determination theory approach. *Motiv. Emot.* 30, 344–360. doi: 10.1007/s11031-006-9051-8
- Savitzky, A., and Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 36, 1627–1639. doi: 10.1021/ac60214a047
- Schultz, W. (2004). Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology. *Curr. Opin. Neurobiol.* 14, 139–147. doi: 10.1016/j.conb.2004.03.017
- Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol.* 57, 87–115. doi: 10.1146/annurev.psych.56.091103.070229
- Schultz, W., Tremblay, L., and Hollerman, J. R. (2000). Reward processing in primate orbitofrontal cortex and basal ganglia. *Cereb. Cortex* 10, 272–284. doi: 10.1093/cercor/10.3.272
- Siebnér, H. R., Limmer, C., Peinemann, A., Bartenstein, P., Drzezga, A., and Conrad, B. (2001). Brain correlates of fast and slow handwriting in humans: a PET-performance correlation analysis. *Eur. J. Neurosci.* 14, 726–736. doi: 10.1046/j.0953-816x.2001.01694.x
- Steele, J. D., and Lawrie, S. M. (2004). Segregation of cognitive and emotional function in the prefrontal cortex: a stereotactic meta-analysis. *Neuroimage* 21, 868–875. doi: 10.1016/j.neuroimage.2003.09.066
- Tricomi, E. M., Delgado, M. R., and Fiez, J. A. (2004). Modulation of caudate activity by action contingency. *Neuron* 41, 281–92. doi: 10.1016/S0896-6273(03)00848-1
- Turner, R. S., Desmurget, M., Grethe, J., Crutcher, M. D., and Grafton, S. T. (2003). Motor subcircuits mediating the control of movement extent and speed. *J. Neurophysiol.* 90, 3958–3966. doi: 10.1152/jn.00323.2003
- Voorn, P., Vanderschuren, L. J., Groenewegen, H. J., Robbins, T. W., and Pennartz, C. M. (2004). Putting a spin on the dorsal-ventral divide of the striatum. *Trends Neurosci.* 27, 468–474. doi: 10.1016/j.tins.2004.06.006
- Weber, R., Ritterfeld, U., and Mathiak, K. (2006). Does playing violent video games induce aggression? Empirical evidence of a functional magnetic resonance imaging study. *Media Psychol.* 8, 39–60. doi: 10.1207/S1532785XMEP0801\_4
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychol. Rev.* 92, 548–573. doi: 10.1037/0033-295X.92.4.548
- Xue, G., Lu, Z., Levin, I. P., Weller, J. A., Li, X., and Bechara, A.

- (2009). Functional dissociations of risk and reward processing in the medial prefrontal cortex. *Cereb. Cortex* 19, 1019–1027. doi: 10.1093/cercor/bhn147
- Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., Chappelow, J. C., and Berns, G. S. (2004). Human striatal responses to monetary reward depend on saliency. *Neuron* 42, 509–517. doi: 10.1016/S0896-6273(04)00183-7
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 28 March 2013; accepted: 28 May 2013; published online: 13 June 2013.
- Citation: Kätsyri J, Hari R, Ravaja N and Nummenmaa L (2013) Just watching the game ain't enough: striatal fMRI reward responses to successes and failures in a video game during active and vicarious playing. *Front. Hum. Neurosci.* 7:278. doi: 10.3389/fnhum.2013.00278
- Copyright © 2013 Kätsyri, Hari, Ravaja and Nummenmaa. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



# Context counts! social anxiety modulates the processing of fearful faces in the context of chemosensory anxiety signals

Dirk Adolph<sup>1\*</sup>, Lukas Meister<sup>2</sup> and Bettina M. Pause<sup>2</sup>

<sup>1</sup> Department of Psychology, Ruhr-University, Bochum, Germany

<sup>2</sup> Department of Experimental Psychology, Heinrich-Heine-University, Düsseldorf, Germany

## Edited by:

Martin Klasen, RWTH Aachen University, Germany

## Reviewed by:

Ruthger Righart, Institute for Stroke and Dementia Research, Germany  
Caroline Huart, Université Catholique de Louvain, Belgium

## \*Correspondence:

Dirk Adolph, Department of Psychology, Clinical Psychology, and Psychotherapy, Ruhr-University, Universitätsstrasse 150, D-44801 Bochum, Germany  
e-mail: dirk.adolph@rub.de

During emotion perception, context is an important source of information. Whether contextual cues from modalities other than vision or audition influence the perception of social emotional information has not been investigated. Thus, the present study aimed at testing emotion perception and regulation in response to fearful facial expressions presented in the context of chemosensory stimuli derived from sweat of anxious individuals. In groups of high (HSA) and low socially anxious (LSA) participants we recorded the startle reflex (Experiment I), and analysed event-related potentials (ERPs; Experiment II) while they viewed anxious facial expressions in the context of chemosensory anxiety signals and chemosensory control stimuli. Results revealed that N1/P1 and N170 amplitudes were larger while late positive potential (LPP) activity was smaller for facial expressions presented in the context of the anxiety and the chemosensory control stimulus as compared to facial expressions without a chemosensory context. Furthermore, HSA participants were highly sensitive to the contextual anxiety signals. They showed enhanced motivated attention allocation (LPP, Study II), as well as larger startle responses toward faces in the context of chemosensory anxiety signals than did LSA participants (Study I). Chemosensory context had no effect on emotion regulation, and both LSA and HSA participants showed effective emotion regulation (Study I and II). In conclusion, both anxiety and chemosensory sport context stimuli enhanced early attention allocation and structural encoding, but diminished motivated attention allocation to the facial expressions. The current results show that visual and chemosensory information is integrated on virtually all levels of stimulus processing and that socially anxious individuals might be especially sensitive to chemosensory contextual social information.

**Keywords:** chemosensory anxiety signals, context, emotion perception, emotion regulation, social anxiety, event-related-potentials, startle reaction

## INTRODUCTION

In social perception, context is an important source of information. For example, in everyday life the perception of a facial expression is almost always accompanied by a diverse range of contextual information, helping people to extract the social meaning of the situation (Aviezer et al., 2008). Accordingly, several studies have demonstrated that emotional context-information affects face perception and accompanying emotional responses (Kim et al., 2004; Schwarz et al., 2012). Although in everyday life a wide variety of contextual information is available, most studies investigating context effects uses visual or acoustic context information only. Whether also cues from other modalities influence the perception of emotional facial expressions has yet to be determined.

For example, chemosensory signals have been shown to modulate a wide variety of emotional responses, and its processing has been demonstrated to be largely independent of the allocation of attentional resources (Pause, 2012). Despite this, knowledge about their potency as context signals is rare. Thus, the main aim

of the present study was to use chemosensory signals as context cues to investigate their impact on emotional responding toward facial expressions.

In detail, to investigate the potency to modulate emotional responding of the context stimulus, we measured defensive motivation during the perception of an anxiety related chemosensory context stimulus while viewing a fearful facial expression by means of the startle reflex (Experiment I). To elucidate the time course of central nervous processing event-related potentials (ERPs) in response to the face stimuli were analyzed (Experiment II). Because in everyday life people must often control their emotional responses to effectively adjust to the social environment (Gross et al., 2006), we investigated if the chemosensory context has an influence on the outcome of an emotion regulation task (Experiments I and II).

Emotion regulation refers to the extrinsic and intrinsic processes responsible for monitoring, evaluating, and modifying emotions (Thompson, 1994). One example is the cognitive reappraisal of emotion eliciting situations (Gross, 2002). Self-reported

emotions, as well as physiological responses to threatening pictures including heart rate and electrodermal activity (Gross, 2002), brain electrical activity (Moser et al., 2009), neuronal responses in the amygdala (Ochsner et al., 2002), and the affect modulated startle-reflex (Jackson et al., 2000) can be significantly enhanced or reduced using cognitive reappraisal. To date, no study has tested whether chemosensory context information influences people's ability to regulate emotional responses.

However, beside emotion regulation, several studies indicate that, comparable to visual stimuli, also the delivery of social chemosensory information can alter emotional responses. For example, chemosensory signals of anxiety alter emotion related neuronal activity (Prehn-Kristensen et al., 2009) and enhance withdrawal related motivation (i.e., the startle reflex) in human perceivers (Prehn et al., 2006). Initial evidence suggests that chemosensory stimuli may also constitute powerful context cues for face perception (Li et al., 2007). Contextual chemosensory anxiety signals diminish the perceptual acuity of visual safety cues (happy facial expressions) (Pause et al., 2004), while the perceptual acuity of fear from ambiguous facial expression (morphs between happy and fearful facial expressions) is enhanced (Zhou and Chen, 2009). Moreover, motivated attention, as indicated by the late positive potential (LPP) within the ERP, directed toward neutral faces is enhanced when they are presented in the context of chemosensory stress signals (Rubin et al., 2012).

In social perception—including face perception and the perception of social chemosensory signals—social anxiety plays a modulating role. Social anxiety is characterized by abnormal processing of social threat information, involving processing biases in attending to, interpreting and remembering it (Hirsch and Clark, 2004). Accordingly, socially anxious individuals show deviant processing of single social fear relevant cues, including chemosensory anxiety signals, with examples being enhanced startle reactivity (Pause et al., 2009) and faster processing of chemosensory anxiety signals than non-anxious individuals (Pause et al., 2010). This suggests an attentional bias comparable to that observed with pictorial stimuli (see for example Kolassa and Miltner, 2006; Mühlberger et al., 2009).

Interestingly, social anxiety may also play a mediating role in the processing of threatening contextual information accompanying face perception, as it has been shown that threatening semantic information about a target facial expression enhances emotional responding in socially anxious individuals (Schwarz et al., 2012). Thus, converging evidence suggest that socially anxious individuals show sustained sensitivity toward various kinds of social signals, including chemosensory and visual signals of anxiety, as well as emotional context information. Therefore, in the present study, we compared a group of low socially anxious individuals (LSA) with a group of high socially anxious individuals (HSA).

## THE PRESENT STUDY

Two experiments assessed emotional reactivity and emotion regulation toward anxious facial expressions in the context of social chemosensory signals or control stimuli. The social chemosensory stimuli were chosen to be either congruent with the foreground picture stimulus (chemosensory anxiety signals derived

from donors in an anxiety provoking situation) or incongruent with the facial stimulus (chemosensory exercise control stimuli). Furthermore, the anxiety signals were derived in a natural anxiety provoking situation, that is, the waiting for an oral examination at the university to reach an academic degree.

As dependent measures, the present study assessed the time course of stimulus processing including its early perceptual and attention sensitive (N1/P1), face specific (N170) and late motivational attention-related components (LPP), as well as the motivational/behavioral relevance of the stimuli (startle reflex). The startle reflex can be considered as a direct readout of activation of a defense system responsible to protect the organism from threat (see Bradley et al., 2001). The startle response is potentiated with the presence of a threatening stimulus (e.g., Bradley et al., 2001), and this potentiation has been argued to mirror the switch from orientation toward a meaningful stimulus to defense motivation as described by Sokolov (1963). In this light, startle potentiation can be characterized as the effect of motivational priming for action reflecting the defense system's general behavioral mobilization (Lang et al., 1997).

Study I assessed the motivational relevance of chemosensory context stimuli by withdrawal related motor behavior, emotional reactivity and emotion regulation using the startle reflex. Because startle responses elicited after target stimulus offset can reliably distinguish controls from phobics (Globisch et al., 1999), in the present study startle probes were elicited during and after the presentation of the target stimuli. If HSA participants are indeed especially sensitive to social stimuli, they should show sustained responsiveness even after stimulus offset. Study II assessed the time course of stimulus processing with ERPs. In detail, an ERP component related to the early structural encoding of the stimuli (N1/P1, N170), as well as a component associated with the allocation of motivated attention (LPP) toward the stimuli were assessed. Previous research has verified that ERPs to emotional facial expressions are sensitive to the modulational effects of contextual information (e.g., the N170, see Righart and De Gelder, 2006, 2008).

We hypothesized that participants would show enhanced processing of faces in the context of chemosensory anxiety signals by showing larger startle amplitudes, enhanced early and late brain potentials, as well as less effective emotion regulation. We also predicted, that this effect would be most pronounced in socially anxious individuals.

## EXPERIMENT I

### METHODS

#### Participants

Forty non-smoking female students participated in the study. They were recruited from the Heinrich-Heine-University of Düsseldorf, reported having a regular menstrual cycle, not using any medication including oral contraceptives and not suffering from mental and physical diseases. In addition none of them suffered from general hyposmia (three alternative forced choice test including one bottle with phenyl-ethyl-alcohol, 1:100, diluted in 1, 2-Propanediol, and two bottles containing the non-odorous solvent only). They were classified as either LSA (LSA,  $n = 20$ ) or HSA (HSA,  $n = 20$ ) based on their trait social anxiety scores

(SIAS, Stangier et al., 1999). Participants scoring 22 or higher ( $>1.5$  SD above the mean of the standard sample) were defined as HSA, those scoring 16 and lower ( $<0.5$  SD above the mean of the standard sample) were defined as LSA. As a result, HSA participants scored well above the suggested cut-off score of 30 for social phobia ( $M = 34.05$ ,  $SD = 9.12$ ), while mean scores for the LSA group were within the normal range ( $M = 10.0$ ,  $SD = 3.96$ ; group comparison:  $p < 0.001$ ). In contrast, HSA participants scored within the normal range on both trait anxiety ( $M = 42.05$ ,  $SD = 9.12$ , STAI-X2, Laux et al., 1981) and self-reported depressive feelings [ $M = 8.51$ ,  $SD = 5.66$  Depressions Skala (DS); Von Zerssen and Koeller, 1976], while LSA participants scored low on both questionnaires (STAI:  $M = 35.60$ ,  $SD = 4.50$ ; DS:  $M = 3.75$ ,  $SD = 2.45$ ; group comparisons for both questionnaires:  $p < 0.01$ ). Both groups scored within the medium range for empathy (LSA:  $M = 30.10$ ,  $SD = 5.70$ ; HSA:  $29.20$ ,  $SD = 5.37$ ; Paulus, 2009), and the frequency of everyday-life use of reappraisal [LSA:  $M = 4.73$ ,  $SD = 0.86$ ; HSA:  $4.93$ ,  $SD = 0.73$ , Emotion Regulation Questionnaire (ERQ); Abler and Kessler, 2009]. The two groups did not differ for age,  $p > 0.10$  ( $M = 24.95$ ,  $SD = 5.73$ , range 19–45). All participants were paid for their participation and gave written informed consent. The study was approved by the ethics committee of the German Psychological Society (DGPs).

### Stimuli

**Chemosensory stimuli.** To collect the chemosensory stimuli, axillary sweat was sampled from 20 male students of European descent from the University of Düsseldorf<sup>1</sup>. All 20 donors donated sweat during both a natural anxiety provoking, and an exercise control condition (an important oral examination at the university, ergometer exercise). The donors' age ranged from 22 to 30 ( $M = 24.9$ ,  $SD = 2.5$ ). Their body mass index was within the normal range (range: 19.6–27.3,  $M = 23.2$ ,  $SD = 1.9$ ), and all reported to have a regular sleep-wake-cycle. All described themselves as healthy, especially with respect to hormonal, neurological, immunological, cardiological, and diseases of the axillae. They were within the normal range for trait anxiety (assessed with the STAI,  $M = 36.85$ ,  $SD = 7.04$ ). All participants donated sweat from both axillae for 90 min during an anxiety and a sport control donation situation using cotton pads (Ebelin Maxi Pads, dm-drugstore, Germany) while following a well-established sampling protocol (Pause et al., 2004, 2009; Prehn et al., 2006; Prehn-Kristensen et al., 2009). In detail, during an interview session, the donors gave written informed consent and were instructed to refrain from eating garlic, onions, asparagus, or spicy food, not to use deodorants and to wash their armpits exclusively with an unperfumed medical soap (Eubos®, Dr. Hobein GmbH, Germany) within 24 h prior to donation. In addition, to control for physiological arousal, the donors' heart rate was sampled during the interview session (baseline) and during the anxiety provoking and the sport control sampling condition using

a mobile pulse monitor (R4 Plus, Omron, Germany). The anxiety condition consisted of 90 min of waiting for an important oral examination at the university in order to assess an academic degree (subjective importance,  $M = 8.29$ ,  $SD = 0.87$ , scale range 0–10). Briefly before the donors entered the examination, they gave ratings of their current emotional state using the self-assessment manikin (SAM) (Bradley and Lang, 1994) (valence:  $-4$  to  $4$ , arousal:  $1-9$ , dominance:  $1-9$ ), and the intensities of the six basic emotions (Ekman and Friesen, 1971) (10 cm visual analogue scales). In addition, the donors' heart rate was recorded. The sport control condition consisted of ergometer exercise and took place on average 6 ( $SD = 4.13$ ) days after the anxiety condition, while the time of day was held constant (There was a mean difference of  $M = 83.75$ ,  $SD = 85.65$  minutes between the beginning of the anxiety and the sport control donation situation). To keep the physiological arousal comparable between the anxiety and the sport control condition, during ergometer exercise, donors' heart rate was held at the individual level that was recorded during the anxiety condition (using a mobile heart rate monitor, T 31, Polar, Germany). Briefly before the end of the sport control condition, the donors' emotional experience was assessed in the same way as in the anxiety condition.

During the anxiety condition, the donors described themselves as feeling more unpleasant, more aroused and less dominant (SAM ratings), as well as more anxious and less happy (basic emotions), than during the sport control condition (Table 1). There were no differences in ratings of disgust, sadness, surprise, or anger between the donation conditions. During the sport control condition the heart rate did not differ from the anxiety condition,  $p = 0.792$  (anxiety condition:  $M = 91.25$ ,  $SD = 22.07$  beats per minute, sport control condition,  $M = 90.95$ ,  $SD = 19.61$  beats per minute). However, both heart rates were higher than during baseline recording ( $M = 68.80$ ,  $SD = 11.22$ ), both  $p < 0.001$ .

After all participants finished the sport control condition, the sweat samples were pooled with distinction to the respective donation condition (sport, anxiety) and stored at  $-20^{\circ}\text{C}$ . The

**Table 1 | Sweat donors' self-reported intensities of basic emotions and SAM.**

	Anxiety condition	Sport-control condition	Comparison
	$M$ ( $SD$ )	$M$ ( $SD$ )	$P$
Anxiety	6.68(1.69)	0.46(0.58)	$<0.001$
Happiness	3.89(2.37)	7.09(2.26)	$<0.001$
Anger	1.74(1.42)	1.18(1.54)	n.s.
Disgust	0.90(1.28)	0.66(1.23)	n.s.
Sadness	1.78(1.91)	1.05(1.06)	n.s.
Surprise	3.04(2.18)	2.65(2.45)	n.s.
SAM <sub>valence</sub>	$-0.05$ (1.76)	2.20(1.28)	$<0.001$
SAM <sub>arousal</sub>	7.35(0.88)	3.85(1.60)	$<0.001$
SAM <sub>dominance</sub>	4.50(1.70)	6.70(1.31)	$<0.01$

Basic emotions range: 0–10 cm visual analogue scale, SAM<sub>valence</sub> range  $-4-4$ , SAM<sub>arousal</sub> and SAM<sub>dominance</sub> range: 1–9.

<sup>1</sup>In order to obtain homogenous chemosensory stimulus material the sweat samples were collected from males only. We have previously shown that the effects of chemosensory stimuli are independent of the sex of the donor (Pause et al., 2009).

quantity of the sweat samples was largely the same for both condition (62 g for the sport control condition, 65 g for the anxiety condition). For the experiment, the homogenized samples were divided into small portions (1.2 g each).

**Visual stimuli.** The visual stimulus material consisted of 14 pictures from 7 male actors (AM05, AM08, AM10, AM14, AM17, AM19, AM22) showing anxious facial expressions with averted gazes to the left and right and were selected from the Karolinska Directed Emotional Faces set (KDEF, Lundqvist et al., 1997)<sup>2</sup>. These stimuli have been shown to reliably elicit emotional responses in women (Adolph and Alpers, 2010).

### Stimulus presentation

The chemosensory stimuli were presented via a modified oxygen mask covering the nose and the mouth using a constant-flow (50 ml/s) 5-channel olfactometer (Lorig et al., 1999; Prehn-Kristensen et al., 2009) including five glass bottles. Two bottles were filled with 1.2 g of cotton pad (homogenized sweat samples either from the anxiety or the sport condition). The startle-eliciting stimulus was a 104 dB/A white noise burst (50 ms, rise-time <1 ms), presented through earplugs (ER4-14A Etymotic Research, USA), and calibrated using a high precision sound-level meter (Bruel & Kjaer, Denmark). Visual stimuli were shown on a 19" monitor in a visual angle of 27 by 22°. Stimulus timing was controlled with the Presentation® software (Version14, Neurobehavioral Systems, USA).

### Individual stimulus validation and odor detection session

For the present experiment participants needed to perceive the facial stimuli as fear inducing. This was tested in an individual session that took place no more than 7 days prior to the main experiment. During this session participants rated their emotional experience toward the pictures using the 6 basic emotions (10 cm visual analogue scales) and the valence and arousal scales of the SAM. As a result, for all participants, the most prominent reported emotion elicited by the pictures was anxiety ( $M = 5.53$ ,  $SD = 1.66$ ), differing significantly from the ratings of the other basic emotions (anger, disgust, sadness, happiness  $p < 0.01$ , surprise,  $p = 0.06$ ). Consequently, no participants had to be excluded. Furthermore, participants rated their own emotional experience toward the pictures as negative (SAM valence,  $M = -1.13$ ,  $SD = 1.01$ ) and mildly arousing (SAM arousal  $M = 4.79$ ,  $SD = 1.37$ ). No differences emerged between HSA and LSA individuals in basic emotions or feelings according to the SAM-scale (all  $p > 0.10$ ), with an exception for sadness [HSA > LSA,  $t_{(38)} = 2.13$ ,  $p = 0.04$ ].

During the same session participants also rated the chemosensory stimuli for intensity, pleasantness, unpleasantness, and familiarity (1 = not at all, 9 = extremely), and their own emotional experience toward the stimuli using the valence and arousal scales of the SAM. Results showed, that they perceived the

stimuli as moderately intense ( $M = 5.35$ ,  $SD = 1.68$ ), unpleasant ( $M = 4.64$ ,  $SD = 1.60$ ), and familiar ( $M = 4.74$ ,  $SD = 1.77$ ), and as low in pleasantness ( $M = 3.19$ ,  $SD = 1.43$ ). Participants rated their own emotional response toward the chemosensory stimuli as mildly negative and mildly arousing (SAM valence:  $M = -0.64$ ,  $SD = 1.15$ , SAM Arousal:  $M = 4.63$ ,  $SD = 1.51$ ). Ratings did not differ between chemosensory anxiety and sport stimuli (all  $p > 0.10$ ) and between high and LSA individuals (all  $p > 0.10$ ).

Of the participants twenty-six (65%) were able to differentiate the chemosensory stimuli from pure cotton pad (two correct detections for each stimulus within three-alternative forced choice tests including cotton pads from either condition, and two non-used cotton pads, all administered via the olfactometer for 2 s).

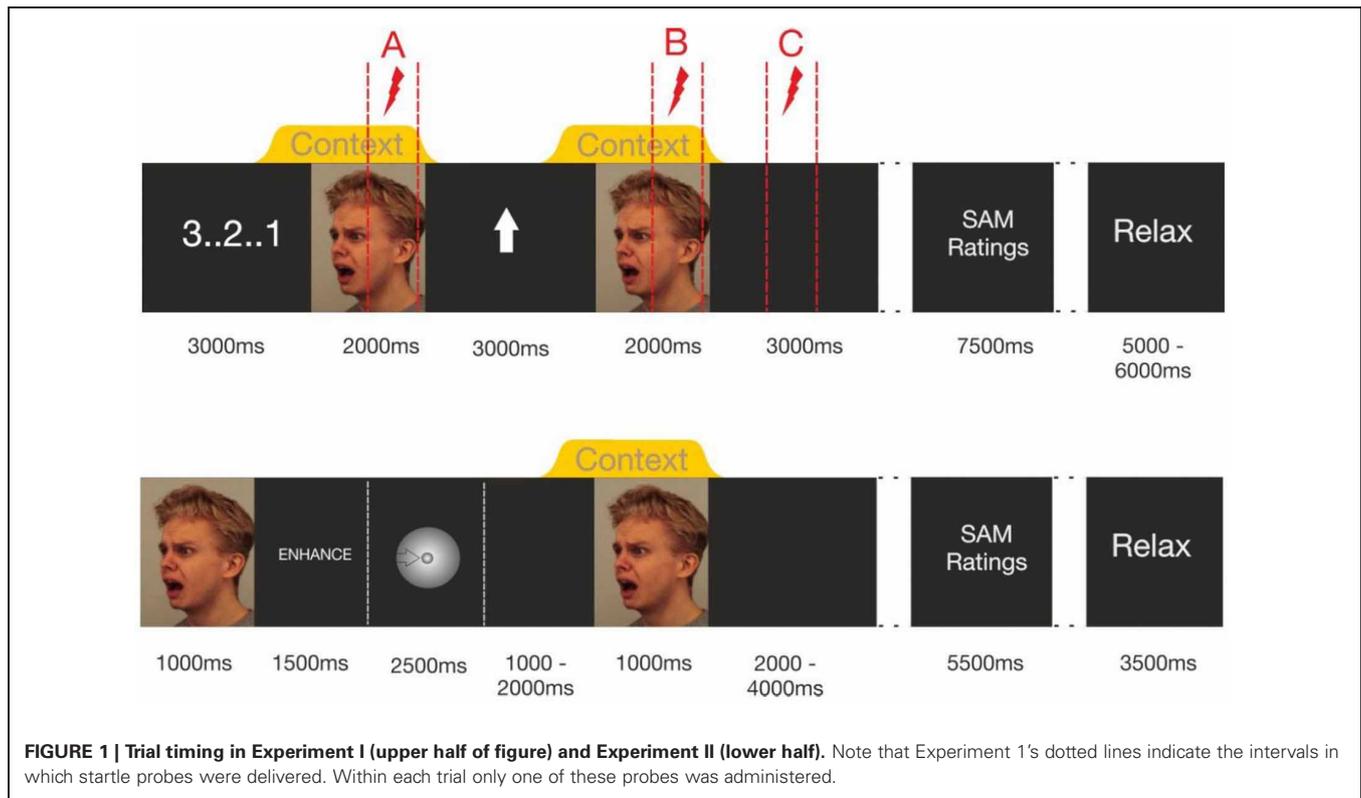
### Experimental session

The experimental session was largely identical to that used in a previous study (Adolph and Pause, 2012). In brief, participants were seated in front of a computer screen, electrodes were attached and the order of stimulus delivery was explained. After participants signaled understanding of stimulus timing, they received detailed instructions on how to breathe during stimulus delivery. Inhalation was monitored using breathing belts (see Data Recording). The session did commence only, after the participants could control their breathing completely. Participants were then given detailed instructions to use cognitive linguistic emotion regulation strategies as used in our previous study. After verbatim instructions, participants were given two practice examples of how to accomplish emotion regulation in response to the face stimuli.

Upon providing verbatim emotion regulation instructions, participants received detailed instructions on how to use the emotion regulation strategies within the experimental protocol. Only after participants signaled full understanding of the emotion regulation procedure, they practiced at least 10 learning trials of each condition. The data recording began, when the participants reported successful emotion regulation in each condition and full compliance to the emotion regulation protocol. Upon start of the first emotion regulation block, participants received 8 startle probes to induce startle habituation. The visual and chemosensory stimuli were then presented in four blocks, counterbalancing the emotion regulation strategy and the chemosensory context (enhance/anxiety; enhance/sport; down-regulate/anxiety; down-regulate/sport) across participants. Each block consisted of 21 trials, with 10-min breaks between the blocks. **Figure 1** shows the sequence of stimulus presentations.

At the beginning of each trial, a visual countdown was presented, followed by the presentation of a fearful facial expression (2000 ms baseline stimulus), in the context of either a chemosensory anxiety signal, or a chemosensory sport stimulus, beginning 500 ms before the onset of and co-terminating with the face stimulus. Then, a visual regulation cue was presented for 3000 ms, an upward-pointing arrow signaling "enhance," or a downward-pointing arrow signaling "down-regulate." The same picture-chemosensory context combination was then presented again (target stimulus). Then, a black screen (3000 ms) followed

<sup>2</sup>Stimuli with averted gaze were chosen because it has been demonstrated that averted fearful expression elicit more negative affect in the perceiver than fearful expressions with direct gaze (Hess et al., 2007).



by the valence and arousal scales of the SAM (3500 ms each with a 500 ms break between the scales) was presented. Thereafter, the participants were free to relax for an interval varying randomly between 5 and 6 s. To prevent habituation effects, two identical chemosensory stimuli never followed each other in consecutive trials. All 7 picture stimuli were delivered to all participants, and the sequence of stimuli was randomized across trials within one block of stimulus presentations.

During each trial one acoustic startle probe was presented to assess emotional responding. Shown in **Figure 1** (upper half) is that probes could occur at three different positions during the trials (Probe A = baseline: 1000–1900 ms after onset of the baseline stimulus, Probe B = target picture perception: 1000–1900 ms after the onset of the target stimulus, Probe C = target picture offset: 2000–2900 ms after the offset of the target stimulus) [see **Figure 1** (upper half)]. During each block, seven startle probes were presented at each probe position resulting in 21 startle-probes per block (enhance anxiety, enhance sport control, down-regulate anxiety, down-regulate sport control). Trials including different startle-probe positions were equally distributed within the blocks. Each stimulus was probed once at each probe position (A, B, C) in each block.

After the last emotion regulation block had ended participants completed a questionnaire in which they were asked to freely describe the strategies they used to enhance and down-regulate their emotions.

#### Data recording

The startle eyeblink was recorded from the orbicularis oculi muscle beneath the left eye using two Ag/AgCl electrodes (inner

diameter 5 mm). Participants breathing cycles were assessed with two respiration belts (Brain Products, Germany) placed around participants' abdomen and thorax. The physiological data were amplified (22bit Quick-Amp, Brain Products, Germany) and recorded with BrainVision Recorder Software (Brain Products, Germany), sampled at 2000 Hz, and filtered on-line using a 50 Hz notch filter. Off-line, the raw EMG was high- (28 Hz, 24 dB/octave) and low-pass filtered (500 Hz, 24 dB/octave) (Van Boxtel et al., 1998).

#### Data reduction

Of the eyeblink responses 3.3% were rejected because they were recorded neither during an increase in inhalation nor briefly (200 ms) after the inhalation maximum, 1.0% because the blink onset occurred during baseline. The remaining trials were rectified and smoothed (20 ms moving average). The startle data were baseline corrected (0–20 ms after startle probe onset), and the startle-response was scored as the maximum deflection within 30–150 ms after startle probe onset. Non-responses (amplitudes  $\leq 2 \times$  the largest amplitude within the baseline interval; 1.2% responses) were scored as 0. Outlier values differing more than two standard deviations from the condition average were excluded (1.7% of responses) (Blumenthal et al., 2005). Due to excessive differences in startle amplitude the startle-responses were z-standardized within each participant and across conditions.

#### Data analysis

For each picture-chemosensory stimulus combination, the baseline emotional response was assessed as the response to

Startle-Probe A within each of the four experimental block. In addition, the regulated emotional response was assessed as the response to startle probes B and C, resulting in one baseline emotional response, and two regulated emotional responses for each of the four blocks. Thus, analysis of variance (ANOVA) including one between subject factor, 2 (*Group*: HSA, LSA) and the within subjective factors, 2 (*Context*: anxiety, sport control)  $\times$  2 (*Emotion regulation*: enhance, down-regulate)  $\times$  3 (*Time*: A = baseline, B = target picture perception, C = target picture offset) were run. Statistical analyses were performed using SPSS 18, and Cohen's effect-size  $f$  was calculated. Huynh–Feldt corrections of degrees of freedom were applied, and corrected  $p$ -values are reported. Subsequent nested effects (Page et al., 2003) and  $t$ -tests were calculated. An alpha level of 5% was used for all statistical tests.

For SAM ratings 2 (*Group*)  $\times$  2 (*Context*)  $\times$  2 (*Emotion Regulation*) ANOVAs were run. Cohen's effect-size  $f$  was calculated, Huynh–Feldt corrections were applied, and corrected  $p$ -values are reported.

## RESULTS

### Emotion regulation strategies

The individual answers of the 40 participants to the post experimental questionnaire on emotion regulation strategies initially were classified by two independent raters. Overall agreement was high between the raters on the single emotion regulation strategies. To down-regulate, the majority of participants reported rationalizing or reinterpreting the expression ( $N = 17$ , 42.5% of all participants). For example, participants reported imagining the face as a comic strip or a photo, or imagined that the expression was triggered by something not dangerous. Most of the remaining participants reported focusing on possible positive aspects or outcome of an imagined situation corresponding to the expression ( $N = 16$ , 40% of all participants). For example, they imagined an assault but that the offender was arrested. Seven participants (3.8%) reported to use strategies other than these two. For example, that they were just trying to keep detached from the person on the photo.

To enhance their emotional response, the majority of participants tried to feel what the person feels on the photo ( $N = 24$ , 60% of all participants). Most of the remaining participants reported focusing on negative aspects or outcome of an imagined situation for the person on the photo ( $N = 14$ , 35% of all participants). For example, participants reported imagining being the victim of an assault together with the person on the photo. A total of 2 participants (5%) reported the use of other strategies, for example, to simply concentrate more on the respective expression. Results of  $\chi^2$  tests indicate that the frequency of use of the different regulation strategies did not differ between HSA and LSA participants (all  $p > 0.10$ ).

### Effects of chemosensory context

**Ratings.** Independent of the emotion regulation strategy, the participants felt more negative when the faces were presented along with a chemosensory anxiety cue, than with the chemosensory sport cue,  $F_{(1, 38)} = 8.07$ ,  $p = 0.007$ ,  $f = 0.46$  (Main effect for *Context*) (Table 2).

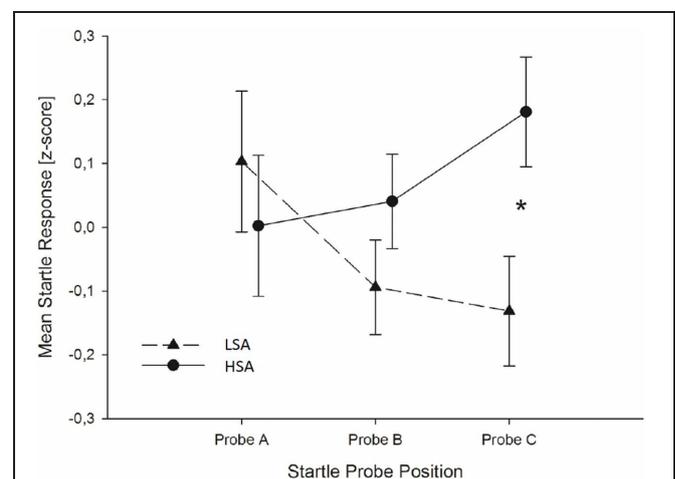
**Startle reflex.** Independently of the emotion regulation strategy, HSA individuals showed larger startle magnitudes toward the faces presented in the context of chemosensory anxiety cues than the LSA participants, especially toward startles presented at probe position C,  $F_{(2, 76)} = 3.36$ ,  $p = 0.040$ ,  $f = 0.30$  [Interaction *Context* by *Time* by *Group*, nested effects: *Group* by *Context* within probe C:  $F_{(1, 38)} = 4.69$ ,  $p = 0.037$ ; *Group* within probe C within HSA:  $F_{(1, 38)} = 6.60$ ,  $p = 0.014$ ] [Figure 2]. No differences between HSA and LSA participants occurred for faces presented in the context of chemosensory sport control stimuli ( $p > 0.10$ ).

To further explore the differences in startle responses toward faces presented in the context of anxiety signals, we calculated

**Table 2 | Emotion regulation effects on subjective ratings and the startle reflex toward anxious faces in the context of chemosensory anxiety or sport-control stimuli in Experiment I (Startle).**

	Chemosensory sport control		Chemosensory anxiety	
	Enhance	Down-regulate	Enhance	Down-regulate
	$M_{(SD)}$	$M_{(SD)}$	$M_{(SD)}$	$M_{(SD)}$
SAM <sub>valence</sub>	-1.70(0.78)	-0.36(1.10)	-1.80(0.72)	-0.45(1.08)
SAM <sub>arousal</sub>	6.14(1.20)	4.28(1.36)	6.16(1.20)	4.42(1.39)
Startle <sub>probeA</sub>	0.05(1.20)	-0.22(0.56)	0.10(0.68)	0.01(0.65)
Startle <sub>probeB</sub>	0.11(0.60)	-0.09(0.61)	0.05(0.57)	-0.10(0.51)
Startle <sub>probeC</sub>	0.24(0.74)	-0.19(0.61)	0.13(0.59)	-0.08(0.54)

SAM<sub>valence</sub> range: -4-4; SAM<sub>arousal</sub> range: 1-9, Startle responses are given as z-score.



**FIGURE 2 | In Experiment I, HSA participants (dashed lines) showed larger startle magnitudes than LSA participants (solid lines) toward the anxious facial expression presented in the context of the chemosensory anxiety signal during startle position C. Furthermore, for HSA participants startle responses did not differ between startle probes A, B, and C, while for LSA participants startle responses were smaller during startle probe B, and probe C, than probe A, indicating startle habituation in LSA, but not in HSA participants. \*0.040.**

habituation of startle responses for HSA and LSA individuals. Results show, that for LSA individuals startle responses elicited during the anxiety context habituated rapidly within the trials. In detail, startle responses elicited during baseline stimuli perception (probe A) were larger than those elicited during target stimulus perception (probe B),  $t_{(19)} = 2.022$ ,  $p = 0.029$ , one-tailed, and larger than those elicited during the late emotion regulation interval (probe C),  $t_{(19)} = 1.794$ ,  $p = 0.045$ , one-tailed. For socially anxious individuals startle responses elicited during the anxiety context did not habituate. That is, startle responses do not differ between baseline, target stimulus perception and late emotion regulation interval, all  $p > 0.10$ , one-tailed (see **Figure 2**).

### Effects of emotion regulation strategies

**Self-Ratings.** After down-regulating their emotions, participants rated their own emotional experience as neutral (**Table 2**) and significantly less negative than after enhancing their emotion,  $F_{(1, 38)} = 57.60$ ,  $p < 0.001$ ,  $f = 1.23$  (Main effect for *Emotion Regulation*). Similarly, after down-regulating, participants rated their experienced arousal as neutral and significantly lower, as after enhancing their emotion,  $F_{(1, 38)} = 68.61$ ,  $p < 0.001$ ,  $f = 1.34$  (Main effect for *Emotion Regulation*). There were no more significant ANOVA effects concerning the ratings (all  $p > 0.10$ ).

**Startle-reflex.** Regardless of *Context* and *Group*, participants exhibited smaller startle magnitudes toward probes in the down-regulation ( $M = -0.112$ ,  $SD = 0.288$ ) as compared to the enhance condition ( $M = 0.112$ ,  $SD = 0.288$ ),  $F_{(1, 38)} = 5.99$ ,  $p = 0.019$ ,  $f = 0.40$  (Main effect for *Emotion Regulation*) suggesting successful emotion regulation. To clarify whether this significant differences between the enhance- and the down-regulate condition was due to successful enhancement or successful down-regulation of emotions, the startle responses during emotion regulation were compared to the baseline responses (collapsed over *Context*, as well as early and late emotion regulation interval). These analysis show that enhancement of emotions was successful,  $t_{(39)} = 1.98$ ,  $p = 0.027$ ,  $d = 0.32$  (one-tailed), and that down-regulation of emotions tended to be effective,  $t_{(39)} = -1.52$ ,  $p = 0.069$ ,  $d = 0.25$  (one-tailed).

## DISCUSSION

In Experiment I, emotion perception and regulation was assessed in response to anxious facial expressions presented in the context of chemosensory anxiety signals. In line with the hypothesis, when presented in the context of chemosensory anxiety signals, the faces were rated as more negative than when presented in the context of chemosensory sport stimuli. Moreover, when a startle response was elicited during face presentations in the chemosensory anxiety context, it was elevated in HSA participants compared to LSA participants. These results show that chemosensory anxiety related context information is capable of altering behaviorally relevant emotional responses (i.e., withdrawal related motor behavior and self-report) toward socially relevant visual stimuli. Thus, the results are in line with findings

of altered visual social perception through chemosensory anxiety signals (Pause et al., 2004; Zhou and Chen, 2009). The fact that the effect of anxiety relevant chemosensory context stimuli was especially pronounced in HSA individuals extends previous work showing a hyperreactivity of HSA individuals toward facial (Blair et al., 2008) and chemosensory signals of anxiety (for a comprehensive discussion see Pause et al., 2009). Furthermore, the results are in line with previous research showing that neutral faces presented in a negative self-evaluative semantic context affects neural responses in HSA individuals more strongly than in healthy controls (Schwarz et al., 2012). The present data extend these results and show for the first time that socially anxious individuals might be especially prone to the impact of non-semantic (threatening) chemosensory context information on the perception of emotional faces. Furthermore, the fact that startle responses differentiate between HSA and LSA individuals mainly in response to startle probe C, that is briefly after target stimulus offset, further underlines the sensitivity of socially anxious individuals toward even weak social cues.

Both, HSA and LSA participants were able to effectively regulate their emotions toward the faces: they exhibited smaller startle-responses, felt less negative and less aroused when down regulating, than when enhancing their emotions. Thus, while emotion regulation toward social cues in HSA participants has been demonstrated before (Goldin et al., 2009), the present results show that also defensive motivation toward socially relevant stimuli can be regulated effectively. Thus, chemosensory context had no influence on the participants' ability to regulate their emotions. Contextual chemosensory anxiety signals have been shown to be especially effective sources of information when the facial information is ambiguous (Zhou and Chen, 2009) or incongruent (Pause et al., 2004). Because in the present study, the participants perceived all faces as clearly negative and anxiety-inducing (individual screening session), the congruent chemosensory information did not add any new information relevant to accomplish the emotion regulation task and might therefore have had no influence. This would imply that visual and chemosensory communication channels constitute specialized independent communication systems, integrating only under circumstances of perceptual uncertainty, or when further information is needed. Thus, the current results suggest that salient visual foreground information can be affected by top-down neuronal control and that contextual chemosensory anxiety cues do alter the general emotional significance of this visual information rather than interacting directly with top-down control mechanisms.

However, the present results leave open the question of whether the effects found in Experiment I have their foundation in earlier processes of stimulus perception, like the structural encoding of the faces (N170), or the allocation of attention (N1/P1) toward the faces. To determine this, in Experiment II an EEG was recorded and the impact of the chemosensory context stimuli (anxiety, sport control, pure cotton pad) was evaluated on early automatic structural encoding of (N1/P1, N170) and late motivational (LPP) attention allocation toward the stimuli.

## EXPERIMENT II

### METHODS

#### Participants

Thirty-six non-smoking female students (different from those in Experiment I) were classified (according to the SIAS) as either non-socially anxious (LSA, scores  $< 16$ ,  $n = 18$ ,  $M = 11.61$ ;  $SD = 3.36$ ) or high-socially anxious (HSA, scores  $> 22$ ,  $n = 18$ ;  $M = 31.22$ ;  $SD = 8.32$ ; group comparison:  $p < 0.001$ ). The mean score of the HSA participants was above the suggested cut-off of 30 for social phobia (Stangier et al., 1999). All participants reported a regular menstrual cycle. Out of the 36 participants, 16 ( $N = 8$  HSA) reported to use hormonal contraceptives. All reported that they used no medication, suffered from no mental and physical diseases or general hyposmia. All participants scored low on social desirability ( $< 5$  on the Lie scale of the EPI, Eggert and Ratschinski, 1983), supporting the validity of the self-report data. HSA participants scored within normal range for trait anxiety (STAI) and depressed feelings (DS), while LSA participants scored low on both questionnaires (STAI:  $M = 35.50$ ;  $SD = 5.23$ , DS:  $M = 5.44$ ,  $SD = 2.81$ , group difference for both questionnaires:  $p < 0.001$ ). Both groups scored within the medium range for the frequency of everyday-life use of reappraisal (ERQ), and for empathy (SPF). The two groups did not differ in the frequency with which they used reappraisal in everyday life,  $p > 0.20$ , in empathic feelings,  $p > 0.20$  or in age,  $p > 0.20$  ( $M = 23.72$ ,  $SD = 4.86$ , range 19–42). All participants were paid for participation and gave written informed consent. The study was approved by the ethics committee of the DGPs.

#### Stimulus material and stimulus presentation

Chemosensory stimulus material was the same as in Experiment I. In addition, pure cotton pad was introduced as a control stimulus. Prior to usage the pure cotton pads were treated in the same way as the anxiety and sport cotton pads: They were pooled, divided into small portions (1.2 g each) and stored at  $-20^{\circ}\text{C}$ . The chemosensory stimuli were presented with a constant-flow (50 ml/s) 5-channel olfactometer and stimuli were presented in three counterbalanced blocks (enhance, down-regulate, watch) of 60 trials each.

The visual stimulus material, consisted of 60 pictures from 30 male actors showing anxious facial expressions with averted gazes to the left and right were chosen from the KDEF set (KDEF, Lundqvist et al., 1997)<sup>2</sup>. The large number of pictures was necessary to prevent habituation effects due to repeated presentation of the face stimuli. Stimulus timing was controlled with the Presentation<sup>®</sup> software (Version 14, Neurobehavioral Systems, USA).

#### Individual stimulus validation and odor detection session

As in Experiment I, participants were asked to judge the chemosensory stimuli for intensity, pleasantness, unpleasantness and familiarity (10 cm visual analogue scales) during an individual stimulus validation session which took place 7 days prior to the main experiment (see Experiment I section Individual Stimulus Validation and Odor Detection Session for details). Tables 3 and 4 shows participants' stimulus ratings. Chemosensory anxiety signals were perceived as more intense

**Table 3 | Mean intensity, pleasantness, unpleasantness, and familiarity ratings of the chemosensory stimuli in Experiment II (EEG).**

	Anxiety		Sport		Cotton pad control	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Intensity	4.64	1.68	3.39	1.66	2.72	1.75
Pleasantness	2.61	1.48	2.78	1.64	2.64	2.05
Unpleasantness	3.47	2.35	2.56	1.76	2.03	1.52
Familiarity	3.69	1.95	3.17	2.08	2.64	2.02

Range 1–9.

**Table 4 | Valence and arousal ratings of the chemosensory stimuli in Experiment II (EEG).**

	Anxiety		Sport		Cotton pad control	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Valence	-0.81	1.06	-0.36	1.42	-0.11	1.14
Arousal	4.72	1.52	4.81	1.51	4.25	1.50

SAM Valence range -4–4, SAM Arousal Range 1–9.

than sport,  $t_{(35)} = 3.38$ ,  $p = 0.002$ , and cotton pad control stimuli,  $t_{(35)} = 5.15$ ,  $p < 0.001$  [main effect stimulus  $F_{(2, 68)} = 13.96$ ,  $p < 0.001$ ,  $f = 0.64$ ]. They were also perceived as more unpleasant than chemosensory sport,  $t_{(35)} = 2.21$ ,  $p = 0.034$ , and cotton pad control stimuli,  $t_{(35)} = 3.64$ ,  $p = 0.001$  [main effect stimulus,  $F_{(2, 68)} = 7.57$ ,  $p = 0.001$ ,  $f = 0.47$ ], and as more familiar than cotton pad control,  $t_{(35)} = 2.72$ ,  $p = 0.010$ , but not than chemosensory sport stimuli,  $t_{(35)} = 1.85$ ,  $p = 0.073$  [main effect stimulus,  $F_{(2, 68)} = 4.57$ ,  $p = 0.015$ ,  $f = 0.37$ ]. Intensity ( $p = 0.068$ ), unpleasantness ( $p = 0.073$ ) and familiarity ratings ( $p = 0.149$ ) between sport and cotton pad control did not differ. There were no differences in pleasantness ratings between any of the stimuli.

Afterwards, participants specified their feelings of pleasantness and arousal (SAM) in response to the chemosensory stimuli. They rated themselves as feeling more unpleasant (SAM valence) when perceiving the chemosensory anxiety signals compared to cotton pad control stimuli,  $t_{(35)} = 2.50$ ,  $p = 0.017$  [main effect stimulus  $F_{(2, 68)} = 3.33$ ,  $p = 0.042$ ,  $f = 0.31$ ]. No more differences were found between HSA and LSA participants concerning the ratings.

Of the participants 19 (53%) were able to differentiate both chemosensory stimuli from cotton pad control (two correct detections for each stimulus within three-alternative forced choice tests including cotton pads from either condition, and two non-used cotton pads, administered via the olfactometer for 2.5 s).

Due to the large number of different facial expressions used in the main experiment (60), individual judgments for the face stimuli were discarded.

#### Experimental session

First electrodes were attached. Then participants received detailed breathing instructions and practiced correct inhalation until

they signaled that correct breathing occurred without any effort. Then, stimulus timing was explained and emotion regulation instructions were given. **Figure 1** (lower half) shows the stimulus timing. At the beginning of each trial, an anxious facial expression was presented for 1 s to prepare the participant for the upcoming emotion regulation task. Then the written emotion regulation instruction was presented for 1.5 s (enhance, down-regulate, or watch) followed by an exhalation cue. It consisted of a ball decreasing continuously in size across a period of 2.5 s. After the exhalation, the participants started with the inhalation. During the inhalation period (randomly 1–2 s after participants started inhaling), the chemosensory context stimulus was presented for 2.5 s. One second after the onset of the context stimulus the facial expression was presented again for 1.5 s. Participants were instructed to keep inhaling until the end of the picture presentation. During the inter stimulus interval (ISI, duration random between 11 and 13 s), participants rated their current emotional state for valence and arousal (SAM). Mean trial duration was 20 s. After the presentation of 30 trials (10 min) a 5 min break was included. During each block, the 60 facial expressions were presented in random order, and were paired with either a chemosensory anxiety ( $n = 20$  trials), sport stimulus ( $n = 20$  trials), or cotton pad control ( $n = 20$  trials)<sup>3</sup>. Chemosensory stimuli were equally distributed within blocks, and the same chemosensory stimulus did not occur during more than three consecutive trials.

Participants received the same emotion regulation instructions as in Experiment I. In addition they were told to perceive the stimuli only passively during the watch block. Then they were instructed in how to use emotion regulation instructions during the task. In brief, with onset of the regulation instruction participants had to think about a regulation strategy. With the onset of the target picture they then had to begin regulating and to continue regulating until the onset of the SAM rating scales. Finally, they practiced at least 10 learning trials per experimental condition. The experiment started only after participants signaled full understanding of, and compliance with, the instructions.

### Data recording

The EEG was recorded with Ag/AgCl electrodes (inner diameter 6 mm) from 25 scalp locations (AF7, FP1, FPz, FP2, AF8, F7, F3, Fz, F4, F8, T7, C3, Cz, C4, T8, P7, P3, Pz, P4, P8, PO7, O1, Oz, O2, PO8) using an electrode cap (EasyCap GmbH, Germany) in reference to the average across all electrodes. In addition both earlobes were recorded. Two electrodes were placed near the right eye (3 cm above, inside the vertical pupil axis and 1.5 cm below, outside the vertical pupil axis) for the recording of vertical and horizontal eye movements. The impedance of the electrodes was kept below 10 k $\Omega$ .

The physiological data were recorded, amplified, and filtered with the BrainVision Recorder software (Brain Products GmbH, Munich, Germany) using a sampling rate of 250 Hz, a low-pass filter of 40 Hz (24 dB/octave) and a 50 Hz notch filter. Offline, EEG signals were re-referenced to linked ear lobes and high

pass filtered (0.04 Hz, 24 dB/octave), afterwards corrected for eye movements (Gratton et al., 1983) and baseline-corrected (0–200 ms before picture onset). Subsequently, trials contaminated with artifacts (due to sweating, movements, or pronounced alpha-activity: 0.25%) and insufficient inhalation of the chemosensory stimuli (begin of inhalation > 300 ms before picture onset or end of inhalation < 700 ms after picture onset: 3.5%) were eliminated. Prior to averaging, in order to ease the component' detection, they were again low-pass filtered (20 Hz, 24 dB/octave).

### Data analysis

The N1 amplitude was quantified as the maximum peak at frontopolar, frontal and central electrode sites (70–140 ms), the P1 as the maximum peak over parietal and occipital electrode sites (70–140 ms). The N170 amplitude was analyzed as minimum peak over parietal and occipital electrode sites (130–180 ms). The LPP were extracted from all electrodes (LPP mean activity: 400–600 ms).

ERPs were subjected to repeated measure mixed model ANOVA. For the N1 component the ANOVA included the between subject factor *Group* (HSA, LSA) and the within subject factors *Context* (chemosensory anxiety, chemosensory sport, cotton pad control), *Emotion Regulation* (enhance, down-regulate, watch), *Sagittal* electrode sites (frontopolar, frontal, central), and *Transversal* electrode sites (lateral left, left, midline, right, lateral right). For the N170 and the P1 (detected at parietal and occipital sites) the factor *Sagittal* had 2 levels (parietal, occipital), while for the LPP (detected at all electrodes) it had 5 levels (frontopolar, frontal, central, parietal, occipital). For reasons of brevity, effects including electrode factors are presented without follow-up tests.

Mean ratings of valence and arousal were calculated within participants according to the conditions and were subjected to a repeated measures mixed model ANOVA including the between subject factor *Group*, and the within subject factors *Emotion Regulation* (enhance, down-regulate, watch), and *Context* (anxiety, sport, cotton pad control).

Cohen's effect-size  $f$  was calculated. Huynh–Feldt corrections of degrees of freedom were applied, and corrected  $p$ -values are reported. Subsequent nested effects (Page et al., 2003) and  $t$ -tests were calculated. An alpha level of 5% was used for all statistical tests.

## RESULTS

### Effects of electrode positions

The N1 amplitude was most pronounced at frontopolar and midline electrode sites [main effects for *Transversal*,  $F_{(2, 68)} = 19.58$ ,  $p < 0.001$ ,  $f = 0.76$  and *Sagittal*,  $F_{(2, 68)} = 26.67$ ,  $p < 0.001$ ,  $f = 0.89$  with largest amplitudes central and frontal midline electrodes [Fz/Cz, interaction *Sagittal* by *Transversal*,  $F_{(8, 272)} = 13.70$ ,  $p < 0.001$ ,  $f = 0.63$ ], while the P1 showed largest amplitudes at parieto-central electrodes [Pz/Oz, interaction *Sagittal* by *Transversal*,  $F_{(4, 136)} = 9.25$ ,  $p = 0.001$ ,  $f = 0.41$ ]. The N170 amplitude was most pronounced at right lateral electrode sites [main effect for *Transversal*,  $F_{(4, 136)} = 22.36$ ,  $p < 0.001$ ,  $f = 0.81$ ], with largest amplitudes were observed over P8 [interaction

<sup>3</sup>The relatively small number of trials per condition was necessary in order to prevent the chemosensory stimuli from habituation.

*Sagittal* by *Transversal*,  $F_{(4, 136)} = 6.92$ ,  $p = 0.001$ ,  $f = 0.45$ ). Finally, the LPP was most pronounced over parietal and occipital electrode sites [main effect *Sagittal*,  $F_{(4, 136)} = 32.60$ ,  $p < 0.001$ ,  $f = 0.98$ ], and central and right electrode sites [main effect for *Transversal*,  $F_{(4, 136)} = 16.02$ ,  $p < 0.001$ ,  $f = 0.69$ ] with largest potentials over Pz and Oz [interaction *Sagittal* by *Transversal*,  $F_{(16, 544)} = 7.86$ ,  $p < 0.001$ ,  $f = 0.48$ ].

### Effects of the chemosensory context

**Ratings.** Participants reported feeling more aroused when the faces were presented in the context of chemosensory anxiety signals ( $M = 5.17$ ,  $SD = 1.14$ ),  $t_{(35)} = 2.19$ ,  $p = 0.035$ , and chemosensory sport ( $M = 5.08$ ,  $SD = 1.01$ ),  $t_{(35)} = 2.35$ ,  $p = 0.024$ , compared to the cotton pad control stimuli ( $M = 4.98$ ,  $SD = 0.98$ ). Arousal ratings between faces presented in the

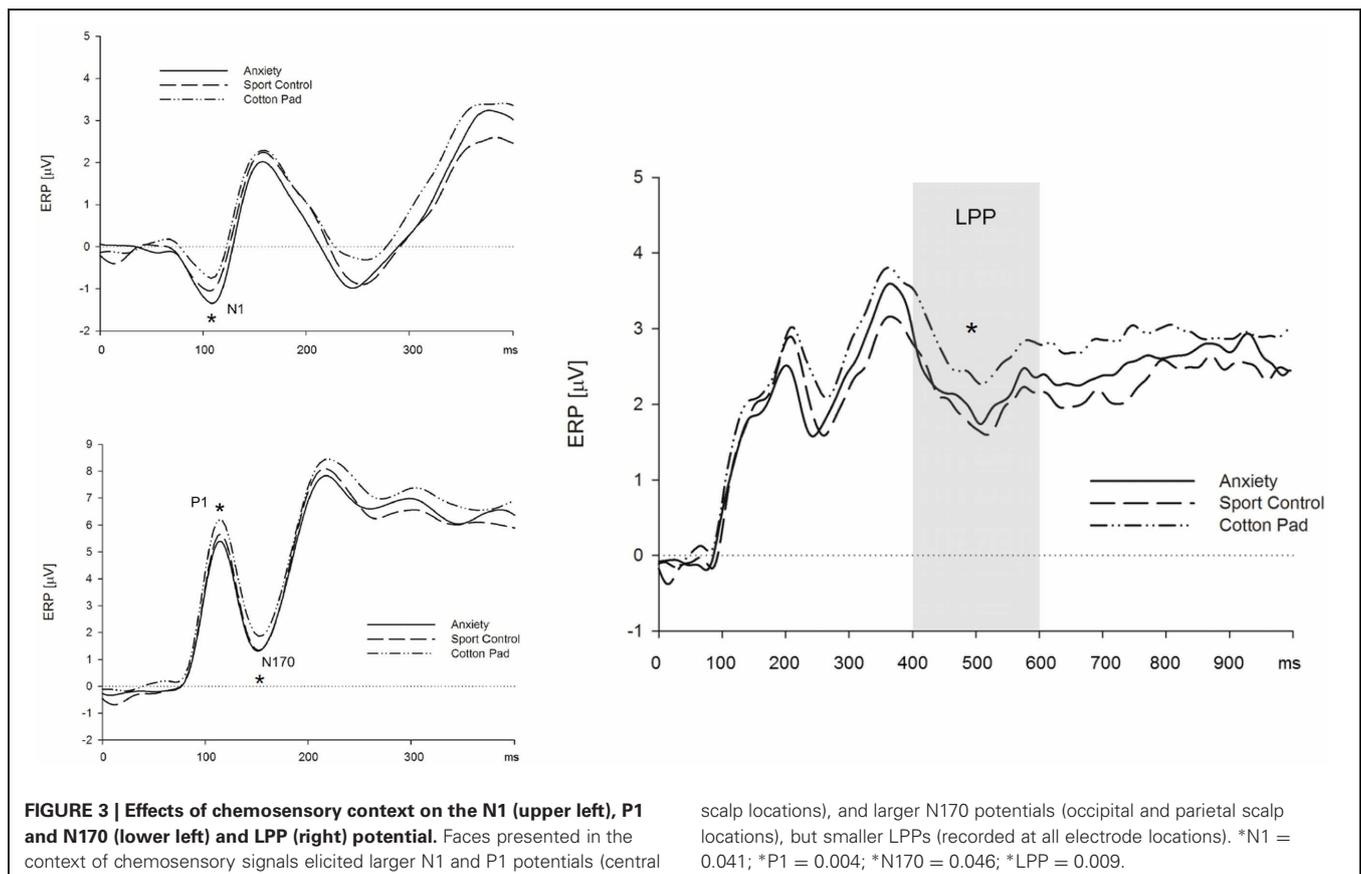
context of chemosensory anxiety signals and sport stimuli did not differ,  $p > 0.10$  [main effect for *Context*,  $F_{(2, 68)} = 3.83$ ,  $p = 0.041$ ,  $f = 0.34$ ] (**Table 5**).

**Early ERP components (N1/P1 and N170).** Above central electrodes, the N1 appeared with larger amplitudes in response to faces presented in the context of chemosensory anxiety signals,  $t_{(35)} = 2.71$ ,  $p = 0.010$ , and sport stimuli,  $t_{(35)} = 1.99$ ,  $p = 0.054$  as compared to faces presented in the context of control stimuli (**Figure 3**) [interaction *Sagittal* by *Context*,  $F_{(4, 136)} = 2.99$ ,  $p = 0.041$ ,  $f = 0.30$ , nested effects, context within central electrode sites,  $F_{(2, 68)} = 5.54$ ,  $p = 0.008$ ,  $f = 0.40$ ]. N1 amplitudes for faces presented in the context of anxiety or sport signals did not differ,  $p > 0.10$ . The P1 amplitudes were larger in the context of chemosensory anxiety ( $p = 0.005$ ) and chemosensory sport

**Table 5 | Emotion regulation effects on subjective ratings for anxious faces in the context of chemosensory anxiety, sport-control or cotton pad stimuli in Experiment II (EEG).**

	Cotton pad			Sport			Anxiety		
	Enhance	Down-regulate	Watch	Enhance	Down-regulate	Watch	Enhance	Down-regulate	Watch
	$M_{(SD)}$	$M_{(SD)}$	$M_{(SD)}$	$M_{(SD)}$	$M_{(SD)}$	$M_{(SD)}$	$M_{(SD)}$	$M_{(SD)}$	$M_{(SD)}$
$SAM_{valence}$	-1.29(0.72)	-0.43(0.72)	-0.65(0.50)	-1.31(0.78)	-0.49(0.68)	-0.74(0.55)	-1.29(0.98)	-0.42(0.74)	-0.81(0.79)
$SAM_{arousal}$	5.75(1.06)	4.32(1.20)	4.88(1.27)	5.79(1.12)	4.46(1.23)	4.99(1.24)	5.92(1.19)	4.49(1.32)	5.10(1.36)

$SAM_{valence}$  range: -4-4;  $SAM_{arousal}$  range: 1-9.



stimuli ( $p = 0.006$ ) as compared to faces presented in the context of control stimuli [main effect for *Context*,  $F_{(2, 68)} = 6.02$ ,  $p = 0.004$ ,  $f = 0.36$  (**Figure 3**)]. Amplitudes for faces presented in the context of anxiety or sport signals did not differ,  $p > 0.10$ . Similar to the N1/P1 deflections, N170 amplitudes were larger for faces presented in the context of chemosensory anxiety,  $t_{(35)} = 2.38$ ,  $p = 0.023$ , and sport signals,  $t_{(35)} = 2.04$ ,  $p = 0.049$  as compared to faces presented in the context of control stimuli [**Figure 3**, main effect for *Context*,  $F_{(2, 68)} = 3.21$ ,  $p = 0.046$ ,  $f = 0.31$ ]. Amplitudes for faces presented in the context of anxiety or sport signals did not differ,  $p > 0.10$  (**Figure 3**).

**Late positive potential.** The LPP was larger for faces presented in the context of control stimuli, as compared to those presented alongside with anxiety signals,  $t_{(35)} = 2.33$ ,  $p = 0.026$ , and sport stimuli,  $t_{(35)} = 2.96$ ,  $p = 0.006$ . [**Figure 3**, main effect for *Context*,  $F_{(2, 68)} = 5.04$ ,  $p = 0.009$ ,  $f = 0.38$ ]. The LPP did not differ between faces presented in the context of anxiety signals and sport stimuli,  $p > 0.10$  (**Figure 3**).

### Effects of social anxiety

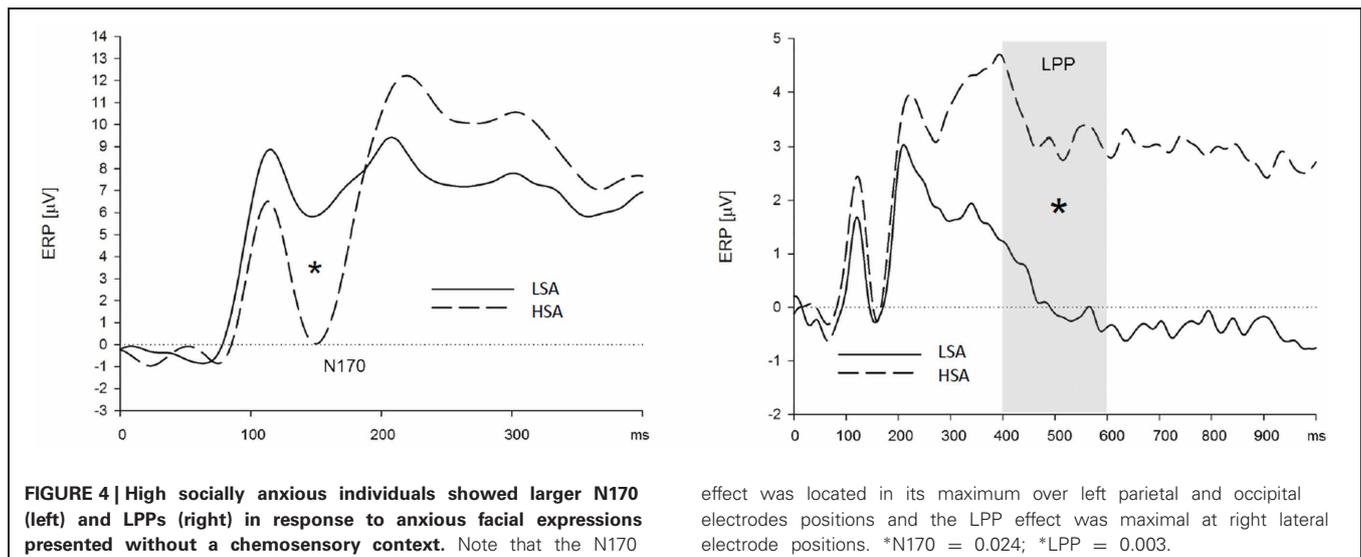
**Early ERP components (N1/P1 and N170).** HSA participants showed larger N170 amplitudes than LSA participants, when they were instructed to watch and to down-regulate their emotions, viewing faces in the context of cotton pad control stimuli (see **Figure 4**), at left, and midline electrode sites [interaction *Group by Transversal by Emotion Regulation by Context*,  $F_{(16, 544)} = 2.45$ ,  $p = 0.007$ ,  $f = 0.27$ , nested effects: *Group by Emotion Regulation by Transversal* within cotton pad *Context*,  $F_{(8, 272)} = 2.66$ ,  $p = 0.024$ ,  $f = 0.28$ , *Group by Transversal* within watch,  $F_{(4, 136)} = 3.66$ ,  $p = 0.031$ ,  $f = 0.33$ , *Group* within left electrode sites within watch,  $F_{(1, 34)} = 5.91$ ,  $p = 0.021$ ,  $f = 0.42$ , *Group* within midline electrode sites within watch,  $F_{(1, 34)} = 4.88$ ,  $p = 0.034$ ,  $f = 0.38$ , *Group by Transversal* within down-regulate,  $F_{(4, 136)} = 3.90$ ,  $p = 0.026$ ,  $f = 0.34$ , *Group* within central electrode sites within down-regulate,  $F_{(1, 34)} = 5.18$ ,  $p = 0.029$ ,  $f = 0.39$ ]. During the enhance condition, there were no differences

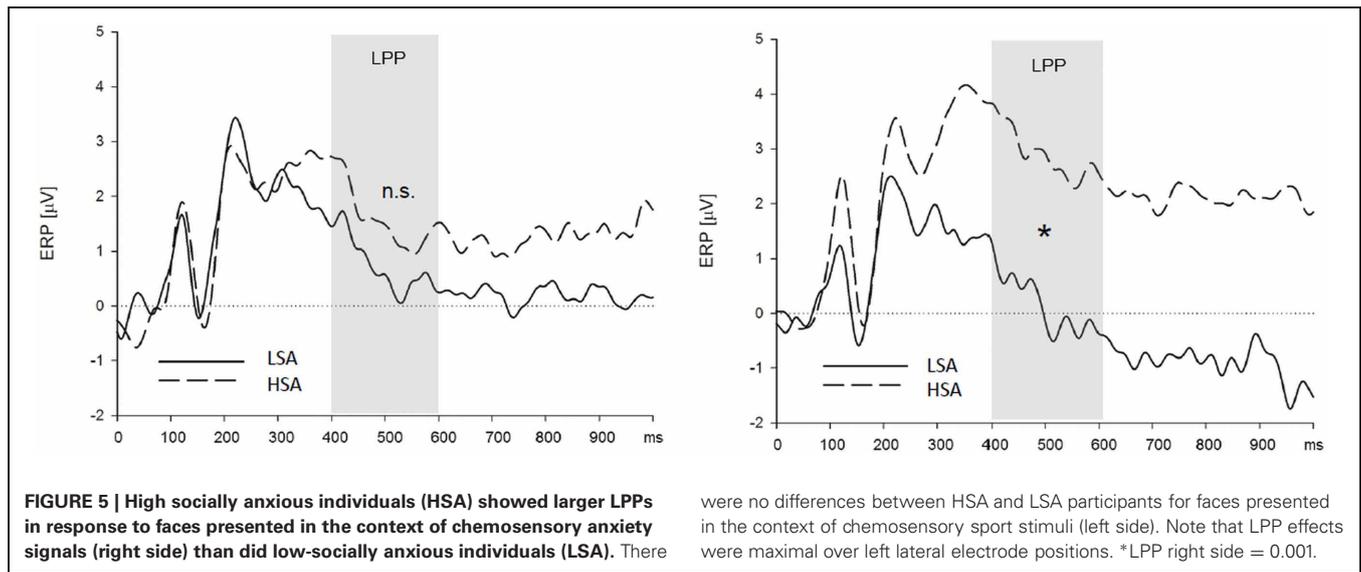
between HSA and LSA participants,  $p > 0.10$ . There were no differences between HSA and LSA participants concerning the N1/P1 component.

**Late positive potential.** HSA participants showed larger LPPs than LSA participants in the context of the chemosensory control and the chemosensory anxiety stimuli: When the facial expressions were presented in the context of the cotton pad control stimuli, HSA participants showed larger LPPs in the watch, and as a trend, in the down-regulate condition than LSA participants [**Figure 4**, interaction *Group by Emotion Regulation by Context* by *Transversal*,  $F_{(16, 544)} = 1.90$ ,  $p = 0.066$ ,  $f = 0.24$ , nested effects: *Group* within right lateral electrode sites within watch,  $F_{(1, 34)} = 9.87$ ,  $p = 0.003$ ,  $f = 0.54$ , *Group* within right lateral electrode sites within down-regulate,  $F_{(1, 34)} = 3.60$ ,  $p = 0.066$ ,  $f = 0.33$ ]. In addition, HSA participants showed larger LPPs during the watch (HSA:  $M = 2.89$ ,  $SD = 2.41$ ; LSA:  $M = 0.13$ ,  $SD = 2.15$ ), and the enhance condition (HSA:  $M = 2.90$ ,  $SD = 3.10$ ; LSA:  $M = 0.60$ ,  $SD = 2.97$ ) toward anxious facial expressions in the context of chemosensory anxiety signals [**Figure 5**, nested effects: *Group by Emotion Regulation by Context* within right lateral electrode sites,  $F_{(4, 136)} = 3.98$ ,  $p = 0.004$ ,  $f = 0.34$ , *Group by Context* within watch within right lateral electrode sites,  $F_{(2, 68)} = 5.81$ ,  $p = 0.005$ ,  $f = 0.41$ , *Group* within watch within chemosensory anxiety within lateral right electrode sites,  $F_{(1, 34)} = 13.17$ ,  $p = 0.001$ ,  $f = 0.62$ , *Group by Context* within enhance within lateral right electrode sites,  $F_{(2, 68)} = 5.42$ ,  $p = 0.007$ ,  $f = 0.40$ , *Group* within enhance within chemosensory anxiety within right lateral electrode sites,  $F_{(1, 34)} = 5.19$ ,  $p = 0.029$ ,  $f = 0.39$ ].

### Effects of emotion regulation

**Ratings.** Concerning the valence ratings, all participants described themselves as feeling less negative during the down-regulate [LSA:  $t_{(17)} = 5.31$ ,  $p < 0.001$ , HSA:  $t_{(17)} = 2.42$ ,  $p = 0.027$ ] and during the watch condition [LSA:  $t_{(17)} = 2.17$ ,  $p = 0.045$ , HSA:  $t_{(17)} = 3.04$ ,  $p = 0.007$ ] than during the enhance condition [main effect *Emotion Regulation*,  $F_{(2, 68)} =$





24.53,  $p < 0.001$ ,  $f = 0.85$ ] (for descriptive statistics see **Table 5**). However, when down regulating their emotion LSA [ $t_{(17)} = 2.17$ ,  $p = 0.045$ ] but not HSA participants ( $p > 0.10$ ) described themselves as feeling less negative as compared to the watch condition [interaction *Emotion Regulation* by *Group*  $F_{(2, 68)} = 3.68$ ,  $p < 0.043$ ,  $f = 0.33$ ].

Concerning the arousal ratings, all participants described themselves feeling less aroused during the down-regulate [LSA:  $t_{(17)} = 6.28$ ,  $p < 0.001$ , HSA:  $t_{(17)} = 3.97$ ,  $p < 0.001$ ] and during the watch condition [LSA:  $t_{(17)} = 3.71$ ,  $p = 0.002$ , HSA:  $t_{(17)} = 3.12$ ,  $p = 0.006$ ] than during the enhance condition [main effect *Emotion Regulation*,  $F_{(2, 68)} = 36.52$ ,  $p < 0.001$ ,  $f = 1.04$ ]. Like for self-reported valence, when down regulating their emotion LSA,  $t_{(17)} = 5.16$ ,  $p < 0.001$ , but not HSA participants,  $p > 0.10$ , described themselves as feeling less aroused than in the watch condition [interaction *Emotion Regulation* by *Group*,  $F_{(2, 68)} = 3.87$ ,  $p < 0.027$ ,  $f = 0.34$ ].

**Early ERP components (N1 and N170).** Participants showed larger N1 amplitudes when they were instructed to enhance ( $M = -3.36$ ,  $SD = 1.71$ ), as compared to the instruction to down-regulate their emotions ( $M = -2.90$ ,  $SD = 1.51$ ),  $t_{(35)} = 2.40$ ,  $p = 0.022$  [main effect *Emotion Regulation*,  $F_{(2, 68)} = 3.00$ ,  $p = 0.056$ ,  $f = 0.81$ ]. Amplitudes did not differ between the enhance and the watch condition ( $M = -3.29$ ,  $2.21$ ),  $p > 0.10$ , and between the down-regulate and watch condition,  $p = 0.073$ . There were no effects of emotion regulation on the N170 component.

**Late positive potential.** Because previous studies show emotion regulation effects mainly for the LPP, the interaction *Group* by *Emotion Regulation* by *Context* by *Transversal*, which was significant as a trend,  $F_{(16, 544)} = 1.90$ ,  $p = 0.066$ ,  $f = 0.24$ , was further explored. Results indicate that the LPP in response to the faces varied with emotion regulation instruction in LSA participants only. LSA participants showed larger LPPs when they were instructed to enhance their emotion elicited by faces

presented in the context of the cotton pad stimuli, as ( $M = 2.28$ ,  $SD = 3.14$ ) than in the watch condition ( $M = 0.144$ ,  $SD = 2.39$ ) in lateral right electrode sites,  $t_{(17)} = 2.51$ ,  $p = 0.023$ . [nested effects for interaction *Group* by *Emotion Regulation* by *Context* by *Transversal*: interaction *Group* by *Emotion Regulation* by *Transversal* within cotton pad context,  $F_{(8, 272)} = 3.07$ ,  $p = 0.016$ ,  $f = 0.30$ , *Group* by *Emotion Regulation* within *Transversal*,  $F_{(2, 68)} = 5.03$ ,  $p = 0.009$ ,  $f = 0.38$ , *Emotion Regulation* within right electrode sites within LSA participants,  $F_{(2, 68)} = 4.43$ ,  $p = 0.015$ ,  $f = 0.36$ , *Emotion Regulation* within right electrode sites within HSA participants,  $p > 0.10$ ]. There were no differences between the enhance and down-regulate ( $M = 1.26$ ,  $SD = 2.35$ ),  $p > 0.10$ , and between the watch and down-regulate conditions,  $p = 0.098$ .

## GENERAL DISCUSSION

Two experiments investigated withdrawal related motor behavior (Experiment I) and ERP correlates (Experiment II) of perception and regulation of facial expressions presented in the context of human chemosensory signals (sport, anxiety) in a group of HSA and a group of LSA individuals. In Experiment I, regardless of emotion regulation or social anxiety, startle responses and emotion regulation effects toward faces presented in the context of chemosensory sport or anxiety stimuli did not differ from each other, suggesting a more powerful influence of the salient visual foreground information as compared to the contextual chemosensory cues on withdrawal related motor behavior. However, although presented at threshold level, HSA individuals as compared to LSA individuals showed a hyperreactivity in withdrawal related motor behavior in the context of chemosensory anxiety signals.

In line with the results of Experiment I, Experiment II also found a preferential processing of contextual chemosensory sport and anxiety information, as indexed by reduced elaborative processing (LPP), compared to the cotton pad control stimuli. It has already been shown that olfactory and visual information are integrated on a neuronal level (Gottfried and

Dolan, 2003), and cross-modal integration has been demonstrated for chemosensory signals and facial expressions (Pause et al., 2004; Zhou and Chen, 2009). Moreover, a recent study demonstrated that the perception of chemosensory information (sport/anxiety), elicits large P3 amplitudes (Pause et al., 2010), suggesting that the processing of this information depends on the allocation of additional neuronal resources. Thus, the additional chemosensory anxiety context information in the present study most likely have distracted neuronal resources from the elaborative processing of the concurrently presented facial expressions, leading to reduced late ERPs toward the faces. This data are consistent with previous reports showing preferential processing of olfactory information in a direct contrast with visual stimuli (Royet et al., 2000; Adolph and Pause, 2012), and reports showing the importance of chemosensory information for social interaction (McClintock, 1971; Kaitz et al., 1987; Wedekind and Furi, 1997; Jacob et al., 2002; Preti et al., 2003). Consistently with the literature, this suggests that the contextual chemosensory information is processed preferentially.

Interestingly, in contrast to the results for late ERPs, larger early (N1/P1 and N170) ERPs for facial expressions were found when they were presented in a social chemosensory context, suggesting an enhancement of early stimulus processing stages for the faces by human chemosensory signals. The present data are in accord with previous research showing enhanced P1 amplitudes for fearful and angry faces as compared to neutral faces (e.g., Batty and Taylor, 2003; Kolassa and Miltner, 2006). Because the P1 has been shown to be attention-sensitive (see Mangun, 1995), the current data suggest that the faces presented in context of chemosensory stimuli received more attentional processing than those presented without a chemosensory context. Interestingly, in a recent study, participants showed faster response times and larger P1 amplitudes toward visual stimuli presented at a location previously cued with emotional prosody (Brosch et al., 2009). The current results extend these findings and show also that human chemosensory signals can enhance early perceptual processing of concurrently presented visual stimuli. Thus, the present and previous data suggest that emotional context information serves to guide perceptual processing, probably through initiation of early attention-associated processes, leading to higher vigilance toward concurrently presented stimuli.

Ratings indicate that the chemosensory anxiety signals were perceived as more intense, unpleasant, and familiar as the sport signals and as cotton pad control. Therefore, it cannot be completely ruled out that some of the observed effects on ERPs occurred because the context stimuli were perceived differently. However, overall, the chemosensory stimuli were described as low in intensity, and as only mildly unpleasant. The subjective emotional responses toward them were described as rather neutral. Furthermore, while differences in ERP effects were observed for anxiety and sport signals in comparison to the cotton pad control stimuli, differences in subjective ratings were evident for anxiety in comparison to the sport and the cotton pad control stimuli. Finally, in line with previous reports, the effects of chemosensory stimuli occurred largely independently of conscious stimulus

processing. Only 50% of the participants were able to consciously distinguish the chemosensory stimuli from cotton pad control. Moreover, differences in stimulus ratings were observed between the chemosensory anxiety and both control stimuli (i.e., cotton pad and chemosensory sport stimuli) only. In contrast, differences in ERPs were observed between cotton pad control and the two chemosensory stimuli (i.e., anxiety, sport), but not between the cotton pad control and the chemosensory sport stimuli. Thus, taken together it seems unlikely that the observed ERP effects are due to the differences in the cognitive evaluation of the chemosensory stimuli.

Enhanced LPPs in socially anxious individuals are found for faces presented in the context of chemosensory anxiety signals. In line with this, HSA participants exhibited larger withdrawal related motor behavior in response to chemosensory anxiety signals than did LSA participants (Pause et al., 2009). In addition, HSA participants compared to LSA participants showed enhanced neuronal processing of the fearful expressions presented without a chemosensory context. This is reflected in enhanced early (N170) and late (LPP) ERPs in HSA participants. Previous studies have shown enhanced automatic guidance of motivated attention (Schupp et al., 2004) toward fearful faces in social anxiety (Mühlberger et al., 2009), and socially anxious individuals have been shown to respond to angry or fearful faces with increased amygdala activation (Straube et al., 2004; Phan et al., 2006).

Our results extend these findings and suggest that even components related to the early structural encoding (N170) of fearful facial expressions are enhanced in socially anxious individuals. In addition, the observed enhanced LPPs in HSA participants indicate an enhanced elaborative processing of fearful facial expressions compared to LSA participants (see also Kolassa and Miltner, 2006; Moser et al., 2008). Thus, converging evidence from previous research and the current study suggest a general processing bias in favor of threatening (angry, fearful) faces and chemosensory signals of anxiety in socially anxious participants, as indexed by deviant stimulus processing during late elaborative and early processing stages.

Correspondingly, no emotion regulation effects on the LPP were found for HSA participants in response to the fearful facial expressions. Maybe this is reflected in a ceiling effect of emotional engagement in HSA participants toward fearful faces that could not be altered using cognitive emotion regulation. This assumption is also supported by the fact that HSA participants showed large LPPs during the watch and the down-regulate conditions compared to LSA participants. These findings provide evidence that during emotion regulation, motivational attention (as reflected by the LPP) is deficient in social anxiety disorders. This is important in terms of theories focusing on attentional biases in social phobia: The reduced ability to distract attention from the feared stimulus might be one source of the attentional biases found frequently in social phobia.

Interestingly, as revealed in Experiment I, despite their hyper-reactivity toward the faces presented in the anxiety context, and consistent with previous reports (Goldin et al., 2009) we found no evidence for impaired emotion regulation of the startle response

in HSA participants. These results indicate a dissociation of the impact of emotion regulation on early attention related visual stimulus processing stages (Experiment II) and on the initiation of behavioral action tendencies (Experiment I). Thus, socially anxious individuals, although impaired in voluntarily regulating motivated attention toward fear relevant stimuli, are not impaired in the later regulation of withdrawal related action tendencies. Interestingly, no differences were found in the self-reported frequency of use of regulation strategies in everyday life and in the post experimental questionnaire between HSA and LSA participants. Thus, because they are frequently confronted with their feared situation, HSA participants may have simply developed more effective regulation strategies and thus are able to overcome their initial hyperreactivity toward the social cues in the present study. Indeed, initial evidence suggest that social phobics show less signal change in emotion regulation related brain areas during cognitive reappraisal, but show no impairment in emotion regulation outcome (Goldin et al., 2009) suggesting that comparable emotion regulation outcome to that of healthy controls is accompanied with the allocation of fewer neuronal resources in socially anxious individuals.

The current findings extend the existing literature and show that socially anxious individuals have a processing bias (Hirsch and Clark, 2004) not only toward visual social signals of threat (Merckelbach et al., 1989; Stein et al., 2002; Straube et al., 2004; Kolassa and Miltner, 2006; Phan et al., 2006; Blair et al., 2008; Moser et al., 2008; Mühlberger et al., 2009), but also in response to social chemosensory signals of anxiety. This processing bias involves both early attentional, as well as late behaviorally relevant information processing. Interestingly, initial evidence shows that social anxiety might also be accompanied by an enhanced vigilance toward chemosensory signals of aggression/dominance (Adolph et al., 2010). This suggests that the processing bias in social anxiety toward social threat information may be generalized to multiple social communication channels. This view is also supported by findings of increased activation of emotion processing brain areas in social phobics toward threatening (angry) prosody (Quadflieg et al., 2008).

Taken together, findings from the literature and the current results suggest a specific multichannel sensitivity of socially anxious individuals toward threat related social information. These findings have important implications. Etiological models suggest that information-processing biases play a central role for the development and maintenance of the disorder (Clark and Wells, 1995). Specifically, it has been argued that socially anxious individuals fail to habituate during social encounters and exhibit continued subjective distress, which may lead to subsequent avoidance, being implicated in the maintenance of the disorder (Beidel et al., 1985). The observed processing biases toward social threat stimuli, especially in terms of contextual chemosensory information, may in part mediate this failure in habituation to the social situation. This assumption is also supported by the fact that HSA participants compared to LSA participants showed significant larger withdrawal related motor behavior under perceptual uncertainty (Experiment I), that is, briefly after stimulus offset. This suggests a sustained hypervigilance and hyperreactivity even after the threatening situation

is over. Moreover, in HSA individuals the startle responses in the context of chemosensory anxiety signals did not habituate within the trials, while they did so for LSA individuals. Thus, therapeutic interventions may profit from incorporating chemosensory, visual, and acoustic threat signals into therapeutic treatments.

As in previous studies, results show that LSA participants could successfully regulate their emotions, as indicated by their ratings of emotional experience, and by the LPP results. In detail, participants rated themselves to feel less negative, and less aroused when down regulating their emotions, while they described themselves as feeling more negative and more aroused when enhancing their emotions. In line with previous reports (Moser et al., 2009), is that the LPP was larger when LSA participants were instructed to enhance their emotions in response to anxious expressions presented in the context of control stimuli (cotton pad), as compared to the watch condition, indicating effective enhancement of emotional responses to fearful facial expressions. We did not find the expected reduction of the LPP in the down regulation condition, as reported previously (Hajcak and Nieuwenhuis, 2006; Moser et al., 2006, 2009). This could be due to the nature of the facial stimuli. Emotional facial expressions are often described as only mildly arousing (Britton et al., 2006; Alpers et al., 2011) as compared to the highly arousing emotional scenes used in other emotion regulation studies. Moreover during the experiment participants were confronted with a large number of trials and thus habituation of emotional responses cannot be ruled out. This may have caused a rather low emotional involvement of the LSA participants, leading to the present null results for the down regulation condition. However, in general, the present results indicate that emotions elicited by threatening social stimuli can be manipulated using cognitive linguistic emotion regulation strategies.

Within the N1 latency range ERPs were larger during the instruction to enhance than to down-regulate emotions. The N1 component is especially sensitive to selective attention (Hillyard et al., 1998). Interestingly, results from a recent emotion regulation study using eye tracking show that selective attention was controlled by the participants differently depending on whether the regulatory goal was to decrease or increase emotions (Van Reekum et al., 2007), suggesting that in the present study attention may have been allocated automatically, i.e., without conscious control, depending on the regulatory goal. In contrast to the N1 results, the face-specific N170 component was not affected by emotion regulation. Early responses at central scalp locations (N1 in the present study) index general aspects of selective attention, while ERPs in the latency range of the N170 reflect modality-specific processing stages (Van Voorhis and Hillyard, 1977). Thus, the results observed for the N170 and the N1 in the present study may arise from distinct aspects of perceptual stimulus processing, and suggest that the structural encoding of facial expressions (N170) may not necessarily rely on the allocation of attentional resources. Taken together, the present study shows that also ERPs as early as the N1 are affected by emotion regulation. Furthermore, they support the view that attention selection may be a frequently used emotion regulation strategy in everyday life (Gross et al., 2006). The data from Experiment II

are also in line with the findings that withdrawal related motor behavior toward faces in the context of chemosensory signals can be regulated successfully. Thus, in terms of emotion regulation, the present study extends previous reports showing that cognitive linguistic emotion regulation strategies are generally useful in regulating visual and olfactory emotional cues, and shows that early attention related stimulus processing and motor behavior toward chemosensory context cues can be regulated effectively.

One shortcoming of the present study could be that we did not assess the basic emotions elicited by the chemosensory stimuli in the perceivers. Therefore, we cannot completely rule out the possibility that at least some of the participants might have experienced disgust while perceiving the stimuli. However, across studies it has been shown that in general emotional reactions to chemosensory stimuli are rather weak and that people cannot label the stimuli in terms of basic emotions, probably because of rather low detection rates of the chemosensory stimuli (for an overview see Pause, 2012). However, although the effect was rather weak, one study has shown that when participants are asked to decide which of the 6 basic emotions fits best to the emotional state of the donors, chemosensory anxiety stimuli are more often described to smell anxiety-like (Pause et al., 2009). Thus, on basis of the literature, it seems rather unlikely that the participants experienced disgust in the present study. Additionally, the data show that participants described their emotions as only mildly unpleasant ( $M = -0.6$ , on a scale ranging from  $-4 =$  very unpleasant to  $+4 =$  very pleasant) and these ratings did not differ between the chemosensory conditions, suggesting only very little emotional involvement at all.

Finally, because only fearful facial expressions were used in the present study, only a main effect of chemosensory context could be tested, not an interaction between facial expression

and chemosensory context. Thus, future studies should include neutral facial expressions to address the question of differential effects of chemosensory context on the processing of emotional and neutral facial expressions (for example on N1/P1 or N170 amplitudes).

## CONCLUSION

In sum, the present results show that social chemosensory information constitute powerful context cues. They are capable of altering the processing of emotional facial expressions in guiding motivated attention and in altering withdrawal related motor behavior. Moreover, the present study shows for the first time that socially anxious individuals at risk for social phobia are especially sensitive toward contextual social chemosensory anxiety signals. They show enhanced withdrawal related motor behavior, no habituation (Study I) and enhanced allocation of attention toward faces presented in the context of these signals as compared to the low-socially anxious individuals. An enhanced vigilance and excitability as well as a lack of habituation toward threatening social context cues may form one basis for the processing bias of threatening social information in socially anxious individuals and may thus be an important factor for the maintenance of the disorder. Thus, people suffering from social phobia may profit from incorporating social context cues into therapeutic treatments.

## ACKNOWLEDGMENTS

This study was supported by a grant of the German Research Foundation, DFG (PA937/3-1) to Bettina M. Pause. The authors would like to thank Sonja Meyer and Dominique Schaub for their help with the data collection and S. Lloyd Williams for his help with language editing and many valuable comments on an earlier version of the manuscript.

## REFERENCES

- Abler, B., and Kessler, H. (2009). Emotion regulation questionnaire - eine deutsche Version des ERQ von Gross and John. *Diagnostica* 55, 144–152. doi: 10.1026/0012-1924.55.3.144
- Adolph, D., and Alpers, G. W. (2010). Valence and arousal: a comparison of two sets of emotional facial expressions. *Am. J. Psychol.* 123, 209–219.
- Adolph, D., and Pause, B. M. (2012). Different time course of emotion regulation towards odors and pictures: are odors more potent than pictures? *Biol. Psychol.* 91, 65–73. doi: 10.1016/j.biopsycho.2012.05.003
- Adolph, D., Schlosser, S., Hawighorst, M., and Pause, B. M. (2010). Chemosensory signals of competition increase the skin conductance response in humans. *Physiol. Behav.* 101, 666–671. doi: 10.1016/j.physbeh.2010.08.004
- Alpers, G. W., Adolph, D., and Pauli, P. (2011). Emotional scenes and facial expressions elicit different psychophysiological responses. *Int. J. Psychophysiol.* 80, 173–181. doi: 10.1016/j.ijpsycho.2011.01.010
- Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., et al. (2008). Angry, Disgusted, or Afraid?: studies on the malleability of emotion perception. *Psychol. Sci.* 19, 724–732. doi: 10.1111/j.1467-9280.2008.02148.x
- Batty, M., and Taylor, M. J. (2003). Early processing of the six basic facial emotional expressions. *Cogn. Brain Res.* 17, 613–620. doi: 10.1016/S0926-6410(03)00174-5
- Beidel, D. C., Turner, S. M., and Dancu, C. V. (1985). Physiological, cognitive and behavioral aspects of social anxiety. *Behav. Res. Ther.* 23, 109–117. doi: 10.1016/0005-7967(85)90019-1
- Blair, K., Shaywitz, J., Smith, B. W., Rhodes, R., Geraci, M., Jones, M., et al. (2008). Response to emotional expressions in generalized social phobia and generalized anxiety disorder: evidence for separate disorders. *Am. J. Psychiatry* 165, 1193–1202. doi: 10.1176/appi.ajp.2008.07071060
- Blumenthal, T. D., Cuthbert, B. N., Filion, D. L., Hackley, S., Lipp, O. V., and Van Boxtel, A. (2005). Committee report: guidelines for human startle eye-blink electromyographic studies. *Psychophysiology* 42, 1–15. doi: 10.1111/j.1469-8986.2005.00271.x
- Bradley, M. M., and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* 25, 49–59. doi: 10.1016/0005-7916(94)90063-9
- Bradley, M. M., Codispoti, M., Cuthbert, B. N., and Lang, P. J. (2001). Emotion and motivation I: defensive and appetitive reactions in picture processing. *Emotion* 1, 276–298. doi: 10.1037/1528-3542.1.3.276
- Britton, J. C., Taylor, S. F., Sudheimer, K. D., and Liberzon, I. (2006). Facial expressions and complex IAPS pictures: common and differential networks. *Neuroimage* 31, 906–919. doi: 10.1016/j.neuroimage.2005.12.050
- Brosch, T., Grandjean, D., Sander, D., and Scherer, K. R. (2009). Cross-modal emotional attention: emotional voices modulate early stages of visual processing. *J. Cogn. Neurosci.* 21, 1670–1679. doi: 10.1162/jocn.2009.21110
- Clark, D. M., and Wells, A. (1995). “A cognitive model of social phobia,” in *Social phobia: Diagnosis, Assessment, and Treatment*, eds R. G. Heimberg, M. R. Liebowitz, D. A. Hope, and F. R. Schneier (New York, NY: Guilford Press Guilford Press), 69–93.
- Eggert, D., and Ratschinski, G. (1983). *Eysenck-Persönlichkeits-Inventar E-P-I*. Göttingen: Hogrefe.
- Ekman, P., and Friesen, W. V. (1971). Constants across cultures in the face and emotion. *J. Pers. Soc. Psychol.* 17, 124–129. doi: 10.1037/h0030377
- Globisch, J., Hamm, A. O., Esteves, F., and Ohman, A. (1999). Fear appears fast: temporal course of startle reflex

- potentiation in animal fearful subjects. *Psychophysiology* 36, 66–75. doi: 10.1017/S0048577299970634
- Goldin, P. R., Manber, T., Hakimi, S., Canli, T., and Gross, J. J. (2009). Neural bases of social anxiety disorder: emotional reactivity and cognitive regulation during social and physical threat. *Arch. Gen. Psychiatry* 66, 170–180. doi: 10.1001/archgenpsychiatry.2008.525
- Gottfried, J. A., and Dolan, R. J. (2003). The nose smells what the eye sees: crossmodal visual facilitation of human olfactory perception. *Neuron* 39, 375–386. doi: 10.1016/S0896-6273(03)00392-1
- Gratton, G., Coles, M. G., and Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalogr. Clin. Neurophysiol.* 55, 468–484. doi: 10.1016/0013-4694(83)90135-9
- Gross, J. J. (2002). Emotion regulation: affective, cognitive, and social consequences. *Psychophysiology* 39, 281–291. doi: 10.1017/S0048577201393198
- Gross, J. J., Richards, J. M., and John, O. P. (2006). “Emotion regulation in everyday life,” in *Emotion Regulation in Couples and Families: Pathways to Dysfunction and Health*, eds D. K. Snyder, J. A. Simpson, and J. N. Hughes (Washington, DC: American Psychological Association), 13–35.
- Hajcak, G., and Nieuwenhuis, S. (2006). Reappraisal modulates the electrocortical response to unpleasant pictures. *Cogn. Affect. Behav. Neurosci.* 6, 291–297. doi: 10.3758/CABN.6.4.291
- Hess, U., Adams, R. B. Jr., and Kleck, R. E. (2007). Looking at you or looking elsewhere: the influence of head orientation on the signal value of emotional facial expressions. *Motiv. Emot.* 31, 137–144. doi: 10.1007/s11031-007-9057-x
- Hillyard, S. A., Vogel, E. K., and Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 353, 1257–1270. doi: 10.1098/rstb.1998.0281
- Hirsch, C. R., and Clark, D. M. (2004). Information-processing bias in social phobia. *Clin. Psychol. Rev.* 24, 799–825. doi: 10.1016/j.cpr.2004.07.005
- Jackson, D. C., Malmstadt, J. R., Larson, C. L., and Davidson, R. J. (2000). Suppression and enhancement of emotional responses to unpleasant pictures. *Psychophysiology* 37, 515–522. doi: 10.1111/1469-8986.3740515
- Jacob, S., McClintock, M. K., Zelano, B., and Ober, C. (2002). Paternally inherited HLA alleles are associated with women’s choice of male odor. *Nat. Genet.* 30, 175–179. doi: 10.1038/ng830
- Kaitz, M., Good, A., Rokem, A. M., and Eidelman, A. I. (1987). Mothers’ recognition of their newborns by olfactory cues. *Dev. Psychobiol.* 20, 587–591. doi: 10.1002/dev.420200604
- Kim, H., Somerville, L. H., Johnstone, T., Polis, S., Alexander, A. L., Shin, L. M., et al. (2004). Contextual modulation of amygdala responsiveness to surprised faces. *J. Cogn. Neurosci.* 16, 1730–1745. doi: 10.1162/0898929042947865
- Kolassa, I.-T., and Miltner, W. H. R. (2006). Psychophysiological correlates of face processing in social phobia. *Brain Res.* 1118, 130–141. doi: 10.1016/j.brainres.2006.08.019
- Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (1997). “Motivated attention: affect, activation, and action,” in *Attention and Orienting: Sensory and Motivational Processes*, eds P. J. Lang, R. F. Simons, and M. T. Balaban (Mahwah, NJ: Erlbaum), 97–135.
- Laux, L., Schaffner, P., Glanzmann, P., and Spielberger, C. D. (1981). *Das State-Trait Angst Inventar (STAI)*. Weinheim: Beltz Testgesellschaft.
- Li, W., Moallem, I., Paller, K. A., and Gottfried, J. A. (2007). Subliminal smells can guide social preferences. *Psychol. Sci.* 18, 1044–1049. doi: 10.1111/j.1467-9280.2007.02023.x
- Lorig, T., Elmes, D., Zald, D., and Pardo, J. (1999). A computer-controlled olfactometer for fMRI and electrophysiological studies of olfaction. *Behav. Res. Methods* 31, 370–375. doi: 10.3758/BF03207734
- Lundqvist, L. O., Flykt, A., and Öhman, A. (1997). *The Karolinska Directed Emotional Faces - KDEF*. Stockholm: Department of Clinical Neuroscience, Psychology section, Karolinska Institutet.
- Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology* 32, 4–18. doi: 10.1111/j.1469-8986.1995.tb03400.x
- McClintock, M. (1971). Menstrual synchrony and suppression. *Nature* 229, 244–245. doi: 10.1038/229244a0
- Merckelbach, H., Van Hout, W., Van Den Hout, M. A., and Mersch, P. P. (1989). Psychophysiological and subjective reactions of social phobics and normals to facial stimuli. *Behav. Res. Ther.* 27, 289–294. doi: 10.1016/0005-7967(89)90048-X
- Moser, J. S., Hajcak, G., Bukay, E., and Simons, R. F. (2006). Intentional modulation of emotional responding to unpleasant pictures: an ERP study. *Psychophysiology* 43, 292–296. doi: 10.1111/j.1469-8986.2006.00402.x
- Moser, J. S., Huppert, J. D., Duval, E., and Simons, R. F. (2008). Face processing biases in social anxiety: an electrophysiological study. *Biol. Psychol.* 78, 93–103. doi: 10.1016/j.biopsycho.2008.01.005
- Moser, J. S., Krompinger, J. W., Dietz, J., and Simons, R. F. (2009). Electrophysiological correlates of decreasing and increasing emotional responses to unpleasant pictures. *Psychophysiology* 46, 17–27. doi: 10.1111/j.1469-8986.2008.00721.x
- Mühlberger, A., Wieser, M. J., Herrmann, M. J., Weyers, P., Troger, C., and Pauli, P. (2009). Early cortical processing of natural and artificial emotional faces differs between lower and higher socially anxious persons. *J. Neural Transm.* 116, 735–746. doi: 10.1007/s00702-008-0108-6
- Ochsner, K. N., Bunge, S. A., Gross, J. J., and Gabrieli, J. D. (2002). Rethinking feelings: an fMRI study of the cognitive regulation of emotion. *J. Cogn. Neurosci.* 14, 1215–1229. doi: 10.1162/089892902760807212
- Page, M. C., Braver, S. L., and Mackinnon, D. P. (2003). *Levine’s Guide to SPSS for Analysis of Variance*. Mahaw, NJ: Lawrence Erlbaum Associates.
- Paulus, C. (2009). Der saarbrücker persönlichkeitsfragebogen, spf (iri) zur messung von empathie. Available online at: [http://bildungswissen.schaften.uni-saarland.de/personal/paulus/homepage/files/SPF-IRI\\_V6.0.pdf](http://bildungswissen.schaften.uni-saarland.de/personal/paulus/homepage/files/SPF-IRI_V6.0.pdf)
- Pause, B. (2012). Processing of body odor signals by the human brain. *Chemosens. Percept.* 5, 55–63. doi: 10.1007/s12078-011-9108-2
- Pause, B. M., Adolph, D., Prehn-Kristensen, A., and Ferstl, R. (2009). Startle response potentiation to chemosensory anxiety signals in socially anxious individuals. *Int. J. Psychophysiol.* 74, 88–92. doi: 10.1016/j.ijpsycho.2009.07.008
- Pause, B. M., Lubke, R. K., Laudien, J. H., and Ferstl, R. (2010). Intensified neuronal investment in the processing of chemosensory anxiety signals in non-socially anxious and socially anxious individuals. *PLoS ONE* 5:e10342. doi: 10.1371/journal.pone.0010342
- Pause, B. M., Ohrt, A., Prehn, A., and Ferstl, R. (2004). Positive emotional priming of facial affect perception in females is diminished by chemosensory anxiety signals. *Chem. Senses* 29, 797–805. doi: 10.1093/chemse/bjh245
- Phan, K. L., Fitzgerald, D. A., Nathan, P. J., and Tancer, M. E. (2006). Association between amygdala hyperactivity to harsh faces and severity of social anxiety in generalized social phobia. *Biol. Psychiatry* 59, 424–429. doi: 10.1016/j.biopsycho.2005.08.012
- Prehn, A., Ohrt, A., Sojka, B., Ferstl, R., and Pause, B. M. (2006). Chemosensory anxiety signals augment the startle reflex in humans. *Neurosci. Lett.* 394, 127–130. doi: 10.1016/j.neulet.2005.10.012
- Prehn-Kristensen, A., Wiesner, C., Bergmann, T. O., Wolff, S., Jansen, O., Mehdorn, H. M., et al. (2009). Induction of empathy by the smell of anxiety. *PLoS ONE* 4:e5987. doi: 10.1371/journal.pone.0005987
- Preti, G., Wysocki, C. J., Barnhart, K. T., Sondheimer, S. J., and Leyden, J. J. (2003). Male axillary extracts contain pheromones that affect pulsatile secretion of luteinizing hormone and mood in women recipients. *Biol. Reprod.* 68, 2107–2113. doi: 10.1095/biolreprod.102.008268
- Quadflieg, S., Mohr, A., Mentzel, H.-J., Miltner, W. H. R., and Straube, T. (2008). Modulation of the neural network involved in the processing of anger prosody: the role of task-relevance and social phobia. *Biol. Psychol.* 78, 129–137. doi: 10.1016/j.biopsycho.2008.01.014
- Righthart, R., and De Gelder, B. (2006). Context influences early perceptual analysis of faces—an electrophysiological study. *Cereb. Cortex* 16, 1249–1257. doi: 10.1093/cercor/bhj066
- Righthart, R., and De Gelder, B. (2008). Rapid influence of emotional scenes on encoding of facial expressions: an ERP study. *Soc. Cogn. Affect. Neurosci.* 3, 270–278. doi: 10.1093/scan/nsn021
- Royet, J. P., Zald, D., Versace, R., Costes, N., Lavenne, F., Koenig, O., et al. (2000). Emotional responses to pleasant and unpleasant olfactory, visual, and auditory stimuli: a positron emission tomography study. *J. Neurosci.* 20, 7752–7759.
- Rubin, D., Botanov, Y., Hajcak, G., and Mujica-Parodi, L. R. (2012). Second-hand stress: inhalation

- of stress sweat enhances neural response to neutral faces. *Soc. Cogn. Affect. Neurosci.* 7, 208–212. doi: 10.1093/scan/nsq097
- Schupp, H. T., Ohman, A., Junghofer, M., Weike, A. I., Stockburger, J., and Hamm, A. O. (2004). The facilitated processing of threatening faces: an erp analysis. *Emotion* 4, 189–200. doi: 10.1037/1528-3542.4.2.189
- Schwarz, K. A., Wieser, M. J., Gerdes, A. B. M., Mühlberger, A., and Pauli, P. (2012). Why are you looking like that? How the context influences evaluation and processing of human faces. *Soc. Cogn. Affect. Neurosci.* 8, 438–445. doi: 10.1093/scan/nss013
- Sokolov, Y. N. (1963). *Perception and the Conditioned Reflex*. New York, NY: Macmillan.
- Stangier, U., Heidenreich, T., Berardi, A., Golbs, U., and Hoyer, J. (1999). Die erfassung sozialer phobie durch die Social Interaction Anxiety Scale (SIAS) und die Social Phobia Scale (SPS). *Z. Klin. Psychol. Psychother.* 28, 28–36. doi: 10.1026//0084-5345.28.1.28
- Stein, M. B., Goldin, P. R., Sareen, J., Zorrilla, L. T. E., and Brown, G. G. (2002). Increased amygdala activation to angry and contemptuous faces in generalized social phobia. *Arch. Gen. Psychiatry* 59, 1027–1034. doi: 10.1001/archpsyc.59.11.1027
- Straube, T., Kolassa, I. T., Glauer, M., Mentzel, H. J., and Miltner, W. H. (2004). Effect of task conditions on brain responses to threatening faces in social phobics: an event-related functional magnetic resonance imaging study. *Biol. Psychiatry* 56, 921–930. doi: 10.1016/j.biopsych.2004.09.024
- Thompson, R. A. (1994). Emotion regulation: a theme in search of definition. *Monogr. Soc. Res. Child Dev.* 59, 25–52. doi: 10.2307/1166137
- Van Boxtel, A., Boelhouwer, A. J. W., and Bos, A. R. (1998). Optimal EMG signal bandwidth and interelectrode distance for the recording of acoustic, electrocutaneous and photic blink reflexes. *Psychophysiology* 35, 690–697. doi: 10.1111/1469-8986.3560690
- Van Reekum, C. M., Johnstone, T., Urry, H. L., Thurow, M. E., Schaefer, H. S., Alexander, A. L., et al. (2007). Gaze fixations predict brain activation during the voluntary regulation of picture-induced negative affect. *Neuroimage* 36, 1041–1055. doi: 10.1016/j.neuroimage.2007.03.052
- Van Voorhis, S., and Hillyard, S. A. (1977). Visual evoked potentials and selective attention to points in space. *Percept. Psychophys.* 22, 54–62. doi: 10.3758/BF03206080
- Von Zerssen, D., and Koeller, D.-M. (1976). *Paranoid-Depressivitäts-Skala/Depressivitäts-Skala PD-S/D-S*. Göttingen: Hogrefe.
- Wedekind, C., and Furi, S. (1997). Body odour preferences in men and women: do they aim for specific MHC combinations or simply heterozygosity? *Proc. R. Soc. Lond. B Biol. Sci.* 264, 1471–1479. doi: 10.1098/rspb.1997.0204
- Zhou, W., and Chen, D. (2009). Fear-related chemosignals modulate recognition of fear in ambiguous facial expressions. *Psychol. Sci.* 20, 177–183. doi: 10.1111/j.1467-9280.2009.02263.x

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 08 March 2013; accepted: 30 May 2013; published online: 19 June 2013.

Citation: Adolph D, Meister L and Pause BM (2013) Context counts! social anxiety modulates the processing of fearful faces in the context of chemosensory anxiety signals. *Front. Hum. Neurosci.* 7:283. doi: 10.3389/fnhum.2013.00283  
Copyright © 2013 Adolph, Meister and Pause. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



# Perceiving blocks of emotional pictures and sounds: effects on physiological variables

Anne-Marie Brouwer<sup>1\*</sup>, Nelleke van Wouwe<sup>2</sup>, Christian Mühl<sup>3</sup>, Jan van Erp<sup>1</sup> and Alexander Toet<sup>1</sup>

<sup>1</sup> Department of Perceptual and Cognitive Systems, TNO, Soesterberg, Netherlands

<sup>2</sup> Department of Neurology, Vanderbilt University, Nashville, TN, USA

<sup>3</sup> INRIA Bordeaux Sud-Ouest, Talence Cedex, France

## Edited by:

Martin Klasen, RWTH Aachen University, Germany

## Reviewed by:

Thomas Baumgartner, University of Basel, Switzerland

Tobias Brosch, University of Geneva, Switzerland

## \*Correspondence:

Anne-Marie Brouwer, Department of Perceptual and Cognitive Systems, TNO, PO Box 23, Kampweg 5, 3769 ZG Soesterberg, Netherlands  
e-mail: anne-marie.brouwer@tno.nl

Most studies on physiological effects of emotion-inducing images and sounds examine stimulus locked variables reflecting a state of at most a few seconds. We here aimed to induce longer lasting emotional states using blocks of repetitive visual, auditory, and bimodal stimuli corresponding to specific valence and arousal levels. The duration of these blocks enabled us to reliably measure heart rate variability as a possible indicator of arousal. In addition, heart rate and skin conductance were determined without taking stimulus timing into account. Heart rate was higher for pleasant and low arousal stimuli compared to unpleasant and high arousal stimuli. Heart rate variability and skin conductance increased with arousal. Effects of valence and arousal on cardiovascular measures habituated or remained the same over 2-min intervals whereas the arousal effect on skin conductance increased. We did not find any effect of stimulus modality. Our results indicate that blocks of images and sounds of specific valence and arousal levels consistently influence different physiological parameters. These parameters need not be stimulus locked. We found no evidence for differences in emotion induction between visual and auditory stimuli, nor did we find bimodal stimuli to be more potent than unimodal stimuli. The latter could be (partly) due to the fact that our bimodal stimuli were not optimally congruent.

**Keywords:** valence, arousal, heart rate, skin conductance, sensory modality

## INTRODUCTION

Identifying an individual's emotions through (neuro) physiological correlates is desirable in a wide range of situations. Examples are continuous and non-interfering evaluation of products like software (Hazlett and Benedek, 2007), improving communication between humans and computers (Picard, 1997) and monitoring patients suffering from phobia or anxiety or trainees in virtual reality environments (Lang et al., 1998; Repetto et al., 2009; Brouwer et al., 2011).

Even though studies on autonomic responses to emotion reported heterogeneous results (Cacioppo et al., 2000; Kreibig, 2010), in an extensive review of the literature Kreibig (2010) makes the case that recordings from the autonomic nervous system can indeed inform us about specific experienced emotions. According to Stemmler (2004) this is to be expected because emotions have distinct goals which require specific sympathetic "fight or flight" and parasympathetic "rest and digest" autonomic responses in order to prepare the body for the appropriate actions. A number of other models at varying conceptual levels have been proposed to clarify the link between emotion and physiological responses (see for a comprehensive review Kreibig, 2011). All of these models view emotions as the instigator of the physiological changes that in turn adapt the organism to (planned) action.

When setting out to study physiological correlates of emotions, one needs to induce emotion in the experimental participant. A relatively straightforward way to affect emotional state is by

showing emotion inducing pictures. The International Affective Picture System (IAPS) has been developed in 1988 (Lang et al., 1988) and since then used in numerous studies. Each picture in this database was rated by large groups of participants for arousal (ranging from calm to excited), valence (ranging from pleasant to unpleasant) and dominance (ranging from in control to being dominated). Almost all studies investigating physiological responses to emotion inducing pictures use stimulus locked variables; variables that are defined and measured with respect to the moment that the stimulus appeared. For example, heart rate and skin conductance are aligned to image onset to analyze acceleration or peaks, respectively. These variables describe states of a few seconds after stimulus onset. However, in many situations where one would like to measure (neuro)physiological correlates of emotion such as those described in the first paragraph, there are no clear stimuli to lock responses to, or it would be impractical to relate (neuro)physiological variables to specific stimuli. Also, measuring emotional states over longer periods of time than a few seconds would be desirable. A final advantage of measuring stimulus-unlocked physiological responses over longer intervals is that certain potentially informative physiological variables cannot be reliably determined over a few seconds. In particular, intervals of minimally 1 minute are necessary to determine (high frequency) heart rate variability meaningfully (Task Force, 1996; Berntson et al., 1997). An exception to studies using physiological responses locked to emotional stimuli is a study by Baumgartner

et al. (2006a). They induced three different emotional states (happiness, sadness, and fear) by using 70-s sequences of emotional pictures and/or 70-s classical musical excerpts. Heart rate, skin conductance, and respiration variables were found to reflect these three emotional states to some extent. In the current study we aim to induce emotional states by presenting stimuli of specific valence and arousal levels in blocks. We examine heart rate variability, heart rate and skin conductance without using information about stimulus onset. A potential difficulty of using blocks of discrete stimuli to induce emotional states is that observers' responses (emotion) may habituate. Therefore we also examine potential effects of valence and arousal over time. A study by Bradley et al. (1996), in which blocks of unpleasant, pleasant, and neutral images were presented, indicates that at least for stimulus locked measures effects of valence (possibly confounded by arousal) remained constant over time for skin conductance and heart rate, and even increased for facial electromyography.

Like visual stimuli, auditory stimuli can be used to induce emotional states. As mentioned above, Baumgartner et al. (2006a) did not only use pictures, but also musical excerpts to induce happiness, sadness, and fear. They compared subjective and physiological variables between modalities. While involvement was higher for music than for pictures, as reflected by subjective involvement ratings and physiological arousal measures, subjectively reported experienced emotion overlapped better with the intended emotion for pictures than for music. Stimulus material was not chosen in order to specifically vary levels of arousal or valence which perhaps led to few effects of emotion on neuro-physiological measures (reduced skin conductance responses for happy compared to sad and fear conditions, and increased respiration rate for fear and happiness compared to sadness when music was involved). Modality did not affect the physiological distinguishability of the three emotions except for the mentioned interaction effect on respiration rate indicating a larger difference in respiration rate between emotions as evoked by music than pictures. Music differs from pictures in more than modality or modality-related aspects. One difference is that a clear inherent meaning as associated with pictures lacks for music. This may explain that emotion as intended by music in Baumgartner et al. (2006a) was less well recognized for music than for pictures. Analogue to the IAPS, Bradley and Lang (2000) developed the International Affective Digitized Sounds (IADS) database that contains acoustic stimuli rated for arousal, valence, and dominance. These stimuli are relatively short in duration and while some of them are (very short) musical excerpts, the large majority of these sounds are associated with inherent meaning (e.g., a gun shot or the sound of a cheering crowd). Bradley and Lang (2000) found that acoustic stimuli from the IADS produced qualitatively similar physiological reactions as those elicited by visual stimuli from the IAPS. However, they concluded that the effect of acoustic stimuli was often weaker than of visual stimuli by comparing their auditory results to results in other studies examining the effect of visual stimuli on physiological measures. Possible reasons mentioned by Bradley and Lang (2000) for the presumed difference are the specific exemplars of stimuli used (the particular sounds tested may have been less emotional than the particular pictures), effects due to stimulus modality (e.g., more extensive inputs from

visual cortex to other areas in the brain compared to auditory) and reasons concerned with the dynamic nature of sounds versus the static nature of images. While a picture is recognized and processed within an instant, information from sound varies over time and needs to accrue in order to be interpreted. However, while these reasons have been suggested why sounds may have a weaker effect than visual stimuli, it has still not unequivocally been proven that such an effect really exists. We here test whether within a single group of observers rather than different groups, sounds, and pictures with approximately equal scores on valence and arousal do indeed differ with respect to their effect on physiological responses. Moreover, we investigate whether elicited emotions and their physiological correlates increase when audio and visual stimuli are combined. Stronger effects could be caused through a type of summation (Nickerson, 1973) or when interaction of the modalities brings the emotional effect to a next level (superadditive effect; Stein and Meredith, 1993). In an fMRI study, Baumgartner et al. (2006b) showed increased activity in emotion processing brain structures when visual emotional stimuli were combined with congruent musical excerpts compared to visual stimuli alone. Consistent with this, subjective and physiological measures in the study by Baumgartner et al. (2006b) reflected more involvement (arousal) for pictures combined with music compared to pictures alone. While these measures did not significantly differ between bimodal and auditory conditions, EEG alpha power measures suggested strongest activation for bimodal compared to the other conditions.

To summarize, in the current study we use blocks of visual, auditory, and bimodal stimuli to induce certain valence and arousal levels. We determine effects of stimulus modality, valence and arousal, as well as their interaction, on heart rate, heart rate variability, and skin conductance without locking variables to stimulus onset. In the following, we present a short overview of the principles behind these dependent variables and how they have been found to be affected by emotional stimuli in previous valence- and arousal-related (stimulus locked) research.

#### CARDIOVASCULAR MEASURES

Heart rate and its variability are affected by activation and suppression of both the sympathetic and parasympathetic nervous systems. Heart rate variability can be divided along three frequency bands, reflecting three main sources (Mulder, 1988; Veltman and Gaillard, 1998): slow changes (0.02–0.06 Hz) caused by processes like temperature regulation, mid-range changes (0.07–0.14 Hz) related to resonance in the veins caused by the blood pressure regulation, and fast changes (0.15–0.50 Hz) reflecting breathing. Effects of the (rather slow) sympathetic system are visible only in the low and mid frequency bands while effects of the (fast) (Berger et al., 1989) parasympathetic system can be observed in all three bands. Under normal resting conditions, heart rate is carefully adapted to blood pressure such as to keep blood pressure around a certain set point. This adaptation lessens under particular circumstances, such as an increase in mental workload (Mulder, 1980; Aasman et al., 1987), therewith decreasing heart rate variability. Grossman and Taylor (2007) propose that parasympathetically modulated heart rate variability facilitates gas exchange and closely interacts with behavioral,

respiratory, and cardiac parasympathetic mechanisms. When the parasympathetic system is suppressed, this adjustment is less tight and heart rate variability decreases. Being affected by both the sympathetic and parasympathetic system, as well as many other physiological processes, associations between heart rate measures and affective reports have been heterogeneous.

With respect to heart rate, recall of both pleasant and unpleasant memories correlate positively with heart rate acceleration (Vrana and Lang, 1990; Cuthbert et al., 2003; Rainville et al., 2006), suggesting that arousal influences heart rate. Using images of the IAPS, Lang et al. (1993) also found a modest positive effect of arousing images on heart rate acceleration, though they reported to not have found this in a previous pilot study. Rather than a positive effect of arousal on heart rate, Ritz et al. (2005) reported heart rate deceleration when scary or happy IAPS pictures were viewed. Also, Bradley and Lang (2000) found that heart rate deceleration was greater when listening to high arousal unpleasant sounds than when listening to low arousal unpleasant sounds. This arousal effect was not seen for pleasant sounds. These three studies aside, most perception studies show valence rather than arousal effects, where pleasant stimuli correlate with higher heart rate acceleration than unpleasant stimuli (Hare et al., 1970; Libby et al., 1973; Winton et al., 1984; Greenwald et al., 1989; Bradley et al., 1990; Lang et al., 1993, 1998; Bradley and Lang, 2000; Anttonen and Surakka, 2005; Codispoti and De Cesarei, 2007; Sokhadze, 2007).

A recent review on studies that examined the association of heart rate variability and work stress concluded that reported work stress is associated with lower heart rate variability (Chandola et al., 2010). Studies on heart rate variability and emotions are mostly dealing with fear or anxiety (George et al., 1989; Friedman and Thayer, 1998; Rao and Yeregani, 2001) where heart rate variability decreases with increased levels of fear. In a study where participants relived emotions, Rainville et al. (2006) found that besides fear, also sadness and happiness decreased high frequency heart rate variability. In contrast to these studies that suggest a negative relation between heart rate variability and arousal, studies in which emotional visual stimuli were used, report increased heart rate variability for erotic images (Ritz et al., 2005) as well as for aversive visual stimuli (Sokhadze, 2007). Whereas studies on mental workload focus their analyses on mid-frequency heart rate variability (reflecting both sympathetic and parasympathetic control), studies on emotions focus on the high frequency band (only parasympathetic).

### SKIN CONDUCTANCE

Electrical skin conductance varies with the moisture level of the skin. Since the sweat glands are controlled by the sympathetic part of the autonomous nervous system (Roth, 1983), skin conductance measures can be taken to indicate arousal. Indeed, a large number of studies found an increase in skin conductance with arousal (independent of valence) (Tucker and Williamson, 1984; Winton et al., 1984; Greenwald et al., 1989; Bradley et al., 1990; Tremayne and Barry, 1990, 2001; Cook et al., 1991; Boucsein, 1992, 1999; Barry and Sokolov, 1993; Khalfa et al., 2002). As Table 1 in Chanel et al. (2009) indicates, skin conductance measures are perhaps the most popular physiological signal in studies trying

to classify emotional states on the basis of (neuro)physiological signals. Arousal seems more closely associated with increases in skin conductance than heart rate (Barry and Sokolov, 1993; Croft et al., 2004; Wilkes et al., 2010). Skin conductance responses vary with rated arousal in emotional/neutral picture viewing tasks (Lang et al., 1993, 1998; Greenwald et al., 1989).

### AIM AND HYPOTHESES

We here aim to manipulate emotional state using blocks of visual, auditory, and bimodal stimuli and determine its effect on physiological responses. Stimuli will be presented in 2-min blocks, corresponding to specific valence and arousal levels. Using variables for which stimulus onset does not need to be known, we examine effects of stimulus modality, valence, and arousal on cardiovascular measures heart rate and heart rate variability as observed over the first minute (1 minute is needed to reliably determine heart rate variability), and on skin conductance over the first half a minute (for skin conductance shorter intervals are sufficient). The remaining duration of the stimulus block is used to examine the course of possible effects over time.

Previous (stimulus locked) research as described in the two sections above lead us to expect that in our study, heart rate will be higher in pleasant stimuli blocks compared to unpleasant blocks, and skin conductance will be higher for high arousing stimulus blocks compared to low arousing blocks. Arousal may affect heart rate variability. We do not expect valence effects on heart rate variability and skin conductance. Valence and arousal effects may be least explicit for auditory stimuli and strongest for bimodal stimuli. This would be reflected by interaction effects between valence and arousal levels and modality.

## MATERIALS AND METHODS

### PARTICIPANTS

Six female and five male participants were recruited through the participant pool of TNO (the research institute where the study was conducted). The participant pool mainly consists of (former) students of a nearby university. Participants were between 20 and 27-years old with a mean age of 23.1 and a standard deviation of 2.0 years. None of them stated to suffer or have suffered from neurological disorders like epilepsy or cerebral hemorrhage, mental illnesses, diabetes, drugs, or alcohol addiction. Participants received a monetary reward to make up for their travel and time. The study is in accordance with the Declaration of Helsinki and has been approved of by the local ethics committee. All participants signed an informed consent form prior to taking part in the experiment.

### APPARATUS

Images were displayed on a 19" Dell 1907FTP LCD screen and audio was presented through Dell AS501 stereo sound bar speakers.

To record ECG, self-adhesive electrodes were attached after having cleaned the contact area with alcohol wipes. The reference electrode was placed on the manubrium of the sternum; the ECG channel electrode was placed at the left, fifth intercostal space; the

ECG ground electrode was placed 5–8 cm below the ECG channel electrode. Recording frequency was 512 Hz.

Skin conductance was recorded by custom-made equipment. The fingertips of the index- and middle-finger were attached to steel plate electrodes. A small voltage (0.5 V) was applied across the electrodes and the resultant current flow was recorded at 512 Hz.

## STIMULI

Eight pictures and eight sounds were selected from the IAPS and IADS for each of five emotional blocks. These emotional blocks were unpleasant—low arousal, unpleasant—high arousal, pleasant—low arousal, pleasant—high arousal and neutral—low arousal. **Table 1** displays the stimulus numbers of the stimuli used. **Table 2** gives the means and standard deviations of the valence and arousal scores as indicated by the IAPS and IADS technical reports (Lang et al., 2005; Bradley and Lang, 2007) for each of the five emotional blocks. Arousal and valence significantly differed between emotional blocks that intended to vary these values (Wilcoxon rank sum tests: all  $p$ -values < 0.01). Furthermore, the stimuli were selected in such a way that the valence and arousal scores were comparable in value between emotional blocks that correspond in level of valence and arousal (e.g., arousal values were approximately the same in unpleasant—low arousal and pleasant—low arousal blocks as verified by Wilcoxon rank sum tests: all  $p$ -values > 0.1). In the neutral blocks, arousal was inherently low. The valence and arousal scores of the pictures were comparable to those of sounds, only the valence of the low arousal—unpleasant pictures was lower than the low arousal—unpleasant sounds (Wilcoxon rank sum test  $p = 0.01$ —all other comparisons > 0.05). For the bimodal condition, an effort was made to pair IAPS and IADS stimuli in the

most congruent way as possible (e.g., the “aimed gun” picture was paired to the “gun shot” sound). The combinations of specific visual and auditory stimuli for the bimodal conditions are presented in **Table 1**.

## DESIGN AND PROCEDURE

Participants were tested individually in the laboratory. Participants were asked to sit still, and to watch and listen to the stimuli. Stimuli were presented in three modality blocks (visual, auditory, and bimodal). The order of these blocks was picked randomly for each participant. The modality blocks were separated by a 15-s rest period. Each modality block consisted of five emotional blocks (unpleasant—low arousal, unpleasant—high arousal, pleasant—low arousal, pleasant—high arousal and neutral—low arousal), in randomized order. The emotional blocks were separated by a 15-s rest period. Each emotional block consisted of two repetitions of 8 stimuli, in randomized order. Each stimulus was presented for 6 s. Stimuli were separated by blank screens with a jittered duration (average duration 1500 ms). After the stimulus presentation and physiological recordings were over, participants rated the stimuli they had observed on a SAM rating scale for arousal and valence [as used in Lang et al. (2005), Bradley and Lang (2007)]. This was done by presenting them with the stimuli again on a laptop, separately for each modality block.

## SIGNAL PROCESSING

The ECG signal was filtered by a 2–200 Hz band-pass 2-sided Butterworth filter. For each participant, we then determined the median RRI of the first 60 s of each of the 15 stimulus blocks (3 modality  $\times$  5 emotional blocks). RRI is the interval between successive heart beats or more precisely, the interval between subsequent R-peaks in the ECG. The peak detection algorithm used

**Table 1 | Numbers and names of the visual (IAPS) (Lang et al., 2005) and auditory (IADS) (Bradley and Lang, 2007) stimuli used in each of the five blocks (negative—low arousal, positive—high arousal, negative—high arousal, positive—high arousal, neutral—low arousal).**

neg-lower		pos-lower		neg-highar		pos-highar		neutral-lower	
vis	aud	vis	aud	vis	aud	vis	aud	vis	aud
2141	242	1463	110	2730	255	4660	202	1390	152
Grieving fem	Female cough	Kittens	Baby	Native boy	Vomit	Erotic couple	Erotic fem	Bees	Tropical
3230	241	2208	813	3060	625	8030	415	1560	114
Dying man	Male cough	Bride	Wedding	Mutilation	May day	Skier	Count down	Hawk	Cattle
3300	280	4623	221	3150	106	8080	352	2220	251
Disabled child	Woman crying	Romance	Male laugh	Mutilation	Growl	Sailing	Sports crowd	Male face	Nose blow
8230	283	5910	601	6250	289	8180	353	2635	113
Boxer	Fight	Fireworks	Colonial music	Aimed gun	Gun shot	Cliff divers	Base ball	Cowboy	Cows
9120	611	7330	721	9050	501	8185	360	3210	729
Oil fires	Battle taps	Ice cream	Beer	Plane crash	Plane crash	Sky divers	Roller coaster	Surgery	Paper
9520	423	8040	816	9250	600	8200	311	4613	364
Kids	Injury	Diver	Guitar	War victim	Bike wreck	Water skier	Crowd	Condom	Bar
9611	699	8120	820	9252	711	8400	365	7620	410
Plane crash	Bomb	Athlete	Funk music	Dead body	Siren	Rafters	Party	Jet	Helicopter
9301	250	8496	220	9921	244	8501	367	9411	722
Toilet	Male sneeze	Water slide	Boy laugh	Fire	Man wheeze	Money	Casino	Boy	Walking

*Stimuli that were presented simultaneously in the bimodal condition are in the same row of each block.*

**Table 2 | Valence and arousal scores with their standard deviations of stimuli used in each of the five emotional blocks as previously reported in the IAPS (Lang et al., 2005) and IADS (Bradley and Lang, 2007) documentation and as currently reported by our participants.**

	Valence				Arousal			
	Score		SD		Score		SD	
	Prev	Curr	Prev	Curr	Prev	Curr	Prev	Curr
<b>VISUAL BLOCK</b>								
neg-lowar	2.60	2.35	1.66	0.71	5.39	4.63	2.20	0.89
pos-lowar	7.33	7.48	1.59	0.64	5.44	5.33	2.30	1.02
neg-highar	2.29	1.84	1.65	0.54	6.64	6.23	2.15	1.20
pos-highar	7.47	7.13	1.58	0.53	6.75	5.86	2.08	1.15
neutral	5.12	4.85	1.74	0.50	5.06	4.35	2.00	0.81
<b>AUDITORY BLOCK</b>								
neg-lowar	3.18	3.17	1.81	0.40	5.68	4.82	1.81	0.80
pos-lowar	6.98	7.07	1.86	0.64	5.61	5.44	1.86	0.98
neg-highar	2.73	2.92	1.68	0.55	6.80	5.40	1.68	1.01
pos-highar	7.09	6.73	1.84	0.64	6.88	5.94	1.84	1.05
neutral	4.88	5.26	1.76	0.43	5.48	4.05	1.76	1.10
<b>BIMODAL BLOCK</b>								
neg-lowar	2.30		0.53		5.05		0.98	
pos-lowar	7.48		0.71		5.70		1.14	
neg-highar	2.05		0.51		6.13		0.89	
pos-highar	7.14		0.62		5.99		0.79	
Neutral	4.68		0.53		4.16		1.21	

For the bimodal block, we only present valence and arousal scores from our own participants.

to identify these peaks required an R-peak to occur at least 222 ms after the previous one (corresponding to a maximum allowed heart rate of 270 b/m). The first R-peak in an ECG trace needed to be between 1 and 5 mV (as measured between the R-peak and the subsequent S-valley) while subsequent peaks were identified as such if they crossed a threshold starting at the height of the just identified peak and then exponentially decreasing over time to an asymptote of 1 mV. This procedure proved to reliably detect heart beats as indicated by visual inspection of the raw ECG signal with labeled peaks. The inverse of the RRIs provided us with the heart rate. Heart rate variability was computed as the root mean squared successive difference (RMSSD: Goedhart et al., 2007) between the RRIs, normalized by dividing this value by the mean RRI. This measure reflects high frequency heart rate variability.

The skin conductance signal was filtered by a 30 Hz low-pass 2-sided Butterworth filter. For each participant, we determined skin conductance over the first 30 s of each of the 15 stimulus blocks by averaging the inverse of the recorded signal. Subsequently, each value was baselined by subtracting skin conductance averaged over the 10 s preceding the block (i.e., during rest).

In order to indicate the course of the different effects over time, we repeated the analyses of ECG and skin conductance over different time intervals: for ECG also the second half of each block was analyzed and for skin conductance also the second, third and fourth quarter.

## STATISTICAL ANALYSIS

For each dependent variable (heart rate, RMSSD, and skin conductance), a repeated measure ANOVA was performed on data from the unpleasant—low arousal, unpleasant—high arousal, pleasant—low arousal, pleasant—high arousal blocks with modality (3 levels), arousal (2 levels) and valence (2 levels) as independent variables. We chose an alpha level of 0.05.

For the variables that were significantly affected by arousal and/or valence, we computed the size of the effect for different time windows for each participant, averaged over modality (because modality turned out not to affect the responses). This was done by subtracting the value of the variable during low arousal of that during high arousal, and the same for low and high valence. To test for changes of these differences over time, we compared the differences between the first and the last time window using paired *t*-tests.

Data from the neutral conditions were not analyzed here.

## RESULTS

### RATINGS

We analyzed the ratings as provided by our participants for the visual and auditory stimuli in the same way as we analyzed the previously reported ratings by Lang et al. (2005) and Bradley and Lang (2007) when constructing the stimulus set (section Stimuli).

Valence significantly differed between emotional blocks that were intended to vary these values (Wilcoxon rank sum tests: all  $p$ -values  $< 0.01$ ). High arousal and low arousal low valence images significantly differed in arousal ( $p < 0.01$ ). However, the difference between high and low arousal failed to reach significance for high valence images, low valence sounds and high valence sounds. Still, effects of arousal were found on the physiological measures as described in section Effects of Modality, Valence and Arousal on Physiological Variables and Arousal and Valence Effects Over Time.

Arousal did not differ between emotional blocks that were intended to induce equal arousal, both within and between modalities. Consistent with our intention, valence scores did not differ for all high valence emotional blocks, within and between modalities. However, unpleasant high arousal images were judged as lower in valence than unpleasant low arousal images ( $p = 0.02$ ). Also, unpleasant images were rated as lower in valence than unpleasant sounds ( $p = 0.02$  for low arousal, and  $p < 0.01$  for high arousal). Still, we did not find effects including modality on the physiological measures as described in section Effects of Modality, Valence and Arousal on Physiological Variables and Arousal and Valence Effects Over Time.

Bimodal stimuli did not systematically show more extreme arousal and valence ratings compared to unimodal stimuli. Only for low arousal blocks, the difference in valence between pleasant and unpleasant stimuli tends to be larger for bimodal stimuli compared to either unimodal condition: an average 5.18 difference in bimodal valence scores versus 5.13 for the second largest valence difference which was found for the visual condition. However, a Wilcoxon rank sum test on these differences did not indicate a significant modality effect.

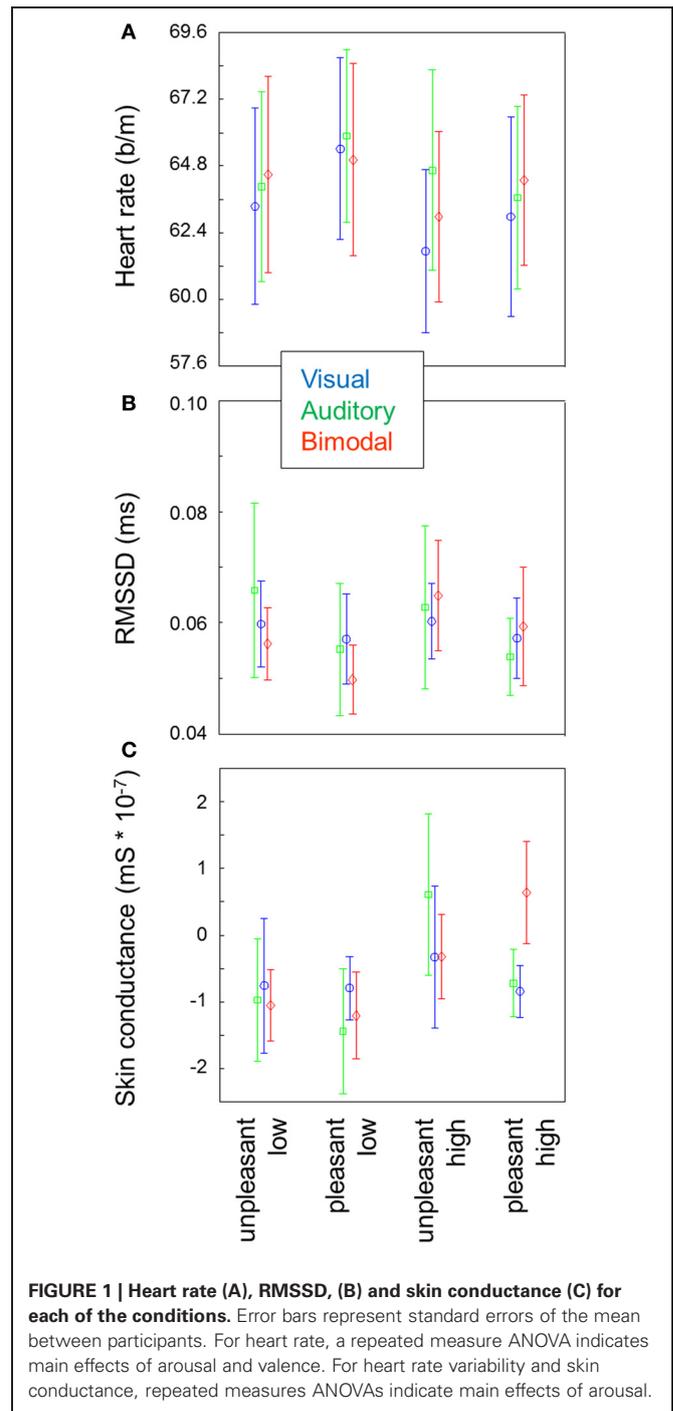
Note that differences between the ratings of our participants and the ones in the studies by Lang et al. (2005) and Bradley and Lang (2007) can be caused by a difference in rating methodology. Whereas Lang and colleagues asked their participants to rate their experienced emotion immediately after stimulus presentation, we presented our participants with the stimuli again at the end of the experiment in order to perform the rating. We modified the rating procedure such as to not disturb the emotion that we intended to elicit through the blocks of stimuli.

#### EFFECTS OF MODALITY, VALENCE, AND AROUSAL ON PHYSIOLOGICAL VARIABLES

Due to technical problems, ECG data were missing for one subject in two conditions and skin conductance data were completely missing for another subject. Data of these subjects are left out in respectively ECG and skin conductance analyses.

**Figure 1A** shows the average heart rate for each of the 15 conditions. Heart rate was lower for high arousal compared to low arousal stimuli [ $F_{(1, 9)} = 25.66$ ,  $p < 0.01$ ], and higher for pleasant stimuli compared to unpleasant stimuli [ $F_{(1, 9)} = 5.59$ ,  $p = 0.04$ ]. There was no effect of modality and no 2- or 3-way interactions between arousal, valence and modality.

Heart rate variability as operationalized by normalized RMSSD was exclusively affected by arousal [ $F_{(1, 9)} = 5.94$ ,  $p = 0.04$ ] where higher variability was found in the high arousal

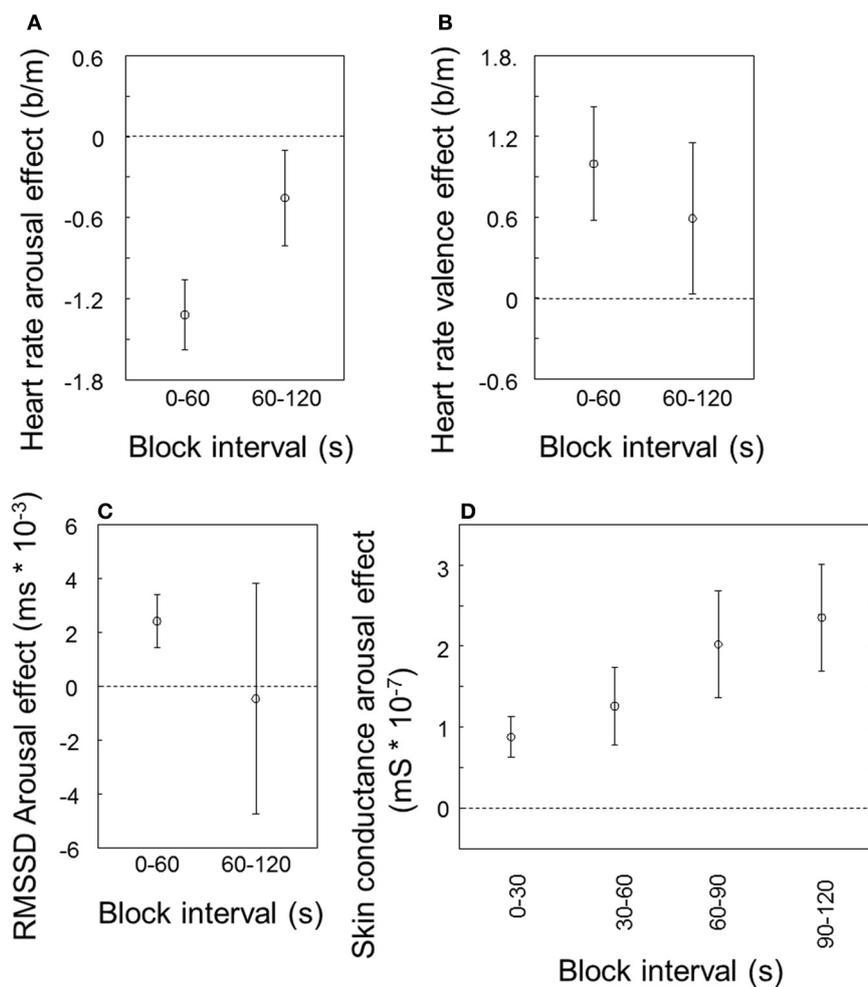


blocks compared to the low arousal blocks (**Figure 1B**). No other main or interaction effects approached significance.

Skin conductance increased with arousal [ $F_{(1, 9)} = 11.83$ ,  $p < 0.01$ ; **Figure 1C**]. No other main or interaction effects approached significance.

#### AROUSAL AND VALENCE EFFECTS OVER TIME

**Figure 2** indicates how the effects of valence and arousal as reported above change over time. The effects of arousal and



**FIGURE 2 | Arousal and valence effects over time, averaged over modalities.** Arousal effect on heart rate (A), valence effect on heart rate (B) and arousal effect on heart rate variability (RMSSD) (C), for the first and

second 60 s of the block; arousal effect on skin conductance for the first through fourth 30 s of the block (D). Error bars represent standard errors of the mean between participants.

valence on heart rate (Figures 2A and B, respectively) tend to be smaller during the second half of the stimulus blocks compared to the first half, only significantly so for arousal [ $t_{(9)} = 2.27$ ,  $p = 0.049$ ]. The effect of arousal on RMSSD (Figure 2C), visible for the first half of the block, has on average disappeared in the second half and dramatically increased in variability between participants. A paired  $t$ -test does not indicate a significant difference between the first and the second half. In contrast to the cardiovascular measures, the effect of arousal on skin conductance becomes stronger rather than (a trend to) weaker (Figure 2D;  $t_{(9)} = 3.06$ ,  $p = 0.01$ ).

#### EFFECT OF OUTLIERS

In order to check for possible effects of outlying values, we identified blocks that differed for more than 3 standard deviations from the mean for heart rate, RMSSD, and skin conductance. There were no outliers for heart rate. For skin conductance, two participants had one or more outlying emotional blocks, and for

RMSSD this was the case for one participant. Excluding these participants respectively from the skin conductance and RMSSD analyses as reported in Arousal and Valence Effects Over Time and did not change the pattern of significant results, indicating that our results cannot be explained by sample outliers.

#### DISCUSSION

We successfully manipulated our observers' emotional state by presenting visual, auditory and bimodal stimuli in blocks, as indicated by effects of valence and arousal on different physiological variables. This was found without locking variables to stimulus onset and despite a large variability in the overall values of the physiological variables between participants (indicated by the large error bars in Figure 1). One physiological variable that has rarely been used in previous emotional perception research, heart rate variability, displayed an effect of arousal. Heart rate variability and skin conductance increased with arousal. Heart rate decreased with arousal and was higher for pleasant stimuli

compared to unpleasant stimuli. Over two minute intervals, cardiovascular effects habituated or tended to habituate whereas the effect on skin conductance increased. We did not find any effect of stimulus modality. We discuss each of these findings below.

### HEART RATE

As described in the Introduction, perception studies that examined the effect of arousal on heart rate produced mixed results: a positive effect (Lang et al., 1993), a negative effect (for unpleasant stimuli—Bradley and Lang, 2000) and no effect (no loading on arousal as indicated by factor analyses—Lang et al., 1998). Studies on recalling emotional (versus neutral) memories generally show increasing heart rate with arousal (Vrana and Lang, 1990; Cuthbert et al., 2003; Rainville et al., 2006). We here find heart rate to decrease with arousal. Generally, heart rate may be expected to increase with arousal in order to get the body ready for action. However, in studies where participants only observe emotional stimuli or events (like ours) another process may dominate. Previous studies showed that allocating attentional resources to a perceived stimulus elicits heart rate deceleration over the first few seconds after stimulus onset (Lacey and Lacey, 1970; Graham, 1992; Codispoti et al., 2001). High arousal probably causes increased information processing or attention and thus a larger drop in heart rate.

The lower heart rate when unpleasant stimuli were presented compared to pleasant stimuli is consistent with the large majority of the perception literature (Hare et al., 1970; Libby et al., 1973; Winton et al., 1984; Greenwald et al., 1989; Bradley et al., 1990; Lang et al., 1993, 1998; Bradley and Lang, 2000; Anttonen and Surakka, 2005; Codispoti and De Cesarei, 2007; Sokhadze, 2007) though there are a few exceptions (Ritz et al., 2005; Dimberg and Thunberg, 2007). It should be noted though that this valence effect on heart rate is not a general finding when looking at the emotion literature as a whole. In her literature review, Kreibig (2010) showed that the effect of valence on heart rate can be in both directions. As with the effect of arousal on heart rate, the type of emotional stimuli seems crucial. Kreibig proposes that an important distinguishing factor is passivity. Emotions involving an element of passivity (e.g. non-crying sadness, contentment) rather result in heart rate decrease in contrast to more “active” emotions (joy, anger). In this vein, one could speculate that unpleasant pictures elicit sadness rather than anger, and pleasant pictures elicit joy rather than contentment resulting in the valence effect on heart rate as reported in the perception literature as well as in the present study.

### HEART RATE VARIABILITY

Most studies on perception of emotional stimuli do not report heart rate variability because physiological variables are analyzed over short time intervals and locked to stimuli. Two studies that did, reported increased high frequency heart rate variability for erotic (Ritz et al., 2005) and aversive (Sokhadze, 2007) images. We also found an increase of heart rate variability with arousal. In contrast, heart rate variability is reported to decrease with arousal in the fields of stress (Chandola et al., 2010) and fear or anxiety (George et al., 1989; Friedman and Thayer, 1998; Rao and Yerevani, 2001). As with heart rate, this difference may be related

to the type of stimuli used and with it, the required action and the exact quality of the experienced emotion. Another factor that may be important here is breathing frequency. Slow, deep breathing produces a sharp increase in heart rate variability (Angelone and Coulter, 1964; Grossman and Taylor, 2007). In our study, participants may have taken a few deep breaths during high and not during low arousal conditions, explaining the high heart rate variability for high arousal. Ritz et al. (2005) showed that emotional pictures can indeed differentially influence respiration. However, they corrected for this in their measure of high frequency heart rate variability and still found an increase in variability with (erotic) arousal pictures.

### SKIN CONDUCTANCE

Our finding that skin conductance increased with stimulus arousal is well in line with results described in the literature (Tucker and Williamson, 1984; Winton et al., 1984; Greenwald et al., 1989; Bradley et al., 1990; Tremayne and Barry, 1990, 2001; Cook et al., 1991; Boucsein, 1992, 1999; Barry and Sokolov, 1993; Khalfa et al., 2002). Several authors stated that arousal is more closely connected to skin conductance than to heart rate (Barry and Sokolov, 1993; Croft et al., 2004; Wilkes et al., 2010). The positive effect of arousal on skin conductance in our study suggests that our stimuli, despite the decrease in heart rate and the increase in heart rate variability, indeed induced different arousal levels as intended.

### EFFECTS OF EMOTIONAL STIMULI OVER LONGER INTERVALS

We found that over two minute intervals, valence and arousal effects in cardiovascular measures (tend to) habituate, whereas the effect of arousal on skin conductance increases. Parasympathetic effects are faster (in the order of milliseconds) than sympathetic effects (in the order of seconds; Rainville et al., 2006; Grossman and Taylor, 2007). However, it is unlikely that this can explain the difference between the courses of the effects over minutes. Another reason might be that while sweat glands are controlled by the sympathetic system, the heart is innervated by both sympathetic and parasympathetic systems, therewith possessing a “brake” that the sweat glands lack. Finally, building a layer of sweat is a relatively slow process. For stimulus locked variables and image blocks of several minutes in duration, Bradley et al. (1996) reported a constant image valence effect on skin conductance and heart rate.

### MODALITIES

Recording responses within a single group of participants, we did not find weaker effects for auditory stimuli compared to visual stimuli as suggested by Bradley and Lang (2000)—even though our match of auditory and visual stimuli was not perfect (see section Stimuli). For heart rate, a variable that Bradley and Lang (2000) specifically reported to be affected relatively little by emotional sounds, there was a trend for visual stimuli to exert the strongest effects (arousal and valence; **Figure 1A**) but it was not close to significance. Bimodal stimuli did not enhance arousal or valence effects over unimodal stimuli; only in heart rate variability (**Figure 1B**) the expected trend was found, but again, this was far from significant. Note that our participants also did not rate

bimodal stimuli more extremely in valence and arousal than unimodal stimuli. Possibly, including more participants would have resulted in (significant) modality effects but at least, our findings show that potential modality effects are weak. In our experiment, bimodal presentation may not have enhanced emotion because the stimuli were not optimally congruent; though not completely off, in most cases they clearly did not originate from the same source. Baumgartner et al. (2006a) used musical stimuli to induce emotions. While music is less comparable to IAPS pictures than IADS sounds, the advantage of music over sounds is that it can more easily be combined with pictures to produce bimodal stimuli that can be considered congruent. Still, Baumgartner and colleagues also did not find stronger involvement (arousal) as reflected by subjective and physiological measures for bimodal stimuli compared to auditory stimuli alone.

## CONCLUSION

In line with Bradley et al. (1996), our results strongly suggest that sustained emotions can be elicited by repeatedly presenting visual and auditory stimuli of similar arousal and valence over time intervals of at least two minutes. This provides a tool for emotion manipulation in emotion research and possibly in treatment or training situations, e.g., where individuals need to learn to optimally function during the experience of negative emotions. Care must be taken not to simply generalize results of studies using passive perception of emotional stimuli to situations with

(other) emotional stimuli that elicit (other) actions than the ones being studied. Another limitation is our relatively small sample size. We measured effects of valence and arousal using physiological variables that were not locked to stimulus onset. This is practical and necessary for measuring emotions in real life situations where there are no specific stimuli to lock responses to or where the exact time of stimulus onset is unknown, such as when evaluating valence or arousal elicited during human-machine interaction or advertisement movies. We found that valence and arousal effects of real life images on heart rate, heart rate variability and skin conductance do not differ from the effects of real life sounds. This finding suggests that emotional auditory stimuli can be used in situations where visual presentation is less practical (e.g., when the visual sensory channel is already occupied). Presenting visual and auditory stimuli together does not enhance the effects. Possibly effects can be enhanced when bimodal stimuli are more congruent, i.e., clearly originate from the same source.

## ACKNOWLEDGMENTS

The authors would like to thank Dennis Coetsier, Hans Veltman, Wouter Vos (all at TNO Human Factors, Netherlands), Ben Mulder and Dick de Waard (both at the University of Groningen, Netherlands) as well as gratefully acknowledge the support of the BrainGain Smart Mix Programme of the Netherlands Ministry of Economic Affairs and Netherlands Ministry of Education, Culture, and Science.

## REFERENCES

- Aasman, J., Mulder, G., and Mulder, L. J. M. (1987). Operator effort and the measurement of heart rate variability. *Hum. Factors* 29, 161–170.
- Angelone, A., and Coulter, N. A. (1964). Respiratory-sinus arrhythmia: a frequency dependent phenomenon. *J. Appl. Physiol.* 19, 479–482.
- Anttonen, J., and Surakka, V. (2005). “Emotions and heart rate while sitting on a chair,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI 2005 (Portland, OR), 491–499, April 2005.
- Barry, R. J., and Sokolov, E. N. (1993). Habituation of phasic and tonic components of the orienting reflex. *Int. J. Psychophysiol.* 15, 39–42. doi: 10.1016/0167-8760(93)90093-5
- Baumgartner, T., Esslen, M., and Jäncke, L. (2006a). From emotion perception to emotion experience: emotions evoked by pictures and classical music. *Int. J. Psychophysiol.* 60, 34–43. doi: 10.1016/j.ijpsycho.2005.04.007
- Baumgartner, T., Lutz, K., Schmidt, C. F., and Jäncke, L. (2006b). The emotional power of music: how music enhances the feeling of affective pictures. *Brain Res.* 1075, 151–164. doi: 10.1016/j.brainres.2005.12.065
- Berger, R. D., Saul, J. P., and Cohen, R. J. (1989). Transfer function analysis of autonomic regulation. I. Canine atrial rate response. *Am. J. Physiol. Heart Circ. Physiol.* 256, H142–H152.
- Berntson, G. G., Bigger, J. T., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., et al. (1997). Heart rate variability: origins, methods, and interpretive caveats. *Psychophysiology* 34, 623–648. doi: 10.1111/j.1469-8986.1997.tb02140.x
- Boucsein, W. (1992). *Electrodermal Activity*. New York, NY: Plenum Press. doi: 10.1007/978-1-4757-5093-5
- Boucsein, W. (1999). Electrodermal activity as an indicator of emotional processes. *Korean J. Sci. Emot. Sensibility* 2, 1–25.
- Bradley, M. M., Cuthbert, B. N., and Lang, P. J. (1990). Startle reflex modification: motion or attention. *Psychophysiology* 27, 513–522. doi: 10.1111/j.1469-8986.1990.tb01966.x
- Bradley, M. M., Cuthbert, B. N., and Lang, P. J. (1996). Picture media and emotion: effects of a sustained affective context. *Psychophysiology* 33, 662–670. doi: 10.1111/j.1469-8986.1996.tb02362.x
- Bradley, M. M., and Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology* 37, 204–215. doi: 10.1111/1469-8986.3720204
- Bradley, M. M., and Lang, P. J. (2007). “The International affective digitized sounds affective ratings of sounds and instruction manual,” *Technical Report B-3 (2nd Edn. IADS-2)*. Gainesville, FL: University of Florida.
- Brouwer, A. M., Neerinx, M. A., Kallen, V. L., van der Leer, L., and ten Brinke, M. (2011). EEG alpha asymmetry, heart rate variability and cortisol in response to virtual reality induced stress. *J. Cyberther. Rehab.* 4, 21–34.
- Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., and Ito, T. A. (2000). “The psychophysiology of emotion,” in *The Handbook of Emotion, 2nd Edn.*, eds M. Lewis, and J. M. Haviland-Jones (New York, NY: Guilford Press), 173–191.
- Chandola, T., Heraclides, A., and Kumari, M. (2010). Psychophysiological biomarkers of workplace stressors. *Neurosci. Biobehav. Rev.* 35, 51–57. doi: 10.1016/j.neubiorev.2009.11.005
- Chanel, G., Kierkels, J. J. M., Soleymani, M., and Pun, T. (2009). Short-term emotion assessment in a recall paradigm. *Int. J. Hum. Comp. Stud.* 67, 607–627. doi: 10.1016/j.ijhcs.2009.03.005
- Codispoti, M., Bradley, M. M., and Lang, P. J. (2001). Affective reactions to briefly presented pictures. *Psychophysiology* 38, 474–478. doi: 10.1111/1469-8986.3830474
- Codispoti, M., and De Cesarei, A. (2007). Arousal and attention: picture size and emotional reactions. *Psychophysiology* 44, 680–686. doi: 10.1111/j.1469-8986.2007.00545.x
- Cook, E. W. III, Hawk, L. H., Davis, T. L., and Stevenson, V. E. (1991). Affective individual differences and startle reflex modulation. *J. Abnorm. Psychol.* 100, 5–13. doi: 10.1037/0021-843X.100.1.5
- Croft, R. J., Gonsalveza, C. J., Gandra, J., Lechema, L., and Barrya, R. J. (2004). Differential relations between heart rate and skin conductance, and public speaking anxiety. *J. Behav. Ther. Exp. Psychiatry* 35, 259–271. doi: 10.1016/j.jbtep.2004.04.012
- Cuthbert, B. N., Lang, P. J., Strauss, C., Drobos, D., Patrick, C. J., and Bradley, M. M. (2003). The psychophysiology of anxiety disorder: fear memory imagery. *Psychophysiology* 40, 407–422. doi: 10.1111/1469-8986.00043
- Dimberg, U., and Thunberg, M. (2007). Speech anxiety and rapid emotional reactions to angry and happy facial expressions. *Scand. J. Psychol.*

- 48, 321–328. doi: 10.1111/j.1467-9450.2007.00586.x
- Friedman, B. H., and Thayer, J. F. (1998). Autonomic balance revisited: panic anxiety and heart variability. *J. Psychonom. Res.* 44, 133–151. doi: 10.1016/S0022-3999(97)00202-X
- George, D. T., Nutt, D. J., Walker, W. V., Porgers, S. W., Adinoff, B., and Linnoila, M. (1989). Lactate and hyperventilation substantially attenuates vagal tone in normal volunteers. *Arch. Gen. Psychiatry* 46, 153–156. doi: 10.1001/archpsyc.1989.01810020055009
- Goedhart, A. D., Van der Sluis, S., Houtveen, J. H., Willemsen, G., and De Geus, E. J. C. (2007). Comparison of time and frequency domains of RSA in ambulatory recordings. *Psychophysiology* 44, 203–215. doi: 10.1111/j.1469-8986.2006.00490.x
- Graham, F. K. (1992). “Attention: the heartbeat, the blink, and the brain. Attention and information processing in infants and adults,” in *Perspectives from Human and Animal research*, eds B. A. Campbell, H. Hayne, and R. Richardson, (Hillsdale, NJ: Lawrence Erlbaum Associates), 3–29.
- Greenwald, M. K., Cook, E. W., and Lang, P. J. (1989). Affective judgment and psychophysiological response: dimensional covariation in the evaluation of pictorial stimuli. *J. Psychophysiol.* 3, 51–64.
- Grossman, P., and Taylor, E. W. (2007). Toward understanding respiratory sinus arrhythmia: relations to cardiac vagal tone, evolution and biobehavioral functions. *Biol. Psychol.* 74, 263–285. doi: 10.1016/j.biopsycho.2005.11.014
- Hare, R. D., Wood, K., Britain, S., and Shadman, J. (1970). Autonomic responses to affective visual stimuli. *Psychophysiology* 7, 408–417. doi: 10.1111/j.1469-8986.1970.tb01766.x
- Hazlett, R. L., and Benedek, J. (2007). Measuring emotional valence to understand the user’s experience of software. *Int. J. Hum. Comp. Stud.* 65, 306–314. doi: 10.1016/j.ijhcs.2006.11.005
- Khalifa, S., Peretz, I., Blondin, J.-P., and Manon, R. (2002). Event-related skin conductance responses to musical emotions in humans. *Neurosci. Lett.* 328, 145–149. doi: 10.1016/S0304-3940(02)00462-7
- Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: a review. *Biol. Psychol.* 84, 394–421. doi: 10.1016/j.biopsycho.2010.03.010
- Kreibig, S. D. (2011). “Emotion, motivation, and cardiovascular response,” in: *How Motivation Affects Cardiovascular Response: Mechanisms and Applications*, eds R. A. Wright and G. H. E. Gendolla (Washington, DC: American Psychological Association), 93–117.
- Lacey, J. I., and Lacey, B. C. (1970). “Some automatic-central nervous system interrelationships,” *Physiological Correlates of Emotion*, ed P. Black (New York, NY: Academic Press), 205–227.
- Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (2005). “International affective picture system (IAPS): Affective ratings of pictures and instruction manual,” in *Technical Report* (Gainesville, FL: University of Florida). doi: 10.1016/j.ijpsycho.2004.05.009
- Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (1998). Emotion, Motivation and anxiety: Brain mechanisms and psychophysiology. *Biol. Psychiatry* 44, 1248–1263. doi: 10.1016/S0006-3223(98)00275-3
- Lang, P. J., Öhman, A., and Vaitl, D. (1988). *The International Affective Picture System Photographic Slides*. Gainesville, FL: Center for Research in Psychophysiology, University of Florida.
- Lang, P. J., Greenwald, M. K., Bradley, M. M., and Hamm, A. O. (1993). Looking at pictures: affective, facial, visceral, and behavioral reactions. *Psychophysiology* 30, 261–273. doi: 10.1111/j.1469-8986.1993.tb03352.x
- Libby, W. L. Jr., Lacey, B. C., and Lacey, J. L. (1973). Pupillary and cardiac activity during visual attention. *Psychophysiology* 10, 210–294. doi: 10.1111/j.1469-8986.1973.tb00526.x
- Mulder, G. (1980). *The Heart of Mental Effort*. Thesis, University of Groningen, Groningen.
- Mulder, L. J. M. (1988). *Assessment of Cardiovascular Reactivity by Means of Spectral Analysis*. Thesis, University of Groningen, Groningen.
- Nickerson, R. S. (1973). Intersensory facilitation of reaction time: energy summation or preparation enhancement. *Psychol. Rev.* 80, 489–509. doi: 10.1037/h0035437
- Picard, R. W. (1997). *Affective Computing*. Cambridge, MA: MIT Press.
- Rainville, P., Bechara, A., Naqvi, N., and Damasio, A. R. (2006). Basic emotions are associated with distinct patterns of cardiorespiratory activity. *Int. J. Psychophysiol.* 61, 5–18. doi: 10.1016/j.ijpsycho.2005.10.024
- Rao, R. K., and Yeregani, V. K. (2001). Decreased chaos and increased non-linearity of heart rate time series in patients with panic disorder. *Auton. Neurosci.* 88, 99–108. doi: 10.1016/S1566-0702(01)00219-3
- Repetto, C., Gorini, A., Vigna, C., Algeri, D., Pallavicini, F., and Riva, G. (2009). The use of biofeedback in clinical virtual reality: the INTREPID project. *J. Vis. Exp.* 12: 1554. doi: 10.3791/1554
- Ritz, T., Thöns, M., Fahrenkrug, S., and Dahme, B. (2005). Airways, respiration, and respiratory sinus arrhythmia during picture viewing. *Psychophysiology* 42, 568–578. doi: 10.1111/j.1469-8986.2005.00312.x
- Roth, W. T. (1983). “A comparison of P300 and the skin conductance response,” in *Tutorials in ERP Research—Endogenous Components*, eds A. W. K. Gaillard and W. Ritter (Amsterdam: North-Holland), 177–199.
- Sokhadze, E. M. (2007). Effects of music on the recovery of autonomic and electrocortical activity after stress induced by aversive visual stimuli. *Appl. Psychophysiol. Biofeedback* 32, 31–50. doi: 10.1007/s10484-007-9033-y
- Stein, B. E., and Meredith, M. A. (1993). *The Merging of the Senses*. Cambridge, MA: MIT Press.
- Stemmler, G. (2004). “Physiological processes during emotion,” in *The Regulation of Emotion*, eds P. Philippot and R. S. Feldman (Mahwah, NJ: Erlbaum), 33–70.
- Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology. (1996). Heart rate variability: standards of measurement, physiological interpretation and clinical use. *Circulation* 93, 1043–1065. doi: 10.1161/01.CIR.93.5.1043
- Tremayne, P., and Barry, R. J. (1990). Repression of anxiety and its effects upon psychophysiological responses to relevant and irrelevant stimuli in competitive gymnasts. *J. Exerc. Sport Psychol.* 12, 327–352.
- Tremayne, P., and Barry, R. J. (2001). Elite pistol shooters: physiological patterning of best vs. worst shots. *Int. J. Psychophysiol.* 41, 19–29. doi: 10.1016/S0167-8760(00)00175-6
- Tucker, D. M., and Williamson, P. A. (1984). Asymmetric neural control systems in human self-regulation. *Psychol. Rev.* 91, 185–215. doi: 10.1037/0033-295X.91.2.185
- Vrana, S. R., and Lang, P. J. (1990). Fear imagery and the startle-probe reflex. *J. Abnorm. Psychol.* 99, 189–197. doi: 10.1037/0021-843X.99.2.189
- Veltman, J. A., and Gaillard, A. W. K. (1998). Physiological workload reaction to increasing levels of task difficulty. *Ergonomics* 41, 656–669. doi: 10.1080/001401398186829
- Wilkes, B. L., Gonsalvez, C. J., and Blaszczyński, A. (2010). Capturing SCL and HR changes to win and loss events during gambling on electronic machines. *Int. J. Psychophysiol.* 78, 265–272. doi: 10.1016/j.ijpsycho.2010.08.008
- Winton, W. M., Putnam, L. E., and Krauss, R. M. (1984). Facial and autonomic manifestations of the dimensional structure of emotion. *J. Exp. Soc. Psychol.* 20, 195–216. doi: 10.1016/0022-1031(84)90047-7

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 18 February 2013; accepted: 04 June 2013; published online: 21 June 2013.

Citation: Brouwer A-M, van Wouwe N, Mühl C, van Erp J and Toet A (2013) Perceiving blocks of emotional pictures and sounds: effects on physiological variables. *Front. Hum. Neurosci.* 7:295. doi: 10.3389/fnhum.2013.00295

Copyright © 2013 Brouwer, van Wouwe, Mühl, van Erp and Toet. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



# On the role of crossmodal prediction in audiovisual emotion perception

Sarah Jessen<sup>1\*</sup> and Sonja A. Kotz<sup>2,3</sup>

<sup>1</sup> Research Group "Early Social Development," Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

<sup>2</sup> Research Group "Subcortical Contributions to Comprehension," Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

<sup>3</sup> School of Psychological Sciences, University of Manchester, Manchester, UK

## Edited by:

Martin Klasen, RWTH Aachen University, Germany

## Reviewed by:

Erich Schröger, University of Leipzig, Germany

Lluís Fuentemilla, University of Barcelona, Spain

## \*Correspondence:

Sarah Jessen, Research Group "Early Social Development," Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstr. 1A, 04103 Leipzig, Germany  
e-mail: jessen@cbs.mpg.de

Humans rely on multiple sensory modalities to determine the emotional state of others. In fact, such multisensory perception may be one of the mechanisms explaining the ease and efficiency by which others' emotions are recognized. But how and when exactly do the different modalities interact? One aspect in multisensory perception that has received increasing interest in recent years is the concept of *cross-modal prediction*. In emotion perception, as in most other settings, visual information precedes the auditory information. Thereby, leading in visual information can facilitate subsequent auditory processing. While this mechanism has often been described in audiovisual speech perception, so far it has not been addressed in audiovisual emotion perception. Based on the current state of the art in (a) cross-modal prediction and (b) multisensory emotion perception research, we propose that it is essential to consider the former in order to fully understand the latter. Focusing on electroencephalographic (EEG) and magnetoencephalographic (MEG) studies, we provide a brief overview of the current research in both fields. In discussing these findings, we suggest that emotional visual information may allow more reliable predicting of auditory information compared to non-emotional visual information. In support of this hypothesis, we present a re-analysis of a previous data set that shows an inverse correlation between the N1 EEG response and the duration of visual emotional, but not non-emotional information. If the assumption that emotional content allows more reliable predicting can be corroborated in future studies, cross-modal prediction is a crucial factor in our understanding of multisensory emotion perception.

**Keywords:** cross-modal prediction, emotion, multisensory, EEG, audiovisual

Perceiving others' emotions is an important component of everyday social interaction. We can gather such information via somebody's vocal, facial, or body expressions, and by the content of his or her speech. If the information obtained by these different modalities is congruent, a correct interpretation appears to be faster and more efficient. This becomes evident at the behavioral level, for instance, in shorter reaction times (Giard and Peronnet, 1999; Sperdin et al., 2009) and higher accuracy (Giard and Peronnet, 1999; Kreifelts et al., 2007), but also at the neural level where clear differences between unisensory and multisensory processing can be observed. An interaction between complex auditory and visual information can be seen within 100 ms (e.g., van Wassenhove et al., 2005; Stekelenburg and Vroomen, 2007) and involves a large network of brain regions ranging from early uni- and multisensory areas, such as the primary auditory and the primary visual cortex (see, e.g., Calvert et al., 1998, 1999; Ghazanfar and Schroeder, 2006) and the superior temporal gyrus (Calvert et al., 2000; Callan et al., 2003), to higher cognitive brain regions, such as the prefrontal cortex and the cingulate cortex (e.g., Laurienti et al., 2003). These data are interpreted to support the assumption of multisensory facilitation.

The fact that multisensory perception leads to facilitation is generally accepted, however, the mechanisms underlying such facilitation, especially for complex dynamic stimuli, are yet to be fully understood. One mechanism that seems to be particularly important in audiovisual perception of complex, ecologically valid information, is cross-modal prediction. In a natural context, visual information typically precedes auditory information (Chandrasekaran et al., 2009; Stekelenburg and Vroomen, 2012). Visual information leads while the auditory one is lagging behind. Thereby, visual information allows generating predictions about several aspects of a subsequent sound, such as the time of its onset and content (e.g., Arnal et al., 2009; Stekelenburg and Vroomen, 2012). Due to this preparatory information flow, the following auditory information processing is facilitated. This mechanism can be seen as an instance of predictive coding as has been discussed for sensory perception in general (see Summerfield and Egner, 2009).

The success and efficiency of cross-modal prediction is influenced by several factors, including attention, motivation, and the emotional state of the observer. Schroeder et al. (2008) for instance suggest an influence of attention on cross-modal prediction in speech perception. In the present paper, however, we

will focus on a different aspect of cross-modal prediction that has largely been neglected: How does the emotional content of the perceived signal influence cross-modal prediction, or, vice versa, what role does cross-modal prediction play in the multisensory perception of emotions? Do emotions lead to a stronger prediction than comparable neutral stimuli or are emotions just another instance of complex salient information?

In the following, we will provide a short overview of recent findings on cross-modal prediction, focusing on electroencephalographic (EEG) and magnetoencephalographic (MEG) results. We will then discuss the role of affective information in cross-modal prediction before outlining necessary further steps to closer investigate this phenomenon.

## CROSS-MODAL PREDICTION

The most common setting, in which cross-modal prediction of complex stimuli is studied, is in audiovisual speech perception (Bernstein et al., 2008; Arnal et al., 2009, 2011). Typically, videos are presented, in which a person is uttering a single syllable. As visual information starts before a sound's onset, its influence on auditory processing can be investigated.

In EEG and MEG studies, it has been shown that the predictability of an auditory signal by visual information affects the brain's response to the auditory information within 100 ms after a sound's onset. Especially the N1 has been studied in this context (e.g., Klucharev et al., 2003; Besle et al., 2004; van Wassenhove et al., 2005), and a reduction of the N1 amplitude has been linked to facilitated processing of audiovisual speech (Besle et al., 2009). Furthermore, the more predictable visual information is, the stronger such facilitation seems to be, as suggested in MEG studies that reported a reduction in M100 latency (Arnal et al., 2009) and amplitude (Davis et al., 2008). Similar results have been obtained in EEG studies; when syllables of different predictability are presented, the syllables with the highest predictability based on visual features lead to the strongest reduction in N1/P2 latency (van Wassenhove et al., 2005).

Cross-modal prediction in complex settings has not only been investigated in speech perception, but also in the perception of other audiovisual events, such as everyday actions (e.g., Stekelenburg and Vroomen, 2007, 2012). Only if sufficiently predictive dynamic visual information is present, a reduction in the auditory N1 can be observed (Stekelenburg and Vroomen, 2007).

Regarding the mechanisms underlying such cross-modal prediction, two distinct pathways have been suggested (Arnal et al., 2009). In a first, indirect pathway, information from early visual areas influences activations in auditory areas via a third, relay area such as the superior temporal sulcus (STS). In a second, direct pathway, a cortico-cortical connection between early visual and early auditory areas is posited without the involvement of any additional area. Interestingly, these two pathways seem to cover different aspects of prediction; while the direct pathway is involved in generating predictions regarding the onset of an auditory stimulus, the indirect pathway rather predicts auditory information at the content-level, for instance, which syllable or sound will be uttered (Arnal et al., 2009). Evidence for a distinction between two pathways also arises from EEG data (Klucharev et al., 2003; Stekelenburg and de Gelder, 2004): while the N1

is assumed to be modulated by predictability of physical stimulus parameters, the P2 seems to be sensitive to the content or the semantic features of the signal (Stekelenburg and Vroomen, 2012).

In recent years, neural oscillations as a crucial mechanism underlying cross-modal prediction have come into focus (e.g., Doesburg et al., 2008; Schroeder et al., 2008; Senkowski et al., 2008; Arnal et al., 2011; Thorne et al., 2011). While the analysis of event-related potentials offers a straight-forward and reliable way to investigate brain responses closely time-locked to a specific event, the analysis of oscillatory activity provides a way to analyze changes in the EEG data with more flexible timing. Furthermore, oscillatory brain activity has been suggested as a potential mechanism to mediate the influence of one brain area onto another (Buzsaki and Draguhn, 2004). Such a mechanism may, for instance, underlie cross-modal prediction, where information from one sensory area affects the activity in a different sensory area (Kayser et al., 2008; Schroeder et al., 2008; Lakatos et al., 2009). In the case of audiovisual prediction, visual information, processed in primary visual areas, thereby has the capacity to prepare auditory areas for incoming auditory information. However, such an operation takes time (Schroeder et al., 2008), and it is therefore essential that visual information precedes the auditory one. Further, it has to provide some information about the upcoming auditory stimulus, such as an expected onset and, preferably, more detailed specification of a sound.

In summary, cross-modal prediction has been extensively studied in audiovisual speech perception and also in the perception of lower-level audiovisual stimuli. Along with an increasing interest in neural oscillations and their function(s) in recent years, new approaches and possibilities to investigate its underlying mechanisms have been developed. However, the role of cross-modal prediction in emotion perception has received hardly any attention. In the following, we will outline what is known regarding the role of emotions in cross-modal predictions.

## EMOTIONS AND CROSS-MODAL PREDICTION

Emotion perception is a case that involves cross-modal prediction. Cross-modal prediction likely contributes to the ease and efficiency with which others' emotions are recognized. One question that arises is whether emotion perception is just one case of cross-modal prediction among others, or whether it differs substantially from cases of non-emotional cross-modal prediction.

Numerous recent studies have investigated the combined perception of emotions from different modalities (e.g., de Gelder et al., 1999; Pourtois et al., 2000, 2002; for a recent review, see Klasen et al., 2012). Emotional faces, bodies, and voices influence each other at various processing stages.

First brain responses to a mismatch between facial and vocal expressions (de Gelder et al., 1999; Pourtois et al., 2000) or also between body and facial expressions (Meeren et al., 2005) can be observed around 100 ms after stimulus onset. Interactions of matching emotional faces and voices are typically observed slightly later, between 200 and 300 ms (Paulmann et al., 2009), though some studies also report interaction effects in the range of the N1 (Jessen and Kotz, 2011). Besides these early effects, interactions between different modalities can be observed at

later processing stages, presumably in limbic areas and higher association cortices (Pourtois et al., 2002; Chen et al., 2010).

However, while the processing of multisensory emotional information has been amply investigated, only recently the dynamic temporal development of the perceived stimuli has come into focus. Classically, most studies used static facial expressions paired with (by its very nature) dynamic vocal expressions (e.g., de Gelder et al., 1999; Pourtois et al., 2000).

While this allows for investigating several aspects of emotion perception under controlled conditions, it is a strong simplification compared to a dynamic multisensory environment. In a natural setting, emotional information usually obeys the same patterns as outlined above: visual information precedes the auditory one. We see an angry face, see a mouth opening, see a breath-intake before we actually hear an outcry or an angry exclamation.

One aspect of such natural emotion perception that cannot be investigated using static stimulus material is the role of prediction in emotion perception. If auditory and visual onsets occur at the same time, we cannot investigate the influence of preceding visual information on the subsequent auditory one. However, two aspects of these studies using static facial expression render them particularly interesting and relevant in the present case.

First, several studies introduced a delay between the onset of a picture and a voice onset in order to differentiate between brain responses to the visual onset and brain responses to the auditory onset (de Gelder et al., 1999; Pourtois et al., 2000, 2002). At the same time, however, such a delay introduces visual, albeit static, information, which allows for the generation of predictions. At which level these predictions can be made depends on the precise experimental setup. While some studies chose a variable delay (de Gelder et al., 1999; Pourtois et al., 2000), allowing for predictions only at the content, but not at the temporal level, others presented auditory information at a fixed delay, which allows for predictions both at the temporal and at a content level (Pourtois et al., 2002). In either case, one can conceive of the results as investigating the influence of static emotional information on subsequent matching or mismatching auditory information.

Second, most studies used a mismatch paradigm, that is, a face and a voice were either of different emotions or one modality was emotional while the other was neutral (de Gelder et al., 1999; Pourtois et al., 2000, 2002). These mismatch settings were then contrasted to matching stimuli, were a face and a voice conveyed the same emotion (or both did not show any emotional information, in a neutral case). While probably not intended by the researchers, such a design may reduce predictive validity to a rather large degree; after the first number of trials, the participant learns that a given facial expression may be followed either by the same or by a different emotion with equal probability. Conscious predictions cannot be made, neither at the content (emotional) level, nor at a more physical level based on facial features. Hence, visual information provides only limited information about subsequent auditory information. Therefore, data obtained from these studies informs us about multisensory emotion processing under conditions, in which predictive capacities are reduced. Note, however, that it is unclear to what extent

one experimental session can reduce the predictions generated by facial expressions, or rather, how much of these predictions are automatic (either innate or due to high familiarity) so that they cannot be overwritten by a few trials, in which they are violated. In fact, the violation responses observed in these studies show that predictions about an upcoming sound are retained to a certain degree. However, some modulation of prediction does seem to take place, as for instance a mismatch negativity can be observed for matching face—voice pairing preceded by a number of mismatching pairings (de Gelder et al., 1999).

The results of these studies are inconsistent with respect to the influence visual information has on auditory information processing. While some report larger N1 responses for matching compared to non-matching face—voice pairings (Pourtois et al., 2000), others do not find differences in the N1 (Pourtois et al., 2002). Instead, they report later differences between matching and non-matching face—voice pairings, for instance in the P2b (Pourtois et al., 2002).

A different approach to investigate the face—voice interaction has been to present emotional facial expressions either alone or combined with matching vocal information (Paulmann et al., 2009). In this study, the onset of visual and auditory information was synchronized, thereby excluding any visual prediction before the sound onset. In such a setting, first effects of emotional information were observed in the P2, showing larger amplitudes for angry compared to neutral stimuli. While the use of matching stimuli presented in either a uni- or a multisensory way provides a promising design to investigate cross-modal prediction, the lack of any audiovisual delay prevents us from drawing any specific conclusions regarding predictive mechanisms.

Overall, visual emotional information does seem to influence auditory processing at a very early stage. However, studies investigating this influence in a natural setting are largely missing.

In two recent EEG-studies, we investigated the interaction between emotional body and voice information by means of video material in order to overcome some of the limitations of previous studies (Jessen and Kotz, 2011; Jessen et al., 2012). Videos, in which actors expressed different emotional states with or without matching vocal expressions were presented. The emotional states “anger” and “fear” were depicted via body-expressions as well as short vocalizations (e.g., “ah”). Furthermore, we included a non-emotional control condition (“neutral”), in which the actor performed a movement that did not express any specific emotion and uttered the same vocalization with a neutral tone of voice. The delay between visual and auditory onsets was different for each stimulus, as the timing of the original recording of the videos was not manipulated. Hence, the vocalization occurred with a variable delay after the actor had started to move. In both studies, we observed smaller N1 amplitudes for emotional compared to neutral stimuli, as well as for audiovisual compared to unisensory auditory stimuli, irrespective of the emotional content. The amplitude reduction for audiovisual stimuli resembles that observed by Stekelenburg and Vroomen (2007) for non-emotional stimuli, supporting the notion that the observed effect can be attributed to predictive visual information. However, we did not find an interaction with emotional content.

While we did not manipulate predictive validity of the visual information in these studies, we were still interested in whether the amount of available visual information influences auditory processing. We therefore correlated the length of the audiovisual delay for each stimulus with the N1 amplitude in response to that stimulus obtained in the audiovisual condition of the experiment reported in Jessen et al. (2012) (Figure 1).

We found a positive correlation for both emotion conditions, that is, the longer the delay between visual and auditory onset, the *smaller* the amplitude of the subsequent N1. The opposite pattern was observed in the neutral condition; the longer the delay, the *larger* the N1 amplitude.

As outlined above, reduced N1 amplitudes in cross-modal predictive settings have commonly been interpreted as increased (temporal) prediction. If we assume that a longer stretch of visual information allows for a stronger prediction, this increase in prediction can explain the reduction in N1 amplitude observed with increasing visual information for emotional stimuli. However, this pattern does not seem to hold for non-emotional stimuli. When the duration of visual information increases, the amplitude of the N1 also increases. Hence, only in the case of emotional stimuli, an increase in visual information seems to correspond to an increase in visual predictability.

Interestingly, this is the case although neutral stimuli, on average, have a longer audiovisual delay (mean delay for stimuli presented in the audiovisual condition: anger: 1032 ms, fear: 863 ms, neutral: 1629 ms), and thus more visual information is available. Therefore, emotional content rather than pure amount of information seems to drive the observed correlation.

Support for the idea that emotional information may have an influence on cross-modal prediction also comes from priming research. The affective content of a prime strongly influences target effects (Carroll and Young, 2005), leading to differences in activation as evidenced by several EEG studies (e.g., Schirmer et al., 2002; Werheid et al., 2005). Schirmer et al. (2002), for instance, observed smaller N400 amplitudes in response to words that matched a preceding prime in contrast to words that violated the prediction. Also, for facial expressions, a decreased ERP

response in frontal areas within 200 ms has been observed in response to primed as compared to non-primed emotion expressions (Werheid et al., 2005).

However, priming studies strongly differ from real multisensory interactions. Visual and auditory information are presented subsequently rather than simultaneously, and typically, visual and auditory stimuli do not originate from the same event. Priming research therefore only allows for investigating prediction at the content level, at which for instance the perception of an angry face primes the perception of an angry voice. It does not allow investigating temporal prediction as no natural temporal relation between visual and auditory information is present.

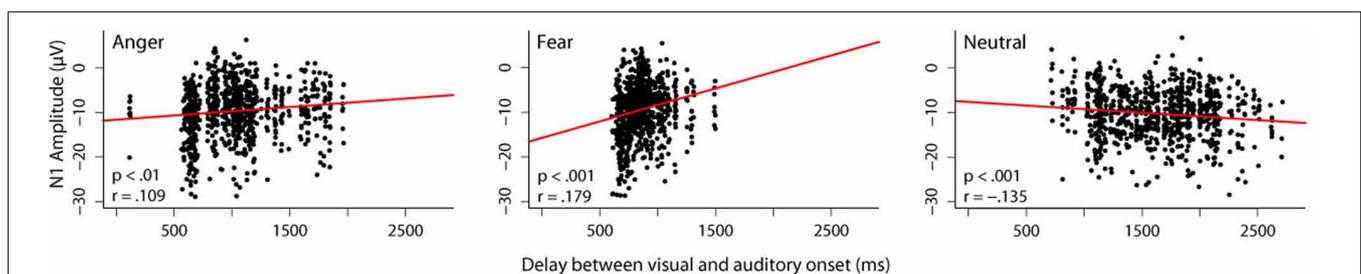
Neither our study referenced above (Jessen et al., 2012) nor the mentioned priming studies were thus designed to explicitly investigate the influence of affective information on cross-modal prediction in naturalistic settings. Hence, the reported data just offer a glimpse into this field. Nevertheless, they highlight the potential role cross-modal prediction may play in the multisensory perception of emotions. We believe that this role may be essential for our understanding of emotion perception, and in the following suggest several approaches suited to illuminate this role.

## FUTURE DIRECTIONS

Different aspects of multisensory emotion perception need to be further investigated in order to understand the role of cross-modal prediction in this context. First, it is essential to establish the influence that emotional content has on cross-modal prediction, especially in contrast to other complex and salient information. Second, it will be necessary to investigate, which aspects of cross-modal prediction are influenced by emotional content. And finally, it is essential to consider how much or how little emotional information is sufficient to influence such predictions. We will take a closer look at all three propositions in the following.

## AFFECTIVE INFLUENCE ON CROSS-MODAL PREDICTION

First, it is necessary to investigate the degree to which affective content influences prediction. The correlation analysis reported above suggests that visual emotions seem to have some influence



**FIGURE 1 | Correlation between audiovisual delay and N1 amplitude.** In one of our studies (Jessen et al., 2012), we presented 24 participants with videos, in which different emotions were expressed by body and vocal expressions simultaneously. The delay between the visual and the auditory onset was different for each stimulus. In order to investigate the influence that a different amount of visual information has on the subsequent auditory processing, we correlated the length of the audiovisual delay with the N1 amplitude separately for each emotion. Trials in which the N1 amplitude differed more than 3 standard deviations from the mean were

excluded from further analysis. Dots represent individual trials. A linear mixed model including the random factor subject and the fixed factors emotion and delay reveals a significant interaction between the fixed factors [ $F_{(1, 2408)} = 33.43, p < 0.0001$ ]. It can be seen that for both emotions, an inverse relation between N1 amplitude and delay exists: the longer the delay, the smaller the N1 amplitude [anger:  $F_{(1, 805)} = 10.98, p < 0.001$ ; fear:  $F_{(1, 773)} = 32.50, p < 0.0001$ ]. The reverse pattern occurs in the neutral condition; here, longer delays correspond to larger N1 amplitudes [ $F_{(1, 784)} = 17.19, p < 0.0001$ ].

on subsequent auditory processing, but further studies are clearly needed.

In order to investigate this aspect, it is crucial to use appropriate stimulus material. Most importantly, such stimulus material has to be dynamic in order to allow for the investigation of temporal as well as content-level predictions. Only dynamic material can cover temporal as well as content predictions and, at the same time, retain the natural temporal relation between visual and auditory onsets. While the use of videos has become increasingly popular in recent years in fMRI studies (e.g., Kreifelts et al., 2007; Pichon et al., 2009; Robins et al., 2009), most EEG (and MEG) studies still rely on static material. One reason for this is probably the very advantage of EEG over fMRI, namely its high temporal resolution. While this allows for close tracking of the time course of information processing, it is also vulnerable to confounds arising from the processing of the preceding visual information. However, this problem can be countered by choosing well-suited control conditions (such as comparably complex and moving non-emotional stimuli). Furthermore, it will be helpful to not exclusively rely on ERP data, but to broaden the analysis to include neural oscillations that can be analyzed in ways less dependent on fixed event onsets (e.g., induced activity, see for instance Tallon-Baudry and Bertrand, 1999). Of particular interest in this context would be the influence emotional visual information has on the phase of oscillatory activity in auditory areas, as well as the relation between low- and high-frequency oscillations. Is, for instance, auditory processing influenced by the phase of the oscillatory activity during visual presentation?

Furthermore, it is necessary to tease apart cross-modal prediction from other forms of multisensory interaction that most likely occur in multisensory emotion perception. Here, it will be essential to manipulate the predictability of the preceding visual information, either at the content level (by for instance using different intensities of emotion expression) or at a temporal level (by providing more or less visual information, see below).

Finally, another important factor may be the role that different types of visual stimuli play, such as facial in comparison to body expressions. Both are visual sources, naturally co-occurring with auditory information, and therefore both can potentially predict auditory information. However, they differ in that facial expressions are more closely linked to vocal utterances. Body expressions, in contrast, may provide more coarse information about emotional states, essential at larger distances. Hence, while facial expressions seem the most obvious candidate, body expressions are not to be forgotten (in fact, the correlation reported above shows brain data in response to body—voice pairings, Jessen et al., 2012).

Insight from these different approaches will allow us to get a general appreciation of how cross-modal prediction influences multisensory emotion perception.

## DIFFERENT PATHWAYS

At a more specific level, one essential question is which aspect of cross-modal prediction can be influenced by emotional content.

One aspect that is highly relevant in this context is the notion of different pathways as outlined by Arnal et al. (2009). For cross-modal emotional prediction, at least three different levels

of prediction become relevant. Predictions may occur at a simple, physical level, comparable to any other stimulus: by the movement of face and body, we can predict when an auditory event onset will occur. This prediction would correspond to the direct pathways posited by Arnal et al. (2009). This direct pathway seems to be involved in cross-modal prediction irrespective of emotional content. Emotions may render temporal predictions possibly even more reliable, as emotional facial expressions are very common, well-rehearsed stimuli and hence may allow for a more precise prediction of the onset in comparison to less frequent stimuli. However, the emotional content itself most likely plays only a minor role in the generation of temporal predictions.

Secondly, predictions may occur at the sound level. Based on the shape of the mouth (and to a certain degree other facial features), predictions can be made regarding the following utterance, be it a word, an interjection, or just a vocalization such as laughter. This type of prediction is specific to complex stimuli, for which the production of a sound can be observed visually, for instance in speech production and actions. When this is not the case, for example, if the button on a radio is pushed, we can predict the sound onset, but not the type of sound we will hear.

For this second type of predictions, emotions are expected to play a more important role, as the content of the vocalization is closely tied to the emotion expressed. Still, they not only predict emotional aspects, but also properties of the upcoming sound that are not mainly related to its emotional quality. Hence predictions specifically related to the affective content are rather a byproduct of general predicting sound features. Nevertheless, quickly determining emotional aspects is essential for fast and efficient emotion processing, and based on this necessity, affective content of the visual signal may lead to a prioritized content processing for sound information.

A third type of prediction is closely related to the prediction of a sound; with respect to cross-modal emotional prediction, we cannot only predict whether an “ah” or an “oh” will occur (as in speech perception), but also whether this “ah” will be uttered in an angry or fearful tone of voice. We can thus predict the emotional content. Both of these latter types of prediction invoke an indirect pathway (Arnal et al., 2009). However, while content prediction can occur in several settings, emotion prediction is specific to human face-to-face interaction.

This last type of predictions, emotion prediction proper, is devoted exclusively to predicting the emotional content of an upcoming signal. Hence, the strongest influence of emotional content is expected to occur at this level.

Nevertheless, in order to better understand cross-modal emotion prediction, it will be necessary to further disentangle the relation between these two types of indirect predictions (i.e., the prediction of speech content such as “ah” and the prediction of emotional content from the tone of voice).

## DURATION OF VISUAL INFORMATION

Another important aspect is the amount of visual information necessary to generate reliable predictions. It has been shown that the delay between the onset of mouth movement and the onset of speech sound typically varies between 100 and 300 ms (Chandrasekaran et al., 2009). Accordingly, most studies using

speech stimuli use an audiovisual delay within that time range (Besle et al., 2004; Stekelenburg and Vroomen, 2007; Arnal et al., 2009). The same holds true for the perception of actions (Stekelenburg and Vroomen, 2007). However, the question arises as to how much delay is actually *necessary* to allow for cross-modal prediction to occur. Stekelenburg and Vroomen (2007), who used speech stimuli with an auditory delay of 160–200 ms as well as action stimuli with an auditory delay of 280–320 ms observed stronger N1 suppression effects for action compared to speech stimuli. They suggested that this difference may be due to the longer stretch of visual information preceding a sound onset. Somewhat shorter optimal delays have been observed using simpler stimulus material and/or more invasive recording. In human EEG, an audiovisual lag of 30 to 75 ms has been found to reliably elicit a phase reset in auditory cortex (Thorne et al., 2011). A similar time window has been found in a study of local field potential in the auditory cortex of macaque monkeys;

the strongest modulation by preceding visual information was observed for a delay between 20 and 80 ms (Kayser et al., 2008).

Hence, providing more visual information may (at least up to some point) allow for a better prediction formation. At the same time, if affective information enhances cross-modal prediction, emotional content may reduce the length of required visual information. Determining the necessary temporal constraints can therefore provide crucial insight onto the effect of emotional information on multisensory information processing.

In summary, we suggest that in order to fully understand multisensory emotion perception, it is essential to take into account the role of cross-modal prediction. It will therefore be necessary to bring together approaches and findings from two flourishing fields that have so far been largely kept separate: cross-modal prediction and emotion perception. Only if we understand the role of prediction, we will be able to fully understand multisensory emotion perception.

## REFERENCES

- Arnal, L. H., Morillon, B., Kell, C. A., and Giraud, A.-L. (2009). Dual neural routing of visual facilitation in speech processing. *J. Neurosci.* 29, 13445–13453. doi: 10.1523/JNEUROSCI.3194-09.2009
- Arnal, L. H., Wyart, V., and Giraud, A. L. (2011). Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nat. Neurosci.* 14, 797–801. doi:10.1038/nn.2810
- Bernstein, L. E., Auer, E. T. Jr., Wagner, M., and Ponton, C. W. (2008). Spatiotemporal dynamics of audiovisual speech processing. *Neuroimage* 39, 423–435. doi: 10.1016/j.neuroimage.2007.08.035
- Besle, J., Bertrand, O., and Giard, M. H. (2009). Electrophysiological (EEG, sEEG, MEG) evidence for multiple audiovisual interactions in the human auditory cortex. *Hear. Res.* 258, 143–151. doi: 10.1016/j.heares.2009.06.016
- Besle, J., Fort, A., Delpuech, C., and Giard, M. (2004). Bimodal speech: early suppressive visual effects in human auditory cortex. *Eur. J. Neurosci.* 20, 2225–2234. doi: 10.1111/j.1460-9568.2004.03670.x
- Buzsaki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304, 1926–1929. doi: 10.1126/science.1099745
- Callan, D. E., Jones, J. A., Munhall, K., Callan, A. M., Kroos, C., and Vatikiotis-Bateson, E. (2003). Neural processes underlying perceptual enhancement by visual speech gestures. *Neuroreport* 14, 2213–2218. doi: 10.1097/00001756-200312020-00016
- Calvert, G. A., Brammer, M. J., Bullmore, E. T., Campbell, R., Iversen, S. D., and David, A. S. (1999). Response amplification in sensory-specific cortices during crossmodal binding. *Neuroreport* 10, 2619–2623. doi: 10.1097/00001756-199908200-00033
- Calvert, G. A., Brammer, M. J., and Iversen, S. D. (1998). Crossmodal identification. *Trends Cogn. Sci.* 2, 247–253. doi: 10.1016/S1364-6613(98)01189-9
- Calvert, G. A., Campbell, R., and Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr. Biol.* 10, 649–657. doi: 10.1016/S0960-9822(00)00513-3
- Carroll, N. C., and Young, A. W. (2005). Priming of emotion recognition. *Q. J. Exp. Psychol.* A 58, 1173–1197. doi: 10.1080/02724980443000539
- Chandrasekaran, C., Trubanova, A., Stillitano, S., Caplier, A., and Ghazanfar, A. A. (2009). The natural statistics of audiovisual speech. *PLoS Comput. Biol.* 5:e1000436. doi: 10.1371/journal.pcbi.1000436
- Chen, Y. H., Edgar, J. C., Holroyd, T., Dammers, J., Thonnessen, H., Roberts, T. P., et al. (2010). Neuromagnetic oscillations to emotional faces and prosody. *Eur. J. Neurosci.* 31, 1818–1827. doi: 10.1111/j.1460-9568.2010.07203.x
- Davis, C., Kislyuk, D., Kim, J., and Sams, M. (2008). The effect of viewing speech on auditory speech processing is different in the left and right hemispheres. *Brain Res.* 1242, 151–161. doi: 10.1016/j.brainres.2008.04.077
- de Gelder, B., Böcker, K. B. E., Tuomainen, J., Hensen, M., and Vroomen, J. (1999). The combined perception of emotion from voice and face: early interaction revealed by human electric brain responses. *Neurosci. Lett.* 260, 133–136. doi: 10.1016/S0304-3940(98)00963-X
- Doesburg, S. M., Emberson, L. L., Rahi, A., Cameron, D., and Ward, L. M. (2008). Asynchrony from synchrony: long-range gamma-band neural synchrony accompanies perception of audiovisual speech asynchrony. *Exp. Brain Res.* 185, 11–20. doi: 10.1007/s00221-007-1127-5
- Ghazanfar, A. A., and Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends Cogn. Sci.* 10, 278–285. doi: 10.1016/j.tics.2006.04.008
- Giard, M. H., and Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J. Cogn. Neurosci.* 11, 473–490. doi: 10.1162/089892999563544
- Jessen, S., and Kotz, S. A. (2011). The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. *Neuroimage* 58, 665–674. doi: 10.1016/j.neuroimage.2011.06.035
- Jessen, S., Obleser, J., and Kotz, S. A. (2012). How bodies and voices interact in early emotion perception. *PLoS ONE* 7:e36070. doi: 10.1371/journal.pone.0036070
- Kayser, C., Petkov, C. I., and Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cereb. Cortex* 18, 1560–1574. doi: 10.1093/cercor/bhm187
- Klasen, M., Chen, Y. H., and Mathiak, K. (2012). Multisensory emotions: perception, combination and underlying neural processes. *Rev. Neurosci.* 23, 381–392. doi: 10.1515/revneuro-2012-0040
- Klucharev, V., Mottonen, R., and Sams, M. (2003). Electrophysiological indicators of phonetic and non-phonetic multisensory interactions during audiovisual speech perception. *Brain Res. Cogn. Brain Res.* 18, 65–75. doi: 10.1016/j.cogbrainres.2003.09.004
- Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., and Wildgruber, D. (2007). Audiovisual integration of emotional signals in voice and face: an event-related fMRI study. *Neuroimage* 37, 1445–1456. doi: 10.1016/j.neuroimage.2007.06.020
- Lakatos, P., O'Connell, M. N., Barczak, A., Mills, A., Javitt, D. C., and Schroeder, C. E. (2009). The leading sense: supramodal control of neurophysiological context by attention. *Neuron* 64, 419–430. doi: 10.1016/j.neuron.2009.10.014
- Laurienti, P. J., Wallace, M. T., Maldjian, J. A., Susi, C. M., Stein, B., and Burdette, J. H. (2003). Cross-modal sensory processing in the anterior cingulate and medial prefrontal cortices. *Hum. Brain Mapp.* 19, 213–223. doi: 10.1002/hbm.10112
- Meeren, H. K. M., van Heijnsbergen, C. C. R. J., and de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proc. Natl. Acad. Sci. U.S.A.* 102, 16518–16523. doi: 10.1073/pnas.0507650102
- Paulmann, S., Jessen, S., and Kotz, S. A. (2009). Investigating the multimodal nature of human communication. Insights from ERPs. *J. Psychophysiol.* 23, 63–76. doi: 10.1027/0269-8803.23.2.63

- Pichon, S., de Gelder, B., and Grèzes, J. (2009). Two different faces of threat. Comparing the neural systems for recognizing fear and anger in dynamic body expressions. *Neuroimage* 47, 1873–1883. doi: 10.1016/j.neuroimage.2009.03.084
- Pourtois, G., Debatisse, D., Despland, P.-A., and de Gelder, B. (2002). Facial expressions modulate the time course of long latency auditory brain potentials. *Brain Res. Cogn. Brain Res.* 14, 99–105. doi: 10.1016/S0926-6410(02)00064-2
- Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B., and Crommelinck, M. (2000). The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport* 11, 1329–1333. doi: 10.1097/00001756-200004270-00036
- Robins, D. L., Hunyadi, E., and Schultz, R. T. (2009). Superior temporal activation in response to dynamic audio-visual emotional cues. *Brain Cogn.* 69, 269–278. doi: 10.1016/j.bandc.2008.08.007
- Schirmer, A., Kotz, S. A., and Friederici, A. D. (2002). Sex differentiates the role of emotional prosody during word processing. *Cogn. Brain Res.* 14, 228–233. doi: 10.1016/S0926-6410(02)00108-8
- Schroeder, C. E., Lakatos, P., Kajikawa, Y., Partan, S., and Puce, A. (2008). Neuronal oscillations and visual amplification of speech. *Trends Cogn. Sci.* 12, 106–113. doi: 10.1016/j.tics.2008.01.002
- Senkowski, D., Schneider, T. R., Foxe, J. J., and Engel, A. K. (2008). Crossmodal binding through neural coherence: implications for multisensory processing. *Trends Neurosci.* 31, 401–409. doi: 10.1016/j.tins.2008.05.002
- Sperdin, H. F., Cappe, C., Foxe, J. J., and Murray, M. M. (2009). Early, low-level auditory-somatosensory multisensory interactions impact reaction time speed. *Front. Integr. Neurosci.* 3:2. doi: 10.3389/neuro.07.002.2009
- Stekelenburg, J. J., and de Gelder, B. (2004). The neural correlates of perceiving human bodies: an ERP study on the body-inversion effect. *Neuroreport* 15, 777–780. doi: 10.1097/00001756-200404090-00007
- Stekelenburg, J. J., and Vroomen, J. (2007). Neural correlates of multi-sensory integration of ecologically valid audiovisual events. *J. Cogn. Neurosci.* 19, 1964–1973. doi: 10.1162/jocn.2007.19.12.1964
- Stekelenburg, J. J., and Vroomen, J. (2012). Electrophysiological correlates of predictive coding of auditory location in the perception of natural audiovisual events. *Front. Integr. Neurosci.* 6:26. doi: 10.3389/fnint.2012.00026
- Summerfield, C., and Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends Cogn. Sci.* 13, 403–409. doi: 10.1016/j.tics.2009.06.003
- Tallon-Baudry, C., and Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn. Sci.* 3, 151–162. doi: 10.1016/S1364-6613(99)01299-1
- Thorne, J. D., De Vos, M., Viola, F. C., and Debener, S. (2011). Cross-modal phase reset predicts auditory task performance in humans. *J. Neurosci.* 31, 3853–3861. doi: 10.1523/JNEUROSCI.6176-10.2011
- van Wassenhove, V., Grant, K. W., and Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proc. Natl. Acad. Sci. U.S.A.* 102, 1181–1186. doi: 10.1073/pnas.0408949102
- Werheid, K., Alpay, G., Jentsch, I., and Sommer, W. (2005). Priming emotional facial expressions as evidenced by event-related brain potentials. *Int. J. Psychophysiol.* 55, 209–219. doi: 10.1016/j.ijpsycho.2004.07.006

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 04 April 2013; accepted: 25 June 2013; published online: 18 July 2013.

Citation: Jessen S and Kotz SA (2013) On the role of crossmodal prediction in audiovisual emotion perception. *Front. Hum. Neurosci.* 7:369. doi: 10.3389/fnhum.2013.00369

Copyright © 2013 Jessen and Kotz. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



# Facial reactions in response to dynamic emotional stimuli in different modalities in patients suffering from schizophrenia: a behavioral and EMG study

**Mariateresa Sestito<sup>1\*</sup>, Maria Alessandra Umiltà<sup>1</sup>, Giancarlo De Paola<sup>2</sup>, Renata Fortunati<sup>2</sup>, Andrea Raballo<sup>3</sup>, Emanuela Leuci<sup>2</sup>, Simone Maffei<sup>2</sup>, Matteo Tonna<sup>2</sup>, Mario Amore<sup>4</sup>, Carlo Maggini<sup>2</sup> and Vittorio Gallese<sup>1\*</sup>**

<sup>1</sup> Unit of Physiology, Department of Neuroscience, University of Parma, Parma, Italy

<sup>2</sup> Psychiatric Division, Department of Neuroscience, University of Parma, Parma, Italy

<sup>3</sup> Department of Mental Health, AUSL of Reggio Emilia, Reggio Emilia, Italy

<sup>4</sup> Psychiatric Division, Department of Neuroscience, University of Genova, Genova, Italy

## Edited by:

Martin Klasen, RWTH Aachen University, Germany

## Reviewed by:

Giuliana Lucci, IRCCS Santa Lucia of Rome, Italy

Lindsay M. Oberman, University of California, San Diego, USA

## \*Correspondence:

Mariateresa Sestito, Unit of Physiology, Department of Neuroscience, University of Parma, Via Volturno 39, I-43100, Parma, Italy  
e-mail: mariateresa.sestito@nemo.unipr.it;

Vittorio Gallese, Section of Physiology, Department of Neuroscience, University of Parma, Via Volturno 39, I-43100, Parma, Italy  
e-mail: vittorio.gallese@unipr.it

Emotional facial expression is an important low-level mechanism contributing to the experience of empathy, thereby lying at the core of social interaction. Schizophrenia is associated with pervasive social cognitive impairments, including emotional processing of facial expressions. In this study we test a novel paradigm in order to investigate the evaluation of the emotional content of perceived emotions presented through dynamic expressive stimuli, facial mimicry evoked by the same stimuli, and their functional relation. Fifteen healthy controls and 15 patients diagnosed with schizophrenia were presented with stimuli portraying positive (laugh), negative (cry) and neutral (control) emotional stimuli in visual, auditory modalities in isolation, and congruently or incongruently associated. Participants were requested to recognize and quantitatively rate the emotional value of the perceived stimuli, while electromyographic activity of Corrugator and Zygomaticus muscles was recorded. All participants correctly judged the perceived emotional stimuli and prioritized the visual over the auditory modality in identifying the emotion when they were incongruently associated (Audio-Visual Incongruent condition). The neutral emotional stimuli did not evoke any muscle responses and were judged by all participants as emotionally neutral. Control group responded with rapid and congruent mimicry to emotional stimuli, and in Incongruent condition muscle responses were driven by what participants saw rather than by what they heard. Patient group showed a similar pattern only with respect to negative stimuli, whereas showed a lack of or a non-specific Zygomaticus response when positive stimuli were presented. Finally, we found that only patients with reduced facial mimicry (Internalizers) judged both positive and negative emotions as significantly more neutral than controls. The relevance of these findings for studying emotional deficits in schizophrenia is discussed.

**Keywords: EMG, emotions, empathy, facial mimicry, schizophrenia, simulation**

## INTRODUCTION

Emotional expressions are widely acknowledged as essential in communicating internal feelings and intentions (Ekman and Oster, 1979). The ability to communicate and understand the emotional states of others and their intentions is a fundamental social skill. Indeed, facial expressions are among the most common and significant emotion stimuli.

To this end, it is well-known that humans react to emotional facial expressions with specific, congruent facial muscle mimicry, which can be reliably measured by electromyography (EMG; e.g., Dimberg, 1982, 1988). For example, pictures of sad facial expressions evoke increased muscle *Corrugator Supercilii* activity, while pictures of happy facial expressions increase muscle *Zygomaticus Major* activity and decrease muscle *Corrugator Supercilii* activity (Lundqvist and Dimberg, 1995; Han et al., 2012). These facial

muscular reactions appear to be spontaneous and automatic (Dimberg and Thunberg, 1998; Dimberg et al., 2000, 2002; Larsen et al., 2003). Many studies demonstrated that facial mimicry contributes to recognition of specific facial expressions (for a review, see Goldman and Sripada, 2005; Niedenthal et al., 2010). Indeed, blocking facial mimicry impairs recognition of facial expression of emotions (Oberman et al., 2007). Furthermore, it has been proposed that mimicry reflects internal embodied simulation of the perceived facial expression in order to facilitate understanding of its emotional meaning (Gallese, 2003, 2005, 2006; Niedenthal, 2007; Halberstadt et al., 2009; Niedenthal et al., 2009) and promoting empathy by means one's facial feedback system (for a review of the facial feedback hypothesis, see Adelman and Zajonc, 1989). A recent EMG study (Dimberg et al., 2011) showed that high empathic people, with respect to low empathic

group, are particularly sensitive in reacting with facial reactions when they look to emotional facial expressions. Moreover, high empathic people rated perceived facial emotional expressions as more intense with respect to low empathic ones.

Historically, affective features of schizophrenia were considered an integral part of the disorder. Bleuler (1950) considered affective disturbance to be a fundamental symptom of schizophrenia, whereas hallucinations and delusions were regarded as accessory symptoms. Studies on patients' facial mimicry in response to emotional stimuli showed that they activate the same muscle of control subjects, but such activation was found to be weaker in patients than in healthy controls (Earnst et al., 1996; Kring and Earnst, 1999). Another study showed, on the other hand, that in contrast to healthy controls, patients diagnosed with schizophrenia demonstrated atypical facial mimicry, which was not associated with any clinical feature of the disorder. The authors of this study suggested that this evidence might account for a low-level disruption contributing to empathy deficits in schizophrenia (Varcin et al., 2010). Similarly, Wolf et al. (2006) found an undecipherable and bizarre mimic pattern within a sample of patients suffering from schizophrenia, called "mimic disintegration" (see Heimann and Spoerri, 1957). Mimic disintegration is defined as the inability to organize specific facial muscle movements as an integrated whole, thus making difficult for observers to decode the emotional state and establish contact or develop a deeper relationship with the patients. Furthermore, many studies investigating everyday life of patients diagnosed with schizophrenia documented an emotional-affective pattern characterized by many negative and few positive experiences, thus making patients' affectivity more negative. Some studies (Mattes et al., 1995; Iwase et al., 1999; Wolf et al., 2004, 2006) found a minor activity of *Zygomaticus* muscle in response to positive stimuli, whereas another study (Sison et al., 1996) found an overall major activation of *Corrugator Supercilii* muscle, interpreted as a sign of the negative attitude showed by patients in everyday life.

Reduced emotional expression (i.e., flat affect) is not only a typical symptom of full-blown schizophrenia (Andreasen, 1984a; Bleuler, 1950). Many findings lend support to the assumption that vulnerability to schizophrenia may be subtly manifested in emotional behavior long before the onset of clinical symptoms. Furthermore, after schizophrenia onset, flat affect increases (Walker et al., 1993). Reduced emotional facial expression could be a disease risk index for high-vulnerability subjects (e.g., Schizotypal Personality patients and first degree relatives) (Phillips and Seidman, 2008). Moreover, previous research on flat affect showed a disjunction between the expression and the experience of emotion in schizophrenia (Bleuler, 1950; Berenbaum and Oltmanns, 1992; Kring et al., 1993; Kring and Neale, 1996; Aghevli et al., 2003; Kring and Earnst, 2003). These studies showed that patients with schizophrenia often reported experiencing strong emotions, but they were significantly less expressive than controls. Thus, observers could note no visible sign of emotion.

The studies using EMG recording to investigate emotional expression in schizophrenia, used different materials and methods. In particular, often non-ecological stimuli, like static images

non-facial stimuli, or fiction movies were used. Many studies indeed highlighted, on the other hand, the importance of dynamic stimuli in the evaluation of emotional expression. A recent study on healthy individuals showed that presentation of dynamic facial expressions evokes stronger EMG responses than static ones. Moreover, participants rated dynamic expressions as more intense than static ones (Rymarczyk et al., 2011).

Emotional facial expression communicates feelings, but is also an important low-level mechanism contributing to the experience of empathy, thereby lying at the core of social interaction. Schizophrenia is associated with pervasive social cognitive impairments that include emotional processing of facial expressions. Despite such disorder might play a crucial role in empathizing deficits and consequently impoverished social skills, previous research on facial expression of emotions in schizophrenia has not yielded unequivocal results. In particular, it remains unaddressed the issue of patients' facial expression as a medium of empathic resonance contributing to the recognition and evaluation of the perceived emotion expressed by others.

The aim of this study was to investigate whether subjective facial mimicry affects the quantitative evaluation of the emotional content of perceived emotions presented through dynamic expressive stimuli, in healthy participants and in patients diagnosed with schizophrenia. To this purpose we employed a novel paradigm by means of which emotional dynamic ecological stimuli were presented in the visual and auditory modalities in isolation and congruently or incongruently associated. This approach enabled us to study the dimensional quality and possible alteration of the emotional responses in these two experimental groups.

## MATERIALS AND METHODS

### PARTICIPANTS

Thirty participants took part to the experiment. Control participants (CNT; ten males, five females, mean age 35.8 years  $SE \pm 2.3$ ) were recruited by public announcement and were blind to the experimental goals. None of them reported the presence of any neurological or psychiatric disorder. Patient group (SZP; ten males, five females, mean age 32.8 years  $SE \pm 1.7$ ) were recruited from the Clinical Psychiatry Institute of the University of Parma. All of them were chronic clinically stable outpatients, mainly diagnosed with schizophrenia, paranoid subtype. Only one patient was diagnosed with a disorganized subtype, one with an undifferentiated subtype and two patients with a residual subtype. Psychiatric diagnosis was established via a structured interview (Structured Clinical Interview for DSM-IV, SCID). Exclusion criteria were the presence of neurological and vascular disorders, dysmetabolic syndrome, alcohol or drugs abuse and mental retardation (Intelligence Quotient score  $< 70$ ). All participants had normal or corrected to normal vision. In addition to being closely matched for gender, the two groups did not differ in age [ $t_{(30)} = -1.06, p > 0.05$ ]. All clinical participants (SZP) were receiving antipsychotic medication (most of them were administered new generation atypical antipsychotics). Since the age of onset and the illness duration indicated that the clinical sample was heterogeneous, for comparing dosages of different drugs we converted doses of medication to chlorpromazine equivalents.

Then we multiplied these equivalents by the time an individual had been on a given dose to obtain cumulative value measured in dose-years. After each dose had been converted to dose-years, the results could be summed to provide a cumulative measure of lifetime exposure (Andreasen et al., 2010).

In order to describe psychopathological features related with schizophrenia, patients were administered a variety of tests: scale for the Assessment of Negative Symptoms (SANS; Andreasen, 1984a), Scale for the Assessment of Positive Symptoms (SAPS; Andreasen, 1984b), Social Anhedonia Scale (SAS; Chapman et al., 1976), Physical Anhedonia Scale (PAS; Chapman et al., 1976). Given that all patients were under medication, we also administered them the Simpson-Angus Extrapyramidal side-effects Scale (Simpson and Angus, 1970), an established, valid and reliable instrument for assessing neuroleptic-induced parkinsonism (Janno et al., 2005). None of them were beyond cut-off value, indicating that SZP participants did not show any significant extrapyramidal side-effect related with drugs assumption. Details about CNT and SZP samples are provided in **Table 1**. Written informed consent was obtained from all participants before entering the study. The local Ethical Committee approved the study.

## STIMULI

Two professional actors (one male and one female) were used for stimuli preparation. Stimuli consisted of 2-s colored video clips showing positive (laugh), negative (cry) and neutral (control) emotions. The neutral video clips showed actors making various faces (i.e., “making a face”) that did not imply any particular emotional content, just that the actors were adopting some specific facial expressions. Actors when performed neutral stimuli always associated the making a face with specific vocalizations. The sound of the neutral stimuli was a vocalization similar to “ahh,” “ohh,” or “eemmh.” Actors’ full face was presented against a gray background. Stimuli consisted of actors’ Laugh (Positive), Cry (Negative) and Control (Neutral) accompanied

by the simultaneously produced sound of laughter, crying and a non-emotional sound, respectively. Half of the stimuli was performed by the male actor, whereas the other half was performed by the female actress. Stimuli were recorded using a digital camera (25 frames/s, 720×576 pixels), with audio digitally recorded at 44.1 kHz. Stimuli were divided into four presentation modalities: Visual only, Audio only, Audio-Visual congruent and Audio-Visual incongruent. Every presentation modality was made of 60 stimuli [24 Laugh (Positive) stimuli, 24 Cry (Negative) stimuli and 12 Control (Neutral) stimuli]. In the Audio modality (A), the sound of the video clips of laugh, cry and control stimuli was extracted from the original video clips and presented alone. In the Video modality (V), only the visual component of video-clips was presented, devoid of any sound. In the Audio-Visual Congruent modality (AVC), the original video clips were presented with both modalities. In the Incongruent Audio-Visual modality (AVI), the video of a given expression was coupled and presented with the audio pertaining to a different video clip performed by the same actor (e.g., audio of laugh with the video of cry, audio of cry with the video of laughs and audio of a given neutral sound with the video of another neutral stimulus). Consequently, in AVI Laugh participants saw an actor crying but heard laughing, in AVI Cry participants saw an actor laughing but heard crying, and in Control condition they saw an actor making an unemotional face while hearing the sound of a different neutral stimulus.

## EXPERIMENTAL PROCEDURE

Participants were individually tested in a sound attenuated laboratory room. They were invited to sit on a comfortable chair in front of a 19-inch computer monitor used for stimuli presentation, located at a distance of 70 cm. Audio tracks were presented at a comfortable sound level (<70 dB) through loudspeakers integrated in the computer monitor. Before starting, participants were invited to relax and refrain from moving during

**Table 1 | Demographic variables and characteristics of Schizophrenia (SZP) and Control (CNT) participants.**

Characteristic	Patients				Controls		
	Mean	SE	Range	Cut-off	Mean	SE	Range
Age (years)	32.80	1.69	25–49		35.80	2.28	25–53
SAPS	26.67	4.17	0–170				
SANS	48.09	4.56	0–125				
PAS	26.36	2.26	0–61	>18			
SAS	19.00	1.74	0–40	>12			
Simpson-Angus Scale	0.36	0.04	0–4	>0.65			
Duration of illness (years)	11.23	1.30	2–24				
Age at first psychosis	22.69	0.66	19–28				
Number of hospitalizations	3.83	0.38	2–7				
Dose of typical and atypical antipsychotics	32.85	4.93					
Dose of atypical antipsychotics	24.84	4.00					
Dose of typical antipsychotics	8.01	1.48					

*Drugs are expressed as the cumulative value measured in dose-years in the form of (chlorpromazine equivalent in mg) × (time on dose measured in years) (Andreasen et al., 2010).*

the experiment. Participants were instructed to carefully listen to and/or watch audiovisual stimuli. After exposure to each stimulus, participants were required to verbally rate how much positive or negative the stimulus was perceived on a Likert scale ranging from  $-3$  (very negative) to  $+3$  (very positive), where 0 indicated lack of perceived emotional content.

The experiment consisted of four experimental blocks of 60 stimuli each presented in randomized order. Each block consisted of one of the four modalities: Audio-Visual Congruent (AVC), Audio-Visual Incongruent (AVI), Audio (A) and Video (V). In every modality three emotional stimuli were presented in randomized sequence: Laugh (Positive), Cry (Negative) and Control (Neutral). A pause was provided at the end of each condition. The order of blocks was counterbalanced among participants. Each trial (**Figure 1**) started with a fixation cross (the “+” symbol) presented for 1000 ms (baseline), immediately followed by the stimulus, which lasted 2000 ms, then followed by a question mark (the “?” symbol). After question mark presentation, participants verbally scored the emotional valence of each stimuli. The experimenter took note of participants’ response in a record sheet and then started manually the next trial. The total duration of the experiment was about 40 min.

#### EMG RECORDING

To measure facial muscle activity, Ag/AgCl surface electrodes (diameter 3 mm) were attached bipolarly over the left (Dimberg and Petterson, 2000) *Zygomaticus major* and the *Corrugator Supercilii* muscle regions (Fridlund and Cacioppo, 1986). In order to reduce the inter-electrode impedance, the participants’ skin was cleaned with alcohol and rubbed with the electrode paste. Continuous electromyography (EMG) recordings from both muscles were simultaneously acquired with a CED Micro 1401 analog-to-digital converting unit (Cambridge Electronic Design, Cambridge, UK). The EMG signal was amplified ( $3000\times$ ), digitized (sampling rate: 2.5 kHz) and stored on a computer for offline analysis.

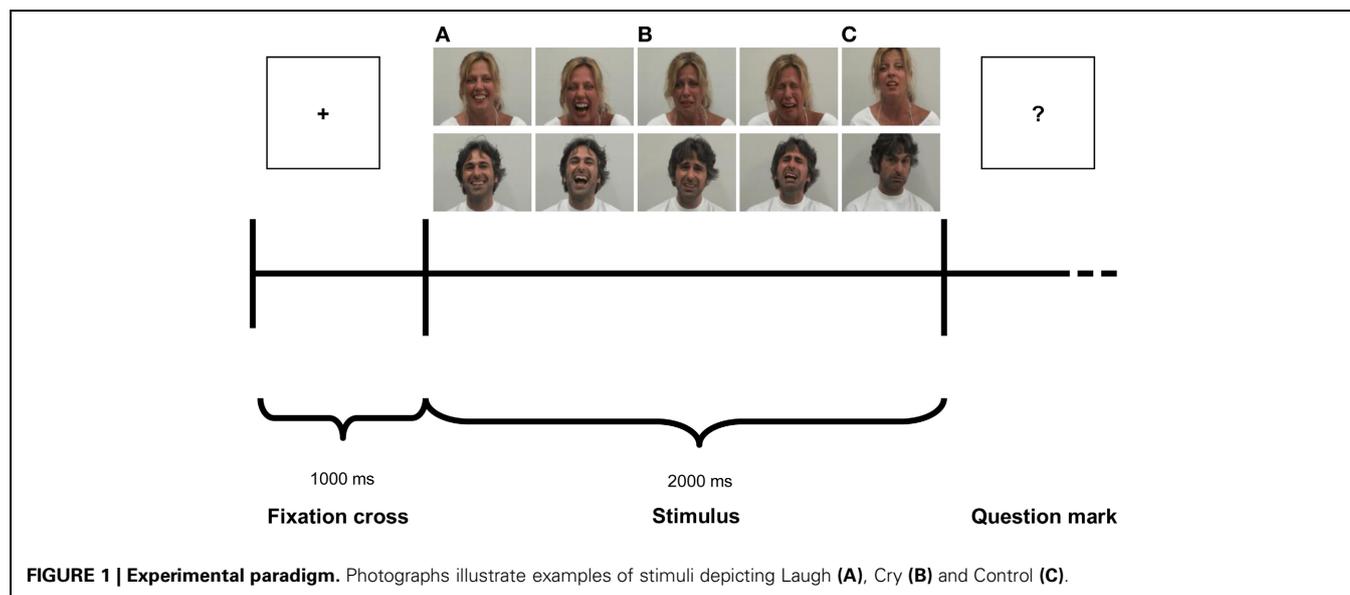
#### DATA AND STATISTICAL ANALYSIS

##### Behavioral rating

The rating score of each participant was averaged on the basis of modality and emotion. The corresponding averaged rating scores were entered into a 4 (Modality: AVC, AVI, A, V)  $\times$  3 (Emotion: Laugh, Cry, Control)  $\times$  2 (Group: SZP and CNT) repeated measures ANOVA, with Modality and Emotion as within-participants factors and Group as between-participants factor.

##### EMG data analysis

Offline, data were submitted to a 50–500 Hz band-pass filter to reduce movement related artifacts and environmental noise, and full-wave rectified. Data were then visually inspected, and data with remaining artifacts were excluded from subsequent analysis [mean percentage of discarded trails: 14.1% for CNT, 12.6% for SZP; *T*-test performed did not show significant differences between groups  $t_{(30)} = 0.5$ ,  $p > 0.6$ ]. In accordance with earlier experiments (e.g., Dimberg et al., 2000), any distinct muscle response to the stimuli was expected to be detectable after 500 ms of exposure. Thus, for each participant and trial, the averaged EMG responses of the two muscles were subdivided in 4 time periods (T1–T4) of 500 ms each. Each time-bin was then normalized with respect to the baseline (i.e., averaged pre-stimulus signal activity lasting 500 ms: from 250 to 750 ms of the 1000 ms total duration of the baseline). Thus, an EMG normalized value above the 100% means an activation of a given muscle with respect to the baseline, whereas an EMG normalized value below the 100% indicate a relaxation of that muscle with respect to the baseline. In order to compare baselines, we performed two ANOVAs, one for each muscle, in which baselines raw data were compared, with Modality (AVC, AVI, A, V) as within-participants factor and Group (SZP, CNT) as between-participants factor. Mean EMG responses were then calculated for each Modality (AVC, AVI, A, V), Emotion (Laugh, Cry, Control) and Period (T1, T2, T3, T4). EMG data were entered into a 4 (Modality: AVC, AVI, A, V)  $\times$  3 (Emotion: Laugh, Cry, Control)  $\times$  4 (Period:



T1: 0–500 ms, T2: 500–1000 ms, T3: 1000–1500 ms, T4: 1500–2000 ms) repeated measures ANOVA, with Modality, Emotion and Period as the within-participants factors and Group (SZP and CNT) as between-participants factor. One separated ANOVA was conducted for each muscle (Corrugator and Zygomaticus).

### Functional relation between EMG and behavioral rating

In order to investigate functional relations between the recorded EMG responses and behavioral rating, we calculated median EMG responses, separately for each group and for each emotion (positive, negative), irrespective of modalities and periods. We excluded from this analysis Control stimuli because they did not evoke any significant EMG response in both muscles (see Results). Regarding positive emotions, we considered for this analysis the following modalities in which we measured (see Results) Zygomaticus muscle activation: AVI Cry, AVC Laugh, A Laugh and V Laugh. Regarding negative emotions, we considered the following modalities in which we measured (see Results) Corrugator muscle activation: AVI Laugh, AVC Cry, A Cry and V Cry.

For each participant, we calculated the median EMG response for each emotion (positive, negative). If this value was equal or greater than the median value calculated separately for positive and negative emotions for the group the participant belonged to, we classified this participant as Externalizer. If, instead, this value was smaller than the median value calculated separately for positive and negative emotions for the group the participant belonged to, we classified this participant as Internalizer (see Kring and Gordon, 1998). Following this procedure, in the CNT group we obtained the median value of 95.14% (8 Externalizers and 7 Internalizers) for positive emotions and the median value of 99.24% (8 Externalizers and 7 Internalizers) for negative emotions. In the SZP group we obtained the median value of 95.15% (6 Externalizers and 9 Internalizers) for positive emotions and the median value of 100% (6 Externalizers and 9 Internalizers) for negative emotions. The corresponding averaged rating scores were entered into a 4 (Modality: AVC, AVI, A, V)  $\times$  2 (Group: SZP and CNT) repeated-measures ANOVAs, with Modality as within-participants factor and Group as between-participants factor. Overall, we ran totally 4 ANOVAs, two in order to analyze behavioral data of the Externalizer cohort (one for each emotion valence: positive, negative) and two in order to analyze behavioral data of the Internalizer cohort (one for each emotion valence: positive, negative).

For all performed analyses, the significance level was set at  $p < 0.05$ . *Post-hoc* comparisons (LSD Fisher test) were applied on all significant main factors and interactions.

## RESULTS

### BEHAVIORAL RESULTS

Results of the repeated-measures ANOVA performed on behavioral rating scores showed that the factor Emotion was significant [ $F_{(2, 56)} = 222.52$   $p < 0.000$ ]. *Post-hoc* comparisons showed that Cry was rated by both groups more negative than Laugh and Control stimuli were considered without any emotional content (all  $p_s < 0.000$ ). Complementing this finding, the interaction between Emotion and Modality [ $F_{(6, 168)} = 156.7$   $p < 0.000$ ] was

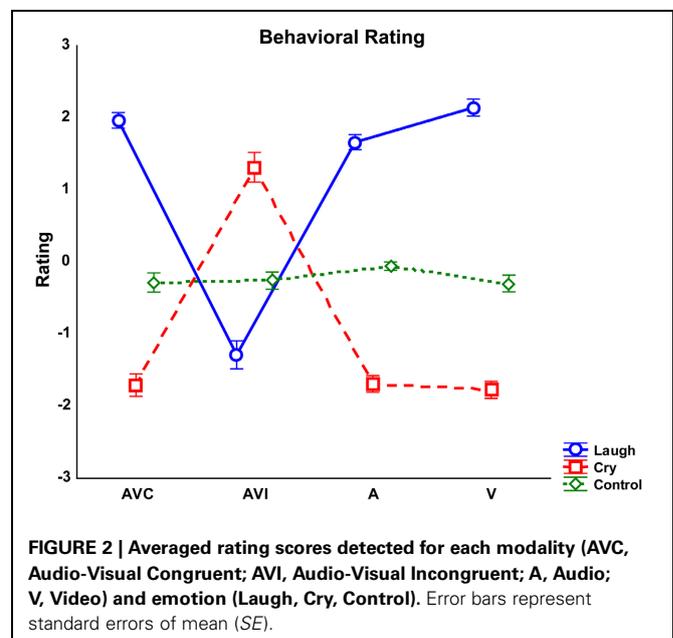
also significant (Figure 2). This interaction was due to the fact that the negative rating scores reported for Laugh stimuli during the AVI modality (in which participants saw cry and heard laugh) differed from the positive rating scores reported for Laugh stimuli in all other modalities (all  $p_s < 0.000$ ). Similarly, the positive rating scores reported for Cry stimuli during the AVI modality (in which participants saw laugh and heard cry) significantly differed from the negative rating scores reported for Cry stimuli in all other modalities (all  $p_s < 0.000$ ). These differences in rating of AVI modality were due to the fact that both groups based their ratings on the emotion they saw (Laugh in AVI Cry and Cry in AVI Laugh), instead of the emotion they heard. *Post-hoc* analysis also showed that Control stimuli were rated as devoid of emotional content in all modalities by both groups (all  $p_s < 0.000$ ). Furthermore, with Laugh stimuli, V modality was rated more positive than A modality ( $p < 0.05$ ).

Results also showed a significant interaction Emotion by Group [ $F_{(2, 56)} = 3.43$   $p < 0.05$ ]. However, *post-hoc* analyses revealed no significant differences between groups (all  $p_s > 0.3$ ).

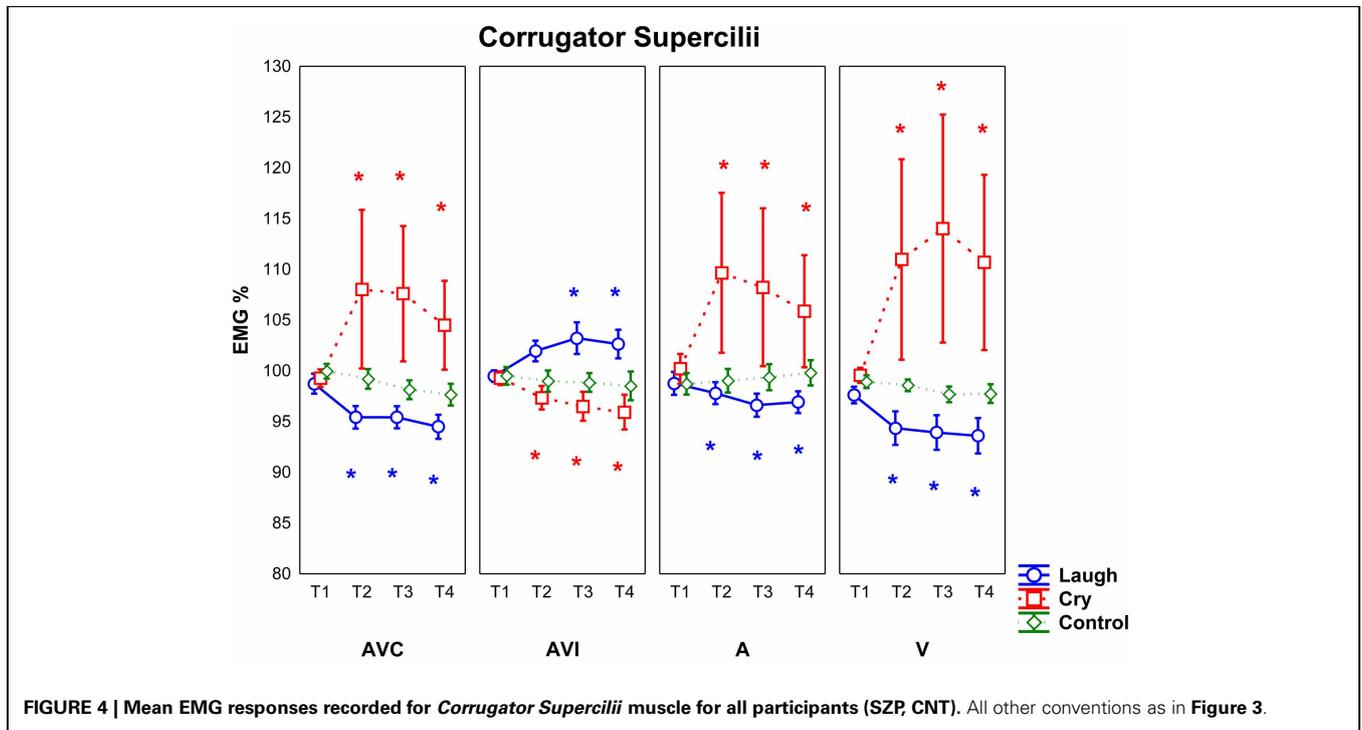
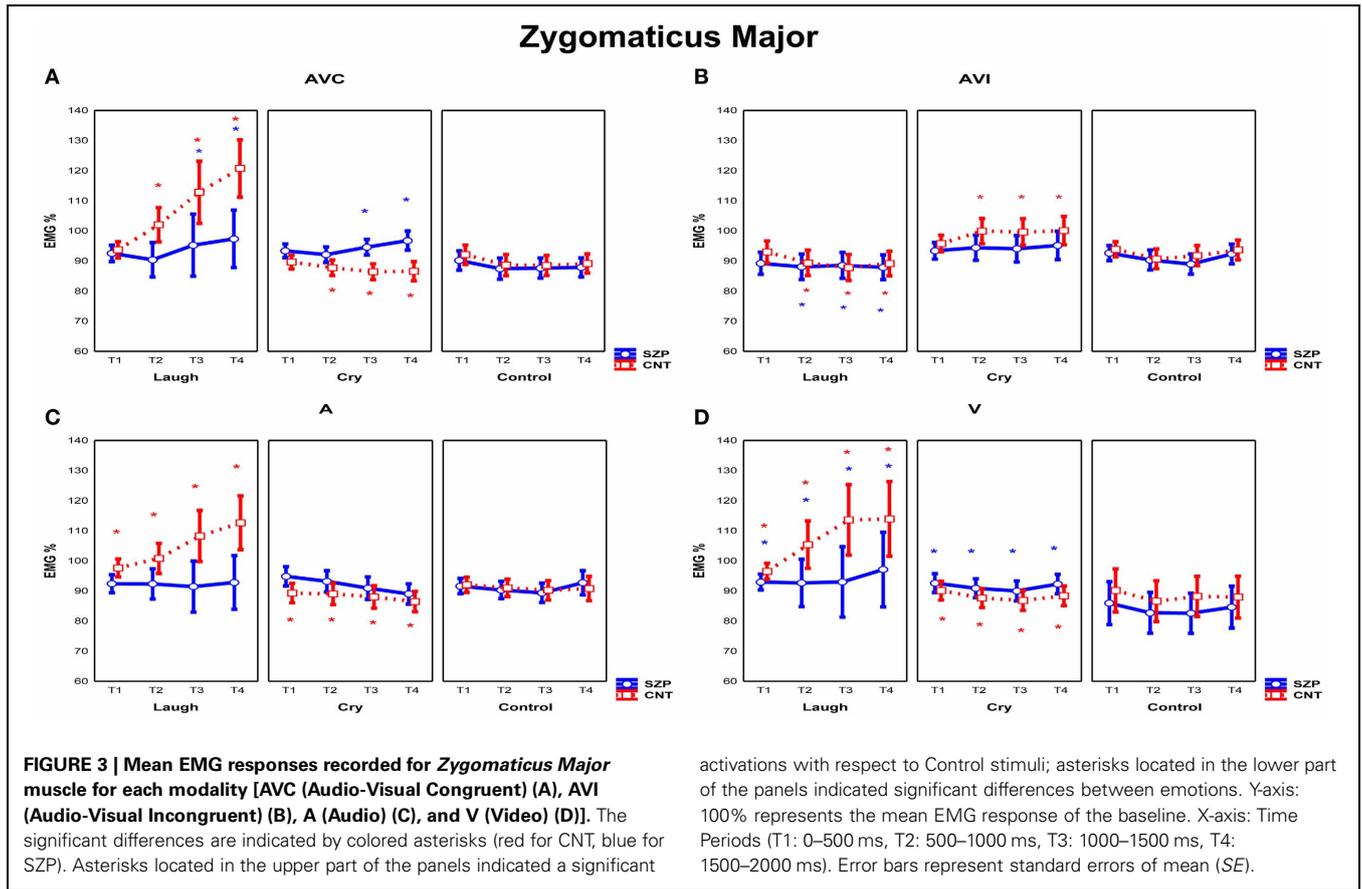
### EMG RESULTS

Two repeated measure ANOVAs, one for each muscle (Corrugator, Zygomaticus), were performed in order to compare baselines between the two groups, with Modality (AVC, AVI, A, V) as within-participants factor and Group (SZP, CNT) as between-participants factor. We found no significant main effect and interactions (all  $p_s > 0.05$ ). These results show that the baselines were not significantly different between the two groups.

Two repeated measures ANOVAs were performed in order to assess Zygomatic Major and Corrugator Supercilii EMG responses during the presentation of the stimuli of positive, negative and neutral facial expressions and/or related sounds in four different modalities (AVC, AVI, A, V) (See Figures 3, 4).



**FIGURE 2 |** Averaged rating scores detected for each modality (AVC, Audio-Visual Congruent; AVI, Audio-Visual Incongruent; A, Audio; V, Video) and emotion (Laugh, Cry, Control). Error bars represent standard errors of mean (SE).



### Zygomaticus Major muscle

The analysis of *Zygomaticus Major* muscle EMG responses revealed a significant main effect of Emotion [ $F_{(2, 56)} = 4.20$   $p < 0.05$ ]. *Post-hoc* showed that during the presentation of Laugh stimuli EMG responses were stronger than during the presentation of Control stimuli ( $p < 0.01$ ).

Furthermore, the interaction Modality by Emotion was also significant [ $F_{(6, 168)} = 4.55$   $p < 0.001$ ]. *Post-hoc* showed that in AVI Cry condition (in which participants saw an actor laughing but heard crying) EMG *Zygomaticus* responses were stronger with respect to A cry and V cry ( $p < 0.05$ ). In AVI Laugh condition (in which participants saw instead an actor crying but heard laughing), EMG *Zygomaticus* responses were weaker than to all other conditions (i.e., AVC laugh, A laugh and V laugh) ( $p < 0.01$ ). In sum, results showed that EMG *Zygomaticus* responses in AVI modality were driven by what participants saw and not by what they heard. Furthermore, *post-hoc* analysis revealed that EMG responses were not modulated in all different modalities by Control stimuli presentation (all  $p_s > 0.05$ ). The interaction Emotion by Period was also significant [ $F_{(6, 168)} = 2.82$   $p < 0.05$ ]. *Post-hoc* comparisons showed that during Laugh stimuli presentation *Zygomatic* EMG responses increased with time (T1 vs. T3  $p < 0.01$ , T1 vs. T4  $p < 0.0001$ , T2 vs. T4  $p < 0.01$ ), whereas no modulation through time periods was found during Cry and Control stimuli presentation (all  $p_s > 0.05$ ).

The interaction Modality by Emotion by Period was also significant [ $F_{(18, 504)} = 3.19$   $p < 0.0001$ ]. Of most interest, a significant interaction of all factors was observed among Modality, Emotion, Period and Group [ $F_{(18, 504)} = 1.68$   $p < 0.05$ ] (**Figure 3**). Since Control stimuli were rated by both groups (SZP, CNT) as neutral and EMG activity was not modulated during perception of these stimuli (all  $p_s > 0.1$ ), we can consider them as effective neutral stimuli for emotion perception. Thus, we compared EMG *Zygomaticus* activity during the presentation of positive (Laugh) and negative (Cry) emotion-related stimuli with neutral (Control) stimuli.

In line with previous literature (Dimberg and Thunberg, 1998), *post-hoc* comparisons revealed that CNT group showed *Zygomaticus* EMG responses when they saw and heard actors laughing in a congruent way (i.e., AVC modality) 500 ms after stimulus onset (T2, T3, T4; all  $p_s < 0.000$ ). By comparing EMG responses to positive and negative stimuli (**Figure 3A**), we found an inhibition in the same temporal periods for the latter ones (T2, T3, T4; all  $p_s < 0.000$ ). During perception of positive stimuli SZP group showed significant EMG activation, occurring later, 1000 ms after stimulus onset (T3, T4; all  $p_s < 0.01$ ). However, the same EMG responses were recorded during both positive and negative stimuli presentations (T3, T4; all  $p_s > 0.8$ ). We thus defined this EMG response as “non-specific activation,” because it appeared independently of the perceived emotion (Laugh, Cry).

As shown in **Figure 3B**, in CNT participants *Zygomaticus* EMG responses occurred when they saw actors laughing but they heard crying (i.e., AVI Cry condition, in which the visual and auditory components of the stimuli are combined in an incongruent way) 500 ms after stimulus onset (T2, T3, T4; all  $p_s < 0.05$ ). By comparing EMG activity recorded during AVI

Laugh condition with that recorded during AVI Cry condition, we found inhibition in the same time periods (T2, T3, T4 all  $p_s < 0.001$ ). In AVI condition, we thus observed that *Zygomatic* muscle activation was driven by what CNT saw and not by what they heard.

SZP participants did not activate *Zygomatic* muscle in this condition (all  $p_s > 0.05$ ). By contrasting EMG activity during AVI Laugh condition with that recorded during AVI Cry condition, inhibition was found (T2, T3, T4; all  $p_s < 0.05$ ). When emotional visual and auditory information contrasted, patients did not activate EMG *Zygomaticus* muscle.

As shown in **Figure 3C**, in CNT group *Zygomaticus* EMG responses occurred when they only heard laughing (i.e., A modality) already at T1 that is, before 500 ms from stimulus onset (T1, T2, T3, T4; all  $p_s < 0.05$ ). By comparing EMG activity recorded during Cry stimuli with that recorded during Laugh stimuli, inhibition during the same time periods was found (T1, T2, T3, T4; all  $p_s < 0.01$ ).

SZP participants did not activate *Zygomatic* muscle in this condition (all  $p_s > 0.4$ ). When positive emotional auditory information was presented alone, patients did not react with any *Zygomatic* EMG responses.

As shown in **Figure 3D**, in CNT group *Zygomaticus* EMG responses occurred when they only saw laugh (i.e., V modality) already at T1 that is, before 500 ms from stimulus onset (T1, T2, T3, T4; all  $p_s < 0.05$ ). Still in V modality, by comparing EMG activity recorded during Cry stimuli with that recorded during Laugh stimuli, inhibition during the same time periods was found (T1, T2, T3, T4; all  $p_s < 0.05$ ).

In SZP group EMG activation occurred already at T1 that is, before 500 ms from stimulus onset (T1, T2, T3, T4; all  $p_s < 0.05$ ) during positive stimuli. However, no difference was found between EMG responses recorded during negative and positive stimuli, because, similarly to AVC modality, *Zygomaticus* responded also during negative stimuli presentation (T1, T2, T3, T4; all  $p_s > 0.05$ ). We define this EMG response as “non-specific activation,” because it appeared independently of the perceived emotion (Laugh, Cry).

### Corrugator Supercilii muscle

The analysis of *Corrugator Supercilii* muscle EMG responses revealed a significant interaction Modality by Emotion [ $F_{(6, 168)} = 3.11$   $p < 0.01$ ]. *Post-hoc* showed that in AVI Cry condition (in which participants saw an actor laughing but heard crying) EMG *Corrugator* responses were weaker than in all other conditions (i.e., AVC cry, A cry and V cry) ( $p < 0.05$ ). In AVI Laugh condition (in which participants saw an actor crying but heard laughing), EMG *Corrugator* responses were stronger only with respect to the V laugh condition ( $p < 0.05$ ). Furthermore, *post-hoc* analysis revealed that EMG responses were not modulated by different modalities during Control stimuli presentation (all  $p_s > 0.7$ ).

Most interestingly, a significant interaction Modality by Emotion by Period was also observed [ $F_{(18, 504)} = 2.78$   $p < 0.001$ ] (**Figure 4**). Since Control stimuli had been rated neutral by both groups (SZP, CNT) and EMG activity was never modulated during this condition (all  $p_s > 0.2$ ), we considered

Control stimuli, also for Corrugator EMG responses, as effective neutral stimuli for emotion perception. We performed the same comparisons already described for Zygomaticus muscle.

*Post-hoc* comparisons revealed that in both groups (SZP, CNT) Corrugator EMG responses occurred when they saw and heard cry (i.e., AVC modality) 500 ms after stimulus onset (T2, T3, T4; all  $p_s < 0.01$ ). Inhibition occurred during positive stimuli presentation in the same time periods (T2, T3, T4; all  $p_s < 0.000$ ).

Both groups activated Corrugator muscle when they saw cry but they heard laugh (i.e., AVI Laugh modality) 1000 ms after stimulus onset (T3  $p < 0.05$ ; T4  $p = 0.05$ ). By comparing EMG activity recorded during AVI Cry with that recorded during AVI Laugh, inhibition 500 ms after stimulus onset was found (T2, T3, T4 all  $p_s < 0.05$ ). In AVI modality, Corrugator muscle activation was driven by what both groups of participants saw and not by what they heard.

Corrugator EMG responses occurred when both groups heard cry (i.e., A modality) 500 ms after stimulus onset (T2, T3, T4; all  $p_s < 0.01$ ), as it happened in AVC modality. EMG activity was inhibited during presentation of Laugh stimuli in the same time periods (T2, T3, T4; all  $p_s < 0.0001$ ).

Corrugator EMG responses occurred when both groups saw cry (i.e., V modality) 500 ms after stimulus onset (T2, T3, T4; all  $p_s = 0.0000$ ), as in AVC and A modalities. EMG activity was inhibited during presentation of Laugh stimuli in the same time periods (T2, T3, T4; all  $p_s < 0.000$ ).

For Corrugator muscle, we found no significant main effects and interactions of Group factor (all  $p_s > 0.05$ ).

#### FUNCTIONAL RELATIONS BETWEEN EMG AND BEHAVIORAL RATING

In order to analyze behavioral data of the Externalizer and Internalizer cohorts, each of which comprised patients and control participants, two ANOVAs for each cohort (one for each emotion: positive, negative) were run.

##### Externalizer cohorts

For both the Externalizer cohorts, no significant Group main effects and interactions were found for positive emotions (all  $p_s > 0.8$ ) as well as for negative emotions (all  $p_s > 0.8$ ). For

positive emotions, only a significant main effect of Modality [ $F_{(3, 42)} = 5.13 p < 0.01$ ] was found. Since no other significant main effects and interactions were detected, this means that participants belonging to the Externalizer cohorts, more responding with congruent facial mimicry to positive and negative emotions, also gave correct behavioral ratings, with no differences between groups.

##### Internalizer cohorts

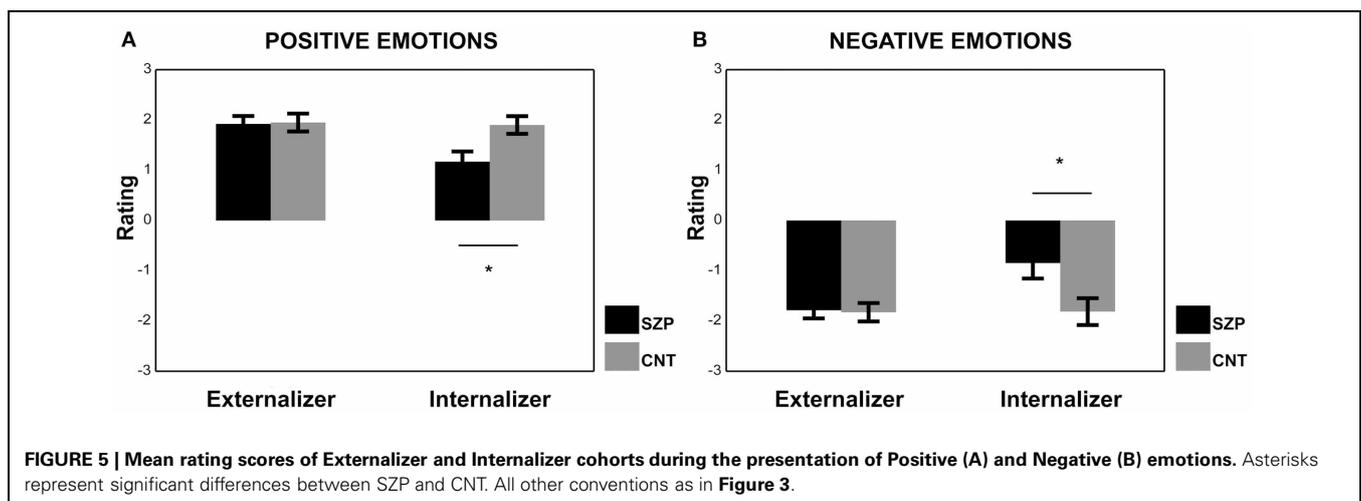
Regarding the results of the ANOVA performed on the Internalizer cohort for positive emotions, we found a significant main effect of Group [ $F_{(1, 12)} = 7.40 p < 0.05$ ]. *Post-hoc* comparisons revealed that CNT group gave more positive ratings than SZP group ( $p < 0.05$ ). The factor Modality was also significant [ $F_{(3, 36)} = 4.77 p < 0.01$ ] and AVI modality received lower ratings with respect to AVC and V modalities ( $p_s < 0.01$ ).

Regarding the results of the ANOVA performed on the Internalizer cohort for negative emotions we found, again, a significant main effect of Group [ $F_{(1, 12)} = 5.63 p < 0.05$ ]. *Post-hoc* comparisons revealed that CNT group gave more negative ratings than SZP group ( $p < 0.05$ ). The factor Modality was also significant [ $F_{(3, 36)} = 4.44 p < 0.01$ ], and AVI modality received the most positive ratings with respect to all other modalities ( $p_s < 0.05$ ).

The Internalizer cohort had EMG below the 100% (i.e., below the baseline value), therefore those participants did not activate their muscles in response to emotional stimuli. Within the Internalizer cohort, we found a significant difference between CNT and SZP groups both for positive and negative emotions. Interestingly, SZP group gave more neutral ratings to perceived positive and negative emotions (see **Figure 5**).

#### DISCUSSION

Our study shows that SZP and CNT participants adequately recognized the emotional quality of the stimuli in all modalities: both groups judged Laugh as positive emotion, Cry as negative emotion and Control stimuli as devoid of any emotional content. Similarly, they were able to score dimensionally the perceived emotions (Laugh, Cry, Control) and, in AVI condition, both



groups privileged the visual over the auditory modality. Indeed, they judged the emotion they saw rather than the emotion they heard.

With respect to EMG recordings, CNT group results are coherent with previous studies (Dimberg and Thunberg, 1998) that documented the role of Zygomaticus muscle in response to positive emotional stimuli and that of Corrugator Supercilii in response to negative ones. Also the timing of the activation of the muscles was in line with previous findings (500 ms after stimulus onset). However, with “single modalities” (A and V) intense EMG responses of Zygomaticus muscle were evoked by positive emotions even before 500 ms from stimulus onset. Notably, the inclusion of Control stimuli enriched the experimental paradigm with an effective neutral stimulus, which proved to evoke no EMG response and be judged by all participants as emotionally neutral. This shows that facial mimicry does not occur indiscriminately whenever one looks at the moving face of someone else, but requires the observation of emotion-specific pattern of facial movements. The same emotional specificity of EMG activation also occurred in Audio modality where only Laugh and Cry sounds were able to evoke EMG activation of the Zygomaticus and Corrugator muscles, respectively.

A further innovative feature of the adopted paradigm was the Incongruent (AVI) condition. Interestingly enough, we discovered that in AVI condition healthy participants reacted with rapid and automatic mimicry following the visual emotional content of the stimuli, while disregarding the auditory expressed contrasting emotion.

Whereas CNT and SZP groups reacted to negative emotional stimuli in the same way, SZP participants did not respond or showed inadequate EMG reactions (a “non-specific response”) to positive emotions. Indeed, in AVI and Audio modalities no EMG activation was found, whereas in AVC and V modalities a non-specific response appeared independently of the perceived emotion (Laugh, Cry).

Several studies investigated how emotional stimuli conveyed by visual and auditory modalities are integrated in healthy population (for a review, see Klasen et al., 2012) and in patients with schizophrenia (de Gelder et al., 2005; de Jong et al., 2009, 2010; Castagna et al., 2013). Studies conducted with healthy participants demonstrated that congruent audiovisual emotions usually yield the highest recognition rates, followed by visual emotions, with the auditory emotions being most difficult to classify (Pourtois et al., 2005; Kreifelts et al., 2007; Collignon et al., 2008; Klasen et al., 2011). In our experiment, both groups prioritized the visual over the auditory modality in identifying the emotion when they were incongruently associated (i.e., AVI modality). Furthermore, only for positive emotions, Video modality was rated by all participants as more intense with respect to Audio modality, whereas the AVC modality was judged as equally intense with respect to single modalities (i.e., Audio and Video). The lack of significant rate advantage of bimodality may be explained by a ceiling effect of high rates in both unimodal conditions (cf. de Gelder and Vroomen, 2000).

A recent study investigating multisensory integration in schizophrenia (Castagna et al., 2013) found that patients were not impaired in basic non-emotional and emotional prosody tasks,

whereas they showed a specific impairment in decoding emotion in a conflicting auditory condition (i.e., when the content of a sentence was not congruent with the emotional tone expressed by the voice) and in a multisensory integration condition (i.e., when complex emotional auditory and visual cues had to be associated). Another study (de Jong et al., 2010) found that in contrast to controls, a stronger impact of facial on vocal emotion perception occurs in patients diagnosed with schizophrenia. Differently from our experiment, all these studies used complex tasks (e.g., cross-modal emotional recognition tasks) that required top-down attentional processes that demanded superior cognitive abilities (i.e., executive functions), which are notoriously impaired in patients suffering from schizophrenia (Castagna et al., 2013). In our experiment, instead, we investigated automatic bottom-up responses to multimodal emotional stimuli presentation.

In our study, we found that CNT and SZP groups similarly reacted with EMG corrugator responses to negative emotional stimuli. Notably, our findings regarding patients’ motor resonance in response to negative stimuli cohere with previous qualitative and quantitative studies documenting how everyday life of patients suffering from schizophrenia is marked by selective biases toward negative emotional experiences which amplify stress-vulnerability and are possibly fostered by persecutory feelings, increased impressionability and oversensitivity to perceived threats. This might be interpreted in line with Kapur’s concept of *aberrant salience*, which posits that positive symptoms of schizophrenia may arise out of “the aberrant assignment of salience to external objects and internal representations” (Kapur, 2003; Van Os and Kapur, 2009). Hence patients’ emotional susceptibility to negative stimuli, resulting in their persistent negative attitude in everyday life might act as a self-perpetuating mechanism of disturbed salience (Mattes et al., 1995; Kring and Earnst, 1999). Thus, positive daily experiences are few and the occasions of showing congruent motor resonance with happy emotions could be consequently uncommon in patients suffering from schizophrenia (Kring and Earnst, 2003; Wolf et al., 2004, 2006; Trèmeau, 2006), hence the lack of specific and congruent responses of Zygomaticus muscle to positive stimuli.

According to previous results (Kring and Neale, 1996; Sison et al., 1996; Aghevi et al., 2003), patients diagnosed with schizophrenia also showed a disjunction between the expression and the behavioral rating of emotions. In the present study we found this dissociation only for positive emotions, where normal emotional rating was not accompanied by congruent EMG responses.

However, by dichotomizing both groups of participants in two cohorts (Externalizers and Internalizers, Kring and Gordon, 1998) according to the intensity of their EMG congruent responses, we found that the patients’ cohort of Internalizers gave more neutral ratings with respect to control group. This means that in patients facial mimicry in response to positive and negative emotions is crucial to correctly judge from a dimensional point of view the perceived emotion. These data cohere with previous findings documenting empathic response deficits in Schizophrenia (Derntl et al., 2009; Varcin et al., 2010) that

may be related with abnormalities in the mirror neurons mechanisms. According to this model, involuntary facial mimicry constitutes an important low-level mechanism contributing to the experience of empathy (for a review, see Singer and Lamm, 2009), *via* processes of simulation and perception-action coupling subserved by activation of the mirror neurons mechanism. In other words, involuntary facial mimicry reflects an embodied simulation of the perceived emotion, which facilitates its understanding (Gallese, 2003, 2005, 2006; Niedenthal, 2007; Halberstadt et al., 2009; Niedenthal et al., 2009) by promoting primary empathic resonance on a bodily level (Gallese, 2001; Preston and De Waal, 2002; Sonnby-Borgström, 2002; Sonnby-Borgström et al., 2003; Oberman et al., 2007). Hence, the disruption of this low-level mechanism may contribute to the well-known empathy deficits in schizophrenia (Varcin et al., 2010). Along similar lines, Dimberg et al. (2011) demonstrated that the ability to react with facial EMG activations to facial expressions and to rate these stimuli as more intense is particularly evident among people with high emotional empathy. Our findings cohere with those of Dimberg et al. (2011), since Internalizer patients neither react with any EMG response, nor rated positive or negative emotional stimuli as significantly more intense.

It should be added that Internalizer healthy participants could correctly score perceived emotions despite their apparent EMG hyporeactivity. This result shows that multi-modal emotion recognition can occur even without full-blown facial mimicry. This might be due to the recruitment of high-level cognitive mechanisms possibly fostered through coping strategies. Facial mimicry might be a necessary condition for *fine-grained emotional evaluation* only for Internalizer patients, who are impaired in correctly judging the intensity of positive and negative emotions.

The interpretation of the current findings, however, should be tempered by some limitations. First, the relatively modest sample size reduced the statistical power. Hence possible group differences, such as regarding the EMG corrugator response to negative stimuli, might not have been detected. Second, all participants with schizophrenia were under antipsychotic medications, which might act as confounders in EMG responses. Nonetheless, since the participants' SAS score (a specific psychometric index sensitive to neuroleptic-induced parkinsonism) was below the cut-off, we are inclined to consider minimal such potential confounder.

It might also be worth noting, that group differences in EMG activation could not be attributed to attentional or motivational factors, for three main reasons. First, all patients were clinically stable (i.e., without hallucinations and similarly flamboyant psychopathology) when underwent the current experiment. Second, the ratings showed that both groups correctly scored the different emotions without significant inter-group differences. Third, the lack of responses that characterized patients was emotion specific, was present only in two modalities, and it was not casually distributed among conditions.

In conclusion, this study provides new evidence on the emotion-specificity of facial mimicry. Further, it demonstrates

that (1) congruent facial mimicry can be evoked multi-modally and that (2) when Video and Audio modalities are incongruently associated, the Video modality prevails on the Audio as a response-trigger. The paradigm also proved sensitive to detect deficits in rapid facial mimicry for positive emotions in patients diagnosed with schizophrenia. We interpreted such deficits in rapid facial mimicry as indicative of a possible low-level impairment of motor resonance mechanisms, which may explain a portion of the empathizing deficits in schizophrenia. This coheres with our finding that the weaker facial mimicry response shown by patients' Internalizer cohort is related to difficulties in correctly judging the intensity of positive and negative emotions.

In our view, these findings could lead to future studies on the nature of emotional deficits in Schizophrenia, capitalizing on the convergence between neuroscience and psychopathology. Indeed, contemporary psychopathological research emphasizes the relevance of disruption of implicit bodily functioning (of which facial mimicry is a crucial component) for the loss of practical immersion in the intersubjective world that constitutes the hallmark of schizophrenia spectrum vulnerability (see Parnas et al., 2002; Stanghellini, 2004; Fuchs, 2005; Parnas, 2011; see also Ebisch et al., 2012; Ferri et al., 2012; Gallese and Ferri, 2013). Therefore, the disturbance of motor resonance revealed in this study, might be implicated in some of the disorders of intersubjective attunement that phenomenologically-oriented psychopathology indicates as core features of Schizophrenia (Minkowski, 1927; Blankenburg, 1971; Parnas and Bovet, 1991; Parnas et al., 2002).

The interpretation of these results grounded on the hypothesis of an impaired functionality of motor resonance mechanisms in patients diagnosed with schizophrenia, should be limited by the lack in our study of direct measures of the entailed underpinning neural mechanisms. Nevertheless, a previous fMRI study carried out by Carr et al. (2003), demonstrated that in healthy participants observation and imitation of emotions activated a similar neural network of brain areas, in which the insula acted as an interface between the premotor component of the mirror mechanism and the limbic system, thus enabling the translation of an observed or imitated facial emotional expression into its internally felt emotional significance. Such results were interpreted by the authors as a mechanism that may mediate the understanding of the emotional state of others, thus contributing to empathy.

Overall, our results provide an encouraging exploratory paradigm to investigate the nature of emotional deficits in Schizophrenia that could be fruitfully coupled with neuroimaging studies aimed to investigate the neural substrate underpinning the deficits in rapid facial mimicry in patients suffering from schizophrenia.

## ACKNOWLEDGMENTS

This work was supported by the EU grants TESIS to Vittorio Gallese. We would like to acknowledge C. Ferrari, F. Tragni, I. Florindo, and M. Sato for their collaboration in stimuli preparation.

## REFERENCES

- Adelmann, P. K., and Zajonc, R. B. (1989). Facial efferece and experience of emotion. *Ann. Rev. Psychol.* 40, 249–280. doi: 10.1146/annurev.ps.40.020189.001341
- Aghevlı, M. A., Blanchard, J. J., and Horan, W. P. (2003). The expression and experience of emotion in schizophrenia: a study of social interactions. *Psychiatry Res.* 119, 261–270. doi: 10.1016/S0165-1781(03)00133-1
- Andreasen, N. C. (1984a). *Scale for the Assessment of Negative Symptoms (SANS)*. Iowa City, IA: University of Iowa.
- Andreasen, N. C. (1984b). *Scale for the Assessment of Positive Symptoms (SAPS)*. Iowa City, IA: University of Iowa.
- Andreasen, N. C., Pressler, M., Nopoulos, P., Miller, D., and Ho, B. C. (2010). Antipsychotic dose equivalents and dose-years: a standardized method for comparing exposure to different drugs. *Biol. Psychiatry* 67, 255–262. doi: 10.1016/j.biopsych.2009.08.040
- Berenbaum, H., and Oltmanns, T. F. (1992). Emotional experience and expression in schizophrenia and depression. *J. Abnorm. Psychol.* 101, 37–44. doi: 10.1037/0021-843X.101.1.37
- Blankenburg, W. (1971). *Der Verlust der natürlichen Selbstverständlichkeit. Ein Beitrag zur Psychopathologie symptomarmer Schizophrenien*. Stuttgart: Enke.
- Bleuler, E. (1950). *Dementia Praecox or the Group of Schizophrenias*. Oxford, UK: International Universities Press. (Original work published in 1911).
- Carr, L., Iacoboni, M., Dubeau, M. C., Mazziotta, J. C., and Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proc. Natl. Acad. Sci. U.S.A.* 100, 5497–5502. doi: 10.1073/pnas.0935845100
- Castagna, F., Montemagni, C., Milani, A. M., Rocca, G., Rocca, P., Casacchia, M., et al. (2013). Prosody recognition and audiovisual emotion matching in schizophrenia: the contribution of cognition and psychopathology. *Psychiatry Res.* 205, 192–198. doi: 10.1016/j.psychres.2012.08.038
- Chapman, L. J., Chapman, J. P., and Raulin, M. L. (1976). Scales for physical and social anhedonia. *J. Abnorm. Psychol.* 85, 374–382. doi: 10.1037/0021-843X.85.4.374
- Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., et al. (2008). Audio-visual integration of emotion expression. *Brain Res.* 1242, 126–135. doi: 10.1016/j.brainres.2008.04.023
- de Gelder, B., and Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cogn. Emot.* 14, 289–311. doi: 10.1080/026999300378824
- de Gelder, B., Vroomen, J., de Jong, S. J., Masthoff, E. D., Trompenaars, F. J., and Hodiomont, P. (2005). Multisensory integration of emotional faces and voices in schizophrenics. *Schizophr. Res.* 72, 195–203. doi: 10.1016/j.schres.2004.02.013
- de Jong, J. J., Hodiomont, P. P., and de Gelder, B. (2010). Modality-specific attention and multisensory integration of emotions in schizophrenia: reduced regulatory effects. *Schizophr. Res.* 122, 136–143. doi: 10.1016/j.schres.2010.04.010
- de Jong, J. J., Hodiomont, P. P., Van den Stock, J., and de Gelder, B. (2009). Audiovisual emotion recognition in schizophrenia: reduced integration of facial and vocal affect. *Schizophr. Res.* 107, 286–293. doi: 10.1016/j.schres.2008.10.001
- Derntl, B., Finkelmeyer, A., Toygar, T. K., Hülsmann, A., Schneider, F., Falkenberg, D. I., et al. (2009). Generalized deficit in all core components of empathy in schizophrenia. *Schizophr. Res.* 108, 197–206. doi: 10.1016/j.schres.2008.11.009
- Dimberg, U. (1982). Facial reactions to facial expressions. *Psychophysiology* 19, 643–647. doi: 10.1111/j.1469-8986.1982.tb02516.x
- Dimberg, U. (1988). Facial electromyography and the experience of emotion. *J. Psychophysiol.* 2, 277–282.
- Dimberg, U., Andréasson, P., and Thunberg, M. (2011). Emotional empathy and facial reactions to facial expressions. *J. Psychophysiol.* 25, 26–31. doi: 10.1027/0269-8803/a000029
- Dimberg, U., and Pettersson, M. (2000). Facial reactions to happy and angry facial expressions: evidence for right hemisphere dominance. *Psychophysiology* 37, 693–696. doi: 10.1111/1469-8986.3750693
- Dimberg, U., and Thunberg, M. (1998). Rapid facial reactions to emotional facial expressions. *Scand. J. Psychol.* 39, 39–45. doi: 10.1111/1467-9450.00054
- Dimberg, U., Thunberg, M., and Elmeled, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychol. Sci.* 11, 86–89. doi: 10.1111/1467-9280.00221
- Dimberg, U., Thunberg, M., and Grunedal, S. (2002). Facial reactions to emotional stimuli: automatically controlled emotional responses. *Cogn. Emot.* 16, 449–471. doi: 10.1080/02699930143000356
- Earnst, K. S., Kring, A. M., Kadar, M. A., Salem, J. E., Shepard, D. A., and Loosen, P. T. (1996). Facial expression in schizophrenia. *Biol. Psychiatry* 40, 556–558. doi: 10.1016/0006-3223(96)00171-0
- Ebisch, S. J. H., Salone, A., Ferri, F., De Berardis, D., Mantini, D., Ferro, F. M., et al. (2012). Out of touch with reality? Social perception in first episode schizophrenia. *Soc. Cogn. Affect. Neurosci.* 8, 394–403. doi: 10.1093/scan/nss012
- Ekman, P., and Oster, H. (1979). Facial expressions of emotion. *Annu. Rev. Psychol.* 30, 527–554. doi: 10.1146/annurev.ps.30.020179.002523
- Ferri, F., Frassinetti, F., Mastrangelo, F., Salone, A., Ferro, F. M., and Gallese, V. (2012). Bodily self and schizophrenia: the loss of implicit self-body knowledge. *Conscious. Cogn.* 21, 1365–1374. doi: 10.1016/j.concog.2012.05.001
- Fridlund, A. J., and Cacioppo, J. T. (1986). Guidelines for human electromyographic research. *Psychophysiology* 23, 567–589. doi: 10.1111/j.1469-8986.1986.tb00676.x
- Fuchs, T. (2005). Corporealized and disembodied minds: a phenomenological view of the body in melancholia and schizophrenia. *Philos. Psychiat. Psychol.* 12, 95–107.
- Gallese, V. (2001). The ‘shared manifold’ hypothesis: from mirror neurons to empathy. *J. Consc. Stud.* 8, 33–50.
- Gallese, V. (2003). The roots of empathy: the shared manifold hypothesis and the neural basis of intersubjectivity. *Psychopathology* 36, 171–180. doi: 10.1159/000072786
- Gallese, V. (2005). Embodied simulation: from neurons to phenomenal experience. *Phenomenol. Cogn. Sci.* 4, 23–48. doi: 10.1007/s11097-005-4737-z
- Gallese, V. (2006). Intentional attunement: a neurophysiological perspective on social cognition and its disruption in autism. *Brain Res.* 1079, 15–24. doi: 10.1016/j.brainres.2006.01.054
- Gallese, V., and Ferri, F. (2013). Jaspers, the body and Schizophrenia. The bodily self. *Psychopathology* (in press).
- Goldman, A. I., and Sripada, C. S. (2005). Simulationist models of face-based emotion recognition. *Cognition* 94, 193–213. doi: 10.1016/j.cognition.2004.01.005
- Halberstadt, J., Winkelman, P., Niedenthal, P. M., and Dalle, N. (2009). Emotional conception: how embodied emotion concepts guide perception and facial action. *Psychol. Sci.* 20, 1254–1261. doi: 10.1111/j.1467-9280.2009.02432.x
- Han, C., Park, G. Y., Wang, S. M., Lee, S. Y., Lee, S. J., Bahk, W. M., et al. (2012). Can botulinum toxin improve mood in depressed patients? *Expert Rev. Neurother.* 12, 1049–1051. doi: 10.1586/ern.12.92
- Heimann, H., and Spoerri, T. (1957). Das Ausdruckssyndrom der mimischen Düsintegration bei chronischen Schizophrenen. *Schweiz. Med. Wochenschr.* 35/36, 1126–1132.
- Iwase, M., Yamashita, K., Takahashi, K., Kajimoto, O., Shimizu, A., Nishikawa, T., et al. (1999). Diminished facial expression despite the existence of pleasant emotional experience in schizophrenia. *Methods Find. Exp. Clin. Pharmacol.* 21, 189–194. doi: 10.1358/mf.1999.21.3.534828
- Janno, S., Holi, M. M., Tuisku, K., and Wahlbeck, K. (2005). Validity of Simpson-Angus Scale (SAS) in a naturalistic schizophrenia population. *BMC Neurol.* 5:5. doi: 10.1186/1471-2377-5-5
- Kapur, S. (2003). Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am. J. Psychiatry* 160, 13–23. doi: 10.1176/appi.ajp.160.1.13
- Klasen, M., Chen, Y. H., and Mathiak, K. A. (2012). Multisensory emotions: perception, combination and underlying neural processes. *Rev. Neurosci.* 23, 381–392. doi: 10.1515/revneuro-2012-0040
- Klasen, M., Kenworthy, C. A., Mathiak, K. A., Kircher, T. T., and Mathiak, K. (2011). Supramodal representation of emotions. *J. Neurosci.* 31, 13635–13643. doi: 10.1523/JNEUROSCI.2833-11.2011
- Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., and Wildgruber, D. (2007). Audiovisual integration of emotional signals in voice and face: an event-related fMRI study. *Neuroimage* 37, 1445–1456. doi: 10.1016/j.neuroimage.2007.06.020
- Kring, A. M., and Earnst, K. S. (1999). Stability of emotional responding in schizophrenia. *Behav. Ther.* 30, 373–388. doi: 10.1016/S0005-7894(99)80016-1
- Kring, A. M., and Earnst, K. S. (2003). “Nonverbal behavior in schizophrenia,” in *Nonverbal Behavior in Clinical Settings*, eds P. Philippot, E. Coats, and R. S. Feldman

- (New York, NY: Oxford University Press), 263–286.
- Kring, A. M., and Gordon, A. H. (1998). Sex differences in emotion: expression, experience, and physiology. *J. Pers. Soc. Psychol.* 74, 686–703. doi: 10.1037/0022-3514.74.3.686
- Kring, A. M., Kerr, S. L., Smith, D. A., and Neale, J. M. (1993). Flat affect in schizophrenia does not reflect diminished subjective experience of emotion. *J. Abnorm. Psychol.* 102, 507–517. doi: 10.1037/0021-843X.102.4.507
- Kring, A. M., and Neale, J. M. (1996). Do schizophrenic patients show a disjunctive relationship among expressive, experiential, and psychophysiological components of emotions? *J. Abnorm. Psychol.* 105, 249–257. doi: 10.1037/0021-843X.105.2.249
- Larsen, J. T., Norris, C. J., and Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology* 40, 776–785. doi: 10.1111/1469-8986.00078
- Lundqvist, L. O., and Dimberg, U. (1995). Facial expressions are contagious. *J. Psychophysiol.* 9, 203–211.
- Mattes, R. M., Schneider, F., Heimann, H., and Birbaumer, N. (1995). Reduced emotional response in schizophrenic patients in remission during social interaction. *Schizophr. Res.* 17, 249–255. doi: 10.1016/0920-9964(95)00014-3
- Minkowski, E. (1927). *La schizophrénie. Psychopathologie des Schizoides et des Schizophrènes*. Paris: Payot.
- Niedenthal, P. M. (2007). Embodying emotion. *Science* 316, 1002–1005. doi: 10.1126/science.1136930
- Niedenthal, P. M., Mermillod, M., Maringer, M., and Hess, U. (2010). The simulation of smiles (SIMS) model: embodied simulation and the meaning of facial expression. *Behav. Brain Sci.* 33, 417–433. doi: 10.1017/S0140525X10000865
- Niedenthal, P. M., Winkielman, P., Mondillon, L., and Vermeulen, N. (2009). Embodiment of emotion concepts. *J. Pers. Soc. Psychol.* 96, 1120–1136. doi: 10.1037/a0015574
- Oberman, L. M., Winkielman, P., and Ramachandran, V. S. (2007). Face to face: blocking facial mimicry can selectively impair recognition of emotional expressions. *Soc. Neurosci.* 2, 167–78. doi: 10.1080/17470910701391943
- Parnas, J. (2011). A disappearing heritage: the clinical core of schizophrenia. *Schizophr. Bull.* 37, 1121–1130. doi: 10.1093/schbul/sbr081
- Parnas, J., and Bovet, P. (1991). Autism in schizophrenia revisited. *Compr. Psychiatry* 32, 7–21. doi: 10.1016/0010-440X(91)90065-K
- Parnas, J., Bovet, P., and Zahavi, D. (2002). Schizophrenic autism: clinical phenomenology and pathogenetic implications. *World Psychiatry* 1, 131–136.
- Phillips, L. K., and Seidman, L. J. (2008). Emotion Processing in Persons at Risk for Schizophrenia. *Schizophr. Bull.* 34, 888–903. doi: 10.1093/schbul/sbn085
- Pourtois, G., de Gelder, B., Bol, A., and Crommelinck, M. (2005). Perception of facial expressions and voices and of their combination in the human brain. *Cortex* 41, 49–59. doi: 10.1016/S0010-9452(08)70177-1
- Preston, S. D., and De Waal, F. B. M. (2002). Empathy: its ultimate and proximate bases. *Behav. Brain Sci.* 25, 1–20.
- Rymarczyk, K., Biele, C., Grabowska, A., and Majczynski, H. (2011). EMG activity in response to static and dynamic facial expressions. *Int. J. Psychophysiol.* 79, 330–333. doi: 10.1016/j.ijpsycho.2010.11.001
- Simpson, G. M., and Angus, G. W. S. (1970). A rating scale for extrapyramidal side effects. *Psychiatr. Scand.* 45, 11–19. doi: 10.1111/j.1600-0447.1970.tb02066.x
- Singer, T., and Lamm, C. (2009). The social neuroscience of empathy. *Ann. N.Y. Acad. Sci.* 1156, 81–96. doi: 10.1111/j.1749-6632.2009.04418.x
- Sison, C. E., Alpert, M., Fudge, R., and Stern, R. (1996). Constricted expressiveness and psychophysiological reactivity in schizophrenia. *J. Nerv. Ment. Dis.* 184, 589–597. doi: 10.1097/00005053-199610000-00002
- Sonnby-Borgström, M. (2002). Automatic mimicry reactions as related to differences in emotional empathy. *Scand. J. Psychol.* 43, 433–443. doi: 10.1111/1467-9450.00312
- Sonnby-Borgström, M., Jönsson, P., and Svensson, O. (2003). Emotional empathy as related to mimicry reactions at different levels of information processing. *J. Nonverbal Behav.* 27, 3–23. doi: 10.1023/A:1023608506243
- Stanghellini, G. (2004). *Disembodied Spirits and Deanimated Bodies: The Psychopathology of Common Sense*. Oxford: Oxford University Press. doi: 10.1093/med/9780198520894.001.0001
- Trèmeau, F. (2006). A review of emotion deficits in schizophrenia. *Dialogues Clin. Neurosci.* 8, 59–70.
- Van Os, J., and Kapur, S. (2009). Schizophrenia. *Lancet* 374, 635–645. doi: 10.1016/S0140-6736(09)60995-8
- Varcin, K. J., Bailey, P. E., and Henry, J. D. (2010). Empathic deficits in schizophrenia: the potential role of rapid facial mimicry. *J. Int. Neuropsychol. Soc.* 16, 621–629. doi: 10.1017/S1355617710000329
- Walker, E. F., Grimes, K. E., Davis, D. M., and Smith, A. J. (1993). Childhood precursors of schizophrenia: facial expressions of emotion. *Am. J. Psychiatry* 150, 1654–1660.
- Wolf, K., Mass, R., Kiefer, F., Eckert, K., Weinhold, N., Wiedemann, K., et al. (2004). The influence of olanzapine on facial expression of emotions in schizophrenia - an improved facial EMG study. *Germ. J. Psychiatry* 7, 14–19.
- Wolf, K., Mass, R., Kiefer, F., Wiedemann, K., and Naber, D. (2006). Characterization of the facial expression of emotions in Schizophrenia Patients: preliminary findings with a new electromyography method. *Can. J. Psychiatry* 51, 335–341.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 April 2013; accepted: 25 June 2013; published online: 23 July 2013.

Citation: Sestito M, Umiltà MA, De Paola G, Fortunati R, Raballo A, Leuci E, Maffei S, Tonna M, Amore M, Maggini C and Gallese V (2013) Facial reactions in response to dynamic emotional stimuli in different modalities in patients suffering from schizophrenia: a behavioral and EMG study. *Front. Hum. Neurosci.* 7:368. doi: 10.3389/fnhum.2013.00368  
Copyright © 2013 Sestito, Umiltà, De Paola, Fortunati, Raballo, Leuci, Maffei, Tonna, Amore, Maggini and Gallese. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



# Experimental and clinical usefulness of crossmodal paradigms in psychiatry: an illustration from emotional processing in alcohol-dependence

Pierre Maurage<sup>1\*</sup> and Salvatore Campanella<sup>2</sup>

<sup>1</sup> Laboratory for Experimental Psychopathology, Faculty of Psychology, Institute of Psychology, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

<sup>2</sup> Laboratory of Psychological Medicine and Addictology, ULB Neuroscience Institute, Université Libre de Bruxelles, Brussels, Belgium

## Edited by:

Martin Klasen, RWTH Aachen University, Germany

## Reviewed by:

Cheryl Grady, University of Toronto, Canada

Benjamin Kreifelts, University of Tübingen, Germany

## \*Correspondence:

Pierre Maurage, Laboratoire de Psychopathologie Expérimentale (LEP), Faculté de Psychologie, 10 Place C. Mercier, B-1348 Louvain-la-Neuve, Belgium  
e-mail: pierre.maurage@uclouvain.be

Crossmodal processing (i.e., the construction of a unified representation stemming from distinct sensorial modalities inputs) constitutes a crucial ability in humans' everyday life. It has been extensively explored at cognitive and cerebral levels during the last decade among healthy controls. Paradoxically however, and while difficulties to perform this integrative process have been suggested in a large range of psychopathological states (e.g., schizophrenia and autism), these crossmodal paradigms have been very rarely used in the exploration of psychiatric populations. The main aim of the present paper is thus to underline the experimental and clinical usefulness of exploring crossmodal processes in psychiatry. We will illustrate this proposal by means of the recent data obtained in the crossmodal exploration of emotional alterations in alcohol-dependence. Indeed, emotional decoding impairments might have a role in the development and maintenance of alcohol-dependence, and have been extensively investigated by means of experiments using separated visual or auditory stimulations. Besides these unimodal explorations, we have recently conducted several studies using audio-visual crossmodal paradigms, which has allowed us to improve the ecological validity of the unimodal experimental designs and to offer new insights on the emotional alterations among alcohol-dependent individuals. We will show how these preliminary results can be extended to develop a coherent and ambitious research program using crossmodal designs in various psychiatric populations and sensory modalities. We will finally end the paper by underlining the various potential clinical applications and the fundamental implications that can be raised by this emerging project.

**Keywords:** crossmodal, emotion, alcohol-dependence, social cognition, face, voice

## INTRODUCTION

Crossmodal processing can be globally defined as the ability to build a unitary representation of one's environment on the basis of stimulations coming from different sensorial modalities (Driver and Spence, 2000). In everyday life, human beings are confronted with a constant flow of multi-sensorial stimulations, and the capacity to integrate them is thus crucial for daily adaptive behaviors like social interactions, spatial attention or perceptuo-motor coordination (Lalanne and Lorenceau, 2004; Campanella and Belin, 2007). The importance of these crossmodal processes has led many researchers to investigate their behavioral and cerebral correlates, and the exploration of crossmodal mechanisms now constitutes a flourishing field in cognitive psychology and neuroscience (Calvert et al., 2001; De Gelder and Bertelson, 2003; Amedi et al., 2005). Indeed, hundreds of studies have been conducted during the last two decades on crossmodality among healthy participants, and huge advances have undeniably been made in understanding the developmental, psychological and cerebral correlates of crossmodality (particularly of face-voice integration). A wide variety of multimodal tasks have been used,

requiring the integration of gender, identity (e.g., Joassin et al., 2011a,b; Love et al., 2011) or emotional features (e.g., Dolan et al., 2001; Pourtois et al., 2005; Ethofer et al., 2006a,b, 2013; Kreifelts et al., 2007, 2009, 2010; Robins et al., 2009; Müller et al., 2011), and these various results led to the identification of several brain areas specifically involved in the multisensory integration (mainly the superior parietal lobule, inferior occipital, middle frontal and superior temporal sulci).

## CROSSMODAL PROCESSES IN PSYCHIATRY: A SERIOUS LACK OF DATA

The exploration of crossmodality in healthy populations has obviously gained a central position in the experimental psychology and neuroscience domains and has come to maturity, as notably illustrated by the development of integrative models structuring the numerous experimental data available (e.g., Campanella and Belin, 2007 for a review). However, there is a patent contrast between the large number of studies exploring the efficient crossmodal processing and the paucity of data currently available concerning the impairment of this processing in pathological populations. Indeed, while scientific knowledge is

classically enriched by results obtained from impaired populations, very few clinical crossmodal research projects have been conducted up to now. Crossmodal integration impairments have been explored in several populations presenting perceptual alterations (e.g., visual or auditory loss, Zupan and Sussman, 2009; Barone, 2010; Massida et al., 2011), but very few studies have focused on crossmodal processing in neurological and psychiatric populations. Among these studies, main results concerned schizophrenia (De Gelder et al., 2005; Pearl et al., 2009; Szycik et al., 2009; Seubert et al., 2010a; Van den Stock et al., 2011), autism (van der Smagt et al., 2007; Mongillo et al., 2008; Foss-Feig et al., 2010; Kwakye et al., 2011) and Alzheimer's disease (Delbeuck et al., 2007), and suggested large-scale crossmodal deficits in these populations, while unimodal processes may be preserved (e.g., De Jong et al., 2009). As these psychopathological states have been described as disconnection syndromes (Friston and Frith, 1995; Melillo and Leisman, 2009), these crossmodal impairments might be related to an alteration of binding abilities needed to connect and integrate the different sensorial inputs. Moreover, as crossmodal integration is crucial in daily life, these alterations could be at least partly responsible for cognitive and social alterations observed in psychiatric states. Up to now however, as only scarce data are available in schizophrenia and autism, and as crossmodality has not been explored in other psychiatric states, many questions have remained unexplored concerning the crossmodal processing in psychiatry.

This striking paradox between the extensive knowledge concerning "normal" crossmodal processing, notably for the integration of emotional stimuli (e.g., De Gelder et al., 1999; Pourtois et al., 2000, 2002; Chen et al., 2010; Kreifelts et al., 2010; Regenbogen et al., 2012; Klasen et al., 2012; Ethofer et al., 2013) and the very few data on impaired crossmodal integration in clinical populations thus constitutes a strong limit of the current knowledge on crossmodality, and voices recently rose to promote the development of crossmodal research among clinical populations, with a double aim. On the one hand, there is a need to obtain a better description of the impairments associated with pathological states, particularly by offering a more ecological and complete evaluation of the deficits. On the other hand, a renewal of the understanding of crossmodal integration among healthy subjects is urgently needed. Indeed, if behavioral deficits in crossmodal processing are observed in a clinical population, comparing the cerebral activations between this population and healthy controls will offer strong insights concerning the brain regions associated with crossmodal processing. As summarized by Laurienti et al. (2005), "the use of clinical populations can add to the battery of study designs available to the imaging scientist investigating multisensory integration." Despite the great promise of this perspective, very little research has up to now attempted to improve our understanding of both pathological states and the mechanisms of multisensory integration in general through crossmodal research in clinical populations.

#### **UNDERLINING THE USEFULNESS OF CROSSMODAL RESEARCH IN PSYCHIATRY**

In view of the present limitations related to crossmodal research in psychiatric populations, the central objective of the present

paper is to underline the usefulness of exploring crossmodal processing among clinical populations, and to help prepare the ground for the expansion of this innovative research topic. First, our recent studies exploring emotional crossmodal processing in alcohol-dependence will be described in order to illustrate the possibilities offered by this research field. Indeed, emotional decoding impairments have been found to be involved in the relapse after detoxification (e.g., Kornreich et al., 2001; Zywiak et al., 2003), and have been extensively investigated using visual or auditory stimulation (e.g., Monnot et al., 2001; Townshend and Duka, 2003, respectively). Capitalizing on these unimodal explorations, our research group has recently conducted several studies using audio-visual bimodal paradigms to increase the ecological validity of the experimental designs. The complementary use of behavioral, electrophysiological and neuroimaging techniques allowed obtaining the first insights concerning multimodal integration in alcohol-dependence. Second, these initial results will be discussed to show how they can (together with preliminary results obtained for multimodal integration in other psychiatric states) be extended to develop a coherent and ambitious research program using various psychiatric populations and sensory modalities. Finally, a conclusive section will underline the various potential clinical applications and the fundamental implications that this emerging project could bring.

#### **EMOTIONAL DEFICITS IN ALCOHOL-DEPENDENCE**

##### **ALCOHOL-DEPENDENCE: A SERIOUS MENTAL HEALTH PROBLEM**

Alcohol-dependence is the most frequent psychiatric diagnosis worldwide and is listed among the three more detrimental health problems (Harper and Matsumoto, 2005). Around 10% of the adult population in Western countries fulfils the diagnosis criteria for alcohol-dependence, and excessive alcohol consumption directly leads to 2.5 million deaths per year worldwide (World Health Organization, 2011). In view of these large-scale deleterious consequences of alcohol-dependence, extensive efforts have been conducted during the last decades to obtain a better understanding of alcohol-dependence at clinical and fundamental levels, particularly concerning the physiological, behavioral and cerebral impairments associated with chronic excessive alcohol consumption. Alcohol-dependence is known to have deleterious effects on many body systems (e.g., hepatic, cardio-vascular or gastrointestinal), but also on the central nervous system (see Oscar-Berman and Marinkovic, 2007 for a review.). Indeed, it has been extensively established that alcohol-dependence leads to major cerebral damage (McIntosh and Chick, 2004; Harper, 2007), particularly affecting white matter (Brooks, 2000; Oscar-Berman and Marinkovic, 2003), but also sub-cortical [e.g., amygdala (Cowen et al., 2004; Fein et al., 2006), insula, thalamus and cerebellum (Szabo et al., 2004; De Bellis et al., 2005)] and cortical [mainly temporal and frontal lobes (Kril et al., 1997; Harper and Matsumoto, 2005; Chanraud et al., 2007)] areas. These brain impairments correlate with the lifetime dose of ethanol consumed (Nicolás et al., 1997). At a functional level, many studies have explored the behavioral correlates of these cerebral effects, and have repeatedly shown impaired performance in a large range of cognitive abilities, ranging from perceptual (e.g., Blusewicz

et al., 1977; Spitzer and Ventry, 1980; Spitzer, 1981) and attentional (e.g., Smith and Oscar-Berman, 1992; Sullivan et al., 1993; Noël et al., 2001) abilities to memory and executive functions (e.g., Bechara et al., 2001; Oscar-Berman et al., 2004; Flannery et al., 2007; Pitel et al., 2007). Nevertheless, in contrast with this extensive exploration of the consequences of alcohol-dependence on cognition, the evaluation of emotional abilities has long been neglected in this pathology.

### THE EMOTIONAL DEFICIT

Emotional states clearly have a major influence on most aspects of our lives, as emotions plays a role in every human's decisions, motivations or social interactions. Moreover it has been repeatedly shown that affective disturbances (e.g., impaired ability to identify or regulate one's own emotional states or to decode other persons' emotions) constitute a central characteristic of most mental diseases from a clinical point of view. Despite this obvious importance of emotions in psychiatry, the interest for experimental exploration of affective impairments in alcohol-dependence only rose during the last decade. While this field is still in its infancy, it already clearly appears that alcohol-dependence is associated with major impairments in a wide range of emotional abilities. Several recent research projects identified an impaired performance in various emotional functions among alcohol-dependent individuals, notably for alexithymia (Taieb et al., 2002; Uzun et al., 2003), emotional intelligence (Riley and Schutte, 2003; Szczepanska et al., 2004; Cordovil de Susa Uva et al., 2010) and empathy (Martinotti et al., 2009; Maurage et al., 2011a). More centrally for the present purpose, a deficit has also been consistently observed for the decoding of the emotions expressed by faces (Oscar-Berman et al., 1990; Frigerio et al., 2002; Clark et al., 2007; Marinkovic et al., 2009; Maurage et al., 2009a, 2011b) and voices (Monnot et al., 2001, 2002; Uekermann et al., 2005). Recently detoxified alcohol-dependent individuals globally over-estimate the intensity of the emotions conveyed by visual and auditory stimuli, have an erroneous interpretation of emotions and are not aware of their deficit (Philippot et al., 1999; Kornreich et al., 2001, 2002). While several contradictory results have emerged, describing a preserved decoding of visual (Uekermann and Daum, 2008) or auditory (Oscar-Berman et al., 1990) stimulations, this emotional decoding deficit is now strongly established as it has been replicated in a wide variety of paradigms and stimulus sets (e.g., morphed or ambiguous faces, Frigerio et al., 2002), and among individuals with various abstinence durations (Townshend and Duka, 2003; Montagne et al., 2006; Foisy et al., 2007a). The causal link between alcohol-dependence and emotional impairment is still unclear as no longitudinal study has up to now directly explored this question, but individuals presenting a high risk of developing alcohol-dependence (i.e., children of alcohol-dependent patients) have strong emotional disturbances, and notably (1) altered activations of the brain areas involved in the elicitation and decoding of emotional feelings, particularly the amygdala (Glahn et al., 2007), (2) reduced social skills abilities and inefficient emotional coping strategies, as explored by self-report questionnaires (Segrin and Menees, 1996), and (3) higher frequency of psychopathological states known to have a strong effect on mood and

emotional states, like depression or anxiety (Sher et al., 1991). It can thus be postulated that these emotional deficits might at least partly precede the appearance of alcohol-dependence. Moreover, while initially described for negative emotions, this impairment has been shown to be generalized to positive affective states if more complex emotional stimuli are used: when more various emotional and mental states are involved in the decoding task, alcohol-dependence leads to a deficit for negative and positive emotions (but not for neutral mental states), as recently shown in a study (Maurage et al., 2011b) exploring the performance of alcohol-dependent individuals at the "Reading the Mind in the Eyes" test (Baron-Cohen et al., 2001). It should also be underlined that this emotional decoding deficit appears specific for emotional features and cannot be fully explained by the more general cerebral and cognitive deficits observed in alcohol-dependence. Indeed, it has been shown that, when emotional decoding is compared with other tasks of similar complexity involving the identification of facial features (e.g., age, gender, or race detection tasks), alcohol-dependent individuals present a global deficit for performance and reaction times in each task. Nevertheless, when the general cognitive impairment is controlled for (using a reaction time subtraction method), alcohol-dependence is no more associated with an alteration for non-emotional face processing tasks but the emotional decoding deficit persists, suggesting that this deficit is specific and not only due to general cognitive alterations (Foisy et al., 2007b; Maurage et al., 2008a). To sum up, it is now clearly recognized that alcohol-dependence is associated with a general impairment for the decoding of the emotional content of stimulations, which is present for faces and voices, but also for other affective stimuli like music (Kornreich et al., 2013) or body postures (Maurage et al., 2009a).

### A MORE GENERAL DEFICIT AFFECTING SOCIAL COGNITION

Importantly, it has been suggested that these emotional decoding impairments might influence social interactions and participate in the maintenance of the pathological state (Walitzer and Dearing, 2006). Indeed, as the development and preservation of adapted social communication is largely based on the ability to correctly express one's own emotional states and to accurately perceive (and react to) those expressed by others (Feldman et al., 1991), the emotional deficits might give rise to impaired interpersonal interactions and could increase the social problems frequently observed in alcohol-dependence. This assumption has not yet been directly tested in alcohol-dependence, but it has been shown that the intensity of emotional decoding alterations (specifically the over-estimation of negative emotions) is strongly correlated with the presence of interpersonal problems (as evaluated by the Inventory of Interpersonal Problems, Horowitz et al., 1988), and particularly with self-control difficulties in social contexts (Maurage et al., 2009a). While it has to be directly tested by prospective studies, this proposal of an involvement of emotional and interpersonal difficulties on the time course of alcohol-dependence recently received further empirical support by means of studies offering a specific exploration of the interpersonal abilities in alcohol-dependence. It has notably been shown that alcohol-dependent individuals have an impaired ability to

understand humor (Uekermann et al., 2007), irony (Amenta et al., 2013) and mental states (Thoma et al., 2013), which are crucial abilities to correctly interact in interpersonal contexts. Moreover, alcohol-dependence is also associated with maladaptive self-beliefs in social context: alcohol-dependent individuals present an over-estimation of the social standards that have to be reached in order to be positively evaluated by others during social interactions, and their inability to actually reach these exaggerated social standards reduces their self-esteem and self-confidence when interacting with others (Maurage et al., 2013a). Finally, alcohol-dependence is also associated with a higher sensitivity to social rejection and ostracism (Maurage et al., 2012a), further underlining the role potentially played by emotional and interpersonal disturbances in this pathology.

Overall, better understanding the wide range of emotional disabilities related to alcohol-dependence (e.g., identifying, expressing and regulating one's own emotions, decoding and correctly reacting to others' emotions) is thus essential at fundamental level, but also for clinical practice, as they might have a role in the maintenance of this pathology, notably by hampering the development of satisfactory interpersonal links, thus potentially leading to social isolation and social stigma (Schomerus et al., 2011a,b). This social isolation could in turn reinforce the excessive alcohol consumption (used as a coping strategy to face isolation) and lead to a vicious circle (e.g., Carton et al., 1999). To sum up, the emotion decoding impairment in alcohol-dependence is now clearly identified and has a significant clinical importance, notably in view of its links with interpersonal problems. However, this deficit has up to now been exclusively explored using paradigms with low ecological validity, namely using only unimodal stimuli (faces or voices). It is thus unclear whether this deficit is maintained, reduced or increased in experimental designs that are closer to real life, specifically when crossmodal stimuli are used.

## CROSSMODAL EMOTIONAL ALTERATIONS IN ALCOHOL-DEPENDENCE

As it has been described in the previous section, the affective deficits related to alcohol-dependence are now clearly established, particularly concerning the alterations in the decoding of emotional faces and voices. However, all these studies were based on unimodal explorations, i.e., on the separate presentations of faces and voices. Nevertheless, sensory events are not experienced in isolation in everyday life, as we are constantly immersed in a flow of multiple sensory cues carrying information from different sensory modalities. Crossmodal processing is thus the rule rather than the exception and this is particularly true for emotions, as the perception and production of emotional states are routinely based on several sensory aspects (e.g., emotional facial expressions and emotional prosody in crossmodal face-voice stimuli). Therefore, while constituting a valuable first exploration, the unimodal investigations of affective processing among alcohol-dependent individuals conducted up to now are insufficient to comprehend the complexity of emotion decoding processing in this population and should be extended to more ecological crossmodal designs. Following this observation, and on the basis of earlier crossmodal studies which explored the integration of

emotional stimuli among healthy controls by means of electrophysiological (e.g., De Gelder et al., 1999; Pourtois et al., 2000, 2002; Chen et al., 2010) and neuroimaging techniques (e.g., Dolan et al., 2001; Ethofer et al., 2006a,b, 2013; Kreifelts et al., 2007, 2009, 2010; Müller et al., 2011), three studies were conducted in our research group to explore, for the first time to the best of our knowledge, the crossmodal emotional decoding in alcohol-dependent participants. Importantly, these studies joining behavioral, electrophysiological and neuroimaging techniques will illustrate the usefulness of a complementary approach combining cognitive psychology and neurosciences approaches to determine the modification of audio-visual emotional decoding in alcohol-dependence. It should also be noted that these three studies are focused on the comparison between recently detoxified alcohol-dependent participants (i.e., individuals diagnosed with alcohol-dependence according to DSM-IV criteria and recruited during their third week of treatment in a detoxification center) and healthy controls paired for age, gender and education. Moreover, alcohol-dependent participants had abstained from alcohol for at least 2 weeks before the experiment took place, thus excluding any potential influence of acute alcohol intoxication on the results observed. Finally, in order to ensure that the emotional decoding deficits were indeed associated with alcohol consumption and not with biasing variables, several control measures were conducted: (1) participants had normal or corrected-to-normal visual and auditory abilities; (2) major medical problems, neurological disease (including epilepsy) and other psychiatric diagnoses (as assessed by an exhaustive psychiatric examination), including polysubstance abuse, constituted exclusion criteria in both groups; (3) subclinical psychiatric comorbidities (particularly depression and anxiety) were controlled for by means of questionnaires [i.e., Beck Depression Inventory (Beck and Steer, 1987) for depression, State-Trait Anxiety Inventory (Spielberger et al., 1983) for anxiety]. These measures were entered as covariables in our statistical analyses to control for the influence of these subclinical psychiatric comorbidities.

## DO ALCOHOL-DEPENDENT INDIVIDUALS PRESENT AN EMOTIONAL CROSSMODAL FACILITATION EFFECT?

The first exploration of crossmodal emotional processing in alcohol-dependence (Maurage et al., 2007a) was based on the elicitation of a "crossmodal facilitation effect." Most classical crossmodal paradigms [e.g., McGurk (McGurk and McDonald, 1976) and ventriloquist (e.g., Alais and Burr, 2004) effects] are based on inhibition effects (i.e., to a deteriorated performance in crossmodal conditions as compared to unimodal). Nevertheless, several more recent studies have developed paradigms leading to a facilitation effect (Calvert et al., 2001; Teder-Sälejärvi et al., 2002), in which congruent bimodal (audio-visual) stimulation leads to better performance (i.e., higher correct response rates and/or shorter reaction times) than unimodal ones. This facilitation effect is considered as the behavioral marker of successful crossmodal integration of stimuli from different modalities (Calvert et al., 2001), and the absence of this facilitation effect would conversely index impaired crossmodal integration. This study was thus based on a design eliciting a facilitation effect to evaluate the presence of this effect among alcohol-dependent individuals.

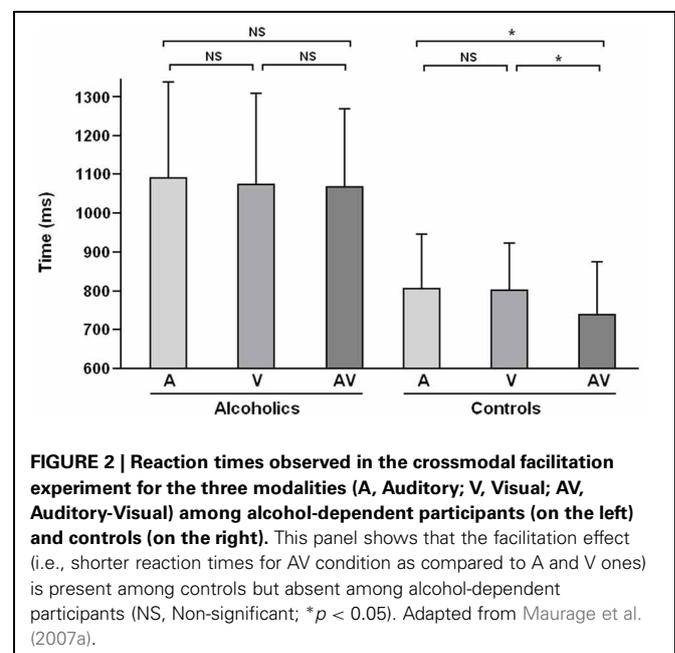
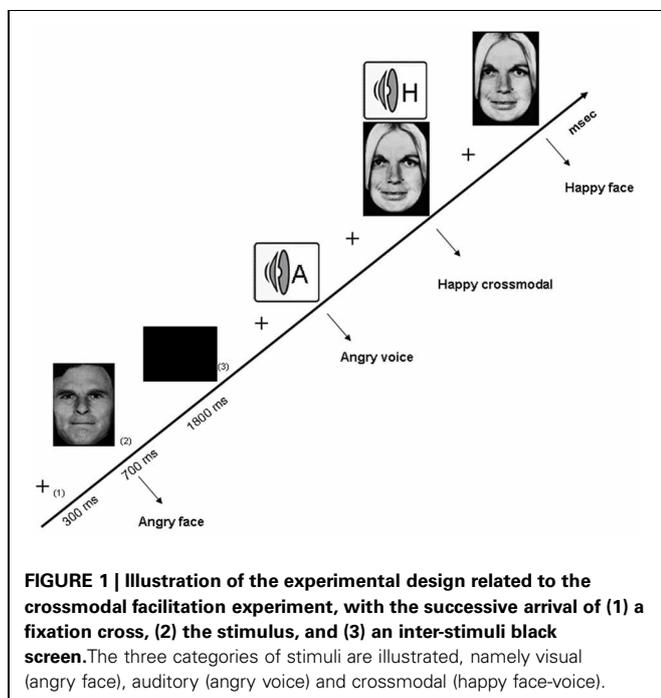
More precisely, an emotion-detection task was used in which participants were presented with emotional facial expressions and voices [i.e., audiotapes enunciating a semantically neutral name with an emotional prosody, taken from a validated battery (Maurage et al., 2007b)] depicting anger or happiness (see **Figure 1** for an illustration of the experimental design). Auditory and visual stimuli were presented separately (unimodal conditions) or simultaneously (crossmodal condition, with emotionally congruent face-voice pairs). Morphed faces were used (40–60% level morphs depicting 40% of happiness and 60% of anger, or conversely) in order to increase the perceptual difficulty of faces and to obtain similar levels of difficulty for vision and audition [as faces are classically processed more rapidly than voices (Ellis et al., 1997; Schweinberger et al., 1997; Joassin et al., 2004)], which is needed to obtain a facilitation effect. Participants (20 alcohol-dependent inpatients and 20 paired controls) had to decide as quickly as possible which emotion was displayed in the stimulus (anger or happiness). While alcohol-dependent individuals were not significantly impaired for crossmodal processing in terms of performance, reaction times results first showed that alcohol-dependent participants were globally slower than controls whatever the experimental condition, which is a classical visuo-motor slowing effect associated with alcohol-dependence (e.g., Fein et al., 1990). But the central outcome of this study is that, as illustrated in **Figure 2**, while control participants showed a clear facilitation effect (i.e., significantly shorter reaction times in the audio-visual condition than in the unimodal auditory and visual), alcohol-dependent individuals did not present this effect, as no differences were observed in this group according to the experimental condition. Alcohol-dependence is thus associated with an absence of crossmodal facilitation effect. As the facilitation effect is the behavioral marker of efficient crossmodal

processing, these results show that alcohol-dependence is associated with impaired auditory-visual integration of complex ecological stimuli. These results constitute the first direct evidence of a crossmodal impairment in alcohol-dependence. Nevertheless, as no control neutral condition was used, it cannot be asserted that this deficit is specific for crossmodal emotional processing as it might also be present for non-emotional crossmodal integration. Moreover, this initial study focused on the behavioral description of the specific deficit for crossmodal processing in alcohol-dependence and did not allow us to explore the cerebral correlates of the deficit. Two complementary studies were thus performed to explore the brain impairments related to this audio-visual integration deficit.

## WHAT ARE THE BRAIN CORRELATES LEADING TO THE ALTERATION OF EMOTIONAL CROSSMODAL PROCESSES IN ALCOHOL-DEPENDENCE?

### At the neurophysiological level

The second study (Maurage et al., 2008b) aimed at describing the brain alterations leading to audio-visual integration impairment, by means of event-related potentials (ERP). ERP record the brain's electrical activity during cognitive tasks with a high temporal resolution and allow us to identify the electrophysiological component associated with the onset of a dysfunction, and then to infer the cognitive stage related to this impairment (Rugg and Coles, 1995). ERP have been widely explored in alcohol-dependent participants during the last decades. The initial explorations (see Hansenne, 2006 for a review) repeatedly described a deficit (reduced amplitude and delayed latency) of the P3b, a long-lasting positive parietal deflection functionally associated with the decisional stage of processing (Polich, 2004, 2007). However, other studies described a deficit in earlier visual ERP components, like P100 (Ogura and Miyazato, 1991), N170 or N200 (Kathmann et al., 1996). These deficits for P100 (linked to early visual processing) and more importantly for N170

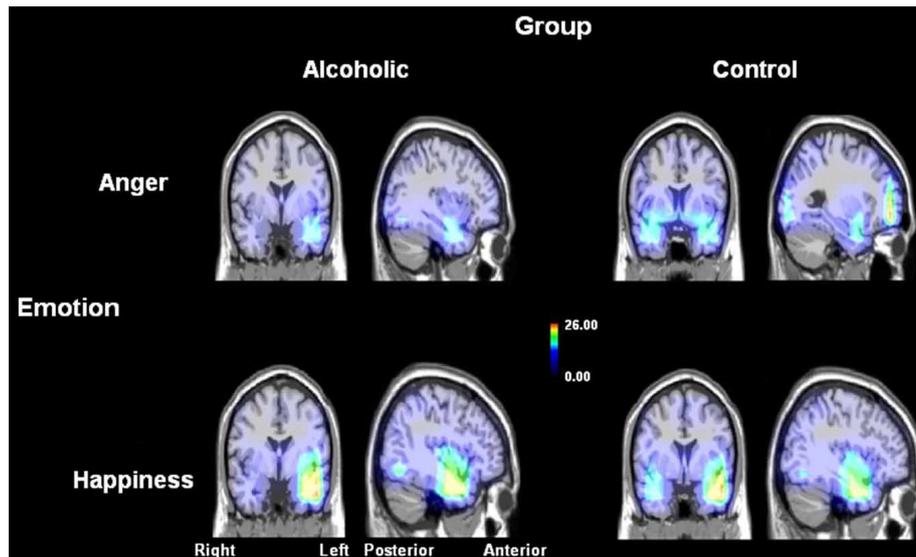


(linked to specific processing of faces) suggest that the impairment in alcohol-dependence begins before the decisional level (P3b), namely at the visuo-spatial level of cognitive processing (Maurage et al., 2007c). Therefore, ERP clearly help to identify the precise stage (e.g., perceptual, attentional or decisional) at which a behavioral deficit originates, and were used here to determine the initial cognitive stage responsible for the crossmodal integration impairment in alcohol-dependence, with the following central question: does the crossmodal deficit start at an early, perceptive stage or only at later processing steps? This second study also explored the potential differential deficit observed for positive (i.e., happiness) vs. negative (i.e., anger) emotions. An emotion detection task was performed by 15 alcohol-dependent participants and 15 paired controls, with visual and auditory stimuli (similar to those presented in the first study) presented separately or simultaneously for 700 ms. Participants had to decide as fast as possible whether the face, voice or face-voice stimulus was an angry, happy or neutral emotional expression. ERP were recorded using a 32 electrode cap to obtain, for each participant, several electrophysiological components of interest (P100, N170–N2, P3b) for each experimental condition (visual, auditory or audio-visual) and each emotion (anger, happiness, neutral). First, alcohol-dependent individuals were less accurate than control participants to identify the emotion depicted in face or voice presented alone, these data being in line with the repeated observation of a deficit in the emotion decoding in alcohol-dependence (e.g., Philippot et al., 1999; Townshend and Duka, 2003; Maurage et al., 2009a). Moreover, alcohol-dependence was associated with slower reaction times, which confirms earlier results (Beatty et al., 1996; Verma et al., 2006) showing a global visuo-motor slowing down in alcohol-dependence, independently of the task or stimuli used. More centrally, the results clearly confirmed the ERP deficits classically observed in alcohol-dependence, as alcohol-dependence was associated with reduced amplitude and delayed latency of the N170/N2 and P3b components for visual and auditory stimulations, thus confirming the ERP alterations repeatedly described in this pathological state (Hansenne, 2006). However, the main result of this study concerned the group differences for the cerebral activations specifically associated with crossmodal processing. Indeed, we used a classical subtraction technique (Teder-Sälejärvi et al., 2002; Joassin et al., 2004) to isolate the electrophysiological activities directly related to visuo-auditory integration, as the auditory (A) and visual (V) unimodal conditions were subtracted from the auditory-visual bimodal condition (AV) using the following formula:  $AV - [A + V]$ . Group comparisons on these specific crossmodal activities showed that alcohol-dependence leads to highly reduced brain activity during integrative processes. Moreover, this deficit is particularly present for anger stimuli, with a strong impairment starting as early as 100 ms after stimulus appearance (while the deficits for happiness and neutral stimuli only appeared after 200–300 ms and were far less marked). Finally, a source location analysis was conducted by means of standardized weighted low-resolution electromagnetic tomography (swLORETA, Palmero-Soler et al., 2007), a technique allowing to accurately reconstruct nearby current sources on the basis of the electroencephalographic data. This analysis showed that the anger crossmodal processing impairment is indexed by a

reduction in frontal activity. These data, shown in **Figure 3**, thus complement the results obtained in the first study by showing (1) that early crossmodal processing of emotional stimulation is impaired in alcohol-dependence, particularly for anger, and (2) that this deficit is associated with a reduction of the electrophysiological activations specifically linked with integrative processes, particularly in frontal areas. However, it should be noted that the electrophysiological deficit found in the present study was not related to reaction times or accuracy alterations, as no specific deficit for emotional stimulations (as compared to neutral ones) was observed among alcohol-dependent individuals at behavioral level. This absence of specific emotional deficit at the behavioral level might be partly explained by the very low sensitivity of the task, which was very easy (leading to a ceiling effect in performance, with more than 90% of correct answers). As this combination between relatively preserved emotional processing at the behavioral level and strongly impaired emotional processing at the cerebral level has been repeatedly observed in electrophysiological and neuroimaging studies in alcohol-dependence (e.g., Maurage et al., 2013b), a potential complementary explanation is that alcohol-dependent individuals might develop alternative strategies to compensate for their deficit. For example, in the present study, they might focus on one sensory modality to compensate for their crossmodal deficit, which might partly hide their emotional decoding deficit. In line with this, it can not be totally excluded that part of the electrophysiological alterations observed here in alcohol-dependent individuals as compared to controls might be due to the use of these alternative strategies. This proposal of the use of alternative strategies to compensate for a deficit in easy tasks should nevertheless be confirmed by specific studies varying the difficulty of the task and the possibility to use these alternative strategies. Moreover, due to their low spatial resolution, ERP are not able to precisely localize the brain areas involved in this integration deficit. Therefore, these results had to be confirmed and complemented by the use of neuroanatomical techniques, which was the central objective of the third study.

#### ***At the anatomical level***

This third study (Maurage et al., 2013b) aimed at precisely locating the cerebral regions responsible for impaired crossmodal processes in alcohol-dependence, by means of functional magnetic resonance imaging (fMRI). On the one hand, alcohol-dependence is known to be associated with major cerebral consequences, particularly in white matter, limbic, temporal and frontal areas. On the other hand, the emotional impairments presented by alcohol-dependent individuals are also well documented, particularly for the decoding of visual or auditory stimulation. Nevertheless, these cerebral and emotional alterations have traditionally been explored separately, and very little is known about the cerebral correlates of emotional impairments in alcohol-dependence. To our knowledge, only a few studies have specifically focused on this topic, comparing the brain activations of recently detoxified alcohol-dependent participants with that of controls during the presentation of emotional scenes (Heinz et al., 2007) or emotional facial expressions (Salloum et al., 2007; Marinkovic et al., 2009). These results show that



**FIGURE 3 | Source reconstruction analysis (using swLORETA software) of the cerebral generators among alcohol-dependent participants (left) and controls (right), for anger (above) and happiness (below)**

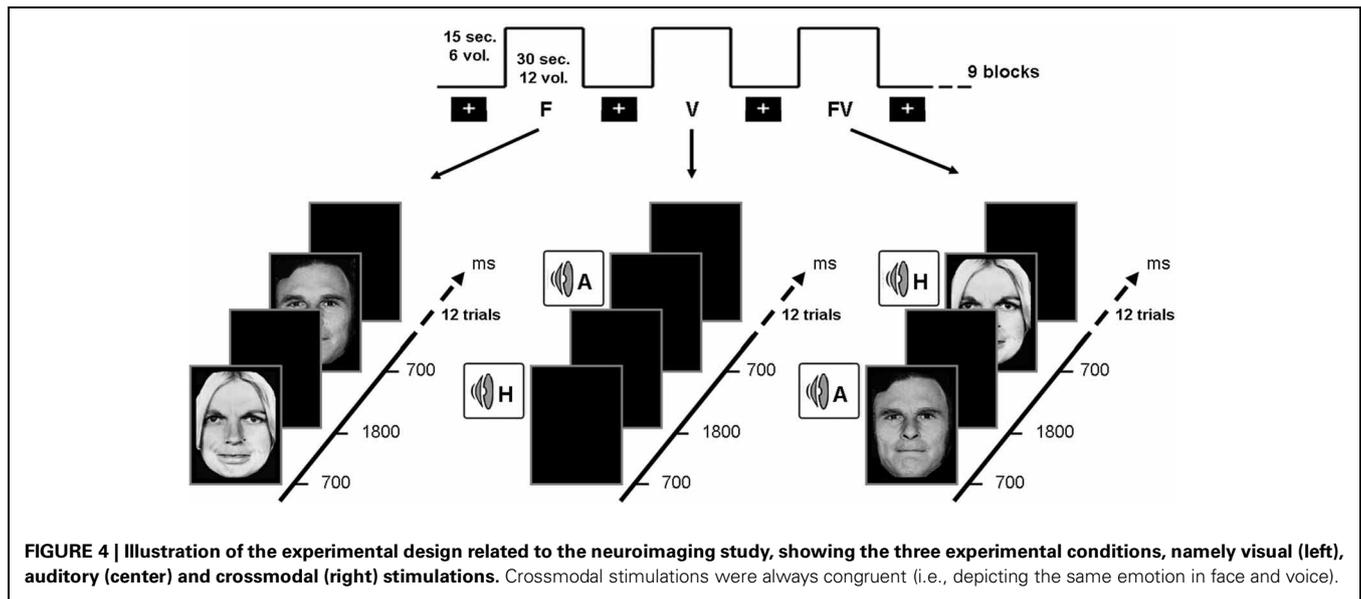
**electrophysiological subtraction  $[AV - (A + V)]$  waves.** Alcohol-dependent individuals exhibited a highly reduced frontal activation as compared to controls for anger stimulations. Adapted from Maurage et al. (2008b).

alcohol-dependence is associated with reduced activity in a wide range of brain areas during the processing of emotional stimuli, encompassing frontal, parietal and temporal regions which are not specifically involved in emotion processing. More precisely, the most statistically significant activity reductions were shown in regions playing a role in emotional processing, particularly the inferior frontal gyrus, anterior cingulate cortex and limbic structures (particularly the amygdala and hippocampus). A more recent study (Schulte et al., 2010) also suggested that alcohol-dependence is associated with white matter abnormalities, thus leading to disconnections between brain areas, and mainly between cortical and limbic structures. As the cortico-limbic connections are central for the processing and interpretation of emotional signals, this white matter deficit could play a major role in the affective disorders observed in alcohol-dependence. Nevertheless, these studies were exclusively based on the presentation of visual emotional stimuli, and the brain correlates of auditory or audio-visual emotional processing remains unexplored.

#### **ALCOHOL-DEPENDENCE LEADS TO SERIOUS CROSSMODAL ALTERATIONS**

On the basis of the two studies presented above and of earlier results suggesting that brain areas dedicated to emotional processing are impaired or disconnected in alcohol-dependence, an fMRI study was conducted to explore the brain correlates of crossmodal emotional processing among alcohol-dependent participants. More precisely, an emotion detection task (illustrated in **Figure 4**) was administered to 12 alcohol-dependent participants and 12 paired controls while their brain activity was recorded using fMRI. The stimuli and task were identical to those presented in the first study, with a binary emotional decision

(anger-happiness) on unimodal (morphed face or voice) or crossmodal (morphed face and voice presented simultaneously) stimulation. Brain activations during unimodal and crossmodal conditions were first computed (by subtracting the activations observed during a rest period without stimulation from those activations), and then the classical  $AV - [A + V]$  comparison was performed to isolate regions specifically involved in the integration of emotional faces and voices in both groups. It should first be underlined that alcohol-dependent individuals, while showing reduced unimodal activations (particularly in the middle frontal gyrus) presented a globally preserved pattern of cerebral activations during the separate processing of faces and voices. However, the central result of this study concerned crossmodal activations and the activations of the specific brain areas related to the integration of audio-visual stimulation. In the control group, the  $AV - [A + V]$  subtraction distinguished two categories of activations: on the one hand, several activations were found in unimodal regions (i.e., superior temporal gyrus for voices and fusiform gyrus for faces), showing that crossmodal stimulations provoke an enhanced activation in cerebral regions specialized in visual or auditory processing, which has been repeatedly observed among healthy participants (e.g., Calvert et al., 1999; Ghazanfar et al., 2005). On the other hand and more importantly, specific multimodal regions were revealed by the subtraction, namely middle frontal gyrus, superior parietal lobule and superior parietal gyrus. This is in line with earlier studies (e.g., Joassin et al., 2011a,b) showing that these brain regions are specifically activated in crossmodal conditions, as they receive multiple inputs from modality-specific regions and integrate them into a unitary and coherent representation of the environment (Rämä and Courtney, 2005; Bernstein et al., 2008). In the alcohol-dependent group however, the only significant activations for crossmodal



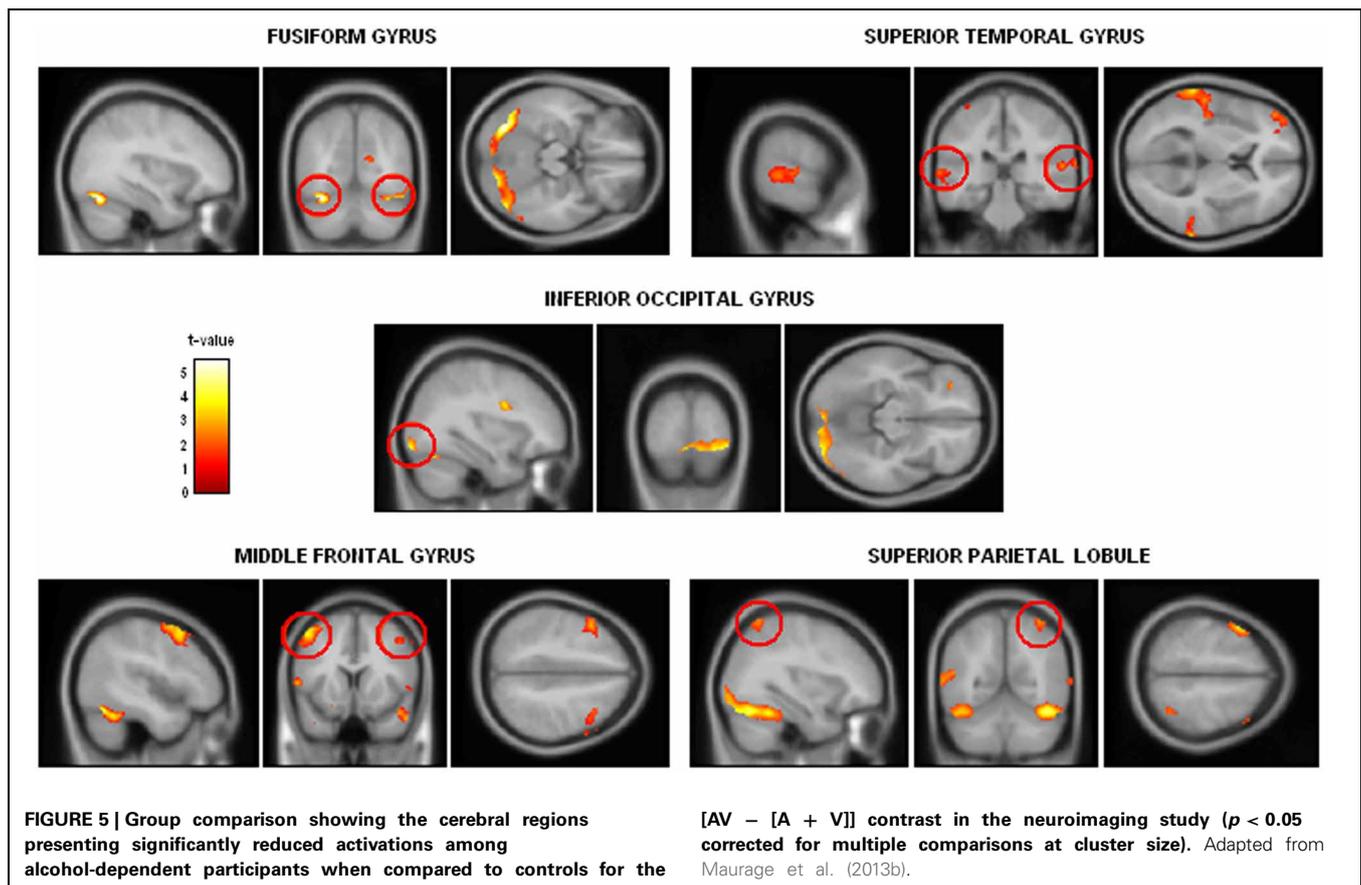
stimulations were found in the unimodal regions cited above (mainly in the auditory regions), with a total absence of activations in the specific crossmodal areas. As further shown in the group comparison for crossmodal activities (Figure 5), it thus appears that alcohol-dependence is associated with a large and specific crossmodal deficit, indexed here by a lack of activation in the regions normally dedicated to the integration of inputs coming from different sensory modalities: alcohol-dependence is therefore associated with serious dysfunctions of the activation and connectivity between the cerebral regions involved in the multimodal perception of the social environment. Finally, psycho-physiological interactions (PPI) analyses allowed us to determine the functional connectivity between unimodal and crossmodal areas. Control participants presented a coherent connectivity pattern with on the one hand increased connectivity within unimodal regions (bilateral fusiform and superior temporal gyri), which confirms the enhanced unimodal connections in crossmodality (Kriegstein and Giraud, 2004), and on the other hand increased connectivity between unimodal and crossmodal areas (inferior occipital gyrus, middle frontal gyrus, superior parietal lobule), underlining the efficient functioning of the crossmodal cerebral network. Conversely, alcohol-dependent individuals did not present this pattern as unimodal areas were partially inter-connected but were not connected with crossmodal ones. As compared to controls, alcohol-dependent individuals showed highly reduced connectivity between unimodal and crossmodal areas, particularly with the middle frontal gyrus. These functional connectivity results thus suggest that the crossmodal deficit in alcohol-dependence could be partly due to disrupted connectivity within the crossmodal network, reducing the connections between unimodal and crossmodal areas. These data are preliminary and will have to be confirmed in future studies using larger groups and alternative experimental designs (Goebel and van Atteveldt, 2009; Love et al., 2011), notably by including neutral stimuli to explore the emotional specificity of

these crossmodal alterations. However, they reinforce our earlier behavioral and electrophysiological results showing an emotional crossmodal processing deficit in alcohol-dependence, and offer the first description of the specific cerebral correlates of this impairment.

## FUNDAMENTAL AND CLINICAL IMPLICATIONS

The three studies presented above present a coherent pattern of results by describing similar specific audio-visual integration impairments in alcohol-dependence. Moreover, the use of different experimental methods and techniques provide complementary views of this impairment from behavioral, electrophysiological and neuroimaging data. Several implications can thus already be outlined on the basis of these results at experimental, fundamental and clinical levels, in order to lay the foundations for potential therapeutic interventions and future experimental investigation of these integrative processes.

First at the experimental level, the observation that the emotional decoding deficit in alcohol-dependence, widely described for unimodal stimulation, is increased in crossmodal stimulation, has shed new light on these earlier results and could influence future studies in the domain. Indeed, as crossmodal situations are omnipresent, our results suggest that earlier studies based on unimodal stimulation (and often using basic stimuli) underestimated the deficits in alcohol-dependence. This crossmodal impairment could also explain the hiatus between the relatively mild deficit frequently observed when presenting unimodal stimuli in experimental situations among alcohol-dependent subjects (e.g., Oscar-Berman et al., 1990; Beatty et al., 2000; Uekermann et al., 2005) and the obvious impairments observed in ecological situations, and notably in clinical observations. The present results could thus lead to a re-evaluation of earlier studies using unimodal stimuli, which probably underestimated the deficit present in real life situations. These results should also encourage future studies to use crossmodal stimulation in order to



correctly evaluate various cognitive and emotional deficits in the processing of social stimuli. More generally, as emotional contexts in everyday life are most often characterized by the simultaneous perception of stimulations from different sensory modalities, our results argue for the development of more ecologically valid experimental designs, for example by means of video clips or virtual reality paradigms. Much progress has already been made in this direction for the evaluation of crossmodal processing among healthy participants (e.g., Vatakis and Spence, 2006; Barkhuysen et al., 2010; Petrini et al., 2010, 2011), and several studies have already used emotional crossmodal movies to explore face-voice integration (e.g., Kreifelts et al., 2010; Ethofer et al., 2013) but it has not been applied to clinical populations yet.

Second, from a more theoretical point of view, the development of experimental work on crossmodal processing in alcohol-dependence and other psychiatric states, could complement the results obtained among healthy participants and help to further renew the knowledge of crossmodal integration in general. Indeed, the exploration of impaired cognitive functions among clinical populations has traditionally been used, in neurology and neuropsychology, in order to add to those observations made among healthy individuals and to give a more exhaustive description of normal functioning. For instance, studies conducted among patients with cerebral lesions

provided the description of double dissociations in memory or attentional systems, thus refining the theoretical models proposed for these systems (e.g., Listerud et al., 2009; Cohn et al., 2010; Barbeau et al., 2011). Specifically for the present topic, our fMRI results, showing that the crossmodal integration impairment in alcohol-dependence is associated with reduced activity in middle frontal gyrus, superior parietal lobule and superior parietal gyrus, confirms that these regions are necessary for a correct integration between faces and voices and thus reinforce the results obtained among healthy participants. This first step underlines the fact that exploring impaired crossmodal processing can offer promising perspectives, notably a better understanding of normal integration functioning.

Third, at the clinical level, the present results clearly confirm earlier data suggesting that emotional impairments (i.e., impaired ability to regulate one's own emotions or to correctly interpret others' emotional states) are a critical deficit in alcohol-dependence. The crossmodal paradigms used highlight that the impairments are more intense when experimental designs are closer to real life emotional situations. It is therefore obvious that emotional perturbations are involved in alcohol-dependence and should thus be considered in clinical settings, notably because they significantly contribute to relapse after detoxification. Indeed, it has been shown by self-report

questionnaires that more than 40% of alcohol-dependent patients identified the presence of emotional disturbances (e.g., persistent negative emotions, depression, anxiety) as being the most important causal factor explaining their relapse (Zywiak et al., 2003). Surprisingly, the experimental studies and theoretical models proposing therapeutic programs for alcohol-dependent individuals have up to now mainly focused on cognitive and behavioral aspects (e.g., development of coping strategies, motivation to change, DiClemente et al., 2003), and the emotional variables have been neglected. However, more recent models of addictions, and particularly the dual-process models (Wiers et al., 2007, 2013; Noël et al., 2010) offered a central position to emotional variables by considering that the addictive states are mostly due to an imbalance between “affective” (involved in automatic-impulsive behaviors) and “reflective” (involved in controlled-deliberate behaviors) systems. The present data, together with earlier results describing emotional alterations among alcohol-dependent individuals, further underline the importance of this affective system and encourage the development of therapeutic approaches focused on the rehabilitation of emotional abilities. Therapeutic programs have recently been developed to improve emotional decoding abilities among clinical populations by specifically training facial expression decoding (e.g., Face Tales Program, Philippot and Power, 2010). Applying them to alcohol-dependence, in complement with the classical rehabilitation programs including psychiatric and psychological therapy, might reduce the relapse rate after detoxification. Future development of these therapeutic proposals should include not only visual emotional stimuli, but also auditory and crossmodal stimulation, in order to develop more realistic emotion decoding rehabilitation programs. More globally, therapeutic interventions could also be improved through communication re-education programs in alcohol-dependence, focusing on crossmodal processing of the expression and identification of emotions in social settings.

Finally, in line with this clinical perspective, the specific crossmodal deficit for anger stimuli described in our electrophysiological study makes particular sense at the therapeutic level. Indeed, many clinical studies (e.g., Bartek et al., 1999; Karno and Longabaugh, 2005) have underlined that alcohol-dependent individuals have considerable difficulties to manage their anger, leading to aggressive behaviors and impulsivity in interpersonal relations. While the links between the regulation one's own anger and the ability to decode the anger expressed by others have not been directly explored yet, these two capacities might be simultaneously impaired in alcohol-dependence and sum up to increase interpersonal problems. However, although some studies have suggested that alcohol-dependence is associated with a relatively specific deficit in anger emotional facial expression decoding (e.g., Philippot et al., 1999), other studies have not replicated these results (Foisy, 2005; Uekermann et al., 2005) and this deficit has not been described for other stimuli (notably auditory prosody). This discrepancy between an obvious clinical deficit and contrasting experimental results could be explained by the fact that previous studies used only unimodal stimuli (mainly emotional facial expressions). These stimuli are artificial because in everyday social

interactions, multimodal stimuli, and mainly auditory-visual stimuli, are much more common. Using more ecologic stimuli, our study established, at the electrophysiological level, the specific crossmodal deficit for anger in alcohol-dependence that has been repeatedly observed at clinical level. The development of future therapeutic programs should thus particularly emphasize the need to take into account this particular deficit for anger expression and decoding among alcohol-dependent individuals.

## PERSPECTIVES FOR FUTURE RESEARCH

As outlined above, our studies have to be considered as a very initial exploratory step in the examination of emotional crossmodal processing among clinical populations. Indeed, these studies focused on a specific clinical population and used a small subset of the possible emotional stimuli and sensory modalities. Nevertheless, when combined with the few previous data sets obtained among other psychiatric populations, these results constitute a reliable and promising basis for the development of an ambitious research program aiming at determining the behavioral and cerebral correlates of impaired crossmodal integration in psychiatry, and finally leading to strong fundamental propositions as well as clinical applications. Before ending this paper, four main directions that should be developed in future research will be proposed, each focusing on the extension of previous results and proposing a diversification of the emotional stimuli used, sensory modalities included and psychiatric populations explored.

## PROMOTING THE USE OF MORE EMOTIONAL EXPRESSIONS

A main limitation of the results presented above is that they only considered a very low number of emotional states (i.e., happy, angry, neutral). A central direction for future research will be to diversify the emotional stimulation used to determine the potential differential deficits associated with different emotional states in alcohol-dependence. It can indeed be hypothesized that alcohol-dependent individuals' emotional crossmodal deficit will vary according to stimulus valence. Our ERP results suggested a specific deficit for anger as compared to happy and neutral stimuli. This specific deficit makes sense at the clinical level and could lead to the development of innovative therapeutic programs. Nevertheless, it is not clear whether this impairment is really limited to anger or is more general, as it could for example be present for every negative emotion. It is thus necessary to develop crossmodal experimental paradigms that explore a broader set of emotions, and particularly of negative ones (e.g., disgust, fear, sadness) to confirm our results and separate the hypotheses of an anger-specific deficit vs. a general deficit for negative emotions. It has also been suggested (e.g., Philippot et al., 1999; Maurage et al., 2008c) that alcohol-dependence could be associated with a particular deficit for decoding and correctly reacting to emotional states which have a high interpersonal value, and particularly which are associated with a social evaluation aspect or moral judgment (e.g., anger, disgust, contempt, e.g., Hutcherson and Gross, 2011), as compared to emotions which express more self-focused feelings (e.g., fear, sadness). Crossmodal paradigms, due to their high ecological

validity, could be very useful in exploring these hypothetical propositions on the differential deficit between emotions presenting high vs. limited social value in alcohol-dependence. A recent study (Lambrecht et al., 2012) has performed this differential exploration between others-oriented and self-oriented emotions among healthy controls, and this experimental approach might be applied in alcohol-dependence to explore the impact of interpersonal value on emotional decoding. More generally, future studies focusing on integration processes in alcohol-dependence should also go beyond the exploration of classical emotion decoding. Indeed, emotional abilities are not limited to this basic emotion decoding as daily life forces us to identify and correctly react to far more various and subtle emotional signals. More complex affective abilities (e.g., empathy, emotional intelligence) are thus also needed to develop and maintain satisfactory interpersonal relations, and several studies have recently shown that alcohol-dependence is associated with impairments for these abilities (e.g., Martinotti et al., 2009; Maurage et al., 2011a,b). These results, in line with recent studies exploring more complex emotional states among healthy people (e.g., Basile et al., 2011; Wagner et al., 2011), underline the need to go further than conventional emotion labels and to use more subtle and ecological paradigms in order to develop a better understanding of emotional impairments in alcohol-dependence. Nevertheless, these studies were conducted by means of unimodal visual stimuli, and thus have a low ecological validity. The use of crossmodal paradigms exploring complex emotional states and affective abilities will enrich these preliminary results by bringing experimental designs closer to daily situations and thus by offering a more valid description of alcohol related emotional and affective disorders. In this view, it would also be particularly important to use dynamic emotional faces synchronized with dynamic voices, as static materials do not capture the liveliness and true form of the facial expressions that typically occur in day-to-day interactions (Sato et al., 2004), and as the emotion content of these “static” non-canonical stimuli is processed by mental strategies and neural events distinct from their more ecologically relevant dynamic counterparts (Kilts et al., 2003).

### GOING BEYOND AUDITORY AND VISUAL MODALITIES

The crossmodal literature puts a strong emphasis on the integration between visual and auditory modalities. This focus is justified by the fact that vision and audition are by far the most dominant sensory modalities among human beings. Nevertheless, the near total absence of data concerning the other senses, and particularly the “chemical senses” (i.e., olfaction and taste) is surprising, as they play an underestimated but crucial role in the daily life of healthy as well as clinical populations (Schiffman, 1997). Indeed, olfactory and gustatory stimulation can also carry a strong emotional valence (e.g., Winston et al., 2005; Greimel et al., 2006; Shepherd, 2006), and exploring the integration between these emotional stimulations and visual or auditory ones could constitute a promising perspective to develop and renew crossmodal integration knowledge. More specifically, olfaction has been shown to play a crucial role in the development and maintenance of alcohol-dependence (e.g., Kareken et al., 2004; Little

et al., 2005), but olfactory processing has up to now been studied very little in this pathology. Recent studies by our group exploring the olfactory abilities associated with excessive alcohol consumption (Maurage et al., 2011c,d) confirmed earlier results (Rupp et al., 2003, 2004, 2006) showing impaired processing of olfactory stimulations in alcohol-dependence, and gave the first insights concerning the cerebral correlates of this deficit (by means of ERP measures). Interestingly, our results showed that olfactory impairments are highly correlated with executive function deficits, and specifically with confabulation problems. These results suggest that these two abilities could rely on the same brain structures (and particularly on the orbitofrontal cortex), and that olfaction measures could be useful to shed new light on the exploration of executive and emotional impairments in alcohol-dependence. This is in line with recent proposals suggesting that olfactory measures could be a reliable cognitive marker in psychiatric disorders (Turetsky et al., 2009; Rupp, 2010). It thus appears that olfaction research is currently becoming a blooming research field among clinical populations. But once again, all these explorations have up to now been limited to unimodal stimulation while in real life situations, olfactory stimulations most often occur in combination with stimulation coming from other sensory modalities. This is particularly true for emotional contexts, and crossmodal explorations combining several senses (beyond audition and vision) are thus urgently needed to develop this new field of research. Very few studies have explored the crossmodal integration of emotional olfactory stimulation, by focusing on the influence of olfactory cues on facial expression decoding (Leppänen and Hietanen, 2003; Seubert et al., 2010a,b). These preliminary results replicated the classical facilitation effect, thus confirming the presence of genuine olfactory-visual integration among healthy participants. Neuroimaging data have also suggested that, while some brain areas (e.g., middle frontal gyrus) could be activated for every crossmodal interaction, independent of the sensory modalities engaged, other structures (mostly the anterior insula) could be specialized for olfactory-visual integration (e.g., Gottfried and Dolan, 2003; Small, 2004). Finally, they showed that schizophrenic patients present an impairment of this olfactory-visual integration, particularly for negative emotional stimuli, which suggests that crossmodal impairments among psychiatric populations could be independent of the sensory modalities involved. On the basis of these innovative explorations, future studies should thus explore, among healthy as well as clinical populations, the correlates of the crossmodal integration between the “chemical senses” and vision or audition. A more ambitious aim could be to go one step further toward ecological validity, by combining more than two sensorial modalities. Indeed, while our emotional experience is frequently based on the simultaneous perception of several sensory modalities, only bimodal stimulation paradigms have been proposed up to now. The recent technical advancements, and notably the growth of virtual reality, could lead to the development of experimental designs stimulating all the senses and thus open new perspectives for crossmodal processing explorations.

## APPLYING CROSSMODAL PARADIGMS TO OTHER PSYCHIATRIC POPULATIONS

The crossmodal studies presented in this chapter exclusively explored alcohol-dependence, and more specifically recently detoxified alcohol-dependent individuals. This population is of course only a specific part of the people presenting alcohol related problems, and more globally of the psychiatric patients. It thus appears important to underline the potential extension of these studies to other populations, in the field of alcohol abuse and dependence, but also in other psychiatric states, with the long term objective of developing a sound and integrative approach of crossmodal processing in clinical populations. Concerning alcohol related problems the literature on cerebral, cognitive and emotional impairments associated with alcohol consumption has classically been focused on long-term alcohol-dependence (namely on the exploration of the impairments due to long term chronic excessive alcohol consumption). Nevertheless, a new field of research has risen during the last decades, aiming at exploring the roots of alcohol addiction, namely the appearance and chronification of the deficits during the development of alcohol-dependence. On the one hand, many studies have been conducted among populations at high risk of becoming alcohol-dependent, mainly among children of alcohol-dependent individuals (Van der Stelt et al., 1994; Lieberman, 2000; Porjesz et al., 2005). These studies have suggested that several cerebral and cognitive impairments could be present before the development of alcohol-dependence, and thus be a causal factor rather than a consequence of excessive alcohol consumption. Our intention is not to go into the details of this important literature, but to underline that these explorations were once again exclusively based on unimodal stimulation. Crossmodal studies among children of alcohol-dependent individuals (notably for emotional abilities, which are still unexplored in this population) could thus give a more ecological and valid exploration of the deficits that are present before the development of alcohol-dependence. On the other hand and more recently, several projects have been conducted concerning the consequences of binge drinking (i.e., the excessive but episodic alcohol consumption, typically observed among adolescents and young adults and considered to be an “entrance door” toward alcohol-dependence, e.g., McCarty et al., 2004; Enoch, 2006). Recent studies have shown that binge drinking leads to cognitive effects (e.g., Giancola, 2002; Townshend and Duka, 2005; Zeigler et al., 2005), and we recently extended these results by suggesting that binge drinking habits rapidly lead to impaired processing of emotional auditory stimulation, and that this alcohol consumption pattern is particularly deleterious for brain functioning (Maurage et al., 2009b, 2012b). Nevertheless, it is still unknown whether these deficits are modified or not when several stimulations are presented together, and crossmodal studies would thus help to extend and clarify these preliminary results. Concerning the exploration of emotional crossmodal processing in psychiatry, it is surprising to notice that very few studies have been conducted among these clinical populations. Many projects have been proposed in recent years in order to explore the visual or auditory decoding of emotional stimulations among a wide variety of psychiatric states, like

depression, autism, anxiety, anorexia nervosa and drug addiction (e.g., Mejias et al., 2005; Mendlewicz et al., 2005; Rossignol et al., 2005; Bhatara et al., 2010), but emotional crossmodal paradigms have only been used in a very limited number of these pathological states. Several studies (De Gelder et al., 2005; De Jong et al., 2009; Pearl et al., 2009; Szyck et al., 2009) have explored the integration of emotional stimulation among schizophrenic patients and consistently described emotional crossmodal deficits in this psychiatric state, notably indexed by reduced audio-visual integration ability and by a vision-audition imbalance (i.e., a dominance of the visual stimulation on the auditory one reducing crossmodal performance). These results, together with those obtained in alcohol-dependence, show that crossmodal processing impairments constitute a crucial aspect of psychiatric states, and should thus encourage the development of emotional crossmodal research among other psychiatric states. This is particularly true among populations which are known to present unimodal emotional decoding deficits, in order to answer the following central question: How does crossmodal integration occur when unimodal outputs are impaired? In other words, do some psychiatric populations manage to compensate their deficit in the processing of unimodal emotional stimuli by taking profit of the simultaneous presentation of two stimulations, or are all psychiatric states associated with increased processing impairments in crossmodal situations, as it has been observed in alcohol-dependence and schizophrenia?

## USING CROSSMODAL TASKS IN REAL CLINICAL SETTINGS

The final objective of this research area is also to offer new ways to manage clinical interventions in real clinical settings. Some preliminary data with clear potential clinical application have already been gathered (Campanella et al., 2010, 2012). Indeed, there is a growing literature base demonstrating that, throughout the information processing stream, a number of early and late neuroelectric features appear to be anomalous in various psychiatric populations. In all of these studies, the primary and most commonly reported finding has been P300 abnormalities (see Hansenne, 2006 for a review). P300 alterations have been highly important in the assessment of the pathophysiological mechanisms responsible for psychopathological states, as it is commonly acknowledged that a reduction of P300 amplitude is: (1) a state marker of depression, i.e., a biological marker that is altered during the disease but that stabilizes after clinical remission (Karaaslan et al., 2003); (2) a trait marker of schizophrenia, i.e., a biological parameter that is changed during and after the disease (Mathalon et al., 2000); and (3) a vulnerability marker of alcoholism, i.e., a biological variable that is altered before the emergence of the disease (high-risk children of alcoholic parents) (e.g., Hill et al., 1999). Such markers, if present, could be used to aid diagnosis, as prognostic elements, or to assist in choosing the most appropriate treatment for psychiatric disorders. They can also enhance our knowledge about the nature and the extent of cognitive damage, and offer deeper theoretical insights into both the aetiology and pathophysiology of the illness. Overall, such markers can improve early detection of illness, and, as such, facilitate more effective and

targeted interventions (see Van der Stelt and Belger, 2007 for a review).

Nevertheless, the clinical sensitivity of the P300 has been hampered by the fact that its parameters (amplitude, latency) are diagnostically unspecific and not reliable enough to be useful for individual patients (Pogarell et al., 2007). Therefore, a current and important challenge for neurophysiologists is to discover novel and appropriate procedures to enhance the applicability and sensitivity of the P3b component in clinical settings. In this view, trying to increase P300 sensitivity, Campanella et al. (2010, 2012) have recently proposed a new variant of the classical oddball procedure using bimodal (visual-auditory) stimulations. The main idea was that, as bimodal face-voice associations require more “complex” interactions between unimodal (sensory) and multimodal (integrative) processes which work in parallel and influence each other, using a bimodal oddball design might enhance the sensitivity of the procedure by increasing the observable P300 differences evoked by “only” unimodal processes. To test this hypothesis, Campanella et al. (2010, 2012) compared two groups of participants: one group was healthy, and the other consisted of people displaying anxious and depressive tendencies. Both groups were submitted to unimodal (visual; auditory) and bimodal oddball tasks. Main results suggested that when the two groups of subjects differ in their subclinical level of comorbid anxiety and depression, unimodal visual and auditory oddball tasks may not allow us to detect this difference using P300 amplitude modulations, but a crossmodal task has greater power to detect even subclinical symptoms. Obviously, these results are preliminary and should be replicated on clinical populations: such experiments are currently underway in our laboratory, with the main aim to be able to index several steps of clinical severity in a pathological state. Moreover, it clearly appears that ERP data should not focus “only” on the P300 component: indeed, for instance, data have shown that (1) P300 deficits are correlated with previous “early” ERP alterations (e.g., Maurage et al., 2007c); and (2) the combination of different ERP components may be helpful to discriminate between different groups of patients. Price et al. (2006) compared and contrasted four electrophysiological endophenotypes -mismatch negativity, P50, P300, and antisaccades-, and analyzed their covariance on the basis of a single cohort of schizophrenic patients, family members and controls, tested with all paradigms. Data showed that the use of an electrophysiological battery has provided novel information on the characteristics of these features in schizophrenia and family member groups. In particular, it has highlighted the heterogeneity of electrophysiological features within these groups and how a combination of features could serve to minimize the impact of such heterogeneity. This outlines the urgent need in further studies to develop multisite guidelines to record a battery of electrophysiological measures that may be compared and used across studies.

## CONCLUSION

The exploration of crossmodal processing among healthy controls has now become an extensive research field: behavioral as well as cerebral correlates of the integration processes between

sensory modalities have been precisely explored among animal and human populations, leading to comprehensive models on this topic. Nevertheless, this maturity of the knowledge concerning “normal” crossmodal processes is in total contradiction with what can be observed in clinical states. Indeed, as outlined above, very little has been done up to now to attempt to understand how these crossmodal processes are impaired among neurological and psychiatric populations, and this astonishing lack of interest has had a detrimental effect on the advances that can be made in this topic.

The main aim of the present paper was thus to underline the urgent need to explore the crossmodal integration abilities among these populations, as progressing in this direction could lead to central implications at clinical and fundamental levels. First, for clinical aspects, using crossmodal designs among clinical populations would lead to a better understanding of the impairments presented by inpatients in real life situations (and notably in emotional contexts). This would provide a more ecological exploration of the cognitive, cerebral and affective deficits in these populations, thus complementing and clarifying earlier results. This could also lead to the development of new therapeutic interventions, using crossmodal clinical settings to rehabilitate impaired abilities (e.g., by means of virtual reality). Second, for fundamental research, while the data obtained among clinical populations have traditionally constituted a strong method to improve the understanding of normal abilities in neuropsychology and neuroscience (with the well-known proposition that exploring an impaired system helps to understand its healthy functioning), this approach has received very little attention in crossmodal processing research. Developing the exploration of integration abilities among clinical populations could shed a new light on the several questions that are still unresolved in this research field.

By describing our research focusing on emotional crossmodal integration in alcohol-dependence, we have only presented here what can be considered as a modest first step toward a real and ambitious research program allowing to precisely describing the crossmodal processing abilities among psychiatric populations. We indeed believe that our work, together with the few studies conducted in schizophrenia, constitute seminal results that should be developed in the future. More specifically, studies to come should extend this exploration of crossmodal processing in at least four main ways, by diversifying the stimuli (i.e., using a wider range of emotional but also non emotional stimuli), the sensory modalities (particularly by including the “chemical senses” in the crossmodal designs), the populations explored (i.e., studying the crossmodal processes among other populations with substance-abuse, but also among other psychiatric and neurologic states), and by adapting experimental paradigms to real needs of current clinical settings. These proposals for future studies are just illustrations of the many prospects offered by this largely unexplored field. In short, everything is still to be done concerning crossmodal processing in psychiatric populations, and our hope is thus that the preliminary data described in the present chapter will open the door to fresh, diverse and complementary studies.

## FUNDING

The Authors are funded by the Belgian Fund for Scientific Research (F. N. R. S., Belgium), but this fund did not exert any editorial direction or censorship on any part of this article.

## REFERENCES

Alais, D., and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.* 14, 257–262.

Amedi, A., von Kriegstein, K., van Atteveldt, N. M., Beauchamp, M. S., and Naumer, M. J. (2005). Functional imaging of human cross-modal identification and object recognition. *Exp. Brain Res.* 166, 559–571. doi: 10.1007/s00221-005-2396-5

Amenta, S., Noël, X., Verbanck, P., and Campanella, S. (2013). Decoding of emotional components in complex communicative situations (irony) and its relation to empathic abilities in male chronic alcoholics: an issue for treatment. *Alcohol. Clin. Exp. Res.* 37, 339–347. doi: 10.1111/j.1530-0277.2012.01909.x

Barbeau, E. J., Pariente, J., Felician, O., and Puel, M. (2011). Visual recognition memory: a double anatomo-functional dissociation. *Hippocampus* 21, 929–934.

Barkhuysen, P., Krahmer, E., and Swerts, M. (2010). Crossmodal and incremental perception of audiovisual cues to emotional speech. *Lang. Speech* 53, 3–30. doi: 10.1177/0023830909348993

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and Plumb, I. (2001). The “Reading the mind in the eyes” test revised version: a study with normal adults, and adults with *Asperger syndrome* or high-functioning autism. *J. Child Psychol. Psychiatry* 42, 241–251. doi: 10.1111/1469-7610.00715

Barone, P. (2010). Is the primary visual cortex multisensory? Comment on “Crossmodal influences on visual perception” by Prof. Ladan Shams. *Phys. Life Rev.* 7, 291–292. doi: 10.1016/j.plrev.2010.07.002

Bartek, J. K., Lindeman, M., and Hawks, J. H. (1999). Clinical validation of characteristics of the alcoholic family. *Nurs. Diagn.* 10, 158–168. doi: 10.1111/j.1744-618X.1999.tb00047.x

Basile, B., Mancini, F., Macaluso, E., Caltagirone, C., Frackowiak, R. S., and Bozzali, M. (2011). Deontological and altruistic guilt: evidence for distinct neurobiological substrates. *Hum. Brain*

*Mapp.* 32, 229–239. doi: 10.1002/hbm.21009

Beatty, W. W., Hames, K. A., Blanco, C. R., Nixon, S. J., and Tivis, L. J. (1996). Visuospatial perception construction and memory in alcoholism. *J. Stud. Alcohol* 57, 136–143.

Beatty, W. W., Tivis, R., Stott, D., Nixon, S. J., and Parsons, O. A. (2000). Neuropsychological deficits in sober alcoholics: influences of chronicity and recent alcohol consumption. *Alcohol Clin. Exp. Res.* 24, 149–154. doi: 10.1111/j.1530-0277.2000.tb04584.x

Bechara, A., Dolan, S., Denburg, N., Hindes, A., Anderson, S. W., and Nathan, P. E. (2001). Decision making deficits, linked to a dysfunction of the ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers. *Neuropsychologia* 39, 376–389. doi: 10.1016/S0028-3932(00)00136-6

Beck, A. T., and Steer, R. A. (1987). *Beck Depression Inventory Manual*. San Antonio, TX: Psychological Corporation.

Bernstein, L. E., Auer, Jr. E. T., Wagner, M., and Ponton, C. W. (2008). Spatiotemporal dynamics of audiovisual speech processing. *Neuroimage* 39, 423–435. doi: 10.1016/j.neuroimage.2007.08.035

Bhatara, A., Quintin, E. M., Levy, B., Bellugi, U., Fombonne, E., and Levitin, D. J. (2010). Perception of emotion in musical performance in adolescents with autism spectrum disorders. *Autism Res.* 3, 214–225. doi: 10.1002/aur.147

Blusewicz, M. J., Dustman, R. E., Schenkenberg, T., and Beck, E. C. (1977). Neuropsychological correlates of chronic alcoholism and aging. *J. Nerv. Ment. Dis.* 165, 348–355. doi: 10.1097/00005053-197711000-00006

Brooks, P. J. (2000). Brain atrophy and neuronal loss in alcoholism: a role for DNA damage? *Neurochem. Int.* 37, 403–412. doi: 10.1016/S0197-0186(00)00051-6

Calvert, G. A., Brammer, M. J., Bullmore, E. T., Campbell, R., Iversen, S. D., and David, A. S. (1999). Response amplification in sensory-specific cortices during crossmodal binding. *Neuroreport* 10, 2619–2623. doi: 10.1097/00001756-199908200-00033

## ACKNOWLEDGMENTS

Pierre Maurage and Salvatore Campanella are Research Associates at the Belgian Fund for Scientific Research (F. N. R. S., Belgium).

Calvert, G. A., Hansen, P. C., Iversen, S. D., and Brammer, M. J. (2001). Detection of auditory visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *Neuroimage* 14, 427–438. doi: 10.1006/nimg.2001.0812

Campanella, S., and Belin, P. (2007). Integrating face and voice in person perception. *Trends Cogn. Sci.* 11, 535–543. doi: 10.1016/j.tics.2007.10.001

Campanella, S., Bruyer, R., Froidbise, S., Rossignol, M., Joassin, F., Kornreich, C., et al. (2010). Is two better than one? A cross-modal oddball paradigm reveals greater sensitivity of the P300 to emotional face-voice associations. *Clin. Neurophysiol.* 121, 1855–1862. doi: 10.1016/j.clinph.2010.04.004

Campanella, S., Delle-Vigne, D., Kornreich, C., and Verbanck, P. (2012). Greater sensitivity of the P300 component to bimodal stimulation in an event-related potentials oddball task. *Clin. Neurophysiol.* 123, 937–946. doi: 10.1016/j.clinph.2011.10.041

Carton, J. S., Kessler, E. A., and Pape, C. L. (1999). Nonverbal decoding skills and relationship wellbeing in adults. *J. Nonverb. Behav.* 23, 91–100. doi: 10.1023/A:1021339410262

Chanraud, S., Martelli, C., Delain, F., Kostogianni, N., Douaud, G., Aubin, H. J., et al. (2007). Brain morphometry and cognitive performance in detoxified alcohol dependents with preserved psychosocial functioning. *Neuropsychopharmacology* 32, 429–438. doi: 10.1038/sj.npp.1301219

Chen, Y. H., Edgar, J. C., Holroyd, T., Dammers, J., Thonnessen, H., Roberts, T. P., et al. (2010). Neuromagnetic oscillations to emotional faces and prosody. *Eur. J. Neurosci.* 31, 1818–1827. doi: 10.1111/j.1460-9568.2010.07203.x

Clark, U. S., Oscar-Berman, M., Shagrin, B., and Pencina, M. (2007). Alcoholism and judgments of affective stimuli. *Neuropsychology* 21, 346–362. doi: 10.1037/0894-4105.21.3.346

Cohn, M., Moscovitch, M., and Davidson, P. S. (2010). Double dissociation between familiarity and recollection in Parkinson’s disease

as a function of encoding tasks. *Neuropsychologia* 48, 4142–4417. doi: 10.1016/j.neuropsychologia.2010.10.013

Cordovil de Susa Uva, M., de Timary, P., Cortesi, M., Mikolajczak, M., du Roy de Blicquy, P., and Luminet, O. (2010). Moderating effect of emotional intelligence on the role of negative affect in the motivation to drink in alcohol-dependent subjects. *Pers. Individ. Differ.* 48, 16–21. doi: 10.1016/j.paid.2009.08.004

Cowen, M. S., Chen, F., and Lawrence, A. J. (2004). Neuropeptides: implications for alcoholism. *J. Neurochem.* 89, 273–285. doi: 10.1111/j.1471-4159.2004.02394.x

De Bellis, M. D., Narasimhan, A., Thatcher, D. L., Keshavan, M. S., Soloff, P., and Clark, D. B. (2005). Prefrontal cortex, thalamus, and cerebellar volumes in adolescents and young adults with adolescent-onset alcohol use disorders and comorbid mental disorders. *Alcohol. Clin. Exp. Res.* 29, 1590–1600. doi: 10.1097/01.alc.0000179368.87886.76

De Gelder, B., and Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends Cogn. Sci.* 7, 460–467. doi: 10.1016/j.tics.2003.08.014

De Gelder, B., Bocker, K. B., Tuominen, J., Hensen, M., and Vroomen, J. (1999). The combined perception of emotion from voice and face: early interaction revealed by human electric brain responses. *Neurosci. Lett.* 260, 133–136. doi: 10.1016/S0304-3940(98)00963-X

De Gelder, B., Vroomen, J., de Jong, S. J., Masthoff, E. D., Trompenaars, F. J., and Hodiamont, P. (2005). Multisensory integration of emotional faces and voices in schizophrenics. *Schizophr. Res.* 72, 195–203. doi: 10.1016/j.schres.2004.02.013

De Jong, J. J., Hodiamont, P. P., Van den Stock, J., and de Gelder, B. (2009). Audiovisual emotion recognition in schizophrenia: reduced integration of facial and vocal affect. *Schizophr. Res.* 107, 286–293. doi: 10.1016/j.schres.2008.10.001

Delbeuck, X., Collette, F., and Van der Linden, M. (2007). Is Alzheimer’s disease a disconnection syndrome? Evidence from a crossmodal audio-visual illusory experiment.

- Neuropsychologia* 45, 3315–3323. doi: 10.1016/j.neuropsychologia.2007.05.001
- DiClemente, C. C., Jordan, L., Marinilli, A. S., and Nidecker, M. (2003). “Psychotherapy in alcoholism treatment,” in *Handbook of Clinical Alcoholism Treatment*, eds B. A. Johnson, P. Ruiz, and M. Galanter (Baltimore, MD: Lippincott, Williams, and Wilkins), 102–110.
- Dolan, R. J., Morris, J. S., and de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proc. Natl. Acad. Sci. U.S.A.* 98, 10006–10010. doi: 10.1073/pnas.171288598
- Driver, J., and Spence, C. (2000). Multisensory perception: beyond modularity and convergence. *Curr. Biol.* 10, 731–735. doi: 10.1016/S0960-9822(00)00740-5
- Ellis, H. D., Jones, D. M., and Mosdell, N. (1997). Intra- and inter-modal repetition priming of familiar faces and voices. *Br. J. Psychol.* 88, 143–156. doi: 10.1111/j.2044-8295.1997.tb02625.x
- Enoch, M. A. (2006). Genetic and environmental influences on the development of alcoholism: resilience vs. risk. *Ann. N.Y. Acad. Sci.* 1094, 193–201. doi: 10.1196/annals.1376.019
- Ethofer, T., Anders, S., Erb, M., Droll, C., Royen, L., Saur, R., et al. (2006a). Impact of voice on emotional judgment of faces: an event-related fMRI study. *Hum. Brain Mapp.* 27, 707–714. doi: 10.1002/hbm.20212
- Ethofer, T., Pourtois, G., and Wildgruber, D. (2006b). Investigating audiovisual integration of emotional signals in the human brain. *Prog. Brain Res.* 156, 345–361. doi: 10.1016/S0079-6123(06)56019-4
- Ethofer, T., Bretschner, J., Wiethoff, S., Bisch, J., Schlipf, S., Wildgruber, D., et al. (2013). Functional responses and structural connections of cortical areas for processing faces and voices in the superior temporal sulcus. *Neuroimage* 76, 45–56. doi: 10.1016/j.neuroimage.2013.02.064
- Fein, G., Bachman, L., Fischer, S., and Davenport, L. (1990). Cognitive impairments in abstinent alcoholics. *West J. Med.* 152, 531–537.
- Fein, G., Landman, B., Tran, H., McGillivray, S., Finn, P., Barakos, J., et al. (2006). Brain atrophy in long-term abstinent alcoholics who demonstrate impairment on a simulated gambling task. *Neuroimage* 32, 1465–1471. doi: 10.1016/j.neuroimage.2006.06.013
- Feldman, R. S., Philippot, P., and Custrini, R. J. (1991). “Social competence and non-verbal behaviour,” in *Fundamentals of Non-verbal Behavior*, ed R. S. Feldman (New York, NY: Cambridge University Press), 329–350.
- Flannery, B., Fishbein, D., Krupitsky, E., Langevin, D., Verbitskaya, E., Bland, C., et al. (2007). Gender differences in neurocognitive functioning among alcohol-dependent Russian patients. *Alcohol. Clin. Exp. Res.* 31, 745–754. doi: 10.1111/j.1530-0277.2007.00372.x
- Foisy, M. L. (2005). *La Reconnaissance des Expressions Faciales émotionnelles dans l'alcoolisme: Spécification de la Nature et de l'origine des Biais évaluatifs*. Unpublished Ph.D. Thesis, Catholic University of Louvain, Louvain-la-Neuve, Belgique.
- Foisy, M. L., Kornreich, C., Fobe, A., D'Hondt, L., Pelc, I., Hanak, C., et al. (2007a). Impaired emotional facial expression recognition in alcohol dependence: do these deficits persist with midterm abstinence? *Alcohol. Clin. Exp. Res.* 31, 404–410. doi: 10.1111/j.1530-0277.2006.00321.x
- Foisy, M. L., Kornreich, C., Petiau, C., Parez, A., Hanak, C., Verbanck, P., et al. (2007b). Impaired emotional facial expression recognition in alcoholics: are these deficits specific to emotional cues? *Psychiatry Res.* 150, 33–41.
- Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W. L., et al. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Exp. Brain Res.* 203, 381–389. doi: 10.1007/s00221-010-2240-4
- Frigerio, E., Burt, D. M., Montagne, B., Murray, L. K., and Perrett, D. I. (2002). Facial affect perception in alcoholics. *Psych. Res.* 113, 161–171. doi: 10.1016/S0165-1781(02)00244-5
- Friston, K. J., and Frith, C. D. (1995). Schizophrenia: a disconnection syndrome? *Clin. Neurosci.* 3, 89–97.
- Ghazanfar, A. A., Maier, J. X., Hoffman, K. L., and Logothetis, N. K. (2005). Multi-sensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J. Neurosci.* 25, 5004–5012. doi: 10.1523/JNEUROSCI.0799-05.2005
- Giancola, P. R. (2002). Alcohol-related aggression during the college years: theories, risk factors and policy implications. *J. Stud. Alcohol* 14, 129–139.
- Glahn, D. C., Lovallo, W. R., and Fox, P. T. (2007). Reduced amygdala activation in young adults at high risk of alcoholism: studies from the Oklahoma family health patterns project. *Biol. Psychiatry* 61, 1306–1309. doi: 10.1016/j.biopsych.2006.09.041
- Goebel, R., and van Atteveldt, N. (2009). Multisensory functional magnetic resonance imaging: a future perspective. *Exp. Brain Res.* 198, 153–164. doi: 10.1007/s00221-009-1881-7
- Gottfried, J. A., and Dolan, R. J. (2003). The nose smells what the eye sees: crossmodal visual facilitation of human olfactory perception. *Neuron* 39, 375–386. doi: 10.1016/S0896-6273(03)00392-1
- Greimel, E., Macht, M., Krumhuber, E., and Ellgring, H. (2006). Facial and affective reactions to tastes and their modulation by sadness and joy. *Physiol. Behav.* 89, 261–269. doi: 10.1016/j.physbeh.2006.06.002
- Hansenne, M. (2006). Event-related brain potentials in psychopathology: clinical and cognitive perspectives. *Psychol. Belg.* 46, 5–36.
- Harper, C. (2007). The neurotoxicity of alcohol. *Hum. Exp. Toxicol.* 26, 251–257. doi: 10.1177/0960327107070499
- Harper, C., and Matsumoto, I. (2005). Ethanol and brain damage. *Curr. Opin. Pharmacol.* 5, 73–78. doi: 10.1016/j.coph.2004.06.011
- Heinz, A., Wrase, J., Kahnt, T., Beck, A., Bromand, Z., Grüsser, S. M., et al. (2007). Brain activation elicited by affectively positive stimuli is associated with a lower risk of relapse in detoxified alcoholic subjects. *Alcohol. Clin. Exp. Res.* 31, 1138–1147. doi: 10.1111/j.1530-0277.2007.00406.x
- Hill, S. Y., Shen, S., Locke, J., Steinhauer, S. R., Konicky, C., Lowers, L., et al. (1999). Developmental delay in P300 production in children at high risk for developing alcohol-related disorders. *Biol. Psychiatry* 46, 970–981. doi: 10.1016/S0006-3223(99)00032-3
- Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureno, G., and Villasenor, V. S. (1988). Inventory of interpersonal problems: psychometric properties and clinical applications. *J. Consult. Clin. Psychol.* 56, 885–892. doi: 10.1037/0022-006X.56.6.885
- Hutcherson, C. A., and Gross, J. J. (2011). The moral emotions: a social-functional account of anger, disgust, and contempt. *J. Pers. Soc. Psychol.* 100, 719–737. doi: 10.1037/a0022408
- Joassin, F., Pesenti, M., Maurage, P., Verreckett, E., Bruyer, R., and Campanella, S. (2011a). Cross-modal interactions between human faces and voices involved in person recognition. *Cortex* 47, 367–376. doi: 10.1016/j.cortex.2010.03.003
- Joassin, F., Maurage, P., and Campanella, S. (2011b). The neural network sustaining the crossmodal processing of human gender from faces and voices: an fMRI study. *Neuroimage* 54, 1654–1661. doi: 10.1016/j.neuroimage.2010.08.073
- Joassin, F., Maurage, P., Bruyer, R., Crommelinck, M., and Campanella, S. (2004). When audition alters vision: an event-related potential study of the cross-modal interactions between faces and voices. *Neurosci. Lett.* 369, 132–137. doi: 10.1016/j.neulet.2004.07.067
- Karaaslan, F., Gonul, A. S., Oguz, A., Erdinc, E., and Esel, E. (2003). P300 changes in major depressive disorders with and without psychotic features. *J. Affect. Disord.* 73, 283–287. doi: 10.1016/S0165-0327(01)00477-3
- Kareken, D. A., Claus, E. D., Sabri, M., Dziedzic, M., Kosobud, A. E. K., Radnovich, A. J., et al. (2004). Alcohol-related olfactory cues activate the nucleus accumbens and the ventromedial area in high-risk drinkers: preliminary findings. *Alcohol. Clin. Exp. Res.* 28, 550–557. doi: 10.1097/01.ALC.0000122764.60626.AF
- Karno, M. P., and Longabaugh, R. (2005). An examination of how therapist directiveness interacts with patient anger and reactance to predict alcohol use. *J. Stud. Alcohol* 66, 825–832.
- Kathmann, N., Soyka, M., Bickel, R., and Engel, R. R. (1996). ERP changes in alcoholics with and without alcohol psychosis. *Biol. Psychiatry* 39, 873–881. doi: 10.1016/0006-3223(95)00289-8
- Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., and Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *Neuroimage* 18, 156–168. doi: 10.1006/nimg.2002.1323
- Klasen, M., Kenworthy, C. A., Mathiak, K. A., Kircher, T. T., and Mathiak, K. (2012). Supramodal representation of emotions. *J. Neurosci.* 31, 13635–13643. doi: 10.1523/JNEUROSCI.2833-11.2011
- Kornreich, C., Blairy, S., Philippot, P., Dan, B., Foisy, M. L., Hess, U., et al. (2001). Impaired emotional facial expression

- recognition in alcoholism compared with obsessive-compulsive disorder and normal controls. *Psychiatry Res.* 102, 235–248. doi: 10.1016/S0165-1781(01)00261-X
- Kornreich, C., Brevers, D., Canivet, D., Ermer, E., Naranjo, C., Constant, E., et al. (2013). Impaired processing of emotion in music, faces and voices supports a generalized emotional decoding deficit in alcoholism. *Addiction* 108, 80–88. doi: 10.1111/j.1360-0443.2012.03995.x
- Kornreich, C., Philippot, P., Foisy, M. L., Blairy, S., Raynaud, E., Dan, B., et al. (2002). Impaired emotional facial expression recognition is associated with interpersonal problems in alcoholism. *Alcohol Alcohol.* 37, 394–400. doi: 10.1093/alcalc/37.4.394
- Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., and Wildgruber, D. (2007). Audiovisual integration of emotional signals in voice and face: an event-related fMRI study. *Neuroimage* 37, 1445–1456. doi: 10.1016/j.neuroimage.2007.06.020
- Kreifelts, B., Ethofer, T., Huberle, E., Grodd, W., and Wildgruber, D. (2010). Association of trait emotional intelligence and individual fMRI-activation patterns during the perception of social signals from voice and face. *Hum. Brain Mapp.* 31, 979–991. doi: 10.1002/hbm.20913
- Kreifelts, B., Ethofer, T., Shiozawa, T., Grodd, W., and Wildgruber, D. (2009). Cerebral representation of non-verbal emotional perception: fMRI reveals audio-visual integration area between voice- and face-sensitive regions in the superior temporal sulcus. *Neuropsychologia* 47, 3059–3066. doi: 10.1016/j.neuropsychologia.2009.07.001
- Kriegstein, K. V., and Giraud, A. L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *Neuroimage* 22, 948–955. doi: 10.1016/j.neuroimage.2004.02.020
- Krill, J. J., Halliday, G. M., Svoboda, M. D., and Cartwright, H. (1997). The cerebral cortex is damaged in chronic alcoholics. *Neuroscience* 79, 983–998. doi: 10.1016/S0306-4522(97)00083-3
- Kwaky, L. D., Foss-Feig, J. H., Cascio, C. J., Stone, W. L., Wallace, M. T. (2011). Altered auditory and multisensory temporal processing in autism spectrum disorders. *Front. Integr. Neurosci.* 4, 1–11. doi: 10.3389/fnint.2010.00129
- Lalanne, C., and Lorenceau, J. (2004). Crossmodal integration for perception and action. *J. Physiol. Paris* 98, 265–279. doi: 10.1016/j.jphysparis.2004.06.001
- Lambrecht, L., Kreifelts, B., and Wildgruber, D. (2012). Age-related decrease in recognition of emotional facial and prosodic expressions. *Emotion* 12, 529–539. doi: 10.1037/a0026827
- Laurienti, P. J., Perrault, T. J., Stanford, T. R., Wallace, M. T., and Stein, B. E. (2005). On the use of super-additivity as a metric for characterizing multisensory integration in functional neuroimaging studies. *Exp. Brain Res.* 166, 289–297. doi: 10.1007/s00221-005-2370-2
- Leppänen, J. M., and Hietanen, J. K. (2003). Affect and face perception: odors modulate the recognition advantage of happy faces. *Emotion* 3, 315–326. doi: 10.1037/1528-3542.3.4.315
- Lieberman, D. Z. (2000). Children of alcoholics: an update. *Curr. Opin. Pediatr.* 12, 336–340. doi: 10.1097/00008480-200008000-00009
- Listerud, J., Powers, C., Moore, P., Libon, D. J., and Grossman, M. (2009). Neuropsychological patterns in magnetic resonance imaging-defined subgroups of patients with degenerative dementia. *J. Int. Neuropsychol. Soc.* 15, 459–470. doi: 10.1017/S1355617709090742
- Little, H. J., Stephens, D. N., Ripley, T. L., Borlikova, G., Duka, T., Schubert, M., et al. (2005). Alcohol withdrawal and conditioning. *Alcohol. Clin. Exp. Res.* 29, 453–464. doi: 10.1097/01.ALC.0000156737.56425.E3
- Love, S., Pollick, F. E., and Latinus, M. (2011). Cerebral correlates and statistical criteria of cross-modal face and voice integration. *Seeing Perceiving* 24, 351–367. doi: 10.1163/187847511X584452
- Marinkovic, K., Oscar-Berman, M., Urban, T., O'Reilly, C. E., Howard, J. A., Sawyer, K., et al. (2009). Alcoholism and dampened temporal limbic activation to emotional faces. *Alcohol. Clin. Exp. Res.* 33, 1880–1892. doi: 10.1111/j.1530-0277.2009.01026.x
- Martinotti, G., Di Nicola, M., Tedeschi, D., Cundari, S., and Janiri, L. (2009). Empathy ability is impaired in alcohol-dependent patients. *Am. J. Addict.* 18, 157–161. doi: 10.1080/10550490802544391
- Massida, Z., Belin, P., James, C., Rouger, J., Fraysse, B., Barone, P., et al. (2011). Voice discrimination in cochlear-implanted deaf subjects. *Hear Res.* 275, 120–129. doi: 10.1016/j.heares.2010.12.010
- Mathalon, D. H., Ford, J. M., and Pfefferbaum, A. (2000). Trait and state aspects of P300 amplitude reduction in schizophrenia: a retrospective longitudinal study. *Biol. Psychiatry* 47, 434–449. doi: 10.1016/S0006-3223(99)00277-2
- Maurage, P., Campanella, S., Philippot, P., Pham, T., and Joassin, F. (2007a). The crossmodal facilitation effect is disrupted in alcoholism: a study with emotional stimuli. *Alcohol Alcohol.* 42, 552–559. doi: 10.1093/alcalc/agm134
- Maurage, P., Joassin, F., Philippot, P., and Campanella, S. (2007b). A validated battery of vocal emotional expressions. *Neuropsychol. Trends* 2, 63–74.
- Maurage, P., Philippot, P., Verbanck, P., Noël, X., Kornreich, C., Hanak, C., et al. (2007c). Is the P300 deficit in alcoholism associated with early visual impairments (P100, N170)? An oddball paradigm. *Clin. Neurophysiol.* 118, 633–644. doi: 10.1016/j.clinph.2006.11.007
- Maurage, P., Campanella, S., Philippot, P., Martin, S., and de Timary, P. (2008a). Face processing in chronic alcoholism: a specific deficit for emotional features. *Alcohol. Clin. Exp. Res.* 32, 600–606. doi: 10.1111/j.1530-0277.2007.00611.x
- Maurage, P., Philippot, P., Joassin, F., Pauwels, L., Pham, T., Alonso Prieto, E. A., et al. (2008b). The auditory-visual integration of anger is impaired in alcoholism: an ERP study. *J. Psychiatry Neurosci.* 33, 111–122.
- Maurage, P., Campanella, S., Philippot, P., Vermeulen, N., Constant, E., Luminet, O., et al. (2008c). Electrophysiological correlates of the disrupted processing of anger in alcoholism. *Int. J. Psychophysiol.* 70, 50–62. doi: 10.1016/j.ijpsycho.2008.05.572
- Maurage, P., Campanella, C., Philippot, P., Charest, I., Martin, S., and de Timary, P. (2009a). Impaired emotional facial expression decoding in alcoholism is also present for emotional prosody and body postures. *Alcohol Alcohol.* 44, 476–485. doi: 10.1093/alcalc/agp037
- Maurage, P., Pesenti, M., Philippot, P., Joassin, F., and Campanella, S. (2009b). Latent deleterious effects of binge drinking over a short period of time revealed by electrophysiological measures only. *J. Psychiatry Neurosci.* 34, 111–118.
- Maurage, P., Grynberg, D., Noël, X., Joassin, F., Philippot, P., Hanak, C., et al. (2011a). Dissociation between affective and cognitive empathy in alcoholism: a specific deficit for the emotional dimension. *Alcohol. Clin. Exp. Res.* 35, 1662–1668.
- Maurage, P., Grynberg, D., Noël, X., Joassin, F., Hanak, C., Verbanck, P., et al. (2011b). The “Reading the Mind in the Eyes” test as a new way to explore complex emotions decoding in alcoholism. *Psych. Res.* 190, 375–378.
- Maurage, P., Callot, C., Philippot, P., Rombaux, P., and de Timary, P. (2011c). Chemosensory event-related potentials in alcoholism: a specific impairment for olfactory function. *Biol. Psychol.* 88, 28–36. doi: 10.1016/j.biopsycho.2011.06.004
- Maurage, P., Callot, C., Chang, B., Philippot, P., Rombaux, P., and de Timary, P. (2011d). Olfactory impairment is correlated with confabulation in alcoholism: towards a multimodal testing of orbitofrontal cortex. *PLoS ONE* 6:e23190. doi: 10.1371/journal.pone.0023190
- Maurage, P., Joassin, F., Philippot, P., Heeren, A., Vermeulen, N., Mahau, P., et al. (2012a). Disrupted regulation of social exclusion in alcohol-dependence: an fMRI study. *Neuropsychopharmacology* 37, 2067–2075. doi: 10.1038/npp.2012.54
- Maurage, P., Joassin, F., Speth, A., Modave, J., Philippot, P., and Campanella, S. (2012b). Cerebral effects of binge drinking: respective influences of global alcohol intake and consumption pattern. *Clin. Neurophysiol.* 123, 892–901. doi: 10.1016/j.clinph.2011.09.018
- Maurage, P., de Timary, P., Moulds, M. M., Wong, Q., Collignon, M., Philippot, P., et al. (2013a). Maladaptive social self-beliefs in alcohol-dependence: a specific bias towards excessive high standards. *PLoS ONE* 8:e58928. doi: 10.1371/journal.pone.0058928
- Maurage, P., Joassin, F., Pesenti, M., Grandin, C., Heeren, A., Philippot, P., et al. (2013b). The neural network sustaining cross-modal integration is impaired in alcohol-dependence: an fMRI study. *Cortex* 49, 1610–1626. doi: 10.1016/j.cortex.2012.04.012
- McCarthy, C. A., Ebel, B. E., Garrison, M. M., DiGiuseppe, D. L., Christakis, D. A., and Rivara, F. P. (2004). Continuity of binge and harmful drinking from late adolescence to early adulthood. *Pediatrics* 114, 714–719. doi: 10.1542/peds.2003-0864-L
- McGurk, H., and McDonald, J. (1976). Hearing lips and seeing

- voices. *Nature* 264, 746–748. doi: 10.1038/264746a0
- McIntosh, C., and Chick, J. (2004). Alcohol and the nervous system. *J. Neurol. Neurosurg. Psychiatry* 75, 16–21. doi: 10.1136/jnnp.2004.045708
- Mejias, S., Rossignol, M., Debatisse, D., Stree, E., Servais, L., Guérit, J. M., et al. (2005). Event-related potentials (ERPs) in ecstasy (MDMA) users during a visual oddball task. *Biol. Psychol.* 69, 333–352. doi: 10.1016/j.biopsycho.2004.11.010
- Melillo, R., and Leisman, G. (2009). Autistic spectrum disorders as functional disconnection syndrome. *Rev. Neurosci.* 20, 111–131. doi: 10.1515/REVNEURO.2009.20.2.111
- Mendlewicz, L., Linkowski, P., Bazelmans, C., and Philippot, P. (2005). Decoding emotional facial expressions in depressed and anorexic patients. *J. Affect. Disord.* 89, 195–199. doi: 10.1016/j.jad.2005.07.010
- Mongillo, E. A., Irwin, J. R., Whalen, D. H., Klaiman, C., Carter, A. S., and Schultz, R. T. (2008). Audiovisual processing in children with and without autism spectrum disorders. *J. Autism. Dev. Disord.* 38, 1349–1358. doi: 10.1007/s10803-007-0521-y
- Monnot, M., Lovallo, W. R., Nixon, S. J., and Ross, E. (2002). Neurological basis of deficits in affective prosody comprehension among alcoholics and fetal alcohol-exposed adults. *J. Neuropsychiatry Clin. Neurosci.* 14, 321–328. doi: 10.1176/appi.neuropsych.14.3.321
- Monnot, M., Nixon, S., Lovallo, W., and Ross, E. (2001). Altered emotional perception in alcoholics: deficits in affective prosody comprehension. *Alcohol. Clin. Exp. Res.* 25, 362–369. doi: 10.1111/j.1530-0277.2001.tb02222.x
- Montagne, B., Kessels, R. P., Wester, A. J., and de Haan, E. H. (2006). Processing of emotional facial expressions in Korsakoff's syndrome. *Cortex* 42, 705–710. doi: 10.1016/S0010-9452(08)70408-8
- Müller, V. I., Habel, U., Derntl, B., Schneider, F., Zilles, K., Turetsky, B. I., et al. (2011). Incongruence effects in cross-modal emotional integration. *Neuroimage* 54, 2257–2266. doi: 10.1016/j.neuroimage.2010.10.047
- Nicolás, J. M., Estruch, R., Salamero, M., Orteu, N., Fernandez-Solà, J., Sacanella, E., et al. (1997). Brain impairment in well-nourished chronic alcoholics is related to ethanol intake. *Ann. Neurol.* 41, 590–598. doi: 10.1002/ana.410410507
- Noël, X., Bechara, A., Brevers, D., Verbanck, P., and Campanella, S. (2010). Alcoholism and the loss of willpower: a neurocognitive perspective. *J. Psychophysiol.* 24, 240–248. doi: 10.1027/0269-8803/a000037
- Noël, X., Van der Linden, M., Schmidt, N., Sferrazza, R., Hanak, C., Le Bon, O., et al. (2001). Supervisory attentional system in nonamnesic alcoholic men. *Arch. Gen. Psychiatry* 58, 1152–1158. doi: 10.1001/archpsyc.58.12.1152
- Ogura, C., and Miyazato, Y. (1991). Cognitive dysfunctions of alcohol dependence using event related potentials. *Arukuru KenkyutoYakubutsu Ison.* 26, 331–340.
- Oscar-Berman, M., Hancock, M., Mildworf, B., Hutner, N., and Weber, D. A. (1990). Emotional perception and memory in alcoholism and aging. *Alcohol. Clin. Exp. Res.* 14, 383–393. doi: 10.1111/j.1530-0277.1990.tb00491.x
- Oscar-Berman, M., Kirkley, S. M., Gansler, D. A., and Couture, A. (2004). Comparisons of Korsakoff and non-Korsakoff alcoholics on neuropsychological tests of prefrontal brain functioning. *Alcohol. Clin. Exp. Res.* 28, 667–675. doi: 10.1097/01.ALC.0000122761.09179.B9
- Oscar-Berman, M., and Marinkovic, K. (2003). Alcoholism and the brain: an overview. *Alcohol. Res. Health* 27, 125–133.
- Oscar-Berman, M., and Marinkovic, K. (2007). Alcohol: effects on neurobehavioral functions and the brain. *Neuropsychol. Rev.* 17, 239–257. doi: 10.1007/s11065-007-9038-6
- Palmero-Soler, E., Dolan, K., Hadschek, V., and Tass, P. A. (2007). SwLORETA: a novel approach to robust source localization and synchronization tomography. *Phys. Med. Biol.* 52, 1783–1800. doi: 10.1088/0031-9155/52/7/002
- Pearl, D., Yodashkin-Porat, D., Katz, N., Valevski, A., Aizenberg, D., Sigler, M., et al. (2009). Differences in audiovisual integration, as measured by McGurk phenomenon, among adult and adolescent patients with schizophrenia and age-matched healthy control groups. *Compr. Psychiatry* 50, 186–192. doi: 10.1016/j.comppsy.2008.06.004
- Petrini, K., Crabbe, F., Sheridan, C., and Pollick, F. E. (2011). The music of your emotions: neural substrates involved in detection of emotional correspondence between auditory and visual music actions. *PLoS ONE* 6:e19165. doi: 10.1371/journal.pone.0019165
- Petrini, K., McAleer, P., and Pollick, F. E. (2010). Audiovisual integration of emotional signals from music improvisation does not depend on temporal correspondence. *Brain Res.* 1323, 139–148. doi: 10.1016/j.brainres.2010.02.012
- Philippot, P., and Power, M. (2010). *FaceTales*. Available online at: <http://www.ips.ucl.ac.be/recherche/projets/FaceTales/en/Home.htm> doi: 10.1016/j.hears.2010.12.010
- Philippot, P., Kornreich, C., Blairy, S., Baerts, I., Den Dulk, A., Le Bon, O., et al. (1999). Alcoholics' deficits in the decoding of emotional facial expression. *Alcohol. Clin. Exp. Res.* 23, 1031–1038.
- Pitel, A. L., Beunieux, H., Witkowski, T., Vabret, F., Guillery-Girard, B., Quinette, P., et al. (2007). Genuine episodic memory deficits and executive dysfunctions in alcoholic subjects early in abstinence. *Alcohol. Clin. Exp. Res.* 31, 1169–1178. doi: 10.1111/j.1530-0277.2007.00418.x
- Pogarell, O., Juckel, G., Norra, C., Leicht, G., Karch, S., Schaaff, N., et al. (2007). Prediction of clinical response to antidepressants in patients with depression: neurophysiology in clinical practice. *Clin. EEG Neurosci.* 38, 74–77. doi: 10.1177/155005940703800208
- Polich, J. (2004). Clinical application of the P300 event-related brain potential. *Phys. Med. Rehabil. Clin. North Am.* 15, 133–161. doi: 10.1016/S1047-9651(03)00109-8
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019
- Porjesz, B., Rangaswamy, M., Kamarajan, C., Jones, K. A., Padmanabhapillai, A., and Begleiter, H. (2005). The utility of neurophysiological markers in the study of alcoholism. *Clin. Neurophysiol.* 116, 993–1018 doi: 10.1016/j.clinph.2004.12.016
- Pourtois, G., de Gelder, B., Bol, A., and Crommelinck, M. (2005). Perception of facial expressions and voices and of their combination in the human brain. *Cortex* 41, 49–59. doi: 10.1016/S0010-9452(08)70177-1
- Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B., and Crommelinck, M. (2000). The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport* 11, 1329–1333. doi: 10.1097/00001756-200004270-00036
- Pourtois, G., Debatisse, D., Despland, P. A., and de Gelder, B. (2002). Facial expressions modulate the time course of long latency auditory brain potentials. *Brain Res. Cogn. Brain Res.* 14, 99–105. doi: 10.1016/S0926-6410(02)00064-2
- Price, G. W., Michie, P. T., Johnston, J., Innes-Brown, H., Kent, A., Clissa, P., et al. (2006). A multivariate electrophysiological endophenotype, from a unitary cohort, shows greater research utility than any single feature in the Western Australian family study of schizophrenia. *Biol. Psychiatry* 60, 1–10. doi: 10.1016/j.biopsycho.2005.09.010
- Rämä, P., and Courtney, S. M. (2005). Functional topography of working memory for face or voice identity. *Neuroimage* 24, 224–234. doi: 10.1016/j.neuroimage.2004.08.024
- Regenbogen, C., Schneider, D. A., Gur, R. E., Schneider, F., Habel, U., and Kellermann, T. (2012). Multimodal human communication-targeting facial expressions, speech content and prosody. *Neuroimage* 60, 2346–2356. doi: 10.1016/j.neuroimage.2012.02.043
- Riley, H., and Schutte, N. S. (2003). Low emotional intelligence as a predictor of substance-use problems. *J. Drug Educ.* 33, 391–398. doi: 10.2190/6DH9-YTOM-FT99-2X05
- Robins, D. L., Hunyadi, E., and Schultz, R. T. (2009). Superior temporal activation in response to dynamic audio-visual emotional cues. *Brain Cogn.* 69, 269–278. doi: 10.1016/j.bandc.2008.08.007
- Rossignol, M., Philippot, P., Douilliez, C., Crommelinck, M., and Campanella, S. (2005). The perception of fearful and happy facial expression is modulated by anxiety: an event-related potential study. *Neurosci. Lett.* 377, 115–120. doi: 10.1016/j.neulet.2004.11.091
- Rugg, M. D., and Coles, M. G. H. (1995). *Electrophysiology of Mind*. Oxford: Oxford University Press.
- Rupp, C. I. (2010). Olfactory function and schizophrenia: an update. *Curr. Opin. Psychiatry* 23, 97–102. doi: 10.1097/YCO.0b013e328336643f
- Rupp, C. I., Fleischhacker, W., Drexler, A., Hausmann, A., Hinterhuber, H., and Kurz, M. (2006). Executive function and memory in relation to olfactory deficits in alcohol-dependent patients. *Alcohol. Clin. Exp. Res.* 30, 1355–1362. doi: 10.1111/j.1530-0277.2006.00162.x
- Rupp, C. I., Fleischhacker, W., Hausmann, H., Mair, D.,

- Hinterhuber, H., and Kurz, M. (2004). Olfactory functioning in patients with alcohol dependence: impairments in odor judgments. *Alcohol Alcohol* 39, 514–519. doi: 10.1093/alcalc/agh100
- Rupp, C. L., Kurz, M., Kemmler, G., Mair, D., Hausmann, A., Hinterhuber, H., et al. (2003). Reduced olfactory sensitivity, discrimination, and identification in patients with alcohol dependence. *Alcohol Clin. Exp. Res.* 27, 432–439. doi: 10.1097/01.ALC.0000057945.57330.2C
- Salloum, J. B., Ramchandani, V. A., Bodurka, J., Rawlings, R., Momenan, R., George, D., et al. (2007). Blunted rostral anterior cingulate response during a simplified decoding task of negative emotional facial expressions in alcoholic patients. *Alcohol Clin. Exp. Res.* 31, 1490–1504. doi: 10.1111/j.1530-0277.2007.00447.x
- Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., and Matsumara, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: an fMRI study. *Cogn. Brain Res.* 20, 81–91. doi: 10.1016/j.cogbrainres.2004.01.008
- Schiffman, M. (1997). Taste and smell losses in normal aging and disease. *JAMA* 278, 1357–1362. doi: 10.1001/jama.1997.03550160077042
- Schomerus, G., Corrigan, P. W., Klauer, T., Kuwert, P., Freyberger, H. J., and Lucht, M. (2011a). Self-stigma in alcohol dependence: consequences for drinking-refusal self-efficacy. *Drug Alcohol Depend.* 114, 12–17. doi: 10.1016/j.drugaldep.2010.08.013
- Schomerus, G., Lucht, M., Holzinger, A., Matschinger, H., Carta, M. G., and Angermeyer, M. C. (2011b). The stigma of alcohol dependence compared with other mental disorders: a review of population studies. *Alcohol Alcohol* 46, 105–112. doi: 10.1093/alcalc/agg089
- Schulte, T., Muller-Oehring, E. M., Pfefferbaum, A., and Sullivan, E. V. (2010). Neurocircuitry of emotion and cognition in alcoholism: contributions from white matter fiber tractography. *Dialogues Clin. Neurosci.* 12, 554–560.
- Schweinberger, S. R., Herholz, A., and Sommer, W. (1997). Recognizing famous voices: influence of stimulus duration and different types of retrieval cues. *J. Speech Lang. Hear. Res.* 40, 453–463.
- Segrin, C., and Menees, M. M. (1996). The impact of coping styles and family communication on the social skills of children of alcoholics. *J. Stud. Alcohol* 57, 29–33.
- Seubert, J., Kellermann, T., Loughead, J., Boers, F., Brensinger, C., Schneider, F., et al. (2010a). Processing of disgusted faces is facilitated by odor primes: a functional MRI study. *Neuroimage* 53, 746–756. doi: 10.1016/j.neuroimage.2010.07.012
- Seubert, J., Loughead, J., Kellermann, T., Boers, F., Brensinger, C. M., and Habel, U. (2010b). Multisensory integration of emotionally valenced olfactory-visual information in patients with schizophrenia and healthy controls. *J. Psychiatry Neurosci.* 35, 185–194. doi: 10.1503/jpn.090094
- Shepherd, G. M. (2006). Smell images and the flavour system in the human brain. *Nature* 444, 316–321. doi: 10.1038/nature05405
- Sher, K. J., Walitzer, K. S., Wood, P. K., and Brent, E. E. (1991). Characteristics of children of alcoholics: putative risk factors, substance use and abuse, and psychopathology. *J. Abnorm. Psychol.* 100, 427–448. doi: 10.1037/0021-843X.100.4.427
- Small, D. A. (2004). Crossmodal integration – insights from the chemical senses. *Trends Neurosci.* 27, 123–124. doi: 10.1016/j.tins.2004.01.002
- Smith, M. E., and Oscar-Berman, M. (1992). Resource-limited information processing in alcoholism. *J. Stud. Alcohol* 53, 514–518.
- Spielberger, D. C., Gorsuch, R. L., Lushene, R. L., Vagg, P. R., and Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Spitzer, J. B. (1981). Auditory effects of chronic alcoholism. *Drug Alcohol Depend.* 8, 317–335. doi: 10.1016/0376-8716(81)90041-7
- Spitzer, J. B., and Ventry, I. M. (1980). Central auditory dysfunction among chronic alcoholics. *Arch. Otolaryngol.* 106, 224–229. doi: 10.1001/archotol.1980.00790280032006
- Sullivan, E. V., Mathalon, D. H., Zipursky, R. B., Kerstein-Tucker, Z., Knight, R. T., and Pfefferbaum, A. (1993). Factors of the Wisconsin Card sorting test as measures of frontal lobe function in schizophrenia and in chronic alcoholism. *Psychiatry Res.* 46, 175–199. doi: 10.1016/0165-1781(93)90019-D
- Szabo, Z., Owonikoko, T., Peyrot, M., Varga, J., Mathews, W. B., Ravert, H. T., et al. (2004). Positron emission tomography imaging of the serotonin transporter in subjects with a history of alcoholism. *Biol. Psychiatry* 55, 766–771. doi: 10.1016/j.biopsych.2003.11.023
- Szczepanska, L., Baran, J., and Mikolajczyk-Boba, M. (2004). Connection between personality and emotional intelligence in groups of patients after suicidal attempts and ethanol dependent persons. *Przegląd Lekki* 61, 287–291.
- Szycik, G. R., Münte, T. F., Dillo, W., Mohammadi, B., Samii, A., Emrich, H. M., et al. (2009). Audiovisual integration of speech is disturbed in schizophrenia: an fMRI study. *Schizophr. Res.* 110, 111–118. doi: 10.1016/j.schres.2009.03.003
- Taieb, O., Corcos, M., Loas, G., Speranza, M., Guillaud, O., Perez-Diaz, F., et al. (2002). Alexithymia and alcohol dependence. *Ann. Med. Interne (Paris)* 153, 51–60.
- Teder-Sälejärvi, W. A., McDonald, J. J., Di Russo, F., and Hillyard, S. A. (2002). An analysis of audiovisual crossmodal integration by means of event-related potentials (ERP) recordings. *Cogn. Brain Res.* 14, 106–114. doi: 10.1016/S0926-6410(02)00065-4
- Thoma, P., Winter, N., Juckel, G., and Roser, P. (2013). Mental state decoding and mental state reasoning in recently detoxified alcohol-dependent individuals. *Psychiatry Res.* 205, 232–240. doi: 10.1016/j.psychres.2012.08.042
- Townshend, J. M., and Duka, T. (2003). Mixed emotions: alcoholics' impairments in the recognition of specific emotional facial expressions. *Neuropsychologia* 41, 773–782. doi: 10.1016/S0028-3932(02)00284-1
- Townshend, J. M., and Duka, T. (2005). Binge drinking, cognitive performance and mood in a population of young social drinkers. *Alcohol Clin. Exp. Res.* 29, 317–325. doi: 10.1097/01.ALC.0000156453.05028.F5
- Turetsky, B. I., Hahn, C. G., Borgmann-Winter, K., and Moberg, P. J. (2009). Scents and nonsense: olfactory dysfunction in schizophrenia. *Schizophr. Bull.* 35, 1117–1113. doi: 10.1093/schbul/sbp111
- Uekermann, J., and Daum, I. (2008). Social cognition in alcoholism: a link to prefrontal cortex dysfunction? *Addiction* 103, 726–735. doi: 10.1111/j.1360-0443.2008.02157.x
- Uekermann, J., Channon, S., Winkel, K., Schlebusch, P., and Daum, I. (2007). Theory of mind, humour processing and executive functioning in alcoholism. *Addiction* 102, 232–240. doi: 10.1111/j.1360-0443.2006.01656.x
- Uekermann, J., Daum, I., Schlebusch, P., and Trenckmann, U. (2005). Processing of affective stimuli in alcoholism. *Cortex* 41, 189–194. doi: 10.1016/S0010-9452(08)70893-1
- Uzun, O., Ats, A., Cansever, A., and Ozsahin, A. (2003). Alexithymia in male alcoholics: study in a Turkish sample. *Compr. Psychiatry* 44, 349–352. doi: 10.1016/S0010-440X(03)00009-9
- Van den Stock, J., de Jong, S. J., Hodiamont, P. P., and de Gelder, B. (2011). Perceiving emotions from bodily expressions and multisensory integration of emotion cues in schizophrenia. *Soc. Neurosci.* 6, 537–547. doi: 10.1080/17470919.2011.568790
- van der Smagt, M. J., van Engeland, H., and Kemner, C. (2007). Brief report: can you see what is not there? Low-level auditory-visual integration in autism spectrum disorder. *J. Autism Dev. Disord.* 37, 2014–2019. doi: 10.1007/s10803-006-0346-0
- Van der Stelt, O., and Belger, A. (2007). Application of electroencephalography to the study of cognitive and brain functions in schizophrenia. *Schizophr. Bull.* 33, 955–970. doi: 10.1093/schbul/sbm016
- Van der Stelt, O., Gunning, W. B., Snel, J., Zeef, E., and Kok, A. (1994). Children of alcoholics: attention, information processing and event-related brain potentials. *Acta Paediatr.* 404, 4–6. doi: 10.1111/j.1651-2227.1994.tb13375.x
- Vatakis, A., and Spence, C. (2006). Audiovisual synchrony perception for speech and music assessed using a temporal order judgment task. *Neurosci. Lett.* 393, 40–44. doi: 10.1016/j.neulet.2005.09.032
- Verma, R. K., Panda, N. K., Basu, D., and Raghunathan, M. (2006). Audiovestibular dysfunction in alcohol dependence. Are we worried? *Am. J. Otolaryngol.* 27, 225–228. doi: 10.1016/j.amjoto.2005.09.005
- Wagner, U., N'Diaye, K., Ethofer, T., and Vuilleumier, P. (2011). Guilt-specific processing in the prefrontal cortex. *Cereb. Cortex* 21, 2461–2470. doi: 10.1093/cercor/bhr016
- Walitzer, K. S., and Dearing, R. L. (2006). Gender differences in alcohol and substance use relapse. *Clin. Psychol. Rev.* 26, 128–148. doi: 10.1016/j.cpr.2005.11.003
- Wiers, R. W., Bartholow, B. D., van den Wildenberg, E., Thush, C.,

- Engels, R. C., Sher, K. J., et al. (2007). Automatic and controlled processes and the development of addictive behaviors in adolescents: a review and a model. *Pharmacol. Biochem. Behav.* 86, 263–283. doi: 10.1016/j.pbb.2006.09.021
- Wiers, R. W., Gladwin, T. E., Hofmann, W., Salemink, E., and Ridderinkhof, K. R. (2013). Cognitive bias modification and cognitive control training in addiction and related psychopathology: mechanisms, clinical perspectives, and ways forward. *Clin. Psych. Sci.* 1, 192–212. doi: 10.1177/2167702612466547
- Winston, J. S., Gottfried, J. A., Kilner, J. M., and Dolan, R. J. (2005). Integrated neural representations of odor intensity and affective valence in human amygdala. *J. Neurosci.* 25, 8903–8907. doi: 10.1523/JNEUROSCI.1569-05.2005
- World Health Organization. (2011). *Global Status Report on Alcohol and Health*. Geneva. Available online at: [http://www.who.int/substance\\_abuse/publications/global\\_alcohol\\_report](http://www.who.int/substance_abuse/publications/global_alcohol_report). Accessed on 05 June 2013.
- Zeigler, D. W., Wang, C. C., Yoast, R. A., Dickinson, B. D., McCaffree, M. A., Robinowitz, C. B., et al. (2005). The neurocognitive effects of alcohol on adolescents and college students. *Prev. Med.* 40, 23–32. doi: 10.1016/j.ypmed.2004.04.044
- Zupan, B., and Sussman, J. E. (2009). Auditory preferences of young children with and without hearing loss for meaningful auditory-visual compound stimuli. *J. Commun. Disord.* 42, 381–396. doi: 10.1016/j.jcomdis.2009.04.002
- Zywiak, W. H., Westerberg, V. S., Connors, G. J., and Maisto, S. A. (2003). Exploratory findings from the reasons for drinking questionnaire. *J. Subst. Abuse Treat.* 25, 287–292. doi: 10.1016/S0740-5472(03)00118-1
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 21 March 2013; accepted: 05 July 2013; published online: 25 July 2013.  
Citation: Maurage P and Campanella S (2013) Experimental and clinical usefulness of crossmodal paradigms in psychiatry: an illustration from emotional processing in alcohol-dependence. *Front. Hum. Neurosci.* 7:394. doi: 10.3389/fnhum.2013.00394  
Copyright © 2013 Maurage and Campanella. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



# Food related processes in the insular cortex

Sabine Frank<sup>1,2</sup>, Stephanie Kullmann<sup>1,2,3,4</sup> and Ralf Veit<sup>1\*</sup>

<sup>1</sup> Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Tübingen, Germany

<sup>2</sup> fMEG Center, University of Tübingen, Tübingen, Germany

<sup>3</sup> Institute for Diabetes Research and Metabolic Diseases of the Helmholtz Center Munich at the University of Tübingen, Tübingen, Germany

<sup>4</sup> German Center for Diabetes Research, Neuherberg, Germany

## Edited by:

Martin Klasen, RWTH Aachen University, Germany

## Reviewed by:

Giuseppina Rota, University of Pisa, Italy

Nils B. Kroemer, Technische Universität Dresden, Germany

Anne Schienle, Karl-Franzens-University, Austria

## \*Correspondence:

Ralf Veit, Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Otfried Müller Strasse 47, 72076 Tübingen, Germany  
e-mail: ralf.veit@uni-tuebingen.de

The insular cortex is a multimodal brain region with regional cytoarchitectonic differences indicating various functional specializations. As a multisensory neural node, the insular cortex integrates perception, emotion, interoceptive awareness, cognition, and gustation. Regarding the latter, predominantly the anterior part of the insular cortex is regarded as the primary taste cortex. In this review, we will specifically focus on the involvement of the insula in food processing and on multimodal integration of food-related items. Influencing factors of insular activation elicited by various foods range from calorie-content to the internal physiologic state, body mass index or eating behavior. Sensory perception of food-related stimuli including seeing, smelling, and tasting elicits increased activation in the anterior and mid-dorsal part of the insular cortex. Apart from the pure sensory gustatory processing, there is also a strong association with the rewarding/hedonic aspects of food items, which is reflected in higher insular activity and stronger connections to other reward-related areas. Interestingly, the processing of food items has been found to elicit different insular activation in lean compared to obese subjects and in patients suffering from an eating disorder (anorexia nervosa (AN), bulimia nervosa (BN)). The knowledge of functional differences in the insular cortex opens up the opportunity for possible noninvasive treatment approaches for obesity and eating disorders. To target brain functions directly, real-time functional magnetic resonance imaging neurofeedback offers a state-of-the-art tool to learn to control the anterior insular cortex activity voluntarily. First evidence indicates that obese adults have an enhanced ability to regulate the anterior insular cortex.

**Keywords:** insular cortex, food, gustatory, neurofeedback, obesity, weight loss, eating disorders

## THE INSULAR CORTEX—FROM NEUROANATOMY TO FUNCTION

The insular cortex is embedded in the lateral sulcus of the mammalian brain. On the basis of cytoarchitectonic studies using myelin staining techniques, the insula can be subdivided in three major compartments according to the laminar structure, referred to as the anterior ventral agranular, dorsal anterior dysgranular, and posterior granular part of the insular cortex (Mesulam and Mufson, 1985; Gallay et al., 2012). The agranular anterior insula in junction to the caudal orbitofrontal cortex (OFC) and the adjacent frontal operculum has been identified as the primary taste cortex (Rolls, 2006). Besides multiple perceptive inputs of gustational cues (smell, taste, temperature, viscosity, texture) in the anterior insula and hence different pathways, additional granular and dysgranular regions especially the dorsal mid-insula are involved in gustation (De Araujo and Simon, 2009; Kurth et al., 2010). Their close interconnections with the OFC indicate that this part plays a predominant role in the evaluation of motivational states and primary reinforcers (Wager and Barrett, 2004). Also functional connectivity based analyses highlight the anterior part of the insular cortex as a major hub in cerebral pro-

cessing of cognitive, emotional, motivational, and sensory stimuli, and, defines together with the anterior cingulate cortex (ACC) the salience network (Menon and Uddin, 2010). The anterior dysgranular part is superior to the agranular part bounded on the border to the frontal operculum. This part is particularly engaged during tasks requiring executive control, shifting attention, and working memory (Wager and Barrett, 2004). The intermediate part of the insula and its dysgranular laminar structure extending into the parietal operculum is strongly connected with all parts of the insula and is involved in motor, somatosensory, and pain processing (Kurth et al., 2010). Hence, neuroanatomical findings indicate that the insular cortex is an important structure on the transition between allocortex and isocortex, hinting to the involvement in a wide range of sensory, emotional, and cognitive processing of gustatory stimuli.

## FOOD PROCESSING IN THE INSULAR CORTEX

The insular cortex is integrated in a distinct network responsible for the neural control of appetite and the regulation of energy balance. Whereas the hypothalamus represents the major homeostatic player, the insular cortex is integrated in the neural

system which is involved in the processing of external sensory information tightly linked to reward processing (Berthoud, 2011). Therefore, the insular cortex activity also contributes to the hedonic system.

Several neuroimaging studies emphasized the functional contribution of the anterior insula in gustatory perception (Small et al., 2003; Veldhuizen et al., 2011; **Figure 1A**), which is represented in the processing of visually presented (Porubska et al., 2006; Frank et al., 2010), tasted or smelled food stimuli (De Araujo et al., 2003), and also in food craving (Pelchat, 1997; Pelchat et al., 2004). Eating per se is a multimodal experience, including taste, olfaction, smell, and somatosensory inputs (De Araujo and Simon, 2009). As part of the primary taste and primary olfactory cortex (Rolls, 2006; Small, 2010), the anterior insula is also highly responsive to different flavors (Rolls, 2005; Small, 2012; Small and Green, 2012). Sensory food-related inputs are combined in the anterior insula (Small, 2012), resulting in increased activation of this region after stimulation with a specific flavor (Small et al., 1999). Small and Prescott (2005) describe overlapping activation in the anterior insula after independent stimulation with taste and odor cues. Besides the taste component, transferred from the taste buds on the tongue to the primary taste cortex, the aroma of food is also experienced olfactorily via the retronasal route (Ruijschop et al., 2009; Small and Green, 2012).

Also, the texture and viscosity of ingested food is represented in the anterior and mid-insular cortex. Here, the activation changes according to the viscosity of a stimulus (De Araujo and Rolls, 2004; Alonso et al., 2007).

Besides components like taste, aroma and texture, also the amount of fat influences the activity in this gustatory and hedonic region. A recent fMRI study (Frank et al., 2012b), investigating the effect of a high- and low-fat meal on the cerebral blood flow, revealed a differential influence of fat on the mid-anterior insular cortex and the hypothalamus. The activity in the hypothalamus, representing the homeostatic system in the brain, decreased after intake of a high-fat meal, whereas the insular cortex activity increased after intake of the low-fat meal. This suggests an interaction of the homeostatic and the gustatory system, which might be mediated by the fat content of the meal.

The processing of food also includes the internal evaluation of the ingested, seen or smelled nutrients. The evaluative component includes interoceptive awareness, which is as well associated with insular processing (Craig, 2009). On a behavioral level, it was shown that good cardiac awareness, as a marker of interoception, is inversely related to the experienced fullness and myoelectric gastric activity after water load (Herbert et al., 2012). On the neuronal level, gastric distention without actual food intake leads to increased activation in the posterior insular cortex (Wang et al., 2008). Such findings corroborate the integrative function of the insular cortex.

A recent meta-analysis by Brooks et al. (2013) report decreased activation in obese compared to lean subjects in the mid-insular cortex, a region shown to be involved in interoceptive awareness (Simmons et al., 2012). The reduced awareness of the bodily state and, therefore, also for appetite signals of the gut and brain might be a reason for obese to consume more food in order to feel the

interoceptive cues from the body in the same way normal-weight people do (Brooks et al., 2013).

Craig (2005) already proposed laterality differences in interoceptive perception related to emotional processing. In a previous study a stronger impairment in taste functions in patients suffering from a lesion in the left anterior insular cortex compared to patients with a lesion in the right anterior insula (Stevenson et al., 2013), was shown. Furthermore, there is evidence that pleasant odors are rather processed in the left hemisphere and unpleasant odors in the right hemisphere (Henkin and Levy, 2001). However, further evidence is needed to understand possible hemispheric relationships of insular functions in more detail.

## EATING DISORDERS

It has been shown that bulimia nervosa (BN) patients exhibit increased insula activation to high-caloric food pictures in comparison to overweight and normal weight control subjects (Schienle et al., 2009). This difference is possibly due to the enhanced autonomic arousal that appetizing food pictures elicit in BN. Increased insula activation was also shown in anorexia nervosa (AN) patients compared to healthy subjects when contrasting pictures of high- versus low-calorie drinks (Nunn et al., 2011; **Figure 1E**). In contrast, after the ingestion of chocolate milk in a hungry state, AN patients exhibited less activation in the insula than control subjects (Vocks et al., 2011). Of special importance is the change in insula function when women recovered from AN or BN. While AN recovered patients showed a decreased anterior insula activity after drinking sweet tastes (Wagner et al., 2008; Oberndorfer et al., 2013), BN recovered patients revealed an enhanced insula response in relation to weight matched controls (Oberndorfer et al., 2013). The different activation patterns may result from an altered processing of hunger or reward signals and a misinterpretation of internal feeling and feeding states that lead to exaggerated or restricted eating behavior even after treatment.

## DIFFERENTIAL ACTIVATION IN LEAN AND OBESE

Neuroimaging studies investigating food processing, by means of visual stimulation, have shown enhanced insula activity in obese compared to lean subjects (**Figure 1B**). Specifically, obese subjects were found to show higher left anterior and right mid-insular activity compared to lean control subjects in response to food cues (Scharmuller et al., 2012). Also, Rothmund et al. (2007) and Stoeckel et al. (2008) reported enhanced activation in response to high-caloric food pictures in obese women in the anterior insula. In adolescent girls, the activation in the anterior insula correlated positively with the BMI during the orientation to food cues (Yokum et al., 2011). Studies investigating linear relationships of BMI with brain functions showed heightened activity in the anterior insula and the adjacent frontal operculum with increasing BMI (Ziauddeen et al., 2012).

Beside visual stimulation with food items, studies using oral food cues have also shown the insular cortex to be vital for food intake. While hunger resulted in a regional cerebral blood flow (rCBF) increase after administration of 2 mL of drinking water, satiation has been associated with a decrease in insular rCBF, suggesting that the reaction of the insular cortex to sensory expe-

periences is affected by hunger (Del Parigi et al., 2002; **Figure 1C**). However, this decrease after satiation was more pronounced in obese compared to lean subjects (Gautier et al., 2000, 2001). Additionally, obese subjects revealed an enhanced sensory experience in the mid-dorsal insula to a liquid meal after a prolonged fast (DelParigi et al., 2005). Concomitantly, these results point to abnormal gustatory processing in obesity in response to a meal as well as to the sensory processing of food. The combination of a sweetened drink with the stimulation with food pictures also revealed enhanced anterior insula activation in obese subjects (Connolly et al., 2013), supporting the integration of multimodal stimuli in this area. Generally, the anterior insular cortex is highly responsive to food intake and anticipated food intake, a response that is more pronounced in obese (Stice et al., 2008, 2009).

Besides changes in activity, the insular cortex also revealed significant changes in functional connectivity in obese compared to lean subjects during resting-state and in response to food cues (**Figure 1D**). As such, the anterior insula has significant functional connections to several frontal, temporal, and parietal areas, in particular to the OFC, inferior frontal cortex and to the ACC in normal weight subjects (Taylor et al., 2009; Deen et al., 2011). In contrast, obese subjects revealed decreased functional connectivity in the insular cortex during resting-state (Kullmann et al., 2012), and increased functional and effective connectivity in response to food cues especially to striatal regions (Garcia-Garcia et al., 2012; Nummenmaa et al., 2012; Kullmann et al., 2013).

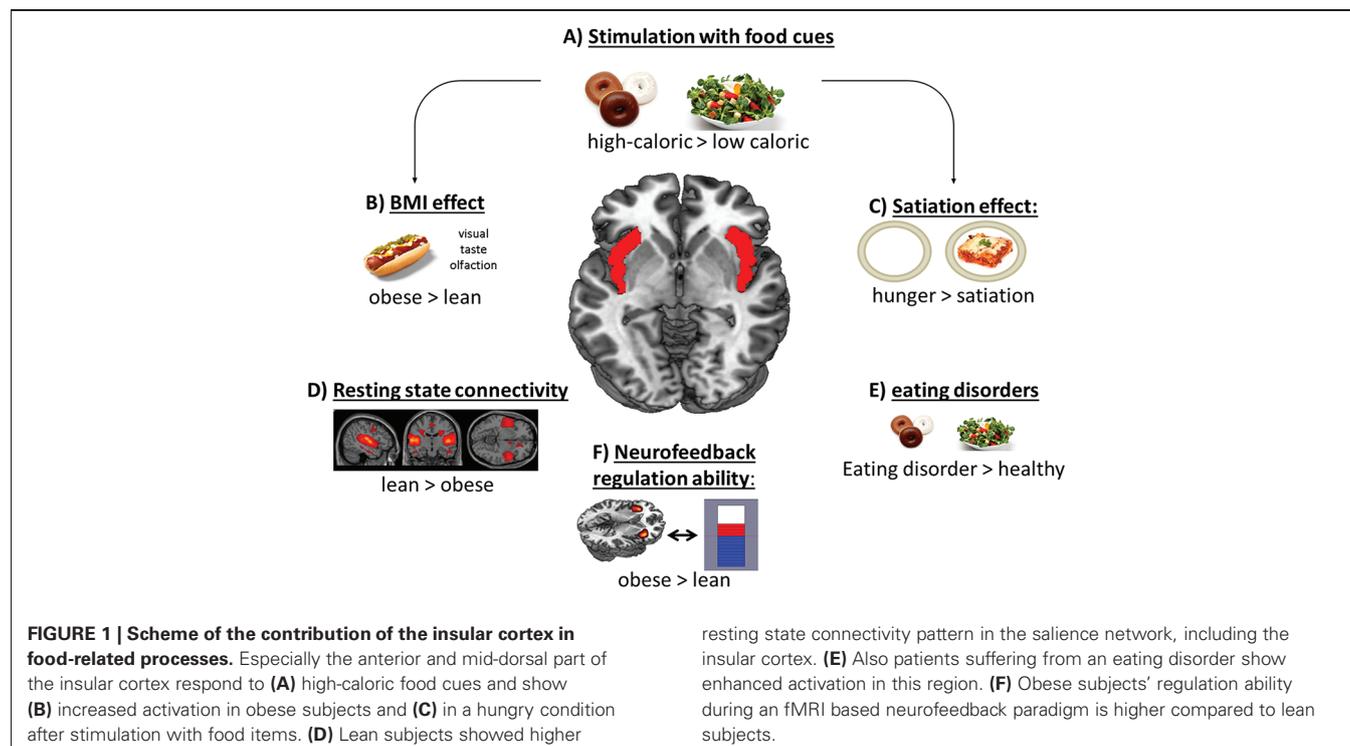
## THE PROBLEM OF WEIGHT LOSS MAINTENANCE

When facing the problem of obesity, one pressing question is how to effectively lose and maintain body weight. Successful weight

loss maintainers show a greater bilateral insula response after orosensory stimulation with food cues (Sweet et al., 2012). Interestingly, the response to visually presented food items in the insular cortex seems to be predictive for the weight loss outcome. Less successful patients in a weight loss program showed higher insular activation pre- and post-treatment (Murdaugh et al., 2012). After successful weight-loss maintenance achieved by bariatric surgery, neuroimaging studies have shown that brain activations after food intake or visual stimulation with food cues are comparable with lean subjects (Van De Sande-Lee et al., 2011; Frank et al., 2013). Also in motivational and reward-related regions (including the insular cortex) stimulation with food pictures showed decreased activation after gastric banding (Bruce et al., 2012).

## NEUROFEEDBACK AS A POSSIBLE THERAPEUTIC APPROACH

Regarding the increasing prevalence for obesity and the frequent failure of weight maintenance after weight loss, new therapeutic approaches are urgently needed. Therefore, it is intriguing to speculate about possible biofeedback strategies. Food-specific electrodermal biofeedback leads to increased food-related self-efficacy and reduced perceived stress (Teufel et al., 2013). Morewedge et al. (2010) reported that food consumption can be reduced by thoughts for food in lean subjects. The focus on food during eating enhances memory for a meal to later time points and reduce later food intake (Higgs and Donohoe, 2011). One innovative approach that might support the effort of obesity treatment is an fMRI-based neurofeedback training, which allows the voluntary regulation of specific brain regions (Birbaumer et al., 2013). Considering the multimodal functions of the insular cortex and its importance for food reward, the anterior insula



seems to be an appropriate target for real-time fMRI (rtfMRI) neurofeedback. In a previous rtfMRI study, addressing the anterior insular cortex as the region of interest (ROI), lean participants learned to regulate this region voluntarily within one day over four training sessions (Caria et al., 2007). In a follow-up study, this group demonstrated that successful regulation compared to no regulation of the anterior insular cortex resulted in increased negative valence ratings of emotional pictures (Caria et al., 2010). Furthermore, it was shown that effective connectivity between the anterior insular cortex and areas involved in emotional processing were strongest in the best regulation session (Veit et al., 2012). In a recent study, we addressed insular neurofeedback training in obese subjects (Frank et al., 2012a). During the training sessions, all obese participants were able to regulate the activity, whereas four out of eleven participants of the lean group were not able to successfully regulate the anterior insula (Figure 1F). Investigating underlying neural connectivity processes, lean regulators in comparison to obese regulators showed stronger functional connectivity in cingular and temporal cortices during regulation. Therefore, lean and obese subjects seem to recruit differential neural networks to perform a voluntary regulation of primary gustatory systems.

## REFERENCES

- Alonso, B. D. C., Marciani, L., Head, K., Clark, P., Spiller, R. C., Rayment, P., et al. (2007). Functional magnetic resonance imaging assessment of the cortical representation of oral viscosity. *J. Texture Stud.* 38, 725–737. doi: 10.1111/j.1745-4603.2007.00122.x
- Berthoud, H. R. (2011). Metabolic and hedonic drives in the neural control of appetite: who is the boss? *Curr. Opin. Neurobiol.* 21, 888–896. doi: 10.1016/j.conb.2011.09.004
- Birbaumer, N., Ruiz, S., and Sitaram, R. (2013). Learned regulation of brain metabolism. *Trends Cogn. Sci.* 17, 295–302. doi: 10.1016/j.tics.2013.04.009
- Brooks, S. J., Cedernaes, J., and Schiöth, H. B. (2013). Increased prefrontal and parahippocampal activation with reduced dorsolateral prefrontal and insular cortex activation to food images in obesity: a meta-analysis of fMRI studies. *PLoS One* 8:e60393. doi: 10.1371/journal.pone.0060393
- Bruce, J. M., Hancock, L., Bruce, A., Lepping, R. J., Martin, L., Lundgren, J. D., et al. (2012). Changes in brain activation to food pictures after adjustable gastric banding. *Surg. Obes. Relat. Dis.* 8, 602–608. doi: 10.1016/j.soard.2011.07.006
- Caria, A., Sitaram, R., Veit, R., Begliomini, C., and Birbaumer, N. (2010). Volitional control of anterior insula activity modulates the response to aversive stimuli. A real-time functional magnetic resonance imaging study. *Biol. Psychiatry* 68, 425–432. doi: 10.1016/j.biopsych.2010.04.020
- Caria, A., Veit, R., Sitaram, R., Lotze, M., Weiskopf, N., Grodd, W., et al. (2007). Regulation of anterior insular cortex activity using real-time fMRI. *Neuroimage* 35, 1238–1246. doi: 10.1016/j.neuroimage.2007.01.018
- Connolly, L., Coveleskie, K., Kilpatrick, L. A., Labus, J. S., Ebrat, B., Stains, J., et al. (2013). Differences in brain responses between lean and obese women to a sweetened drink. *Neurogastroenterol. Motil.* 25, 579–e460. doi: 10.1111/nmo.12125
- Craig, A. D. (2005). Forebrain emotional asymmetry: a neuroanatomical basis? *Trends Cogn. Sci.* 9, 566–571. doi: 10.1016/j.tics.2005.10.005
- Craig, A. D. (2009). How do you feel now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70. doi: 10.1038/nrn2555
- De Araujo, I. E., and Rolls, E. T. (2004). Representation in the human brain of food texture and oral fat. *J. Neurosci.* 24, 3086–3093. doi: 10.1523/jneurosci.0130-04.2004
- De Araujo, I. E., Rolls, E. T., Kringelbach, M. L., Mcglone, F., and Phillips, N. (2003). Taste-olfactory convergence, and the representation of the pleasantness of flavour, in the human brain. *Eur. J. Neurosci.* 18, 2059–2068. doi: 10.1046/j.1460-9568.2003.02915.x
- De Araujo, I. E., and Simon, S. A. (2009). The gustatory cortex and multisensory integration. *Int. J. Obes. (Lond)* 33(Suppl. 2), S34–S43. doi: 10.1038/ijo.2009.70
- Deen, B., Pitskel, N. B., and Pelphrey, K. A. (2011). Three systems of insular functional connectivity identified with cluster analysis. *Cereb. Cortex* 21, 1498–1506. doi: 10.1093/cercor/bhq186
- Del Parigi, A., Gautier, J. F., Chen, K., Salbe, A. D., Ravussin, E., Reiman, E., et al. (2002). Neuroimaging and obesity: mapping the brain responses to hunger and satiation in humans using positron emission tomography. *Ann. N Y Acad. Sci.* 967, 389–397. doi: 10.1111/j.1749-6632.2002.tb04294.x
- Delparigi, A., Chen, K., Salbe, A. D., Reiman, E. M., and Tataranni, P. A. (2005). Sensory experience of food and obesity: a positron emission tomography study of the brain regions affected by tasting a liquid meal after a prolonged fast. *Neuroimage* 24, 436–443. doi: 10.1016/j.neuroimage.2004.08.035
- Frank, S., Laharnar, N., Kullmann, S., Veit, R., Canova, C., Hegner, Y. L., et al. (2010). Processing of food pictures: influence of hunger, gender and calorie content. *Brain Res.* 1350, 159–166. doi: 10.1016/j.brainres.2010.04.030
- Frank, S., Lee, S., Preissl, H., Schultes, B., Birbaumer, N., and Veit, R. (2012a). The obese brain athlete: self-regulation of the anterior insula in adiposity. *PLoS One* 7:e42570. doi: 10.1371/journal.pone.0042570
- Frank, S., Linder, K., Kullmann, S., Heni, M., Ketterer, C., Cavusoglu, M., et al. (2012b). Fat intake modulates cerebral blood flow in homeostatic and gustatory brain areas in humans. *Am. J. Clin. Nutr.* 95, 1342–1349. doi: 10.3945/ajcn.111.031492
- Frank, S., Wilms, B., Veit, R., Ernst, B., Thurnheer, M., Kullmann, S., et al. (2013). Altered brain activity in severely obese women may recover after Roux-en Y gastric bypass surgery. *Int. J. Obes. (Lond)* doi: 10.1038/ijo.2013.60. [Epub ahead of print].
- Gallay, D. S., Gallay, M. N., Jeanmonod, D., Rouiller, E. M., and Morel, A. (2012). The insula of reil revisited: multiarchitectonic organization in macaque monkeys. *Cereb. Cortex* 22, 175–190. doi: 10.1093/cercor/bhr104
- García-García, I., Jurado, M. A., Garolera, M., Segura, B., Sala-Llonch, R., Marques-Iturria, I., et al. (2012). Alterations of the salience network in obesity: a resting-state fMRI study. *Hum. Brain Mapp.* doi: 10.1002/hbm.22104. [Epub ahead of print].

## CONCLUSION

In conclusion, the insular cortex, especially the anterior part, is a multimodal and integrative area for the processing of food-related items. Central gustatory processes are tightly linked to interoception represented in reduced awareness of bodily signals including satiety signals. Therefore, interoception is associated with eating behavior and consequently also with obesity and eating disorders. In fact, multiple functions integrated in the insular cortex correlate and interact with gustatory processes. It has been shown, that obese subjects show higher responses in the anterior insular cortex to food cues independent of the modality (taste, visual). Moreover, rtfMRI guided neurofeedback training of the insular cortex raises the possibility to modify eating behavior.

## ACKNOWLEDGMENTS

This work was supported by the “Kompetenznetz Adipositas (Competence Network for Adiposity)” funded by the German Federal Ministry of Education and Research (FKZ: 01GI1122F). In addition, the authors acknowledge support of the Helmholtz Alliance ICAMED-Imaging and Curing Environmental Metabolic Diseases and Open Access Publishing Fund of Tübingen University.

- Gautier, J. F., Chen, K., Salbe, A. D., Bandy, D., Pratley, R. E., Heiman, M., et al. (2000). Differential brain responses to satiation in obese and lean men. *Diabetes* 49, 838–846. doi: 10.2337/diabetes.49.5.838
- Gautier, J. F., Del Parigi, A., Chen, K., Salbe, A. D., Bandy, D., Pratley, R. E., et al. (2001). Effect of satiation on brain activity in obese and lean women. *Obes. Res.* 9, 676–684. doi: 10.1038/oby.2001.92
- Henkin, R. I., and Levy, L. M. (2001). Lateralization of brain activation to imagination and smell of odors using functional magnetic resonance imaging (fMRI): left hemispheric localization of pleasant and right hemispheric localization of unpleasant odors. *J. Comput. Assist. Tomogr.* 25, 493–514. doi: 10.1097/00004728-200107000-00001
- Herbert, B. M., Muth, E. R., Pollatos, O., and Herbert, C. (2012). Interoception across modalities: on the relationship between cardiac awareness and the sensitivity for gastric functions. *PLoS One* 7:e36646. doi: 10.1371/journal.pone.0036646
- Higgs, S., and Donohoe, J. E. (2011). Focusing on food during lunch enhances lunch memory and decreases later snack intake. *Appetite* 57, 202–206. doi: 10.1016/j.appet.2011.04.016
- Kullmann, S., Heni, M., Veit, R., Ketterer, C., Schick, F., Häring, H. U., et al. (2012). The obese brain: association of body mass index and insulin sensitivity with resting state network functional connectivity. *Hum. Brain Mapp.* 33, 1052–1061. doi: 10.1002/hbm.21268
- Kullmann, S., Pape, A. A., Heni, M., Ketterer, C., Schick, F., Häring, H. U., et al. (2013). Functional network connectivity underlying food processing: disturbed salience and visual processing in overweight and obese adults. *Cereb. Cortex* 23, 1247–1256. doi: 10.1093/cercor/bhs124
- Kurth, F., Zilles, K., Fox, P. T., Laird, A. R., and Eickhoff, S. B. (2010). A link between the systems: functional differentiation and integration within the human insula revealed by meta-analysis. *Brain Struct. Funct.* 214, 519–534. doi: 10.1007/s00429-010-0255-z
- Menon, V., and Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* 214, 655–667. doi: 10.1007/s00429-010-0262-0
- Mesulam, M. M., and Mufson, E. J. (1985). “The insula of reil in man and monkey: architectonics, connectivity and function,” in *Cerebral Cortex*, eds A. Peters and E. O. Jones. (New York: Plenum Press), 179–226.
- Morewedge, C. K., Huh, Y. E., and Vosgerau, J. (2010). Thought for food: imagined consumption reduces actual consumption. *Science* 330, 1530–1533. doi: 10.1126/science.1195701
- Murdaugh, D. L., Cox, J. E., Cook, E. W., 3rd, and Weller, R. E. (2012). fMRI reactivity to high-calorie food pictures predicts short- and long-term outcome in a weight-loss program. *Neuroimage* 59, 2709–2721. doi: 10.1016/j.neuroimage.2011.10.071
- Nummenmaa, L., Hirvonen, J., Hannukainen, J. C., Immonen, H., Lindroos, M. M., Salminen, P., et al. (2012). Dorsal striatum and its limbic connectivity mediate abnormal anticipatory reward processing in obesity. *PLoS One* 7:e31089. doi: 10.1371/journal.pone.0031089
- Nunn, K., Frampton, I., Fuglset, T. S., Torzsok-Sonnevend, M., and Lask, B. (2011). Anorexia nervosa and the insula. *Med. Hypotheses* 76, 353–357. doi: 10.1016/j.mehy.2010.10.038
- Oberndorfer, T. A., Frank, G. K., Simmons, A. N., Wagner, A., Mccurdy, D., Fudge, J. L., et al. (2013). Altered insula response to sweet taste processing after recovery from anorexia and bulimia nervosa. *Am. J. Psychiatry* doi: 10.1176/appi.ajp.2013.11111745. [Epub ahead of print].
- Pelchat, M. L. (1997). Food cravings in young and elderly adults. *Appetite* 28, 103–113. doi: 10.1006/appe.1996.0063
- Pelchat, M. L., Johnson, A., Chan, R., Valdez, J., and Ragland, J. D. (2004). Images of desire: food-craving activation during fMRI. *Neuroimage* 23, 1486–1493. doi: 10.1016/j.neuroimage.2004.08.023
- Porubska, K., Veit, R., Preissl, H., Fritsche, A., and Birbaumer, N. (2006). Subjective feeling of appetite modulates brain activity: an fMRI study. *Neuroimage* 32, 1273–1280. doi: 10.1016/j.neuroimage.2006.04.216
- Rolls, E. T. (2005). Taste, olfactory, and food texture processing in the brain, and the control of food intake. *Physiol. Behav.* 85, 45–56. doi: 10.1016/j.physbeh.2005.04.012
- Rolls, E. T. (2006). Brain mechanisms underlying flavour and appetite. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 361, 1123–1136. doi: 10.1098/rstb.2006.1852
- Rothmund, Y., Preuschhof, C., Bohner, G., Bauknecht, H. C., Klingebiel, R., Flor, H., et al. (2007). Differential activation of the dorsal striatum by high-calorie visual food stimuli in obese individuals. *Neuroimage* 37, 410–421. doi: 10.1016/j.neuroimage.2007.05.008
- Ruijschop, R. M., Boelrijk, A. E., De Graaf, C., and Westerterp-Plantenga, M. S. (2009). Retronasal aroma release and satiation: a review. *J. Agric. Food Chem.* 57, 9888–9894. doi: 10.1021/jf901445z
- Schrammuller, W., Ubel, S., Ebner, F., and Schienle, A. (2012). Appetite regulation during food cue exposure: a comparison of normal-weight and obese women. *Neurosci. Lett.* 518, 106–110. doi: 10.1016/j.neulet.2012.04.063
- Schienle, A., Schafer, A., Hermann, A., and Vaitl, D. (2009). Binge-eating disorder: reward sensitivity and brain activation to images of food. *Biol. Psychiatry* 65, 654–661. doi: 10.1016/j.biopsych.2008.09.028
- Simmons, W. K., Avery, J. A., Barcalow, J. C., Bodurka, J., Drevets, W. C., and Bellgowan, P. (2012). Keeping the body in mind: insula functional organization and functional connectivity integrate interoceptive, exteroceptive, and emotional awareness. *Hum. Brain Mapp.* doi: 10.1002/hbm.22113. [Epub ahead of print].
- Small, D. M. (2010). Taste representation in the human insula. *Brain Struct. Funct.* 214, 551–561. doi: 10.1007/s00429-010-0266-9
- Small, D. M. (2012). Flavor is in the brain. *Physiol. Behav.* 107, 540–552. doi: 10.1016/j.physbeh.2012.04.011
- Small, D. M., and Green, B. G. (2012). “A proposed model of a flavor modality,” in *The Neural Bases of Multisensory Processes*, eds M. M. Murray and M. T. Wallace. (Boca Raton (FL): CRC Press), 717–738.
- Small, D. M., Gregory, M. D., Mak, Y. E., Gitelman, D., Mesulam, M. M., and Parrish, T. (2003). Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron* 39, 701–711. doi: 10.1016/s0896-6273(03)00467-7
- Small, D. M., and Prescott, J. (2005). Odor/taste integration and the perception of flavor. *Exp. Brain Res.* 166, 345–357. doi: 10.1007/s00221-005-2376-9
- Small, D. M., Zald, D. H., Jones-Gotman, M., Zatorre, R. J., Pardo, J. V., Frey, S., et al. (1999). Human cortical gustatory areas: a review of functional neuroimaging data. *Neuroreport* 10, 7–14. doi: 10.1097/00001756-199901180-00002
- Stevenson, R. J., Miller, L. A., and Mcgrillen, K. (2013). The lateralization of gustatory function and the flow of information from tongue to cortex. *Neuropsychologia* 51, 1408–1416. doi: 10.1016/j.neuropsychologia.2013.04.010
- Stice, E., Spoor, S., Bohon, C., Veldhuizen, M. G., and Small, D. M. (2008). Relation of reward from food intake and anticipated food intake to obesity: a functional magnetic resonance imaging study. *J. Abnorm. Psychol.* 117, 924–935. doi: 10.1037/a0013600
- Stice, E., Spoor, S., Ng, J., and Zald, D. H. (2009). Relation of obesity to consummatory and anticipatory food reward. *Physiol. Behav.* 97, 551–560. doi: 10.1016/j.physbeh.2009.03.020
- Stoeckel, L. E., Weller, R. E., Cook, E. W., 3rd, Twieg, D. B., Knowlton, R. C., and Cox, J. E. (2008). Widespread reward-system activation in obese women in response to pictures of high-calorie foods. *Neuroimage* 41, 636–647. doi: 10.1016/j.neuroimage.2008.02.031
- Sweet, L. H., Hassenstab, J. J., McCaffery, J. M., Raynor, H. A., Bond, D. S., Demos, K. E., et al. (2012). Brain response to food stimulation in obese, normal weight, and successful weight loss maintainers. *Obesity (Silver Spring)* 20, 2220–2225. doi: 10.1038/oby.2012.125
- Taylor, K. S., Seminowicz, D. A., and Davis, K. D. (2009). Two systems of resting state connectivity between the insula and cingulate cortex. *Hum. Brain Mapp.* 30, 2731–2745. doi: 10.1002/hbm.20705
- Teufel, M., Stephan, K., Kowalski, A., Kasberger, S., Enck, P., Zipfel, S., et al. (2013). Impact of biofeedback on self-efficacy and stress reduction in obesity: a randomized controlled pilot study. *Appl. Psychophysiol. Biofeedback* doi: 10.1007/s10484-013-9223-8. [Epub ahead of print].
- Van De Sande-Lee, S., Pereira, F. R., Cintra, D. E., Fernandes, P. T., Cardoso, A. R., Garlipp, C. R., et al. (2011). Partial reversibility of hypothalamic dysfunction and changes in brain activity after body mass reduction in obese subjects. *Diabetes* 60, 1699–1704. doi: 10.2337/db10-1614
- Veit, R., Singh, V., Sitaram, R., Caria, A., Rauss, K., and Birbaumer, N. (2012). Using real-time fMRI to learn voluntary regulation of the anterior insula in the presence of

- threat-related stimuli. *Soc. Cogn. Affect. Neurosci.* 7, 623–634. doi: 10.1093/scan/nsr061
- Veldhuizen, M. G., Albrecht, J., Zelano, C., Boesveldt, S., Breslin, P., and Lundstrom, J. N. (2011). Identification of human gustatory cortex by activation likelihood estimation. *Hum. Brain Mapp.* 32, 2256–2266. doi: 10.1002/hbm.21188
- Vocks, S., Herpertz, S., Rosenberger, C., Senf, W., and Gizewski, E. R. (2011). Effects of gustatory stimulation on brain activity during hunger and satiety in females with restricting-type anorexia nervosa: an fMRI study. *J. Psychiatr. Res.* 45, 395–403. doi: 10.1016/j.jpsychires.2010.07.012
- Wager, T. D., and Barrett, L. F. (2004). From affect to control: functional specialization of the insula in motivation and regulation. *Published online at PscExtra.*
- Wagner, A., Aizenstein, H., Mazurkewicz, L., Fudge, J., Frank, G. K., Putnam, K., et al. (2008). Altered insula response to taste stimuli in individuals recovered from restricting-type anorexia nervosa. *Neuropsychopharmacology* 33, 513–523. doi: 10.1038/sj.npp.1301443
- Wang, G. J., Tomasi, D., Backus, W., Wang, R., Telang, F., Geliebter, A., et al. (2008). Gastric distention activates satiety circuitry in the human brain. *Neuroimage* 39, 1824–1831. doi: 10.1016/j.neuroimage.2007.11.008
- Yokum, S., Ng, J., and Stice, E. (2011). Attentional bias to food images associated with elevated weight and future weight gain: an FMRI study. *Obesity* 19, 1775–1783. doi: 10.1038/oby.2011.168
- Ziauddeen, H., Farooqi, I. S., and Fletcher, P. C. (2012). Obesity and the brain: how convincing is the addiction model? *Nat. Rev. Neurosci.* 13, 279–286. doi: 10.1038/nrn3212
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 26 June 2013; accepted: 05 August 2013; published online: 23 August 2013.
- Citation:* Frank S, Kullmann S and Veit R (2013) Food related processes in the insular cortex. *Front. Hum. Neurosci.* 7:499. doi: 10.3389/fnhum.2013.00499
- This article was submitted to the journal *Frontiers in Human Neuroscience*.
- Copyright © 2013 Frank, Kullmann and Veit. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# The sensory channel of presentation alters subjective ratings and autonomic responses toward disgusting stimuli—Blood pressure, heart rate and skin conductance in response to visual, auditory, haptic and olfactory presented disgusting stimuli

Ilona Croy<sup>1,2\*</sup>, Kerstin Laqua<sup>1</sup>, Frank Süß<sup>3</sup>, Peter Joraschky<sup>2</sup>, Tjalf Ziemssen<sup>4</sup> and Thomas Hummel<sup>1</sup>

<sup>1</sup> Department of Otorhinolaryngology, Smell and Taste Clinic, University of Dresden Medical School, Dresden, Germany

<sup>2</sup> Department of Psychosomatic Medicine, University of Dresden Medical School, Dresden, Germany

<sup>3</sup> Department of Occupational and Social Medicine, University of Dresden Medical School, Dresden, Germany

<sup>4</sup> Center of Clinical Neuroscience, Neurological University Clinic, University of Dresden Medical School, Dresden, Germany

## Edited by:

Martin Klasen, RWTH Aachen University, Germany

## Reviewed by:

Ruthger Righart, Institute for Stroke and Dementia Research, Germany  
Tamer Demiralp, Istanbul University, Turkey

## \*Correspondence:

Ilona Croy, Department of Occupational and Environmental Medicine, University of Gothenburg, Medicinargatan 16, Box 414, 405 30 Gothenburg, Sweden  
e-mail: ilona.croy@amm.gu.se

Disgust causes specific reaction patterns, observable in mimic responses and body reactions. Most research on disgust deals with visual stimuli. However, pictures may cause another disgust experience than sounds, odors, or tactile stimuli. Therefore, disgust experience evoked by four different sensory channels was compared. A total of 119 participants received 3 different disgusting and one control stimulus, each presented through the visual, auditory, tactile, and olfactory channel. Ratings of evoked disgust as well as responses of the autonomic nervous system (heart rate, skin conductance level, systolic blood pressure) were recorded and the effect of stimulus labeling and of repeated presentation was analyzed. Ratings suggested that disgust could be evoked through all senses; they were highest for visual stimuli. However, autonomic reaction toward disgusting stimuli differed according to the channel of presentation. In contrast to the other, olfactory disgust stimuli provoked a strong decrease of systolic blood pressure. Additionally, labeling enhanced disgust ratings and autonomic reaction for olfactory and tactile, but not for visual and auditory stimuli. Repeated presentation indicated that participant's disgust rating diminishes to all but olfactory disgust stimuli. Taken together we argue that the sensory channel through which a disgust reaction is evoked matters.

**Keywords:** disgust, olfaction, vision, audition, touch, rating, heart rate, blood pressure

## INTRODUCTION

Disgust is ranked among the basic emotions of humans. One of the most popular theories states that there are six inborn basic emotions in human, which are present across all cultures. These are happiness, anger, disgust, sadness, fear, and surprise (Ekman et al., 1983). Although it has been questioned if emotions can be categorized in such a way [see Barrett et al. (2007) and Panksepp (2007) for ongoing debate], it is undisputed that all healthy persons are able to feel disgust [for overview see Rozin et al. (2000) and Tybur et al. (2009)].

Emotions can be evoked by environmental cues and visual or visual-auditory material often serves as emotional trigger in an experiment [for overview, see for instance Kreibig (2010)]. Coming from olfactory research, we observed that odorous cues have a high potential to evoke disgust. It might even be one of the key functions of the olfactory system to warn about microbial threats by evoking disgust (Stevenson, 2010). There are some studies indicating that although not every emotion can be induced easily using odors, disgust can be evoked reliably by the sense of smell (Alaoui-Ismaili et al., 1997b; Bensafi et al., 2002; Croy et al., 2011).

Based on this, we wondered if the sensory channel of presentation contributes to emotional experience. Darwin already noted 150 years ago that different senses may have a special relation toward disgust. He defined disgust as “something revolting, primary in relation to the senses of taste and smell, as actually perceived or vividly imagined; and secondarily to anything which causes a similar feeling, through the sense of smell, touch, and even eyesight” (Darwin, 1872). Nevertheless, most research on disgust deals with pictures or videos [for overview see Kreibig (2010)].

Emotional stimuli are processed in two stages: First persons orient to the sensory input and process the contextual details. Heart rate (HR) decelerates and skin conductance (SCL) decreases mirroring the parasympathetic-sympathetic co-activated orienting reaction. Then the relevant information is retrieved from memory and the participants implicitly prepare for relevant action (Bradley et al., 2001). For threatening stimuli, orienting is normally followed by a sympathetic driven increase of HR preparing the body for the fight- or flight reaction (Bradley et al., 2001). Disgusting stimuli however might have other behavioral requirements. The typical disgust elicitors are spoiled

food, illness-related stimuli and feces (Rozin and Fallon, 1987; Rozin et al., 2000; Vaitl et al., 2005). Instead of a fast and sympathetic dominated typical fight- or flight reaction another behavior seems reasonable: Going away from the source or removing the source from the body, for instance by vomiting. In fact, disgust causes specific reaction patterns, observable in mimics and typical body reaction up to regurgitation (Rozin and Fallon, 1987) and is accompanied by an increase of skin conductance level and a decrease of heart rate (Vaitl et al., 2005).

Why should disgust reaction differ according to the evoking sensory channel? The senses fulfill different functions, have different neurological pathways and access to explicit memory differs between the senses. Based on those considerations, we hypothesize that disgust response differs depending on the sensory channel of disgust perception.

The main function of disgust is avoidance of disease (Oaten et al., 2009). Therefore, disgust motivates rejection of potential health threatening objects especially from microbial sources, such as found in wounds, spoiled food and organic waste. Typical disgust objects seem to be in near distance: Most microbial threats have to be touched, inhaled or eaten to infiltrate the body powerfully. Consequently, we would expect that proximal senses, such as touch and olfaction can evoke strong disgust and produce enhanced reaction compared to stimuli processed through the sense of vision or audition.

Second, the senses use different neurological pathways. Specialized receptors of each sensory channel transform environmental inputs into electrical signals, which are transported to the related primary sensory cortex. In contrast to other senses, the olfactory system projects ipsilaterally and most fibers bypass the thalamus and project directly into amygdala, piriform cortex, and entorhinal cortex (Gottfried, 2006).

Third, although emotions can occur independently from cognition (Izard, 1992), it is undisputed that cognition influences emotional experience. Because of the organization of working memory, visual and verbal cues can be identified easily (Baddeley and Hitch, 1974). Environmental stimuli processed through the visual and auditory channel may therefore trigger much contextual information. This helps in selecting the appropriate behavioral response. However, odors are not that easy to identify (Jonsson and Olsson, 2003). Accordingly, labeling has a strong influence on emotional rating of odors. For example, participants liked the very same odor significantly less and even processed it in a different way, when it was labeled “body odor” instead of “cheddar cheese” (De Araujo et al., 2005). We hypothesize that a label enhances disgust perception for olfactory, but not for visual and auditory stimuli.

Related to the assumptions above, objects presented by different channels might have a differential potential to stay in memory. If the same disgusting object is presented repeatedly, the reaction to this object might change according to the sensory channel. We argued that verbal and visual stimuli are easy to categorize, which enhances recognition in repeated presentation. A study conducted on aversive (fearful) stimuli for instance showed that autonomic response to aversive pictures decreased after some days (Tabibnia et al., 2008). For non-verbal auditory, olfactory, and tactile stimuli categorization and

therefore recognition may be more difficult. This may lead to slower habituation of emotional response in case of repeated presentation.

Taken together we have reason to assume that the sensory channel of presentation contributes to disgust reactions. Former studies indicate that disgust can be evoked through the olfactory (e.g., Bensafi et al., 2002), tactile (e.g., Hertenstein et al., 2009; Oum et al., 2011), and visual (e.g., Collet et al., 1997) channel as well as through a combination of the visual and auditory channels using film-clips (e.g., Kunzmann and Gruhn, 2005). However, whether disgust perception differs with regard to the sensory channel has—to the best of our knowledge—not been studied yet.

We attempted to systematically evaluate disgust reactions evoked by the visual, auditory, tactile, and olfactory sense<sup>1</sup>. Stimuli of the three most pronounced disgust categories were presented through each of the sensory channels and subjective perception and reactivity of the autonomic nervous system was analyzed. Previous studies examining autonomic reactivity for disgusting stimuli found differences in facial electromyographic activity (i.e., Bensafi et al., 2002), skin potential (i.e., Alaoui-Ismaili et al., 1997a; Collet et al., 1997), systolic blood pressure (i.e., Prkachin et al., 1999; Kunzmann and Gruhn, 2005), and heart rate (i.e., Alaoui-Ismaili et al., 1997a; Rohrmann et al., 2009), for overview see Kreibitz (2010). We concentrated on measurements of skin conductance, systolic blood pressure, and heart rate. In order to minimize a potential bias in stimulus selection for single stimuli, results of all three disgust categories were averaged for each sensory channel. To test the influence of cognition, semantic information was added to one third of the stimuli. The experiment was repeated twice with a subgroup of the same participants.

## MATERIALS AND METHODS

### PARTICIPANTS

A total of 124 healthy people participated in the study, 5 had to be excluded from analysis because of technical problems with the autonomic measurement. Thus, data is analyzed from 119 participants (60 women, 59 men, age range 18–36 years, mean = 22.7 years; standard deviation = 3 years). Most of them were graduate students of the Technical University of Dresden. Completion of a detailed medical history form by each participant enabled confirmation of his or her good physical health. Normal olfactory function was ascertained in all participants using an olfactory screening test (Hummel et al., 2001, 2007).

In order to analyze the influence of repeated presentation, 43 sex matched participants—equally spread around the labeling condition (see below)—took part in two repetitions of the experiment. The study followed the Declaration of Helsinki on Biomedical Research Involving Human Subjects and was approved by the Ethics Committee from the University of Dresden Medical School. All participants provided written

<sup>1</sup>We do not examine the sense of gustation, because the pure sense of gustation only involves the qualities sour, sweet, salty, bitter, and umami. Although bitterness might provide an excellent disgust induction technique, it is not possible to present the aforementioned disgust categories purely through the gustatory channel. In order to keep the design comparable across the senses, we decided not to examine taste.

informed consent. They received a small amount of money for their participation.

## MATERIALS

Disgust and control stimuli were presented through the visual, auditory, tactile, and olfactory channel. To enhance the validity of the experiment, disgusting stimuli of three different categories were chosen for each sensory channel: “spoiled food,” “illness related,” and “feces.” Choice of categories was based on the current literature (Rozin and Fallon, 1987; Rozin et al., 2000; Vaitl et al., 2005). The study did not aim to compare different disgust categories, but different categories were presented to enhance the validity of the experiment and merged for analysis. For control purpose, stimuli with low emotional value were used. The sound of a person writing from the International Affective Digital Sounds database (358) was selected and accordingly a picture of a pencil was used for visual stimulation and pencil for tactile stimulation. As there is no matching odor, we decided to use chocolate odor [Bell flavors and Fragrances (diluted up to 66% in polypropylenglycol)], which is—based on previous experience in our laboratory—perceived as rather neutral.

A description of the stimuli is presented in **Table 1**. The order of the 16 different presentations in total was pseudo-randomized across participants. Each participant received one disgust category labeled while the other two were presented without the participants’ knowing what the stimulus was supposed to be. For instance for participant A all “spoiled food” stimuli had a label, when presented through the four channels. But none of the “illness related” or “feces” stimuli was presented labeled. Order of labeling was randomized across participants. In case of repeated presentation the same stimuli were labeled for each participant.

## RATINGS

Ratings for *arousing and hedonic qualities* of the stimuli was assessed using the Self Assessment Manikins (Lang, 1980), a visual analog rating scale which prompts responses for arousal and valence ratings on a 9-point-scale, whereby “1” means “not at all” pleasant or arousing and “9” means “extremely” pleasant or arousing.

To judge which *emotion* was evoked by the stimuli, participants were asked to state to which degree the presented stimulus evoked the following five basic emotions: happiness, disgust, anger, anxiety, and sadness. Although listed as one of the basic emotions,

we decided not to analyze “surprise.” Among the proposed basic emotions by Ekman et al. (1983), surprise is discussed most controversial and might rather reflect orienting reaction (Posner et al., 2005). Ratings were given on an analog rating scale from 1 to 9 (the emotion is experienced “not at all” to “extremely strong”).

## AUTONOMIC MEASUREMENTS

Recordings were performed in a sitting position in an air-conditioned laboratory. Food intake and consumption of caffeine or nicotine had to be stopped at least 1 h before the examination. Continuous monitoring of heart rate, blood pressure (COLIN, Ohmeda), skin conductance, and respiration was performed using the SUEMPATHY device (SUEmpathie100, SUESS Medizin-Technik, Aue, Germany). The sampling frequency was 512 Hz for each channel. After 5 min of calibration and a test measurement, data acquisition commenced with the beginning of the experiment and lasted until the end of the presentation of all stimuli.

## PROCEDURE

In total each participant received 24 stimuli; 12 control (4 sensory channels  $\times$  3 repetitions) + 12 disgust (4 sensory channels  $\times$  1 labeled disgust category + 4 sensory channels  $\times$  2 unlabeled disgust categories) in one out of 12 predefined orders. To attenuate the possibly aversive effect of repeated presentation of disgusting stimuli, each disgusting stimulus was followed by a control. However, only data from the first occurrence of each control stimulus was rated by the participants and only this one was analyzed with respect to autonomic measurements. Each participant received one of the disgust categories labeled (see Materials).

### *Stimuli were presented in the following way*

The participants sat relaxed in a distance of about 70 cm in front of a monitor on which all instructions were presented. Prior to each stimulus participants read one of the following instructions “You are going to see/hear/touch/smell something”<sup>2</sup>. After this

<sup>2</sup>In case of labeling the instruction was modified. For instance in the disgust category “spoiled food” the label was modified to: (“you are going to see a picture of rotten food/are going to hear a person who ate something spoiled/going to smell rotten food/going to touch rotten food”). For category “illness related” the instruction was changed so that the participants were warned about a “sick person,” in feces category they were warned about “feces.”

**Table 1 | Description of disgust and control stimuli.**

	Vision	Audition	Touch	Olfaction
Spoiled food	Picture of spoiled pasta*	Sound of a woman vomiting*	Rubber pasta*	Carbon disulfide
Illness related	Picture of melanom at the toe	Sound of a women coughing	Tissue smeared with soap*	Artificial sweat
Feces	Picture of feculent toilet	Sound of a person with diarrhea*	Feces made of flour and water*	Odor of feces
Control	Picture of a pencil*	Sound of a person writing	Pencil*	Odor of chocolate

Olfactory stimuli are diluted in polypropylenglycol. In detail: picture of melanom taken from *Duale Reihe Dermatologie*, by courtesy of Thieme Verlag, picture of feculent toilet taken from the International Affective Picture System (IAPS, 9301); sound of coughing women International Affective Digital Sounds (IADS, 242), sound of a person writing IADS (358); the following odors were used: Civette Base 847 (Fragrance Resources, Hamburg, Germany diluted up to 5% in polypropylenglycol), carbon disulfide (order # 180173; Sigma, Deisenhofen, Germany), artificial sweat (2-methyl-3-mercapto-butanol; Unilever, Port Sunlight, UK), Chocolate odor, Bell flavors and Fragrances, (diluted up to 66% in polypropylenglycol). All items marked with \* were made by the experimenters.

the screen went black for 10 s. For the next 10 s the stimuli were presented. Pictures were shown on the monitor, auditory stimuli were given through earphones. During the whole experiment, the participants stretched their left arm out with palms in a supine position. When tactile stimuli were presented, they were placed on the participants palm in order to allow the participants to touch these stimuli, without seeing them. Liquid odors were presented in small glass bottles filled with cotton pads. The experimenter put the opened bottles beneath the nose of the participants. The whole procedure was practiced with each participant before the experiment started. After 10 s of presentation, the stimuli were removed and a black screen was presented for 30 s. After this interval, participants were asked to judge the arousal and valence of the stimuli presented as well as the emotions evoked by the stimuli. The rating scales were presented via monitor and participants told their judgment to the experimenter who typed them in. There was no time limit to provide the subjective ratings. After ratings, the instruction for the next stimulus was presented.

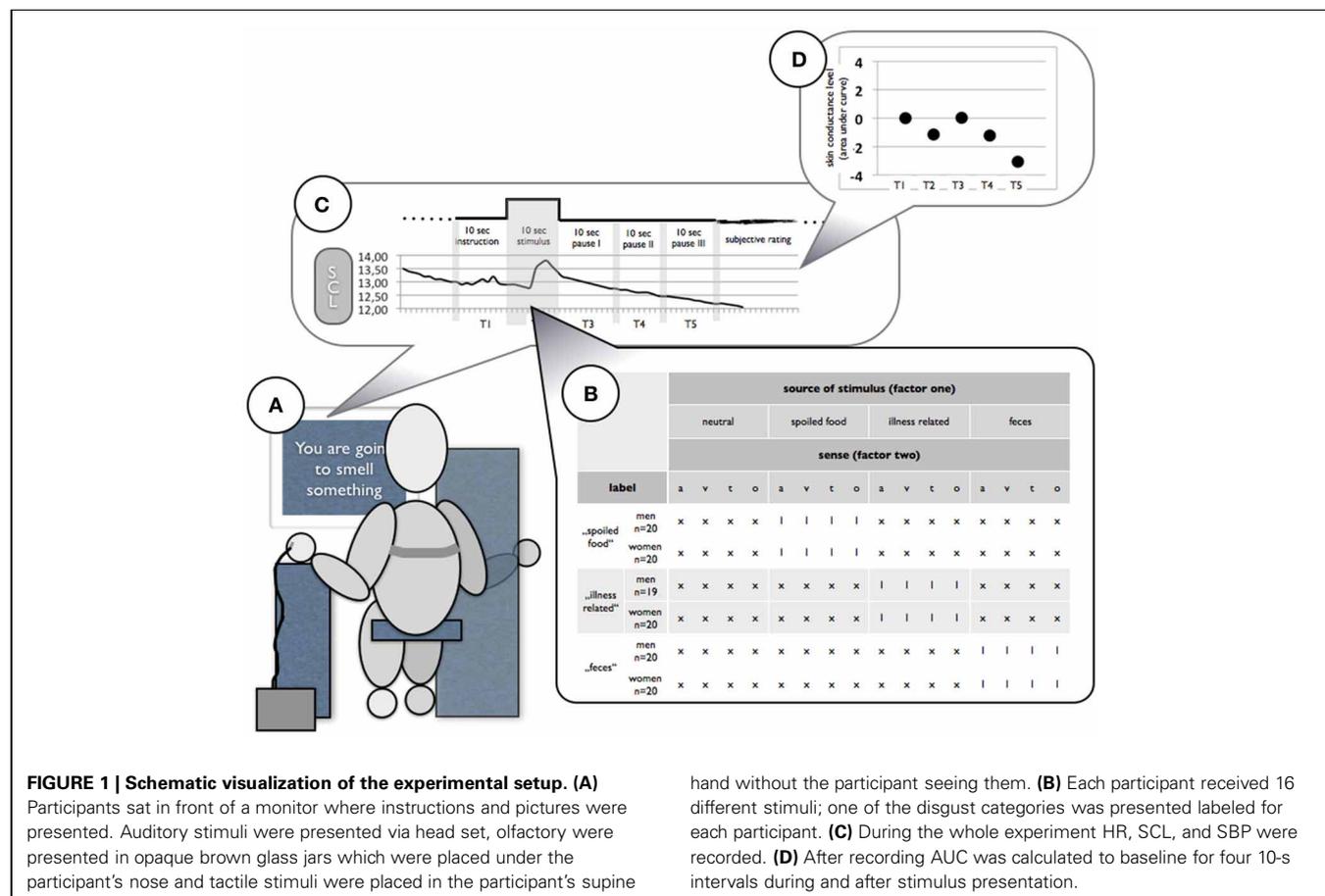
Two repetitions took place about 2 and 4 weeks after the first experiment. Order of stimulus presentation remained unchanged.

## ANALYSIS METHODS

Off-line, a trained observer identified recordings with artifacts which were excluded from further analysis. Afterwards, data

analysis was performed as area under the curve-analysis (AUC). The AUC measure was chosen to reduce large inter-individual variance as it is present in SCL amplitudes for instance. AUC analysis has been shown to be a better predictor of autonomic arousal than conventional analysis (Bach et al., 2010). The mean of data at T1 (10 s before the stimulus onset) served as baseline and AUC for data at T2 (1–10 s; during stimulus presentation), T3 (11–20 s, pause 1 after stimulus presentation), T4 (21–30 s, pause 2 after stimulus presentation), and T5 (31–40 s, pause 3 after stimulus presentation) was calculated according to this baseline (see **Figure 1**). We decided to split data into 10 s intervals in order to see potential autonomic correlation of both reactions: orienting and action preparing (Bradley et al., 2001). In result, a matrix was generated for each measurement (heart rate [HR], systolic blood pressure [SBP], and skin conductance level [SCL]), which encompassed 119 participants and AUC-data at T1, T2, T3, T4, T5 for each of 16 stimuli. For presentation purpose, the AUC was afterwards corrected by the factor 0.1, reflecting the mean increase or decrease in the 10 s interval.

All data were analyzed using the SPSS 19 Software (SPSS Inc., Chicago, IL, USA). Data of the three different disgust categories were combined by averaging the responses to two unlabeled disgust stimuli of each sensory channel. Comparisons between evoked emotions and between control and disgusting



hand without the participant seeing them. **(B)** Each participant received 16 different stimuli; one of the disgust categories was presented labeled for each participant. **(C)** During the whole experiment HR, SCL, and SBP were recorded. **(D)** After recording AUC was calculated to baseline for four 10-s intervals during and after stimulus presentation.

stimuli were performed using ANOVA for repeated measurements with the within-subject-factors “disgust” (control vs. disgust) and “sense” (4). For autonomic data, timeline served as additional covariate representing the 4 measurement points after baseline.

The effect of label was analyzed for disgusting stimuli only, using “label” (2) as within-subject factor in the ANOVA. The effect of repeated stimulation was analyzed for those participants who took part in the experiment three times. Here, the first presentation was compared to the last one for all of the disgusting stimuli in the four different sensory channels. Only responses to unlabeled stimuli were analyzed. Level of significance was set at  $p = 0.05$ . Wherever appropriate, results are presented Bonferroni-corrected to minimize influence of multiple testing. This is indicated by “ $p_{\text{corr}N}$ ,” with “ $N$ ” indicating the number of comparisons for which the  $P$ -value is corrected.

## RESULTS

### RATINGS OF DISGUST, VALENCE, AND AROUSAL

Ratings of evoked basic emotion are visualized in **Figure 2**. For each of the disgusting stimuli, disgust was evoked more than in the matching control [ $t_{(118)} = 8.2\text{--}44.3$ ;  $p_{\text{corr}12} < 0.01$ ] and disgust was the emotion evoked most strongly [ $t_{(118)} = 6.1\text{--}35.6$ ;  $p_{\text{corr}4} < 0.01$ ], except for tactile-“spoiled food.” Here, disgust ratings were not significantly higher than ratings for happiness [ $t_{(118)} = 3.2$ ;  $p_{\text{corr}4} = 0.06$ ]. Together, this justifies merging the disgust categories for the following results section.

Comparison of the sensory channels revealed that seeing disgusting stimuli led to higher ratings of disgust than hearing, touching, or smelling them. This was the case for each of the three disgust categories [ $t_{(118)} = 3.1\text{--}13.1$ ;  $p_{\text{corr}9} < 0.05$ ], except for spoiled food where pictures did not get significantly higher ratings than sounds [ $t_{(118)} = 1.4$ ;  $p_{\text{corr}9} > 0.05$ ].

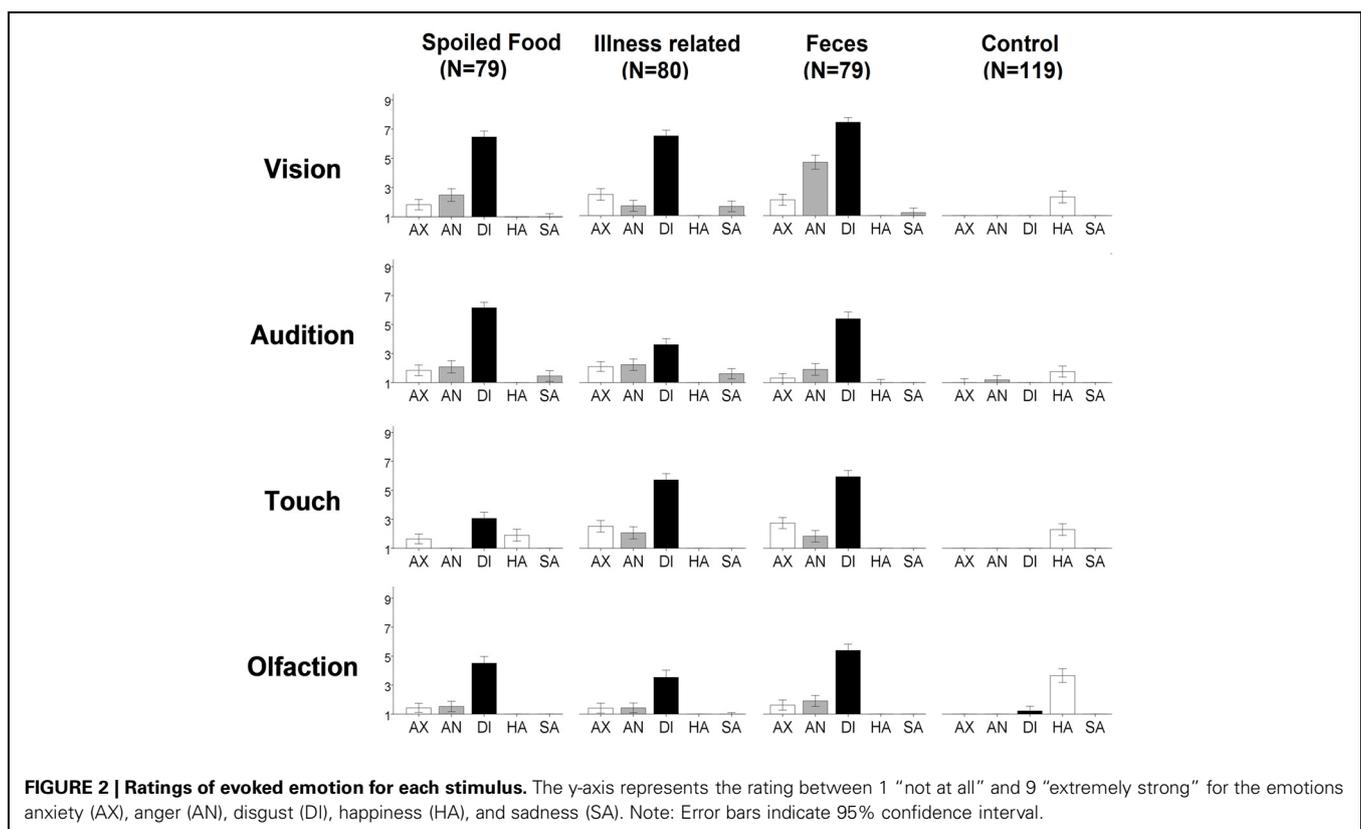
Ratings of hedonic valence and arousal are provided in **Table 2**. A main effect of disgust was revealed for valence and arousal, indicating that the disgusting stimuli were rated as less pleasant and more arousing compared to the controls. This was the case in all of the sensory channels [pleasantness:  $F_{(1, 118)} = 353.7\text{--}624.0$ ,  $p_{\text{corr}3} < 0.01$ ; arousal  $F_{(1, 118)} = 173.4\text{--}244.1$ ;  $p_{\text{corr}3} < 0.01$ ].

Focusing on the disgusting stimuli only, there was a significant main effect of the sensory channel, with visual stimuli being perceived as least pleasant [ $F_{(3, 116)} = 65.4$ ;  $p < 0.01$ ] and most arousing [ $F_{(3, 116)} = 23.2$ ;  $p < 0.01$ ].

### REACTIONS OF THE AUTONOMIC NERVOUS SYSTEM TO DISGUSTING STIMULI

#### Comparison between disgusting and the matching control stimuli

No significant main effect of sensory channel or disgust was revealed for HR but a significant interaction [ $F_{(3, 113)} = 3.6$ ,  $p = 0.01$ , compare **Table 3** and **Figure 3**]. *Post-hoc* test revealed a significant difference in the tactile channel. Here HR decreased in the disgust stimuli compared to the control [ $F_{(3, 113)} = 9.5$ ,  $p < 0.01$ ]. Similarly, for SCL a significant main effect of the sensory channel [ $F_{(3, 113)} = 10.0$ ,  $p < 0.01$ ] and a significant interaction [ $F_{(3, 113)} = 5.9$ ,  $p < 0.01$ ] was found. *Post-hoc* testing revealed



**Table 2 | Ratings of hedonic valence and arousal of the disgusting and control stimuli applied through different senses (N = 119).**

	Vision	Audition	Touch	Olfaction
Disgust	Valence: 2.8 ± 1.2 Arousal: 5.1 ± 1.8	Valence: 3.6 ± 1.1 Arousal: 4.5 ± 1.8	Valence: 4.7 ± 1.4 Arousal: 4.0 ± 1.8	Valence: 4.0 ± 1.3 Arousal: 3.8 ± 1.7
Control	Valence: 6.2 ± 1.3 Arousal: 1.9 ± 1.4	Valence: 5.5 ± 1.4 Arousal: 2.9 ± 1.9	Valence: 6.3 ± 1.4 Arousal: 2.4 ± 1.6	Valence: 6.5 ± 1.9 Arousal: 2.9 ± 1.9

Ratings are given on a 9-point scale with “1” meaning “not at all” pleasant respective arousing and “9” meaning “extremely” pleasant respective arousing. Only participants ratings for stimuli presented unlabeled are given in order to avoid possible label influences.

**Table 3 | Autonomic measurements.**

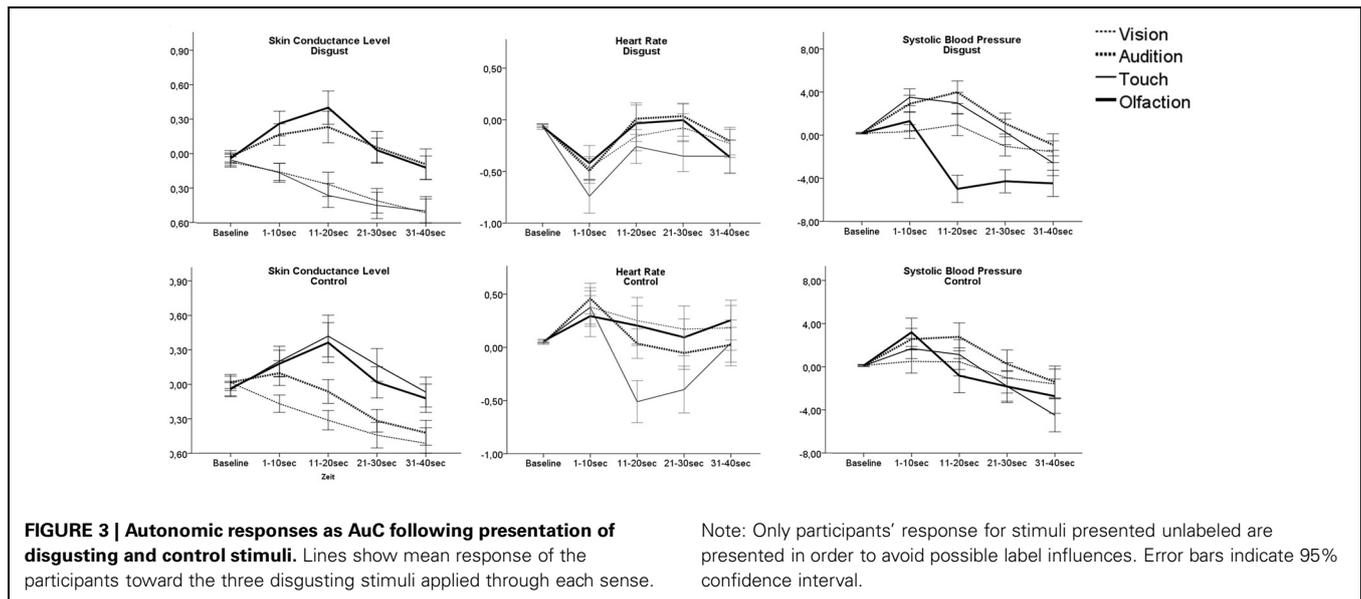
Presentation	Sense	Stimuli	SCL in $\mu$ S		HR in bpm		SBP in mmHG	
			Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
First presentation N = 116	Vision	Disgust	-0.17	0.61	-0.23	0.73	-2.8	44.6
		Control	-0.2	0.62	-0.25	1.13	-3.6	63.8
	Audition	Disgust	0.28	0.75	-0.16	0.76	<b>13.8</b>	<b>50.2</b>
		Control	0	0.64	-0.12	0.84	<b>8</b>	<b>64.3</b>
	Touch	Disgust	<b>-0.2</b>	<b>0.64</b>	<b>-0.43</b>	<b>0.88</b>	<b>8.1</b>	<b>55.3</b>
		Control	<b>0.37</b>	<b>0.88</b>	<b>0.12</b>	<b>1.13</b>	<b>-7.3</b>	<b>68.6</b>
	Olfaction	Disgust	0.34	0.71	-0.20	0.92	<b>-24.9</b>	<b>60.3</b>
		Control	0.3	0.83	-0.21	1.01	<b>-4.7</b>	<b>75.8</b>
Second presentation N = 47	Vision	Disgust	-0.39	0.72	-0.13	0.70	0.8	29
		Control	-0.12	0.43	-0.23	0.76	-1.8	42.4
	Audition	Disgust	0.13	0.71	0.03	0.68	4.4	26.2
		Control	-0.01	0.36	-0.31	0.71	5.2	47.1
	Touch	Disgust	-0.43	0.58	-0.49	0.84	1.8	36.1
		Control	0.15	0.64	0.18	0.93	-1.7	44.3
	Olfaction	Disgust	0.28	0.69	-0.16	0.85	-5.7	31.6
		Control	0.06	0.5	-0.28	0.94	-0.3	45.4
Third presentation N = 43	Vision	Disgust	-0.24	0.63	-0.10	0.66	0	24.4
		Control	-0.04	0.49	-0.09	1.08	4.4	51.8
	Audition	Disgust	0.04	0.65	-0.26	0.65	1.2	27.1
		Control	-0.03	0.43	-0.18	0.99	8.5	36.1
	Touch	Disgust	-0.35	0.66	-0.72	1.03	5	25.3
		Control	0.1	0.45	0.20	1.04	5.2	45.7
	Olfaction	Disgust	0.21	0.75	-0.09	0.70	-2.2	30.1
		Control	-0.02	0.54	-0.24	0.92	-3.3	37.3

In columns the changes of HR, SCL and SBP averaged over the four time points for are provided for the disgusting stimuli and the control stimulus applied through the sensory systems. For the disgust stimuli, only unlabeled presented stimuli are taken into account. Bold numbers indicate that there is a significant difference in autonomic response between the control and the disgusting stimuli in the sensory system ( $p < 0.05$ ).

that tactile disgusting stimuli led to significantly lower SCL compared to control [ $F_{(1, 115)} = 12.3, p < 0.01$ ]. For SBP a significant main effect of the sensory channel [ $F_{(3, 113)} = 39.6, p < 0.01$ ] and a significant interaction [ $F_{(3, 113)} = 22.9, p < 0.01$ ] was found. *Post-hoc* testing revealed that tactile disgust stimuli led to enhanced SBP compared to control [ $F_{(1, 115)} = 21.6, p < 0.01$ ] and auditory disgust stimuli tended to lead to enhanced SBP compared to control [ $F_{(1, 115)} = 3.5, p = 0.06$ ]. Olfactory disgust stimuli on the other side led to reduced SBP compared to control [ $F_{(1, 115)} = 36.7, p < 0.01$ ].

### Comparison of disgusting stimuli between sensory channels

Autonomic reaction patterns differed between the sensory channels for all measurements: SCL [ $F_{(3, 113)} = 98.2, p < 0.01$ ], HR [ $F_{(3, 113)} = 10.4, p < 0.01$ ], and SBP [ $F_{(3, 113)} = 80.4, p < 0.01$ ]. *Post-hoc* testing showed that SCL responses toward olfactory and auditory stimuli differed from responses to tactile and visual stimuli ( $p_{\text{corr6}} < 0.01$ ). Olfactory and auditory disgust stimuli led to a SCL peak 10–20 s after stimulus presentation, while tactile and visual stimuli lead to a slow decrease of SCL. For HR the decrease was strongest for tactile disgust stimuli ( $p_{\text{corr6}} < 0.01$ ), while



there was no significant difference between the other sensory stimuli.

For SBP, responses toward visual stimuli could be differentiated from others ( $p_{\text{corr6}} < 0.01$ ) by leading to little SBP change, while disgusting stimuli applied through the auditory or tactile channel were followed by a slow and strong increase of SBP with compensatory decrease. Olfactory stimuli on the other side were followed by a biphasic reaction with short increase followed by a strong decrease of SBP, which was sign different from SBP response to other sensory evoked disgust ( $p_{\text{corr6}} < 0.01$ ).

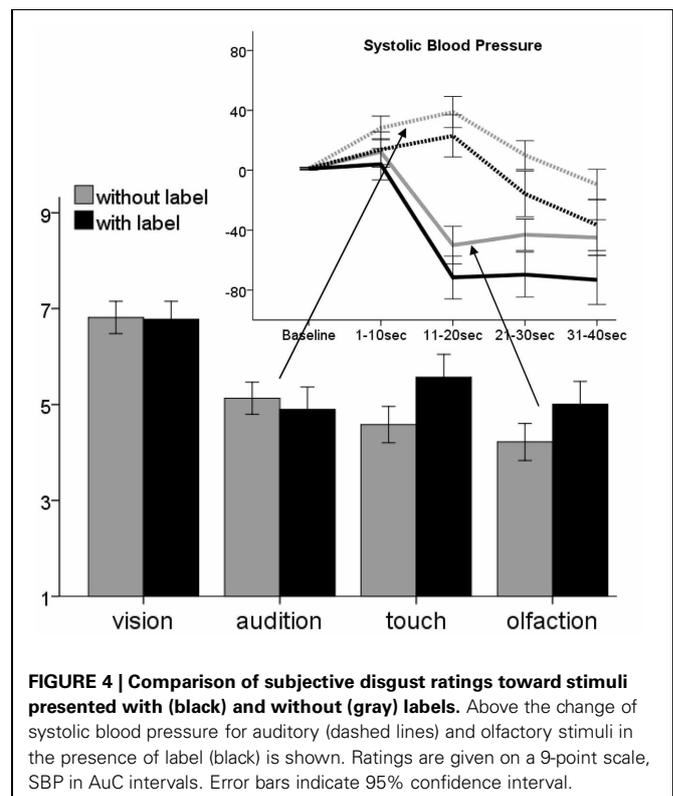
#### INFLUENCE OF LABEL

Labeling led to *enhanced disgust ratings of olfactory* [ $t_{(117)} = 2.6$ ,  $p_{\text{corr4}} = 0.04$ , compare **Figure 4**] and *tactile* [ $t_{(117)} = 3.2$ ,  $p_{\text{corr4}} = 0.01$ ] stimuli. For visual and auditory stimuli no significant influence of labeling was found. Labeling significantly *enhanced HR deceleration in tactile stimuli* [ $F_{(1, 115)} = 6.9$ ,  $p_{\text{corr4}} = 0.04$ , compare **Table S1**], *enhanced SBP decrease following disgusting olfactory stimuli* [ $F_{(1, 115)} = 26.8$ ,  $p_{\text{corr4}} < 0.01$ ] and *diminished SBP reaction toward auditory stimuli* [ $F_{(1, 115)} = 22.5$ ,  $p_{\text{corr4}} < 0.01$ ]. No significant influence of labeling was found for SCL response.

#### INFLUENCE OF REPEATED PRESENTATION

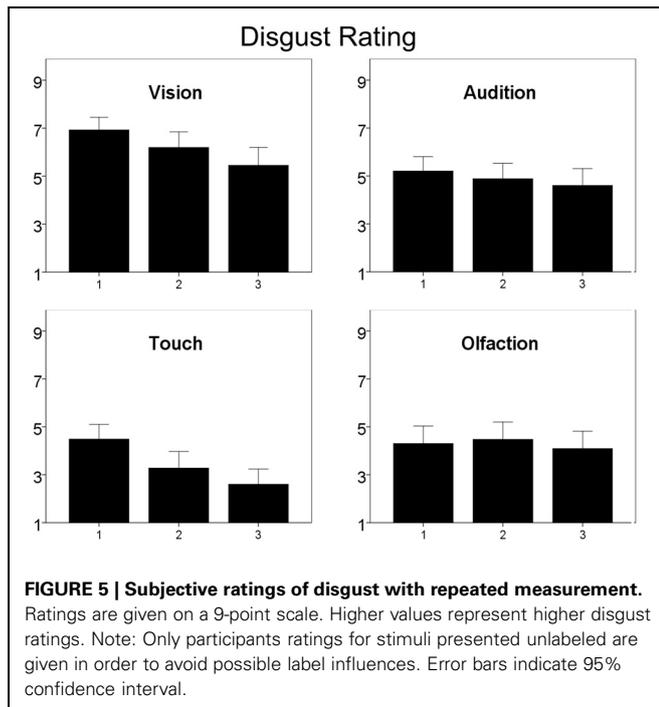
There was a significant main effect of repetition [ $F_{(1, 42)} = 20.2$ ,  $p < 0.01$ , compare **Figure 5**] and a significant interaction between repetition and sensory channel [ $F_{(3, 40)} = 7.8$ ,  $p < 0.01$ ]. *Post-hoc* testing revealed a significant decrease of disgust ratings for visual [ $F_{(1, 42)} = 20.8$ ,  $p < 0.01$ ], tactile [ $F_{(1, 42)} = 31.6$ ,  $p < 0.01$ ], and auditory [ $F_{(1, 42)} = 3.8$ ,  $p = 0.03$ ], but not for olfactory stimuli.

For autonomic measurements either no effect or a diminished response was observed with repeated measurements (compare **Table 3**). In tendency, there was a main effect of repetition for HR [ $F_{(3, 40)} = 3.8$ ,  $p = 0.05$ ] and an interaction between repetition and the sensory channel [ $F_{(3, 40)} = 2.5$ ,  $p = 0.06$ ]. *Post-hoc* tests



revealed a diminished reaction between the first and the third trial for the sense of touch [ $F_{(1, 42)} = 10.1$ ,  $p < 0.01$ ] but not for the other sensory channels.

There was a main effect of repetition for the SCL [ $F_{(3, 40)} = 31.2$ ,  $p < 0.01$ ], indicating that SCL reaction diminished with repeated presentation. However, there was no significant interaction between repetition and the sensory channel.



There was a main effect of repetition for the SBP [ $F_{(3, 40)} = 4.3, p = 0.04$ ], and an interaction between repetition and the sensory channel [ $F_{(3, 40)} = 7.2, p < 0.01$ ]. *Post-hoc* tests revealed a flattening of the SBP curve between the first and the third trial for the sense of olfaction [ $F_{(1, 42)} = 15.5, p < 0.01$ ] and in tendency for audition [ $F_{(1, 42)} = 3.6, p = 0.06$ ].

## DISCUSSION

The study was designed to compare disgust reactions evoked through the visual, auditory, tactile, and olfactory sense. Confirming previous studies (Alaoui-Ismaili et al., 1997a; Collet et al., 1997; Bensafi et al., 2002; Hertenstein et al., 2009; Croy et al., 2011; Oum et al., 2011), the ratings show that disgust can be evoked through the visual, olfactory, and tactile channel. Furthermore, disgust could be evoked through the auditory channel using non-verbal information. To our knowledge this has not been shown before, though the finding is not very surprising.

We assumed that the sensory channel of presentation contributes to disgust reaction. Supporting this, autonomic reaction toward disgusting stimuli differed according to the channel of presentation. Labeling enhanced disgust reaction for olfactory and tactile, but not for visual and auditory stimuli. Furthermore, with repeated measurements participant's disgust rating diminished to all but olfactory applied stimuli. The results are discussed in detail below.

According to Bradley and colleagues autonomic reaction to an emotional cue is biphasic: The initial orienting reaction, indicated by deceleration of HR and increase of SCL, is replaced by an action tendency toward the stimulus (Bradley et al., 2001). We observed a HR deceleration within the first 10 s for all disgust stimuli, potentially reflecting an orienting reaction. An increase

of SCL however, was only observed for disgusting auditory and olfactory stimuli.

After the initial orientation phase, autonomic reaction patterns differed between the senses. SCL decreased for visual and tactile stimuli but showed an increase for olfactory and auditory stimuli. SBP increased for auditory and tactile stimuli, but showed a strong decline for olfactory stimuli. Autonomic responses are highly dependent on context and relevant action tendencies (Van Diest et al., 2009) and the different patterns observed may indicate different action tendencies.

*Olfaction* is strongly linked to food intake and plays a critical role in checking whether food is spoiled or edible. Consequently, people without a sense of smell report more often to have accidentally eaten spoiled food (Croy et al., 2012). In contrast to stimuli that are seen, heard or touched, olfactory stimuli have a relatively high probability to be in the mouth (via the retronasal pathway) or to be about to enter the body. Therefore, odors related to potential harmful substances may evoke a disgust reaction that prepares for vomiting. The HR reduction, indicating a vagal reaction, supports this hypothesis as well as the strong decrease of blood pressure, which has been found to be related to vomiting (Pusch et al., 2002).

For *visual* evoked disgust, autonomic response failed to show a clear effect compared to the controls. After the orienting reaction, visual evoked disgust reaction was mainly characterized by decrease of SCL. The weak autonomic reaction could indicate that visually evoked disgust (at least with the stimuli we used) does not initiate strong action tendencies. A similar slow SCL decrease was previously observed for the presentation of disgusted faces (Collet et al., 1997). In another study however, an increase of SCL in the first seconds following presentation of disgusting pictures is reported (Bradley et al., 2001). An explanation might be that the authors analyzed the maximum SCL amplitude in a given time interval compared to baseline. We used a more conservative approach by analyzing AUC. Interestingly, the relatively weak effects of visual presented disgust stimuli were accompanied by the highest disgust ratings. We assume that visual stimuli evoke stronger memory traces than tactile or olfactory stimuli, because they can be categorized easily (Baddeley and Hitch, 1974). This information contributes to emotional experience (Bradley, 2000) and may enhance disgust ratings.

Autonomic disgust reaction evoked by *auditory* stimuli was characterized by a significant, but relatively low, increase of SCL and SBP, indicating sympathetic activation. This prepares the body for fast reaction and could indicate a weak fight and flight action tendency.

The *tactile* channel has to be interpreted with caution for two reasons: The stimulus characteristics were not obvious at once but changed over time, while the participants touched the object. This may influence experience, as indicated by the ratings: Although the participants experienced tactile objects as clearly disgusting, they were not rated very unpleasant. Furthermore, the autonomic measures of tactile stimuli were influenced by the participants moving the fingers of the non-attached side. Those circumstances might result in altered autonomic reaction and enhanced orienting as indicated by the strong decrease of HR.

The *effect of labeling* supports the differential impact of the senses. Labeling increased disgust ratings and autonomic reaction toward disgusting odors and tactile cues, but not for auditory and visual ones. Labeling adds contextual information. However, for visual and even auditory stimuli labeling presumably did not add more information than that already retrieved from memory. For olfactory and tactile stimuli on the other hand, labeling altered the response. This is in line with previous studies for the sense of smell (De Araujo et al., 2005; Bensafi et al., 2007, 2012) and an interesting finding for the sense of touch.

In accordance with our hypothesis, disgust reaction for *repeated presentation* also differed between the sensory channels: For all but olfactory stimuli disgust ratings decreased. As olfactory stimuli are hard to identify (Jonsson and Olsson, 2003), recognition is difficult. That may explain that emotional response did not decrease with repeated presentation. We hypothesized a similar effect for auditory and tactile cues. However, the effect of labeling suggests that auditory stimuli evoked a lot of context information. For tactile information on the other hand, carefully touching the objects for 10 s could lead to enlarged encoding, which would make recognition easier. The autonomic response toward disgusting stimuli either decreased over time or remained unchanged. No enhanced response was observed with repeated measurement, suggesting that there was no sensitization toward disgusting stimuli.

We are aware of several *limitations* of the study. First, the stimuli differed in intensity and hedonic value: Visual applied stimuli were rated as most unpleasant and arousing. This could either indicate a bias in the choice of stimuli or disgust is in fact more intense if evoked through the sense of vision compared to other senses. Although being rated more arousing, unpleasant and disgusting, visual disgust stimuli did not evoke more anger, sadness, happiness, or anxiety than disgust stimuli applied through the other senses. In order to clarify the influence on autonomic measurements, one has to take care to match stimuli in intensity. This can be done by reducing intensity of the visual stimuli, for instance by reducing contour or color of the pictures. However, ecological validity should be preserved. For the autonomic results

enhanced intensity in the visual stimuli should rather over- than underestimate the effect, but still there was no significant difference in autonomic measurements between the visual control and disgust stimuli.

Second, all of the control stimuli were rated rather positive. This may reflect an inter class bias (Hoyt, 2000), meaning that the disgusting stimuli enhanced the contrast to the control ones and therefore the control stimuli were rated more positive. Third, the autonomic measurements used were not specific for disgust reaction. A facial EMG at the levator nasi muscle could add useful information. And fourth, in order to keep the already complex design as simple as possible, we did not include a control emotion. This might be a problem, as the participants may at some point of the experiment be aware that half of the stimuli are rather disgusting. That could bias their answers toward the disgust category. Future studies should investigate whether threatening or joy evoking stimuli, for instance, are perceived in another way when heard, seen, touched, or smelled.

For disgust, we argue that the sensory channel of presentation contributes to the emotional experience. This might also integrate the controversial findings of autonomic measurements on disgust (Kreibig, 2010). Therefore, research on emotions should pay more attention on the sensory channel, through which emotions are evoked.

## ACKNOWLEDGMENTS

The authors acknowledge support by the German Research Foundation and the Open Access Publication Funds of the TU Dresden. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: [http://www.frontiersin.org/Human\\_Neuroscience/10.3389/fnhum.2013.00510/abstract](http://www.frontiersin.org/Human_Neuroscience/10.3389/fnhum.2013.00510/abstract)

### Table S1 | Effect of labeling on autonomic measurement.

## REFERENCES

- Alaoui-Ismaili, O., Robin, O., Rada, H., Dittmar, A., and Vernet-Maury, E. (1997a). Basic emotions evoked by odorants: comparison between autonomic responses and self-evaluation. *Physiol. Behav.* 62, 713–720. doi: 10.1016/S0031-9384(97)90016-0
- Alaoui-Ismaili, O., Vernet-Maury, E., Dittmar, A., Delhomme, G., and Chanel, J. (1997b). Odor hedonics: connection with emotional response estimated by autonomic parameters. *Chem. Senses* 22, 237–248. doi: 10.1093/chemse/22.3.237
- Bach, D. R., Friston, K. J., and Dolan, R. J. (2010). Analytic measures for quantification of arousal from spontaneous skin conductance fluctuations. *Int. J. Psychophysiol.* 76, 52–55. doi: 10.1016/j.ijpsycho.2010.01.011
- Baddeley, A. D., and Hitch, G. J. (1974). “Working memory,” in *Recent Advances in Learning and Motivation*, ed G. A. Bower (New York, NY: Academic Press).
- Barrett, L. F., Lindquist, K. A., Bliss-Moreau, E., Duncan, S., Gendron, M., Mize, J., et al. (2007). Of mice and men: natural kinds of emotions in the mammalian brain? A response to panksepp and izard. *Perspect. Psychol. Sci.* 2, 297–311. doi: 10.1111/j.1745-6916.2007.00046.x
- Bensafi, M., Croy, I., Phillips, N., Rouby, C., Sezille, C., Gerber, J., et al. (2012). The effect of verbal context on olfactory neural responses. *Hum. Brain Mapp.* doi: 10.1002/hbm.22215. [Epub ahead of print].
- Bensafi, M., Rinck, F., Schaal, B., and Rouby, C. (2007). Verbal cues modulate hedonic perception of odors in 5-year-old children as well as in adults. *Chem. Senses* 32, 855–862. doi: 10.1093/chemse/bjm055
- Bensafi, M., Rouby, C., Farget, V., Bertrand, B., Vigouroux, M., and Holley, A. (2002). Psychophysiological correlates of affects in human olfaction. *Neurophysiol. Clin.* 32, 326–332. doi: 10.1016/S0987-7053(02)00339-8
- Bradley, M. M. (2000). “Emotion and motivation,” in *Handbook of Psychophysiology*, eds T. L. Cacioppo, L. G. Tassinary, and G. G. Bernston (New York, NY: Cambridge University Press), 602–640.
- Bradley, M. M., Codispoti, M., Cuthbert, B. N., and Lang, P. J. (2001). Emotion and motivation I: defensive and appetitive reactions in picture processing. *Emotion* 1, 276–298. doi: 10.1037/1528-3542.1.3.276
- Collet, C., Vernet-Maury, E., Delhomme, G., and Dittmar, A. (1997). Autonomic nervous system response patterns specificity to basic emotions. *J. Auton. Nerv. Syst.* 62, 45–57. doi: 10.1016/S0165-1838(96)00108-7
- Croy, I., Negoias, S., Novakova, L., Landis, B., and Hummel, T. (2012). Learning about the functions of the olfactory system from people without a sense of smell. *PLoS*

- ONE:7:e33365. doi: 10.1371/journal.pone.0033365
- Croy, I., Olgun, S., and Joraschky, P. (2011). Basic emotions elicited by odors and pictures. *Emotion* 11, 1331–1335. doi: 10.1037/a0024437
- Darwin, C. (1872). *The Expression of the Emotions in Man and Animals*. London: John Murray. doi: 10.1037/10001-000
- De Araujo, I. E., Rolls, E. T., Velasco, M. I., Margot, C., and Cayeux, I. (2005). Cognitive modulation of olfactory processing. *Neuron* 46, 671–679. doi: 10.1016/j.neuron.2005.04.021
- Ekman, P., Levenson, R. W., and Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science* 221, 1208–1210. doi: 10.1126/science.6612338
- Gottfried, J. A. (2006). Smell: central nervous processing. *Adv. Otorhinolaryngol.* 63, 44–69. doi: 10.1159/000093750
- Hertenstein, M. J., Holmes, R., McCullough, M., and Keltner, D. (2009). The communication of emotion via touch. *Emotion* 9, 566–573. doi: 10.1037/a0016108
- Hoyt, W. T. (2000). Rater bias in psychological research: when is it a problem and what can we do about it? *Psychol. Methods* 5, 64–86. doi: 10.1037/1082-989X.5.1.64
- Hummel, T., Kobal, G., Gudziol, H., and Mackay-Sim, A. (2007). Normative data for the “Sniffin’ Sticks” including tests of odor identification, odor discrimination, and olfactory thresholds: an upgrade based on a group of more than 3,000 subjects. *Eur. Arch. Otorhinolaryngol.* 264, 237–243. doi: 10.1007/s00405-006-0173-0
- Hummel, T., Konnerth, C. G., Rosenheim, K., and Kobal, G. (2001). Screening of olfactory function with a four-minute odor identification test: reliability, normative data, and investigations in patients with olfactory loss. *Ann. Otol. Rhinol. Laryngol.* 110, 976–981.
- Izard, C. E. (1992). Basic emotions, relations among emotions, and emotion cognition relations. *Psychol. Rev.* 99, 561–565. doi: 10.1037/0033-295X.99.3.561
- Jonsson, F. U., and Olsson, M. J. (2003). Olfactory metacognition. *Chem. Senses* 28, 651–658. doi: 10.1093/chemse/bjg058
- Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: a review. *Biol. Psychol.* 84, 394–421. doi: 10.1016/j.biopsycho.2010.03.010
- Kunzmann, U., and Gruhn, D. (2005). Age differences in emotional reactivity: the sample case of sadness. *Psychol. Aging* 20, 47–59. doi: 10.1037/0882-7974.20.1.47
- Lang, P. J. (1980). “Behavioral treatment and bio-behavioral assessment: computer applications,” in *Technology in Mental Health Care Delivery System*, eds J. B. Sidowski, J. H. Johnson, and T. A. Williams (Norwood, NJ: Ablex), 119–137.
- Oaten, M., Stevenson, R. J., and Case, T. I. (2009). Disgust as a disease-avoidance mechanism. *Psychol. Bull.* 135, 303–321. doi: 10.1037/a0014823
- Oum, R. E., Lieberman, D., and Aylward, A. (2011). A feel for disgust: tactile cues to pathogen presence. *Cogn. Emot.* 25, 717–725. doi: 10.1080/02699931.2010.496997
- Panksepp, J. (2007). Neurologizing the psychology of affects how appraisal-based constructivism and basic emotion theory can coexist. *Perspect. Psychol. Sci.* 2, 281–296. doi: 10.1111/j.1745-6916.2007.00045.x
- Posner, J., Russell, J. A., and Peterson, B. S. (2005). The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Dev. Psychopathol.* 17, 715–734. doi: 10.1017/S0954579405050340
- Prkachin, K. M., Williams-Avery, R. M., Zwaal, C., and Mills, D. E. (1999). Cardiovascular changes during induced emotion: an application of lang’s theory of emotional imagery. *J. Psychosom. Res.* 47, 255–267. doi: 10.1016/S0022-3999(99)00036-7
- Pusch, F., Berger, A., Wildling, E., Tiefenthaler, W., and Krafft, P. (2002). The effects of systolic arterial blood pressure variations on postoperative nausea and vomiting. *Anesth. Analg.* 94, 1652–1655.
- Rohrmann, S., Hopp, H., Schienle, A., and Hodapp, V. (2009). Emotion regulation, disgust sensitivity, and psychophysiological responses to a disgust-inducing film. *Anxiety Stress Coping* 22, 215–236. doi: 10.1080/10615800802016591
- Rozin, P., and Fallon, A. E. (1987). A perspective on disgust. *Psychol. Rev.* 94, 23–41. doi: 10.1037/0033-295X.94.1.23
- Rozin, P., Haidt, J., and McCauley, C. (2000). “Disgust,” in *Handbook of Emotions, 2nd Edn.*, eds M. Lewis and M. Haviland-Jones (New York, NY: Guilford Press), 637–653.
- Stevenson, R. J. (2010). An initial evaluation of the functions of human olfaction. *Chem. Senses* 35, 3–20. doi: 10.1093/chemse/bjp083
- Tabibnia, G., Lieberman, M. D., and Craske, M. G. (2008). The lasting effect of words on feelings: words may facilitate exposure effects to threatening images. *Emotion* 8, 307–317. doi: 10.1037/1528-3542.8.3.307
- Tybur, J. M., Lieberman, D., and Griskevicius, V. (2009). Microbes, mating, and morality: individual differences in three functional domains of disgust. *J. Pers. Soc. Psychol.* 97, 103–122. doi: 10.1037/a0015474
- Vaitl, D., Schienle, A., and Stark, R. (2005). Neurobiology of fear and disgust. *Int. J. Psychophysiol.* 57, 1–4. doi: 10.1016/j.ijpsycho.2005.01.005
- Van Diest, I., Bradley, M. M., Guerra, P., Van Den Bergh, O., and Lang, P. J. (2009). Fear-conditioned respiration and its association to cardiac reactivity. *Biol. Psychol.* 80, 212–217. doi: 10.1016/j.biopsycho.2008.09.006

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 27 March 2013; accepted: 08 August 2013; published online: 03 September 2013.

Citation: Croy I, Laqua K, Süß F, Joraschky P, Ziemssen T and Hummel T (2013) The sensory channel of presentation alters subjective ratings and autonomic responses toward disgusting stimuli—Blood pressure, heart rate and skin conductance in response to visual, auditory, haptic and olfactory presented disgusting stimuli. *Front. Hum. Neurosci.* 7:510. doi: 10.3389/fnhum.2013.00510

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2013 Croy, Laqua, Süß, Joraschky, Ziemssen and Hummel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Sex differences in chemosensation: sensory or emotional?

Kathrin Ohla<sup>1</sup> and Johan N. Lundström<sup>1,2,3\*</sup>

<sup>1</sup> Monell Chemical Senses Center, Philadelphia, PA, USA

<sup>2</sup> Division of Psychology, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

<sup>3</sup> Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA

## Edited by:

Yu-Han Chen, University of New Mexico, USA

## Reviewed by:

Dirk Adolph, Ruhr-University-Bochum, Germany  
Geraldine Coppin, Yale University, USA

Pauline Joussain, Centre de Recherche en Neurosciences de Lyon, France

## \*Correspondence:

Johan N. Lundström, Division of Psychology, Department of Clinical Neuroscience, Karolinska Institutet, Nobels väg 9, 171 65 Solna, Stockholm, Sweden  
e-mail: johan.lundstrom@ki.se

Although the first sex-dependent differences in chemosensory processing were reported in the scientific literature over 60 years ago, the underlying mechanisms are still unknown. Generally, more pronounced sex-dependent differences are noted with increased task difficulty or with increased levels of intranasal irritation produced by the stimulus. Whether differences between the sexes arise from differences in chemosensory sensitivity of the two intranasal sensory systems involved or from differences in cognitive processing associated with emotional evaluation of the stimulants is still not known. We used simultaneous and complementary measures of electrophysiological (EEG), psychophysiological, and psychological responses to stimuli varying in intranasal irritation and odorousness to investigate whether sex differences in the processing of intranasal irritation are mediated by varying sensitivity of the involved sensory systems or by differences in cognitive and/or emotional evaluation of the irritants. Women perceived all stimulants more irritating and they exhibited larger amplitudes of the late positive deflection of the event-related potential than men. No significant differences in sensory sensitivity, anxiety, and arousal responses could be detected. Our findings suggest that men and women process intranasal irritation differently. Importantly, the differences cannot be explained by variation in sensory sensitivity to irritants, differences in anxiety, or differences in physiological arousal. We propose that women allocate more attention to potentially noxious stimuli than men do, which eventually causes differences in cognitive appraisal and subjective perception.

**Keywords:** sex differences, ERPs, trigeminal, olfactory, GSR, emotion

## INTRODUCTION

It is often stated that women have a better sense of smell than men and when sex differences<sup>1</sup> are reported, women tend to outperform men in odor tasks. However, among the plethora of olfactory sensory studies, differences between men and women almost exclusively exist for tasks that involve odor naming and memory retrieval (Cain, 1982; Doty et al., 1985; Oberg et al., 2002). Importantly, reports of sex differences for detection thresholds, an effective measure of sensory sensitivity, are scarce with a few exceptions originating from studies that used odors with a profound biological or cognitive meaning (Koelega and Koster, 1974; Lundstrom et al., 2003). Inspired by the recent demonstration that women tend to be more reactive to stimuli that are perceived as emotional or irritating (Vigil, 2009), we set out to test the hypothesis that sex differences for chemosensory stimuli are predominantly mediated by differences in cognitive or emotional appraisal rather than sensory sensitivity *per se* by means of both psychological and biometric measures.

<sup>1</sup>We are in this report first and foremost interested in effects linked to biological processes rather than gender identity and or societal factors. We are therefore using the term “sex,” a term commonly used when referring to potential differences between men and women based on the underlying biology, rather than “gender,” a term commonly used when referring to sexual (gender) identity.

In our everyday life, few, if any, odors are processed exclusively by the olfactory system. In most cases, the olfactory and trigeminal systems conjointly process odors. The trigeminal system mediates sensations such as burning, cooling, and tingling, even in the absence of an olfactory percept (Laska et al., 1997). In contrast to what is reported for purely olfactory stimulants, reports of sex-dependent differences in trigeminal sensitivity are more robust in that most studies indeed find significant sex differences. Here, women exhibit higher sensory trigeminal sensitivity (Shusterman et al., 2003), better perceptual acuity (Shusterman and Balmes, 1997; Andersson et al., 2011), and better lateralization ability (Stuck et al., 2006) compared to men. Robust sex differences to purely trigeminal stimulation have been reported in event-related potentials (ERPs) studies (Hummel et al., 1998; Lundstrom et al., 2005; Stuck et al., 2006; Scheibe et al., 2009). These studies reported larger amplitudes and shorter latencies of the late positive component (LPC) of the ERPs to the trigeminal compound carbon-dioxide (CO<sub>2</sub>) for women compared to men (Hummel et al., 1998; Lundstrom et al., 2005). However, although the pronounced sex differences for trigeminal stimuli suggest that women’s peripheral trigeminal system is more reactive compared to the sensory system of men, negative mucosa potentials, a non-invasive method to record pain-related electrical potentials from the human respiratory nasal mucosa (Kobal, 1981, 1985), has failed to reveal any sex differences (Frasnelli

and Hummel, 2003; Frasnelli et al., 2007). This indicates that the demonstrated sex differences are not primarily mediated by a difference in peripheral processing.

The LPC has been tied to stimulus assessment and evaluation (Polich and Kok, 1995; Pause et al., 1996). Along those lines, sex-dependent differences of the LPC would indicate that women assess nasal irritants differently than men. Further support for the notion of sex differences in stimulus assessment comes from a recent study demonstrating that women tend to be more reactive to stimuli that are perceived as emotional, unpleasant, or threatening (Vigil, 2009), a finding that has been suggested to be indicative of sex-dependent differences in strategies employed when processing emotional stimuli (Hall et al., 2004; Whittle et al., 2011). Women, in general, also exhibit a larger emotional response to sensory stimuli, including intranasal irritation, than men do (Whittle et al., 2011). Importantly, comparative findings exist for the chemical senses. Women report chemical intolerance to a larger degree than men (Johansson et al., 2005; Berg et al., 2008) and women's general responses to intranasal irritation is to a large extent comparable to individuals suffering from chemical intolerance, so-called multiple chemical sensitivity (MCS) (Andersson et al., 2009, 2011), a diagnosis that has been linked to the cognitive processing of the odor rather than sensory acuity *per se* (Hillert et al., 2007). Interestingly, patients with MCS have been successfully treated with a selective serotonin reuptake inhibitor (Andine et al., 1997) the action of which has been linked to a specific reduction of 5-HT<sub>1a</sub> receptors in the amygdala and insular cortex, both part of the fear processing network (Hillert et al., 2013). Together, these findings suggest that sex-dependent differences for bimodal odors are to some extent linked to the degree of irritation sensation and the cognitive and emotional evaluation of these sensations rather than the sensory processing of the odor alone. In line with that, Ferdenzi et al. (2008) proposed that, starting from young age, women develop a stronger emotional reaction to intranasal sensations than men do. Based on findings from MCS patients, Andersson et al. (2011) recently brought forward the novel hypothesis that a heightened emotional response may render women to allocate more attention toward intranasal stimuli; a potential mechanism mediating previously reported sex-differences.

To assess whether sex-dependent differences in chemosensory processing are primarily mediated by differences in sensory sensitivity or cognitive and emotional appraisal of the stimuli, we used a wide array of measures to capture potential differences in the psychological, sensory, physiological, and neuronal domain. First, we assessed self-reported anxiety before and after stimulation. Furthermore, we measured sensory sensitivity to irritants as well as olfactory discrimination ability. During chemosensory stimulation, ERPs were obtained together with subjective ratings of the stimuli; ERPs provide an ideal tool to distinguish brain processes related to sensory decoding from processes associated with higher cognitive functions, such as attention or memory. At the same time, we measured galvanic skin responses (GSR), a sensitive marker of arousal and emotional responsiveness. As stimulants, we used the mostly odorless CO<sub>2</sub>, which is primarily processed by the trigeminal system with little to no activation of the olfactory system, at high and low irritating concentrations to assess

effects of trigeminal irritation independent of odor. In addition, we presented the bimodal odorant cineol at a concentration that combined high irritation and high odorousness.

We hypothesized that sex-dependent differences are mediated by cognitive processes related to attention and stimulus appraisal rather than by differences in sensitivity of the peripheral sensory system. Accordingly, sex-dependent differences should be observed for the LPC. The finding of sex-related differences in measures of subjective anxiety and/or measures of arousal (GSR) would point to differences in emotional responsiveness; an interaction of anxiety and GSR with the LPC effect would suggest that differences in affective processing can modulate cognitive evaluation of the irritants. Conversely to our hypothesis, sex differences for the early ERP components and for thresholds to the irritant would indicate that sex-dependent effects have a predominantly peripheral origin that is in the receptor organ.

## MATERIALS AND METHODS

### PARTICIPANTS

Thirty-seven healthy (no self-reported nasal, psychiatric, and neurological disorders), right-handed participants completed the study. Out of these, eight were excluded before statistical analyses; four due to technical problems during the experiment and four based on excessive movement artifacts that precluded ERP analyses. Consequently, a total of 29 participants, 14 men (25.1 years old,  $SD = 4.8$ , range = 20–32) and 15 women (25.6 years old,  $SD = 3.8$ , range = 21–33) were included in all analyses except for GSR. For the GSR analyses, one male participant was excluded due to technical problems with the GSR recording. In order to minimize hormonal influences in the participating women, one third of the women was tested during their follicular phase, during their luteal phase, or while being on hormonal birth control, respectively. Menstrual cycle phase was determined based on retrospective calculation from the point of the onset of the last menses (Lundstrom et al., 2006). Note that effect due to menstrual cycle phase was not assessed due to the limited sample size. All participants were paid for participation and provided written informed consent. The study adhered to the revised Declaration of Helsinki and all aspects of the study were approved by the University of Pennsylvania Institutional Review Board.

### OLFACTORY IDENTIFICATION ABILITY AND TRIGEMINAL SENSITIVITY

We assessed participants' olfactory identification ability in order to rule out that any of the participants was anosmic by using the Sniffin' Sticks olfactory identification (ID) test (Kobal et al., 1996; Hummel et al., 1997). The test consists of 16 individual felt-tip pens, each containing a distinct odor that is identified using a four-alternative forced-choice paradigm. Two female participants did not participate in the ID test because they were highly familiar with the test and they had achieved high scores previously. Trigeminal sensitivity was assessed for the bimodal odor menthol with a 2-alternative, forced-choice, nostril-laterality detection threshold task using an ascending staircase with 5 reversals (Frasnelli et al., 2011a). For this, 16 concentrations of menthol (R.J. Reynolds Tobacco Company, CAS 2216-51-5, declared purity >99.97%), a bimodal odorant, ranging from 0.1 to 50% with each concentration reduced by one third, were prepared

in propylene glycol (1,2 propanediol, Fisher Scientific, Acros Organics, CAS 57-55-6, declared purity >99%) and presented in 60 mL amber glass bottles. Sensory sensitivity to menthol rather than to CO<sub>2</sub> or cineole, the two stimuli used in the EEG portion of the study, was assessed in order to avoid familiarization to one of the two stimulants. Recent findings suggest that sensitivity to menthol is highly correlated with sensitivity to cineole (Frasnelli et al., 2011a) leading us to believe that thresholds to menthol represent a valid measure of sensitivity to intranasal irritation in a broader sense.

### SUBJECTIVE ANXIETY MEASURE

As a measure of subjective anxiety, participants completed the State-Trait-Anxiety Inventory, STAI (Spielberger et al., 1970). The test consists of two parts that assess state and trait anxiety; the range of scores is 20–80 where higher scores indicate greater anxiety. While state anxiety refers to the momentary tendency to experience anxiety, trait anxiety is enduring and universal across different situations (Spielberger and Sydeman, 1994). Both, the STAI-S (state anxiety) and STAI-T (trait anxiety) scores were obtained before and after chemosensory stimulation. STAI-S scores were used to assess possible changes in participants' momentary (state) anxiety attributable to the experimental procedures. Because no such changes were observed, we averaged the pre and post STAI scores (for the S and T subtests) for further analyses.

### STIMULI AND PROCEDURES

The experiment was conducted in an air-conditioned room constructed specifically for olfactory testing with a high turnover of the total air volume to limit lingering odors. Participants were seated comfortably while EEG was recorded. Chemosensory stimuli were presented using an air-dilution olfactometer (OM6b, Burghart Messtechnik, Wedel, Germany), which embeds the chemosensory stimuli in a continuous stream of humidified (80%) and heated (36°C) air with a flow rate of 6.1 l/min. The method allows fast rise times of the stimulus (Lorig, 2000) and minimizes somatosensory stimulation from changes in the air flow through the nostrils (Sobel et al., 1998). Stimuli were the non-odorous, trigeminal CO<sub>2</sub> at 50% v/v and 60% v/v, in the following referred to as CO<sub>2</sub>low and CO<sub>2</sub>high, respectively, and the bimodal, olfactory-trigeminal, cineole (Eucalyptol; Sigma-Aldrich, CAS 470-82-6, declared purity 99%) at 50% v/v. The two concentrations of CO<sub>2</sub> were selected based on a pilot study ( $n = 5$ ) where the low concentration produced a tactile but no irritating or stinging sensation and the high concentration produced a clear irritating or stinging sensation. All stimuli were presented monorhinally starting with either the right or left nostril and shifting sides halfway at a scheduled break. The olfactometer was placed in a neighboring room to limit acoustic interference and participants were presented brown noise via isolating in-ear headphones to preclude auditory cues from the olfactometer and the shifting air flow.

Each trial started with a central fixation cross presented on a computer screen for a variable interval of 3–9 s. Within this interval, a chemosensory stimulus was presented for 250 ms. Fixation was replaced by the written instruction to rate stimulus irritation,

odorousness, and pleasantness on a visual analog scale (VAS) ranging from 1 (not at all) to 100 (extremely strong/pleasant) using the right index finger and a mouse. The rating period started 2.5 s after stimulus delivery. A total of 90 trials (30 trials for each stimulus category: CO<sub>2</sub>low, CO<sub>2</sub>high, and cineole) was presented in pseudo-random order with a variable inter-trial interval of 23.5–38.5 s. Participants were instructed to pay attention to the chemosensory stimulus and to avoid any movements and eye blinks. To allow the presentation of chemosensory stimuli independent of the individual's respiratory cycle, all participants were trained in the velopharyngeal breathing technique, a technique that limits the respiratory flow of air through the nasal cavity (Kobal, 1981), and asked to use this breathing throughout the ERPs portion of the experiment.

### PHOTO-IONIZATION DETECTION (PID) BASED TIMING CORRECTION

All mechanical devices exhibit a time-lag between the TTL (transistor-transistor logic) pulse originating from the stimulus computer that initiate stimulus delivery and the actual delivery of the stimulus. This time lag artificially delays the ERP with the corresponding value. We measured the time lag between TTL pulse and arrival of odor molecules at the outlet of the nasal cannula using a fast response miniature photo-ionization detector (PID Mod. 200A, Aurora Scientific inc., Aurora, Ontario, Canada). The sensor has a true frequency response of 330 Hz with a 10–90% rise time of 0.6 ms and the detection limit is 100 ppb (parts per billion) contaminant in air. Onset of the TTL trigger sent and the ongoing PID signal was recorded for 24 continuous stimuli (30 s inter-stimulus interval) per condition using the Powerlab amplifier system (ADInstruments, Colorado Springs, CO) and analyzed using Origin 8.5 (OriginLab, Northampton, MD). Responses were averaged for each condition and latencies from TTL trigger to onset and 50% stimulus concentration—the concentration at which the stimulus approximately starts to be detected—were measured for the three conditions. Averaged measured stimulus onset delays were: cineole 50 ms, CO<sub>2</sub>low 63 ms, and CO<sub>2</sub>high 64 ms. These values were used to temporally adjust the recorded ERP responses to match stimulus onset to the delivery of the stimulant to the receptors rather than to the TTL pulse.

### GALVANIC SKIN RESPONSES

Galvanic skin responses (GSR), a non-invasive measure of autonomic nervous system activity [for a comprehensive overview, please see (Stern et al., 2001)], were recorded from bipolar Ag-AgCl electrodes with a surface of 10 mm<sup>2</sup> according to existing standards (Fowles et al., 1981). The electrodes were placed at the palmar surface of the medial phalanges of the left index and middle fingers. The electrodes were connected to a ML116 GSR amplifier connected to a Powerlab 16/30 system (ADInstruments, Colorado Springs, CO). The amplifier used low constant-voltage AC excitation and automatic zeroing, which reduces electrode polarization artifacts. GSR data were recorded at 200 Hz and analyzed offline using LabChart 7.1 (ADInstruments, Colorado Springs, CO). For analyses, the continuous data were filtered with a 0.01 Hz high-pass filter to remove slow drifts and linear trends. GSR peak amplitudes (in  $\mu$ S) were defined as maximum

amplitudes in a 10 s time window after stimulus onset after baseline (500 ms prior to stimulus onset) subtraction.

## ELECTROPHYSIOLOGICAL RECORDINGS (EEG)

### Data acquisition and preprocessing

Brain electrical activity was recorded continuously with a BioSemi Active-Two amplifier system (BioSemi, Amsterdam, Netherlands) using 32 Ag/AgCl active electrodes mounted in an elastic cap and placed according to the extended 10–20 system and two additional electrodes, CMS (common mode sense) and DRL (driven right leg) to replace the function of conventional ground electrode (<http://www.biosemi.com/faq/cms&drl.htm>). Lateral eye movements were monitored with a bipolar outer canthus montage (horizontal electrooculogram). Vertical eye movements and blinks were monitored with a bipolar montage positioned below and above the right eye (vertical electrooculogram). Data were recorded with a sampling rate of 512 Hz and analog filtered from 0.16 to 100 Hz. The continuous EEG signal was stored on a hard disk for off-line analysis.

EEG data were processed using the open-source EEGLAB toolbox (Swartz Center for Computational Neurosciences, La Jolla, CA; <http://www.sccn.ucsd.edu/eeGLAB/>; Delorme and Makeig, 2004) running under the Matlab environment (The Mathworks, Inc., Natick, Massachusetts, USA) and the Cartool software by Denis Brunet ([brainmapping.unige.ch/cartool](http://brainmapping.unige.ch/cartool)). Data were 0.2 Hz high-pass filtered (0.03 Hz transition band width) and segmented into epochs of 3 s (–1000 to 2000 ms relative to stimulus trigger sent to the olfactometer). After manual rejection of epochs with unique, non-stereotypical artifacts, extended infomax independent component analysis (ICA), as implemented in EEGLAB, was applied to the remaining concatenated single trials. Independent components representing common EEG artifacts, such as eye blinks, were visually identified and removed. Back-projected single trials were again screened for residual artifacts. On average, 1% of all trials were rejected leaving an average of 29 trials per condition for further analyses. Data were re-referenced to the averaged mastoids after artifact rejection and correction, and a 30 Hz low-pass filter (1 Hz transition band width) was applied. Subsequently, the onset time of each of the remaining trials was shifted by the stimulus onset delay to the 50% rise latencies obtained from the PID measurement described above. Finally, the baseline (300 ms prior to stimulus onset) was subtracted.

### Event-related potentials

Event-related potentials (ERPs) were computed for single electrodes before ERPs were averaged across experimental conditions and participants and plotted to visualize the waveform data. Two major ERP deflections were apparent in the grand-averaged ERPs: a minimum (N1) from 200 to 450 ms at centro-lateral electrodes and a slow, positive deflection (LPC) from 400 to 900 ms at centro-parietal electrodes. For statistical analyses, electrodes exhibiting minimum/maximum amplitudes for the N1 and LPC in the grand-averaged waveform were collapsed, a method commonly used to gain statistical power. Then, the minimum peak amplitudes and peak latencies in the 200–450 ms period were extracted at centro-lateral electrodes (FC1, FC2, C3, Cz, C4, Cp1, and Cp2) to characterize the N1. For the LPC, mean instead

of peak amplitudes were extracted because the mean amplitude of a slow potential is a more valid measure. For this, mean amplitudes in a 300 ms time window around the peak of the grand-averaged data, i.e., in the 350–650 ms time period for cineole and in the 460–760 ms time period for CO<sub>2</sub>, were extracted at centro-parietal electrodes (FC1, Fz, FC2, Cz, Cp1, Pz, and Cp2).

### Topographic pattern analyses

It is commonly agreed upon that scalp topographies of the electric field do not change randomly over time, but rather form topographic states that remain stable for periods of several tens of milliseconds; changes in the topography follow from changes in the underlying neural generators (Lehmann et al., 1987). We grouped waveforms into periods of similar topography (also referred to as microstates) using a modified K-means clustering (Pascual-Marqui et al., 1995) as implemented in the Cartool software on the grand-averaged data over the 0 to 1.400 ms interval to identify the predominant maps and their sequence. Model parameters were set such that clusters with a spatial correlation greater than 92% were merged and that each map had to be observed for at least 30 ms. The optimal number of template maps was determined using a combination of criteria: a peak of the modified Krzanowski–Lai criterion (Krzanowski and Lai, 1985) and minimal cross validation. The cluster analysis provides a descriptive means to summarize the ERP data by a limited number of topographic maps.

### STATISTICAL ANALYSES

Statistical analyses were performed with Matlab and SPSS. Initially, we tested for differences between pre- and post-experimental differences in anxiety and submitted the pre-/post-STAI-S and pre-/post-STAI-T scores to Wilcoxon signed-rank tests. This non-parametric test was chosen because the data were not normally distributed. Since pre and post scores were similar for both STAI-S and STAI-T, pre- and post-experimental scores were averaged and used for all further analyses. Then, STAI-S and STAI-T scores and menthol thresholds were submitted to Mann–Whitney *U*-tests in order to test for differences between men and women. Perceptual ratings, GSR, and ERP peak results were submitted to repeated measures ANOVAs with the within subjects factor *stimulant* (CO<sub>2</sub>high, CO<sub>2</sub>low, cineole) and the between subjects factor *sex* (men, women) using SPSS 20.0 (IBM, Armonk, New York, USA). Student's *t*-tests were used for subsequent pairwise comparisons to resolve significant main effects. Huynh–Feld correction for violations of the assumption of sphericity was used when appropriate; uncorrected *F*-values and degrees of freedom and corrected *p*-values are reported. The  $\eta^2$  statistic was adopted to describe the estimated proportion of variance explained by the factors. The alpha level was *a priori* set to 0.05.

## RESULTS

### ELECTROPHYSIOLOGICAL DATA (EEG)

The grand averaged baseline corrected ERPs showed two main deflections at midline electrodes: the N1 with a minimum at around 250 ms for cineole and at 400 ms for CO<sub>2</sub> over the vertex and adjacent lateral electrodes (i.e., the centro-lateral ROI) and the late positive complex (LPC) with a maximum at around

425 ms for cineole and at 565 ms and 585 ms for CO<sub>2</sub>high and CO<sub>2</sub>low, respectively, over centro-parietal electrodes (**Figures 1A, 2B**). When comparing men and women, differences in latencies and amplitudes became apparent (**Figure 1B, Table 1**).

### Sex-dependent differences in ERP responses

Sex-related ERP differences were found for the amplitude of the LPC only. Women demonstrated higher LPC amplitudes than men as indicated in a main effect of sex [ $F_{(1, 27)} = 19.658, p < 0.001, \eta^2 = 0.421$ ]. Student's *t*-tests revealed that women exhibited higher LPC amplitudes than men for CO<sub>2</sub>high [ $t_{(1, 27)} = 3.59, p = 0.001$ ], CO<sub>2</sub>low [ $t_{(1, 27)} = 5.23, p < 0.001$ ], as well as for cineole [ $t_{(1, 27)} = 2.06, p = 0.05$ ].

Topographic patterns were subjected to a cluster analysis after the ERPs to the different stimulants were averaged; five microstates explained 95.85% of the variance in the grand-averaged ERP data from 0 to 1400 ms (**Figure 2**). The topographical voltage maps corresponding to each of the five segments are displayed in **Figure 2B**. The temporal extent of each map is indicated as colored segments under the global field power (GFP) for each sex. While the sequence of map was highly similar between men and women, the timing was shifted toward faster map occurrence in women suggesting that the underlying cortical generators were similar between the sexes. The first deflection, represented by map 2 (see **Figure 2B**), with a minimum over centro-temporal sites, constitutes the N1. After a brief transition (map 4), the late positive complex (LPC or P3) established with a maximum over central and parietal electrodes (map 5).

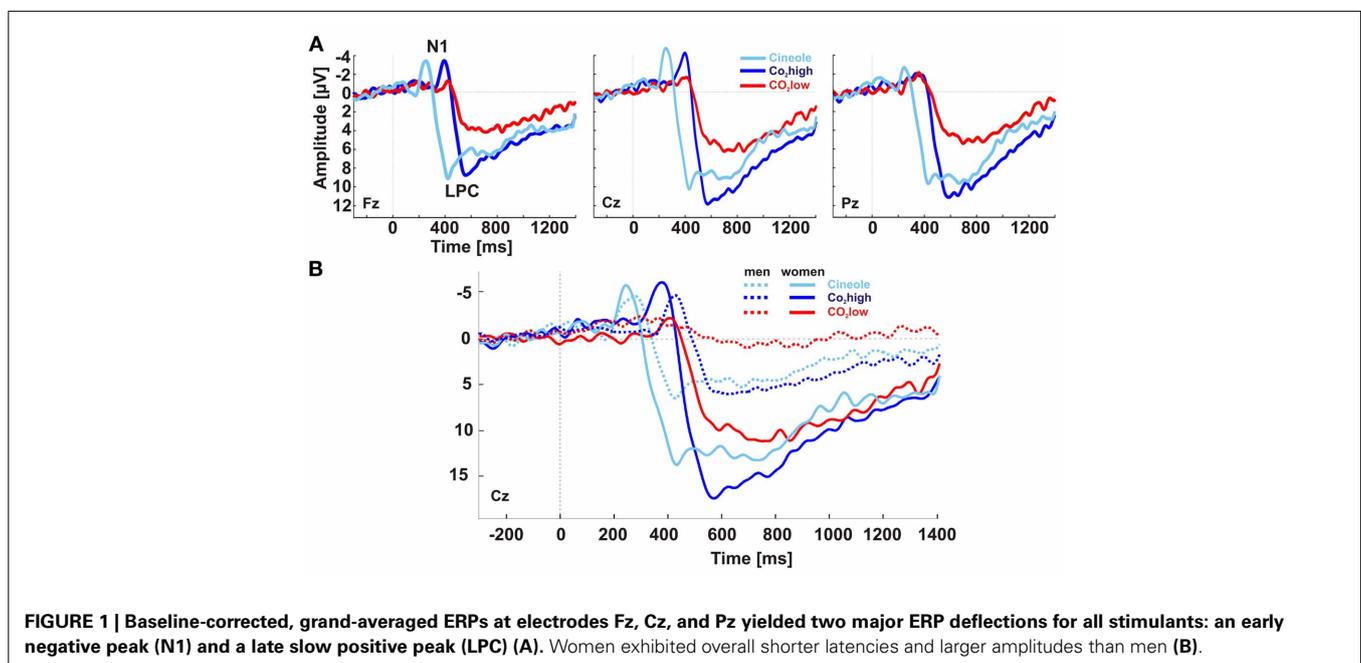
### Stimulant-dependent effects

The ERPs yielded significant differences in response to the three stimulants, as indicated by main effects of stimulant, for N1 peak amplitudes [ $F_{(2, 54)} = 8.87, p < 0.001, \eta^2 = 0.247$ ],

N1 peak latencies [ $F_{(2, 54)} = 17.577, p < 0.001, \eta^2 = 0.394$ ], LPC peak latencies [ $F_{(2, 54)} = 57.865, p < 0.001, \eta^2 = 0.682$ ] and LPC mean peak amplitudes [ $F_{(2, 54)} = 15.068, p < 0.001, \eta^2 = 0.358$ ]. Independent sample paired Student's *t*-tests were subsequently used to resolve the main effects. N1 peaks were most pronounced for CO<sub>2</sub>high and cineole, which were significantly augmented compared to CO<sub>2</sub>low [ $t_{(1, 28)} = 4.591, p < 0.001$  and  $t_{(1, 28)} = 3.235, p < 0.01$ , respectively]. Similarly, LPC mean peak amplitudes were smaller for CO<sub>2</sub>low than for CO<sub>2</sub>high [ $t_{(1, 28)} = 7.344, p < 0.001$ ] as well as for cineole [ $t_{(1, 28)} = 3.895, p = 0.001$ ]. N1 latencies were shorter for cineole than for CO<sub>2</sub>low [ $t_{(1, 28)} = 7.866, p < 0.001$ ] and than

**Table 1 | ERP peak latencies (in ms) and amplitudes (in  $\mu$ V).**

		All		Men		Women	
		Mean	SEM	Mean	SEM	Mean	SEM
N1 amplitude	CO <sub>2</sub> low	-0.336	0.042	-0.356	0.0336	-0.317	0.076
	CO <sub>2</sub> high	-0.531	0.065	-0.555	0.100	-0.509	0.087
	Cineole	-0.488	0.050	-0.537	0.054	-0.442	0.081
N1 latency	CO <sub>2</sub> low	334	21.6	325	34.1	342	27.9
	CO <sub>2</sub> high	368	17.2	396	23.2	343	24.1
	Cineole	264	8.4	269	12.2	259	11.9
LPC amplitude	CO <sub>2</sub> low	0.234	0.056	0.018	0.028	0.435	0.072
	CO <sub>2</sub> high	0.530	0.071	0.311	0.058	0.735	0.100
	Cineole	0.504	0.063	0.378	0.088	0.622	0.079
LPC latency	CO <sub>2</sub> low	586	20.5	604	29.6	568	28.7
	CO <sub>2</sub> high	566	15.5	577	22.2	555	22.1
	Cineole	420	12.4	409	18.3	430	16.9



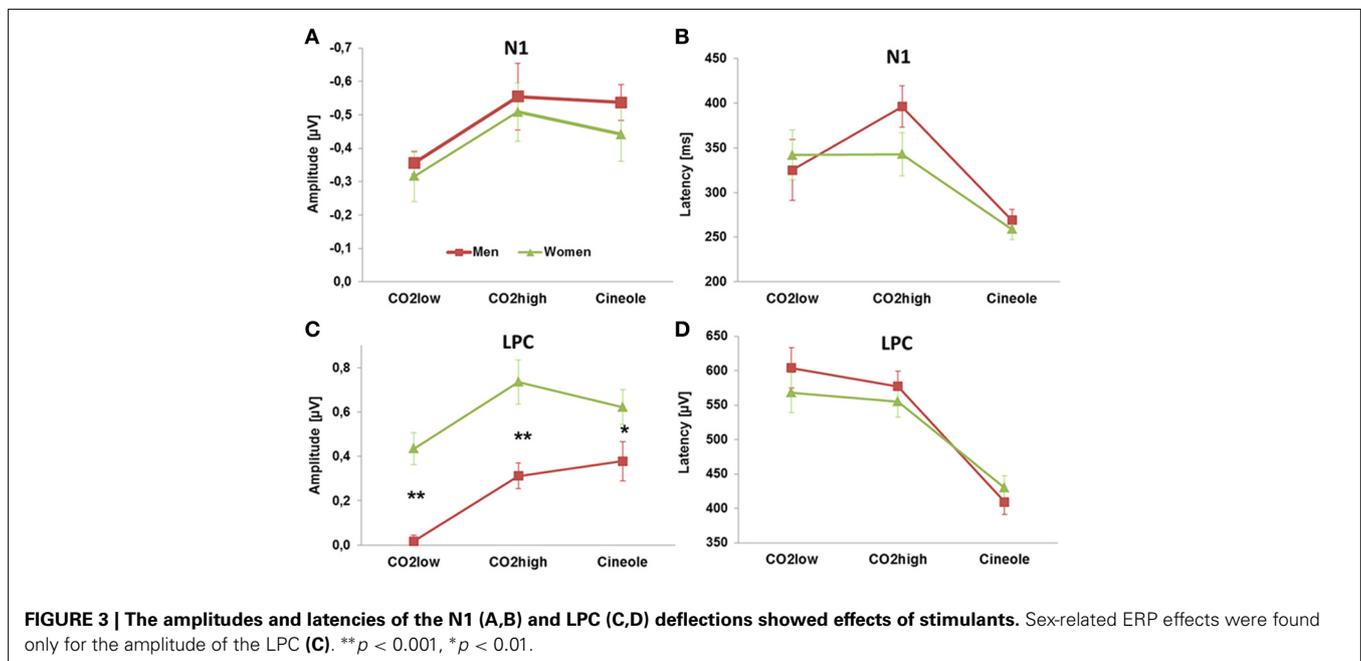
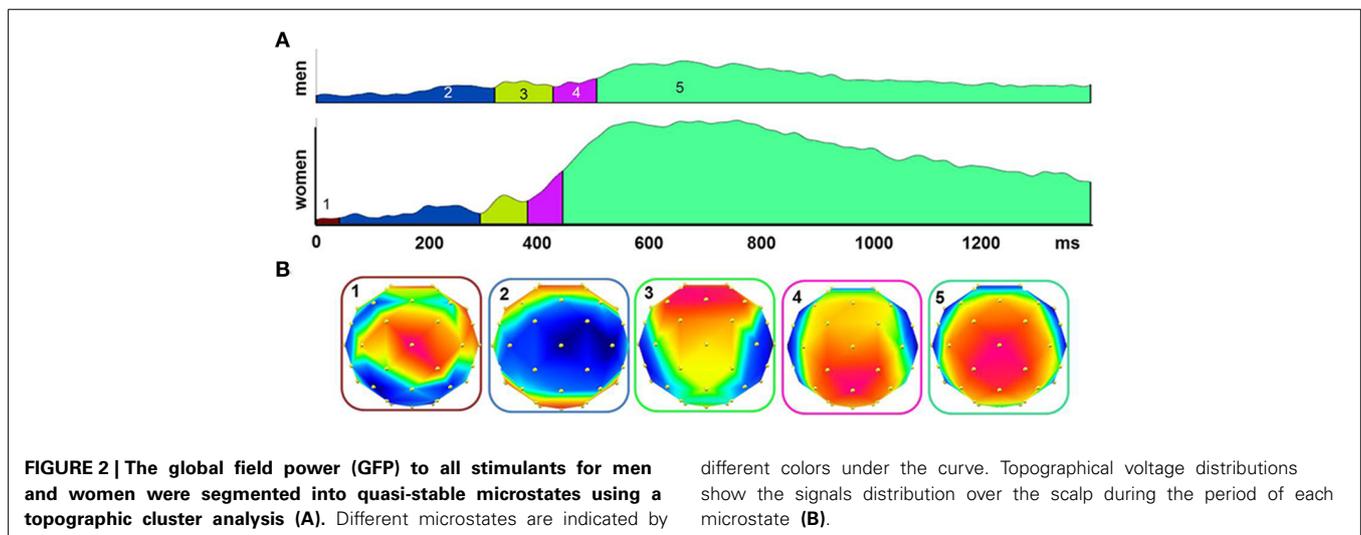
for CO<sub>2</sub>high [ $t_{(1, 28)} = 3.357$ ,  $p < 0.01$ ]. Likewise, LPC latencies were shorter for cineole than both CO<sub>2</sub>low [ $t_{(1, 28)} = 7.876$ ,  $p < 0.001$ ] and for CO<sub>2</sub>high [ $t_{(1, 28)} = 11.378$ ,  $p < 0.001$ ]. **Figure 3** displays the amplitudes and latencies of the N1 and LPC for each stimulant and men and women separately.

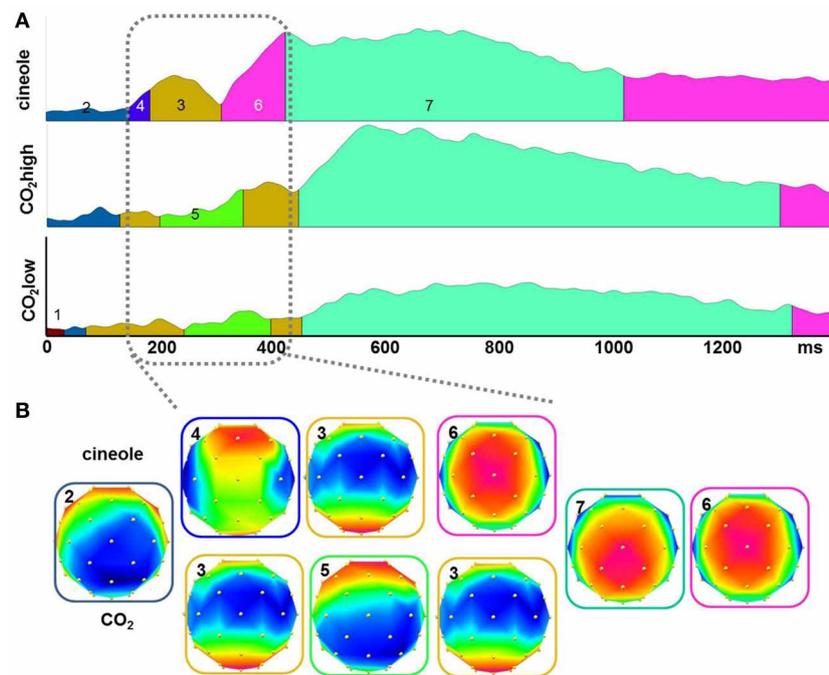
Topographic pattern analyses provided seven microstates accounting for 95.53% of the variance in the grand-averaged ERP data from 0 to 1400 ms of the three experimental conditions. The topographical voltage maps corresponding to each of the seven segments are displayed in **Figure 4B**. The temporal extent of each map is indicated as colored segments under the GFP for each stimulant (**Figure 4A**). Differences in map sequence were apparent between cineole and CO<sub>2</sub> at both intensities; the differences were most prominent during the time period of the N1

deflection and suggest different underlying neuronal generators for the different stimulants.

### BEHAVIORAL DATA

Sensory and behavioral data are summarized in **Table 2**. All participants scored above 11 on the olfactory identification test (mean = 14.3, SEM  $\pm$  0.25, range = 11–16) and we could establish a trigeminal detection sensitivity score in all participants (mean 9.0, SEM  $\pm$  0.45, range = 4.8–15.2). There was no significant sex differences in performance for either odor identification ( $Z = 1.48$ ,  $p = 0.138$ , Mann–Whitney test, 2-tailed) or trigeminal sensitivity ( $Z = 1.004$ ,  $p = 0.315$ , Mann–Whitney test, 2-tailed). Similarly, anxiety scores were similar in men and women for STAI-S ( $Z = 0.153$ ,  $p = 0.879$ ) and STAI-T ( $Z = 1.638$ ,  $p = 0.101$ ; Mann–Whitney test, 2-tailed).





**FIGURE 4 |** The global field power (GFP) from all participants for the three stimulants were segmented into quasi-stable microstates using a topographic cluster analysis (A). Different microstates are indicated by different colors under the curve.

Topographical voltage distributions show the signals distribution over the scalp during the period of each microstate (B). Differences in maps occurrence for the three stimulants were observed during the time period of the N1.

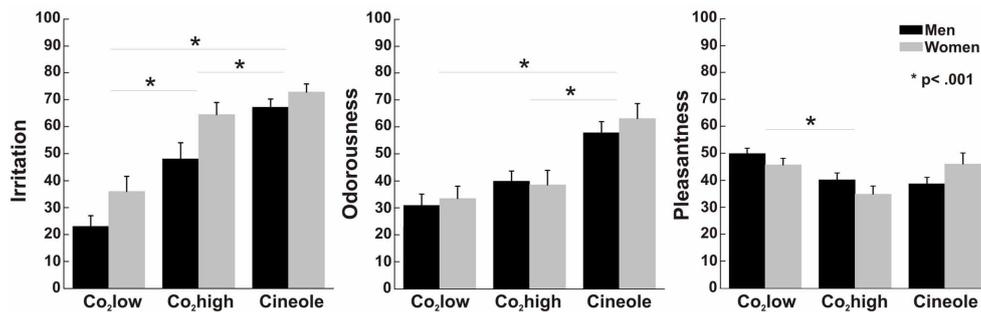
**Table 2 |** Sensory data, anxiety scores, behavioral ratings, and GSR amplitudes (in  $\mu$ S).

		All		Men		Women	
		Mean	SEM	Mean	SEM	Mean	SEM
Odor identification (16-ID)		14.26	0.25	13.86	0.39	14.69	0.29
Menthol threshold		9.00	0.45	8.68	0.68	9.30	0.63
STAI-S*		31.84	1.55	32.54	2.75	31.20	1.64
STAI-T*		34.05	1.62	36.93	2.80	31.37	1.50
Irritation	CO <sub>2</sub> low	29.78	3.64	23.04	3.97	36.08	5.62
	CO <sub>2</sub> high	56.53	3.97	47.94	6.01	64.55	4.51
	Cineole	70.17	2.12	67.22	3.06	72.92	2.85
Odorousness	CO <sub>2</sub> low	31.48	2.98	30.19	4.08	32.69	4.44
	CO <sub>2</sub> high	38.27	3.16	38.80	3.66	37.78	5.18
	Cineole	59.03	3.37	56.33	4.00	61.56	5.40
Pleasantness	CO <sub>2</sub> low	47.15	1.59	49.29	1.98	45.15	2.39
	CO <sub>2</sub> high	37.08	1.98	39.81	2.50	34.54	2.97
	Cineole	41.93	2.45	38.32	2.34	45.30	4.09
GSR	CO <sub>2</sub> low	0.228	0.04	0.239	0.05	0.216	0.07
	CO <sub>2</sub> high	0.312	0.07	0.382	0.09	0.252	0.10
	Cineole	0.428	0.09	0.525	0.13	0.344	0.13

\*Average score before–after.

Figure 5 illustrates participants' ratings to all stimulants. Participants perceived cineole more irritating than CO<sub>2</sub>high [ $t_{(1, 28)} = 3.836, p = 0.001$ ] and CO<sub>2</sub>low [ $t_{(1, 28)} = 10.134, p < 0.001$ ] and CO<sub>2</sub>high was perceived more irritating than CO<sub>2</sub>low [ $t_{(1, 28)} = 8.671, p < 0.001$ ] yielding a main effect of stimulant [ $F_{(1, 27)} = 67.874, p < 0.001, \eta^2 = 0.715$ ]. A significant sex effect [ $F_{(1, 27)} = 5.757, p < 0.05, \eta^2 = 0.176$ ] indicated that women rated the stimulants consistently more irritating (mean = 57.8, SEM = 3.36) than men (mean = 46.1, SEM = 3.52).

Perceived odorousness was similar in men and women [ $F_{(1, 27)} = 0.151, p = 0.701, \eta^2 = 0.006$ ] but it varied for the different stimulants as indicated by a stimulant main effect of Stimulant [ $F_{(1, 27)} = 45.66, p < 0.001, \eta^2 = 0.628$ ]. As expected, cineole was more odorous than CO<sub>2</sub>low [ $t_{(1, 28)} = 7.343, p < 0.001$ ] and CO<sub>2</sub>high [ $t_{(1, 28)} = 6.791, p < 0.001$ ]. Surprisingly, the odorless CO<sub>2</sub>high was rated more odorous than CO<sub>2</sub>low [ $t_{(1, 28)} = 4.058, p < 0.001$ ] despite both stimulants being considered as odorless in their percept. Difficulties to distinguish between odorousness and intensity/irritation, an observation that some participants reported after the experiment, may have contributed to this finding. We therefore calculated Pearson correlation coefficients between odorousness and irritation ratings. Positive correlations were found between odorousness and irritation ratings for CO<sub>2</sub>low ( $r = 0.62, p < 0.001$ ) and CO<sub>2</sub>high ( $r = 0.37, p < 0.05$ ) but not for cineole ( $r = 0.29, p = 0.119$ ), supporting the notion that participants confused odorousness and irritation for the rather odorless CO<sub>2</sub>high.



**FIGURE 5 | Perceptual ratings with error bars representing standard error of the mean.** Women rated all stimuli as more irritating than men but no sex differences were found for either pleasantness and or odorosity. In

all participants, CO<sub>2</sub>low was least irritating, followed by CO<sub>2</sub>high, and then cineole. Cineole was more odorosity and CO<sub>2</sub>low was most pleasant in all participants.

Men and women rated the pleasantness of all stimuli similarly [ $F_{(1, 27)} = 0.066$ ,  $p = 0.799$ ,  $\eta^2 = 0.002$ ]. Independent of sex, ratings varied between the three stimulants as indicated by a main effect of stimulant [ $F_{(1, 27)} = 9.191$ ,  $p = 0.001$ ,  $\eta^2 = 0.254$ ]. In detail, CO<sub>2</sub>low was perceived as more pleasant than CO<sub>2</sub>high [ $t_{(1, 28)} = 6.381$ ,  $p < 0.001$ ], which was rated as least pleasant. A significant interaction between stimulant and sex [ $F_{(1, 27)} = 4.244$ ,  $p < 0.05$ ,  $\eta^2 = 0.136$ ] was found. However, pair-wise  $t$ -tests yielded no significant sex effects for individual odors (all  $t_s < 1.5$ ).

#### GALVANIC SKIN RESPONSES

GSR peak responses were highest for cineole (mean = 0.428, SEM  $\pm$  0.09), intermediate for CO<sub>2</sub>high (mean = 0.312, SEM  $\pm$  0.07), and smallest for CO<sub>2</sub>low (mean = 0.228, SEM  $\pm$  0.04) (see **Table 1**). Cineole elicited significantly higher GSR than CO<sub>2</sub>high [ $t_{(1, 27)} = 2.373$ ,  $p < 0.05$ ] and CO<sub>2</sub>low [ $t_{(1, 27)} = 3.642$ ,  $p = 0.001$ ] and CO<sub>2</sub>high elicited higher GSR than CO<sub>2</sub>low [ $t_{(1, 27)} = 2.551$ ,  $p < 0.05$ ], resulting in a significant main effect of stimulant [ $F_{(2, 52)} = 9.998$ ,  $p < 0.001$ ,  $\eta^2 = 0.278$ ]. Men and women showed similar responses [ $F_{(2, 52)} < 1$ ]. We subsequently assessed, by means of Pearson's correlation coefficients, whether individual anxiety was reflected in the magnitude of the GSR response. Positive correlations were found between GSR and STAI-T scores in women for all stimulants: CO<sub>2</sub>low ( $r = 0.56$ ,  $p < 0.05$ ), CO<sub>2</sub>high ( $r = 0.62$ ,  $p = 0.01$ ), and cineole ( $r = 0.69$ ,  $p < 0.01$ ). However, no such relation was found in men.

#### DISCUSSION

Our results show differential electrophysiological responses to intranasal irritation for women and men: women exhibited significantly increased amplitudes of the late positive ERP potential compared to men. The ERP effect was observed independently of stimulus intensity and of the stimulant used. Yet, women subjectively perceived the stimuli more irritating than men did. Interestingly, men and women were similar with respect to sensory sensitivity, measures of anxiety, and autonomous physiological responses. Consequently, we suggest that women and men process intranasal irritants differently and that this difference is

due to cognitive evaluation of the irritants rather than peripheral differences in sensory sensitivity.

We found increased LPC amplitudes along with higher reported irritation in women as compared to men; importantly, the findings occurred in the absence of sex differences in trigeminal sensitivity and anxiety. Augmented ERP amplitudes to trigeminal and bimodal stimuli in women have been described previously for early (Lundstrom and Hummel, 2006) and late potentials (Olofsson and Nordin, 2004; Lundstrom and Hummel, 2006). In most cases, sex-related effect of the ERPs have been interpreted as, or associated with, heightened sensitivity of women compared to men (Olofsson and Nordin, 2004; Stuck et al., 2006). The present data, however, show sex-specific effects to nasal irritation of the LPC only, while the N1, a marker of both exogenous and endogenous stimulus processing (Pause and Krauel, 2000), was similar for both sexes. In contrast to previous findings (Frasnelli et al., 2011b), men and women displayed similar thresholds for menthol, an odorous irritant, and similar odor identification abilities in our study. Taken together, our findings suggest an absence of strong sex-dependent differences in sensory sensitivity toward nasal stimulation.

Chemosensory ERPs have been less investigated in comparison to ERPs derived from the non-chemical senses. It is for that reason that the late positive deflections of the ERP appear to be inconsistently labeled and categorized. Particularly, a clear dissociation between the chemosensory P2 and P3 has yet to be made (Pause, 2002). The LPC in our present study is characterized by a voltage distribution with a parietal maximum which is indicative for a P3 (Polich, 2007). We therefore refer to it as LPC, a P3-like deflection. The LPC has been shown to reflect the cognitive processing of a stimulus (Polich and Kok, 1995; Polich, 2007) including involuntary (re)allocation of attention (Yamaguchi and Knight, 1991), context updating in memory (Donchin and Coles, 1988), and event categorization (Kok, 2001). These processes are achieved after perceptual analyses of the stimulus and comparison of the percept against internal memory representations, leading to the notion that the LPC represents the final step of perceptual processing (Verleger, 1988). Considering the overall cognitive characterization of the LPC, our findings of

enhanced LPC amplitudes in women probably reflect differential subjective stimulus evaluation and/or emotional classification. It is prudent to point out that although our data fail to provide evidence for sex-differences in emotional responsiveness to the stimuli, as measured by GSR, it is still possible that differences exist in the emotional classification. It is, however, conceivable that augmented LPC amplitudes in women reflect stronger allocation of attention as a consequence of experience and expectations about the stimuli (Carrion and Bly, 2008). This interpretation is further corroborated by recent findings of Andersson et al. (2011), who demonstrated pronounced sex differences for the LPC with larger amplitudes in women than in men, for both the trigeminal CO<sub>2</sub> and the bimodal amyl acetate when the stimuli were attended to but not when the stimuli were to be ignored.

Variations in chemosensory perception have been linked to personality (Croy et al., 2011) and individual level of arousal (Pribram and McGuinness, 1975). Based on its intricate connection to the pain system, one of the primary functions of the intranasal trigeminal system is to act as a sentinel that senses irritation from odorous and odorless stimulants and to warn the body against potentially noxious stimuli (Hummel and Livermore, 2002). It is therefore reasonable to assume that increased levels of anxiety and also heightened arousal drive the susceptibility and sensitivity to irritants and that this relation is more pronounced during or directly after the presentation of irritants; especially in comparison to non-irritating odor stimuli. We assessed trait anxiety before testing and state anxiety just before and after nasal stimulation. Men and women showed no significant differences in anxiety scores, a finding that rules out that anxiety in general and, more specifically, anxiety succeeding the stimulation contributed to the observed sex-related differences in chemosensory ratings and the LPC. Notably, we cannot exclude sex-related differences in attitudes toward the stimuli within the present study. Women have reported a higher interest in the sense of smell than men and attitudes were associated with self-reported olfactory sensitivity in a recent study (Seo et al., 2011). However, whether these findings can be readily transferred from olfactory to trigeminal stimulants needs to be demonstrated. Recent findings do suggest, however, that sensitivity to pain represents a distinct category that is independent of sensitivity to odors (Hummel et al., 2011). Furthermore, we measured GSR, a sensitive measure of autonomic arousal that has been shown to be tied to emotional responsiveness and attention (Neumann and Blanton, 1970), during the presentation of the irritants. We found no differences between men and women. The fact that we found no sex-related differences in arousal processing, as measured by GSR and relevant personality traits like anxiety, stronger support our hypothesis that chemosensory sex differences result from higher cognitive processes. Although the exact generators of the LPC have yet to be identified, several cortical, frontal, temporal, parietal, and subcortical limbic, and thalamic structures have been implicated in its generation (Polich, 2007), thus signifying the involvement of a complex neuronal network that may eventually manifest in differential reports of subjectively experienced irritation.

The responses to bimodal and trigeminal stimulation and to different intensities within the trigeminal modality have been described in previous studies (for example, Iannilli et al., 2013). However, it is pertinent to discriminate intensity and/or modality specificity from sex-related ERP differences. In order to address this problem, we presented two different stimulants and two different intensities of the same stimulant. We observed no interaction between stimulants and sex and between intensity and sex indicating that our reported sex differences are independent of the class of stimulants and also of stimulus intensity. Sex-independent differences were, however, observed between stimulants and intensities. When comparing the responses to cineole and CO<sub>2</sub>, we observed an apparent latency shift of the waveform toward faster responses for cineole; this effect was significant for the LPC. Shorter latencies together with higher amplitudes have been reported for CO<sub>2</sub> in comparison to the non-irritating, less intense phenyl ethyl alcohol (rose-like smell) (Scheibe et al., 2009). Also, different activations pattern of the sensory processing pathways play likely a role: CO<sub>2</sub> compared to the odorous H<sub>2</sub>S yielded increased activation of the anterior cingulate during the first 140 ms; during the subsequent time period until 320 ms, the orbitofrontal cortex responded stronger to the odor than to the irritant (Iannilli et al., 2013). When comparing CO<sub>2</sub>high and CO<sub>2</sub>low, we observed intensity-dependent shifts of the waveform toward shorter latencies and higher amplitudes for both the N1 and LPC. This observation was expected and is in line with the notion that early deflections of the ERP reflect the processing of sensory properties of a stimulus for non-chemical (Coles and Rugg, 1996) and chemical senses (Ohla et al., 2010). In line with this, Frasnelli et al. (2003) have demonstrated a linear relation between concentrations of CO<sub>2</sub> and the amplitudes of early and late ERP deflections. Here, the shift of the LPC can be seen as the consequence of the earlier and enhanced perceptual analysis of the stimulus. The latency of the LPC has indeed been shown to be indicative of a difference in the time to detect and evaluate a stimulus (Kutas et al., 1977; Magliero et al., 1984).

In the present study, we used a comprehensive array of psychological, sensory, and psychophysiological measures to investigate sex-related differences in the perception of intranasal irritation. Our results show that women process intranasal irritation differently than men; this effect was manifested in increased irritation perception and enlarged ERP amplitudes of the LPC. Importantly, the differences cannot be explained by variation in sensory sensitivity to irritants or differences in anxiety. We propose that women allocate more attention to potentially noxious stimuli than men do, which eventually causes differences in cognitive appraisal.

## ACKNOWLEDGMENTS

Support was provided by the Knut and Alice Wallenberg Foundation (KAW 2012.0141) awarded to Johan N. Lundström. The authors thank Andrea Lordan for data collection and pre-processing of GSR and behavioral data. The Cartool software ([brainmapping.unige.ch/cartool](http://brainmapping.unige.ch/cartool)) has been programmed by Denis Brunet, from the Functional Brain Mapping Laboratory, Geneva, Switzerland, and is supported by the Center for Biomedical Imaging (CIBM) of Geneva and Lausanne.

## REFERENCES

- Andersson, L., Bende, M., Millqvist, E., and Nordin, S. (2009). Attention bias and sensitization in chemical sensitivity. *J. Psychosom. Res.* 66, 407–416. doi: 10.1016/j.jpsychores.2008.11.005
- Andersson, L., Lundberg, C., Astrom, J., and Nordin, S. (2011). Chemosensory attention, habituation and detection in women and men. *Int. J. Psychophysiol.* 79, 316–322. doi: 10.1016/j.ijpsycho.2010.11.008
- Andine, P., Ronnback, L., and Jarvholm, B. (1997). Successful use of a selective serotonin reuptake inhibitor in a patient with multiple chemical sensitivities. *Acta Psychiatr. Scand.* 96, 82–83. doi: 10.1111/j.1600-0447.1997.tb09910.x
- Berg, N. D., Linneberg, A., Dirksen, A., and Elberling, J. (2008). Prevalence of self-reported symptoms and consequences related to inhalation of airborne chemicals in a Danish general population. *Int. Arch. Occup. Environ. Health* 81, 881–887. doi: 10.1007/s00420-007-0282-0
- Cain, W. S. (1982). Odor identification by males and females: predictions vs performance. *Chem. Senses* 7, 129–142. doi: 10.1093/chemse/7.2.129
- Carrion, R. E., and Bly, B. M. (2008). The effects of learning on event-related potential correlates of musical expectancy. *Psychophysiology* 45, 759–775. doi: 10.1111/j.1469-8986.2008.00687.x
- Coles, M. G. H., and Rugg, M. D. (1996). "Event-related brain potentials: an introduction," in *Electrophysiology of Mind: Event-Related Brain Potentials and Cognition*, eds M. D. Rugg and M. G. H. Coles (Oxford: Oxford University Press), 1–27.
- Croy, I., Springborn, M., Lotsch, J., Johnston, A. N., and Hummel, T. (2011). Agreeable smellers and sensitive neurotics—correlations among personality traits and sensory thresholds. *PLoS ONE* 6:e18701. doi: 10.1371/journal.pone.0018701
- Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009
- Donchin, E., and Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behav. Brain Sci.* 11, 357–374. doi: 10.1017/S0140525X00058027
- Doty, R. L., Applebaum, S., Zusho, H., and Settle, R. G. (1985). Sex differences in odor identification ability: a cross-cultural analysis. *Neuropsychologia* 23, 667–672. doi: 10.1016/0028-3932(85)90067-3
- Ferdenzi, C., Coureaud, G., Camos, V., and Schaal, B. (2008). Human awareness and uses of odor cues in everyday life: results from a questionnaire study in children. *Int. J. Behav. Dev.* 32, 422–431. doi: 10.1177/0165025408093661
- Fowles, D. C., Christie, M. J., Edelberg, R., Grings, W. W., Lykken, D. T., and Venables, P. H. (1981). Committee report. Publication recommendations for electrodermal measurements. *Psychophysiology* 18, 232–239. doi: 10.1111/j.1469-8986.1981.tb03024.x
- Frasnelli, J., Albrecht, J., Bryant, B., and Lundstrom, J. N. (2011a). Perception of specific trigeminal chemosensory agonists. *Neuroscience* 189, 377–383. doi: 10.1016/j.neuroscience.2011.04.065
- Frasnelli, J., Hummel, T., Berg, J., Huang, G., and Doty, R. L. (2011b). Intranasal localizability of odorants: influence of stimulus volume. *Chem. Senses* 36, 405–410. doi: 10.1093/chemse/bjr001
- Frasnelli, J., and Hummel, T. (2003). Age-related decline of intranasal trigeminal sensitivity: is it a peripheral event? *Brain Res.* 987, 201–206. doi: 10.1016/S0006-8993(03)03336-5
- Frasnelli, J., Lotsch, J., and Hummel, T. (2003). Event-related potentials to intranasal trigeminal stimuli change in relation to stimulus concentration and stimulus duration. *J. Clin. Neurophysiol.* 20, 80–86. doi: 10.1097/00004691-200302000-00011
- Frasnelli, J., Schuster, B., and Hummel, T. (2007). Subjects with congenital anosmia have larger peripheral but similar central trigeminal responses. *Cereb. Cortex* 17, 370–377. doi: 10.1093/cercor/bhj154
- Hall, G. B., Witelson, S. F., Szechtman, H., and Nahmias, C. (2004). Sex differences in functional activation patterns revealed by increased emotion processing demands. *Neuroreport* 15, 219–223. doi: 10.1097/00001756-200402090-00001
- Hillert, A., Koch, S., and Lehr, D. (2013). Burnout phenomenon exemplified by the teaching profession: paradigms, findings and perspectives of profession-related therapy and prevention approaches. *Nervenarzt* 84, 806–812. doi: 10.1007/s00115-013-3745-4
- Hillert, L., Musabasic, V., Berglund, H., Ciumas, C., and Savic, I. (2007). Odor processing in multiple chemical sensitivity. *Hum. Brain Mapp.* 28, 172–182. doi: 10.1002/hbm.20266
- Hummel, T., Barz, S., Pauli, E., and Kobal, G. (1998). Chemosensory event-related potentials change with age. *Electroencephalogr. Clin. Neurophysiol.* 108, 208–217. doi: 10.1016/S0168-5597(97)00074-9
- Hummel, T., and Livermore, A. (2002). Intranasal chemosensory function of the trigeminal nerve and aspects of its relation to olfaction. *Int. Arch. Occup. Environ. Health* 75, 305–313. doi: 10.1007/s00420-002-0315-7
- Hummel, T., Sekinger, B., Wolf, S. R., Pauli, E., and Kobal, G. (1997). 'Sniffin' sticks': olfactory performance assessed by the combined testing of odor identification, odor discrimination and olfactory threshold. *Chem. Senses* 22, 39–52. doi: 10.1093/chemse/22.1.39
- Hummel, T., Springborn, M., Croy, I., Kaiser, J., and Lotsch, J. (2011). High pain sensitivity is distinct from high susceptibility to non-painful sensory input at threshold level. *Int. J. Psychophysiol.* 80, 69–74. doi: 10.1016/j.ijpsycho.2011.01.012
- Iannilli, E., Wiens, S., Arshamian, A., and Seo, H. S. (2013). A spatiotemporal comparison between olfactory and trigeminal event-related potentials. *Neuroimage* 77, 254–261. doi: 10.1016/j.neuroimage.2012.12.057
- Johansson, A., Bramerson, A., Millqvist, E., Nordin, S., and Bende, M. (2005). Prevalence and risk factors for self-reported odour intolerance: the Skovde population-based study. *Int. Arch. Occup. Environ. Health* 78, 559–564. doi: 10.1007/s00420-005-0616-8
- Kobal, G. (1981). *Electrophysiologische Untersuchungen des Menschlichen Geruchssinns*. Stuttgart: Thieme Verlag.
- Kobal, G. (1985). Pain-related electrical potentials of the human nasal mucosa elicited by chemical stimulation. *Pain* 22, 151–163. doi: 10.1016/0304-3959(85)90175-7
- Kobal, G., Hummel, T., Sekinger, B., Barz, S., Roscher, S., and Wolf, S. (1996). "Sniffin' sticks": screening of olfactory performance. *Rhinology* 34, 222–226.
- Koelega, H. S., and Koster, E. P. (1974). Some experiments on sex differences in odor perception. *Ann. N.Y. Acad. Sci.* 237, 234–246. doi: 10.1111/j.1749-6632.1974.tb49859.x
- Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology* 38, 557–577. doi: 10.1017/S0048577201990559
- Krzanowski, W., and Lai, Y. T. (1985). A criterion for determining the number of groups in a data set using sum of square clustering. *Biometrics* 44, 23–34. doi: 10.2307/2531893
- Kutas, M., McCarthy, G., and Donchin, E. (1977). Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science* 197, 792–795. doi: 10.1126/science.887923
- Laska, M., Distel, H., and Hudson, R. (1997). Trigeminal perception of odorant quality in congenitally anosmic subjects. *Chem. Senses* 22, 447–456. doi: 10.1093/chemse/22.4.447
- Lehmann, D., Ozaki, H., and Pal, I. (1987). EEG alpha map series: brain micro-states by space-oriented adaptive segmentation. *Electroencephalogr. Clin. Neurophysiol.* 67, 271–288. doi: 10.1016/0013-4694(87)90025-3
- Lorig, T. S. (2000). The application of electroencephalographic techniques to the study of human olfaction: a review and tutorial. *Int. J. Psychophysiol.* 36, 91–104. doi: 10.1016/S0167-8760(99)00104-X
- Lundstrom, J. N., Frasnelli, J., Larsson, M., and Hummel, T. (2005). Sex differentiated responses to intranasal trigeminal stimuli. *Int. J. Psychophysiol.* 57, 181–186. doi: 10.1016/j.ijpsycho.2005.01.003
- Lundstrom, J. N., and Hummel, T. (2006). Sex-specific hemispheric differences in cortical activation to a bimodal odor. *Behav. Brain Res.* 166, 197–203. doi: 10.1016/j.bbr.2005.07.015
- Lundstrom, J. N., Hummel, T., and Olsson, M. J. (2003). Individual differences in sensitivity to the odor of 4,16-androstadien-3-one. *Chem. Senses* 28, 643–650. doi: 10.1093/chemse/bjg057
- Lundstrom, J. N., McClintock, M. K., and Olsson, M. J. (2006). Effects of reproductive state on olfactory sensitivity suggests odor specificity. *Biol. Psychol.* 71, 244–247. doi: 10.1016/j.biopsycho.2005.07.001
- Magliero, A., Bashore, T. R., Coles, M. G., and Donchin, E. (1984). On the dependence of P300 latency on stimulus evaluation processes. *Psychophysiology* 21, 171–186. doi: 10.1111/j.1469-8986.1984.tb00201.x
- Neumann, E., and Blanton, R. (1970). The early history of electrodermal research. *Psychophysiology*

- 6, 453–475. doi: 10.1111/j.1469-8986.1970.tb01755.x
- Oberg, C., Larsson, M., and Backman, L. (2002). Differential sex effects in olfactory functioning: the role of verbal processing. *J. Int. Neuropsychol. Soc.* 8, 691–698. doi: 10.1017/S1355617702801424
- Ohla, K., Toepel, U., Le Coutre, J., and Hudry, J. (2010). Electrical neuroimaging reveals intensity-dependent activation of human cortical gustatory and somatosensory areas by electric taste. *Biol. Psychol.* 85, 446–455. doi: 10.1016/j.biopsycho.2010.09.007
- Olofsson, J. K., and Nordin, S. (2004). Gender differences in chemosensory perception and event-related potentials. *Chem. Senses* 29, 629–637. doi: 10.1093/chemse/bjh066
- Pascual-Marqui, R. D., Michel, C. M., and Lehmann, D. (1995). Segmentation of brain electrical activity into microstates: model estimation and validation. *IEEE Trans. Biomed. Eng.* 42, 658–665. doi: 10.1109/10.391164
- Pause, B. M. (2002). “Human brain activity during the first second after odor presentation,” in *Olfaction, Taste, and Cognition*, eds C. Rouby, B. Schaal, D. Dubois, R. Gervais, and A. Holley (Cambridge: Cambridge University Press), 309–323. doi: 10.1017/CBO9780511546389.027
- Pause, B. M., and Krauel, K. (2000). Chemosensory event-related potentials (CSERP) as a key to the psychology of odors. *Int. J. Psychophysiol.* 36, 105–122. doi: 10.1016/S0167-8760(99)00105-1
- Pause, B. M., Sojka, B., Krauel, K., and Ferstl, R. (1996). The nature of the late positive complex within the olfactory event-related potential (OERP). *Psychophysiology* 33, 376–384. doi: 10.1111/j.1469-8986.1996.tb01062.x
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019
- Polich, J., and Kok, A. (1995). Cognitive and biological determinants of P300: an integrative review. *Biol. Psychol.* 41, 103–146. doi: 10.1016/0301-0511(95)05130-9
- Pribram, K. H., and McGuinness, D. (1975). Arousal, activation, and effort in the control of attention. Arousal, activation, and effort in the control of attention *Psychol. Rev.* 82, 116–149. doi: 10.1037/h0076780
- Scheibe, M., Opatz, O., and Hummel, T. (2009). Are there sex-related differences in responses to repetitive olfactory/trigeminal stimuli? *Eur. Arch. Otorhinolaryngol.* 266, 1323–1326. doi: 10.1007/s00405-008-0860-0
- Seo, H. S., Guarneros, M., Hudson, R., Distel, H., Min, B. C., Kang, J. K., et al. (2011). Attitudes toward olfaction: a cross-regional study. *Chem. Senses* 36, 177–187. doi: 10.1093/chemse/bjq112
- Shusterman, D., and Balmes, J. (1997). Measurement of nasal irritant sensitivity to pulsed carbon dioxide: a pilot study. *Arch. Environ. Health* 52, 334–340. doi: 10.1080/00039899709602208
- Shusterman, D., Murphy, M. A., and Balmes, J. (2003). Differences in nasal irritant sensitivity by age, gender, and allergic rhinitis status. *Int. Arch. Occup. Environ. Health* 76, 577–583. doi: 10.1007/s00420-003-0459-0
- Sobel, N., Prabhakaran, V., Desmond, J. E., Glover, G. H., Goode, R. L., Sullivan, E. V., et al. (1998). Sniffing and smelling: separate subsystems in the human olfactory cortex. *Nature* 392, 282–286. doi: 10.1038/32654
- Spielberger, C. D., Gorsuch, R. L., and Lushene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Spielberger, C. D., and Sydeman, S. J. (1994). “The state-trait anxiety inventory and state-trait anger expression inventory,” in *The Use of Psychological Testing for Treatment Planning and Outcome Assessment*, ed M. E. Maruish (Hillsdale, NJ: Lawrence Erlbaum Associates), 292–321.
- Stern, R. M., Ray, W. J., and Quigley, K. S. (2001). *Psychophysiological Recording*. New York, NY: Oxford University Press.
- Stuck, B. A., Frey, S., Freiburg, C., Hormann, K., Zahnert, T., and Hummel, T. (2006). Chemosensory event-related potentials in relation to side of stimulation, age, sex, and stimulus concentration. *Clin. Neurophysiol.* 117, 1367–1375. doi: 10.1016/j.clinph.2006.03.004
- Verleger, R. (1988). The true P3 is hard to see: some comments on Kok’s (1986) paper on degraded stimuli. *Biol. Psychol.* 27, 45–50. doi: 10.1016/0301-0511(88)90004-X
- Vigil, J. M. (2009). A socio-relational framework of sex differences in the expression of emotion. *Behav. Brain Sci.* 32, 375–390. discussion: 391–428. doi: 10.1017/S0140525X09991075
- Whittle, S., Yucel, M., Yap, M. B., and Allen, N. B. (2011). Sex differences in the neural correlates of emotion: evidence from neuroimaging. *Biol. Psychol.* 87, 319–333. doi: 10.1016/j.biopsycho.2011.05.003
- Yamaguchi, S., and Knight, R. T. (1991). Anterior and posterior association cortex contributions to the somatosensory P300. *J. Neurosci.* 11, 2039–2054.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 17 June 2013; accepted: 06 September 2013; published online: 26 September 2013.

Citation: Ohla K and Lundström JN (2013) Sex differences in chemosensation: sensory or emotional? *Front. Hum. Neurosci.* 7:607. doi: 10.3389/fnhum.2013.00607

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2013 Ohla and Lundström. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Avoidant symptoms in PTSD predict fear circuit activation during multimodal fear extinction

Rebecca K. Sripada<sup>1,2\*</sup>, Sarah N. Garfinkel<sup>3</sup> and Israel Liberzon<sup>1</sup>

<sup>1</sup> Department of Psychiatry, University of Michigan, Ann Arbor, MI, USA

<sup>2</sup> Veterans Affairs Center for Clinical Management Research, Department of Veterans Affairs Healthcare System, Ann Arbor, MI, USA

<sup>3</sup> Department of Psychiatry, Brighton and Sussex Medical School, Brighton, UK

## Edited by:

Martin Klasen, Rheinisch-Westfälische Technische Hochschule, Aachen University, Germany

## Reviewed by:

Tanja Jovanovic, Emory University, USA

Gregory J. Quirk, University of Puerto Rico School of Medicine, Puerto Rico

## \*Correspondence:

Rebecca K. Sripada, Department of Psychiatry, University of Michigan, 4250 Plymouth Road, 2702 Rachel Upjohn Building, Ann Arbor, MI 48109, USA  
e-mail: rekaufma@umich.edu

Convergent evidence suggests that individuals with posttraumatic stress disorder (PTSD) exhibit exaggerated avoidance behaviors as well as abnormalities in Pavlovian fear conditioning. However, the link between the two features of this disorder is not well understood. In order to probe the brain basis of aberrant extinction learning in PTSD, we administered a multimodal classical fear conditioning/extinction paradigm that incorporated affectively relevant information from two sensory channels (visual and tactile) while participants underwent fMRI scanning. The sample consisted of fifteen OEF/OIF veterans with PTSD. In response to conditioned cues and contextual information, greater avoidance symptomatology was associated with greater activation in amygdala, hippocampus, vmPFC, dmPFC, and insula, during both fear acquisition and fear extinction. Heightened responses to previously conditioned stimuli in individuals with more severe PTSD could indicate a deficiency in safety learning, consistent with PTSD symptomatology. The close link between avoidance symptoms and fear circuit activation suggests that this symptom cluster may be a key component of fear extinction deficits in PTSD and/or may be particularly amenable to change through extinction-based therapies.

**Keywords:** fear conditioning, avoidance, posttraumatic stress disorder, fMRI, neuroimaging, amygdala, hippocampus

## INTRODUCTION

Posttraumatic stress disorder (PTSD) is a debilitating anxiety disorder that afflicts approximately 7 percent of the general population (Kessler et al., 2005). PTSD is characterized by three symptom clusters: reexperiencing, hyperarousal, and avoidance symptoms (APA, 2000). The avoidance cluster includes avoidance of internal and external reminders of the trauma, failure to recall important aspects of the trauma, loss of interest in significant activities, subjective detachment or estrangement from others, restricted range of affect, and sense of foreshortened future (APA, 2000). Some studies suggest that avoidance symptoms track the diagnosis of PTSD better than either of the other two symptom clusters (North et al., 1999). In addition, clinical research indicates that for individuals with PTSD, avoidance symptoms may be the most detrimental symptoms to psychosocial functioning (Hendrix et al., 1998; Riggs et al., 1998; Ruscio et al., 2002; Kuhn et al., 2003; Samper et al., 2004; Lauterbach et al., 2007; Solomon and Mikulincer, 2007; Malta et al., 2009) and quality of life (Lunney and Schnurr, 2007; Schnurr and Lunney, 2008). Furthermore, early avoidance symptoms may predict subsequent PTSD development (Bryant et al., 2000; North et al., 2012). These multiple lines of evidence suggest that avoidant symptoms might signify a key process in PTSD pathophysiology.

In parallel, convergent evidence suggests that PTSD is associated with various abnormalities in fear associated learning, including greater acquisition of conditioned fear, overgeneralization of conditioning, impaired inhibitory learning, and impaired extinction (Orr et al., 2000; Lissek et al., 2005; Milad et al., 2008,

2009; Jovanovic et al., 2009, 2010; Rougemont-Bucking et al., 2011; Mahan and Ressler, 2012; Lommen et al., 2013). It has been suggested that deficits in fear associated learning may play a role in the development (see Lommen et al., 2013) and maintenance (Mahan and Ressler, 2012) of PTSD, and that abnormalities in the extinction and/or retention of conditioned fear may be particularly salient for the persistence of fear memories in PTSD (Milad et al., 2008, 2009). Few studies to date have probed the neural circuitry underlying fear extinction deficits in PTSD, but the existing evidence suggests key roles for amygdala, hippocampus, and vmPFC in this process (Milad et al., 2005, 2007, 2009; Mahan and Ressler, 2012).

Conceptually, deficits in fear-associated learning have been hypothesized to contribute to the development and maintenance of reexperiencing and hyperarousal symptoms. However, the link between fear-associated learning deficits and other key components of PTSD pathophysiology, such as avoidance symptoms, is not well understood. Animal models suggest that avoidance may stem from fear extinction deficits. For instance, Chen et al. (2012) demonstrated that rats displaying greater fear after conditioning go on to exhibit greater behavioral avoidance over a 4 week period. It is also possible that avoidance symptoms may exacerbate fear extinction deficits by reducing the frequency with which individuals come in contact with feared stimuli, thus providing less opportunity for extinction to occur. For instance, socially anxious individuals with more severe avoidance in early treatment experience greater subsequent fear in later treatment (Aderka et al., 2013). However, no

research has investigated the neurobiological underpinnings of this phenomenon.

In order to probe the brain basis of the link between avoidance and aberrant extinction learning in PTSD, we administered a multimodal classical fear conditioning/extinction paradigm that incorporated affectively relevant information from two sensory channels (visual and tactile) in an fMRI environment. Mild shock was used as the unconditioned stimulus (US), and colored lights were used as the conditioned stimuli (CS+ and CS-). We have previously demonstrated that PTSD patients exhibit impaired extinction recall and greater return of extinguished fear when trauma-relevant stimuli are presented, and that these abnormalities are particularly associated with avoidance symptoms (Garfinkel et al., unpublished). Thus, the current study sought to investigate whether individual differences in avoidance symptoms might influence extinction learning. We hypothesized that in response to conditioned stimuli (CS's) and context, individuals with more severe avoidance symptoms would demonstrate greater activity in brain networks related to emotion processing and fear expression.

## MATERIAL AND METHODS

### SUBJECTS

Fifteen right-handed OEF/OIF veterans with PTSD were recruited from the Ann Arbor Veterans Affairs PTSD Clinic. Mean age was 27.3 (SD = 4.5). Ten patients were married and five were single. Twelve patients were Caucasian, one was Asian, one was African-American, and one was Hispanic. Participants were included as part of a larger sample (Sripada et al., 2012a,b; Garfinkel et al., unpublished) that also included healthy combat-exposed controls. All participants received comprehensive psychiatric assessment with the Mini-International Neuropsychiatric Interview (Sheehan et al., 1998) and the Clinician-Administered PTSD Scale (CAPS; Blake et al., 1995). Mean CAPS score was 75.9 (SD = 17.2). All combat exposure (including index trauma for PTSD participants) took place within 5 years prior to study enrollment. Clinical interviews were performed by experienced masters- or doctoral-level clinicians with extensive training in the CAPS, at a subspecialty clinic specializing in PTSD. Exclusion criteria were as follows: (a) psychosis, (b) history of traumatic brain injury, (c) alcohol or substance abuse or dependence in the past 3 months, (d) any psychoactive medication other than sleep aids, (e) left-handedness, (f) presence of ferrous-containing metals within the body, and (g) claustrophobia. Seven participants also met diagnostic criteria for depression, and one had comorbid panic disorder; however, PTSD was always the primary diagnosis. Two participants were using low-dose trazodone as a sleep aid; no other psychiatric medications were permitted. After a complete description of the study was provided to the participants, written informed consent was obtained. The study was approved by the institutional review boards of the University of Michigan Medical School and the Ann Arbor VA Healthcare System. All procedures took place between August 2008 and July 2010.

### TASK

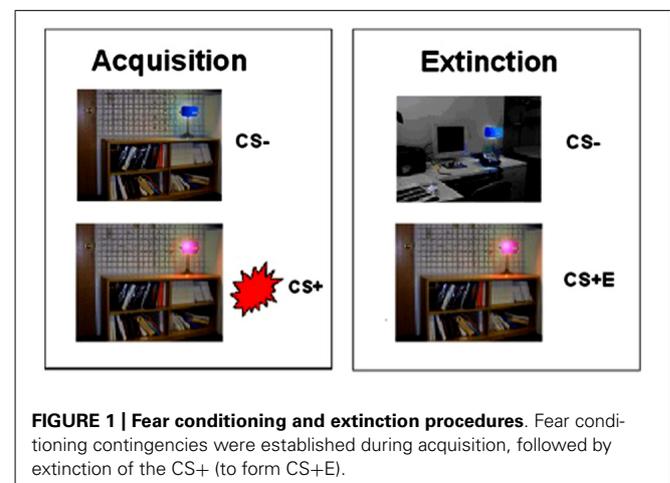
Participants were fear conditioned in a modified version of Milad et al.'s (2005,2007) paradigm. The CS's were colored lights (pink

and blue), presented on a background of an office or library setting (context). The US was an electric shock (500 ms duration pulse sequence) delivered to the index and middle fingers, titrated individually to the level defined as "highly annoying but not painful" (Orr et al., 2000).

Habituation, fear acquisition, and fear extinction all occurred within the scanner in three separate functional runs. Prior to each run, participants were informed that they could receive a shock at any time (see Milad et al., 2007). Habituation involved 12 presentations of the context plus light pairings, and ensured that participants became familiar with stimuli and contexts. During fear acquisition, one context (either the office or library) was presented, counterbalanced on a between-subjects basis. This context remained on the screen for 2–7 s, followed by activation of the light (either pink or blue) for a further 2–7 s, ensuring that the epoch for each context and context + light pairing amounted to 9 s in total. For the CS+, the US was delivered at 60% contingency (10 out of 16 trials), to coincide with CS offset. The other CS was presented 16 times, and was never associated with shock (forming the CS-). The 16 CS- trials were interleaved with the 16 CS+ trials. Each trial was followed by the presentation of a white fixation cross on a black background, jittered for a duration of 12–18 s. Fear extinction followed fear acquisition, and involved a switch in context (from office to library or vice versa). During extinction, the stimuli formerly associated with shock (CS+) were presented in the absence of shock (CS+E), interspersed with presentations of the CS- (see Figure 1). 16 presentations of the CS+E were interleaved with 16 CS- presentations. The 32 trials of fear acquisition and 32 trials of extinction learning were blocked into the first 16 (early) trials and the last 16 (late) trials. In order to isolate successfully acquired (fully learned) conditioning, we restricted our analysis of the acquisition phase to the *late* acquisition phase. Conversely, to maintain a focus on extinction learning rather than extinction retention, our analysis of the extinction phase was restricted to *early* extinction.

### fMRI DATA ACQUISITION

Scans were collected on a 3.0 Tesla General Electric Signa® Excite™ scanner (Milwaukee, WI, USA). After subjects were



**FIGURE 1 | Fear conditioning and extinction procedures.** Fear conditioning contingencies were established during acquisition, followed by extinction of the CS+ (to form CS+E).

positioned in the scanner, a T1-weighted low resolution structural image was acquired approximately parallel to the AC-PC line [gradient recall echo sequence (GRE), repetition time (TR) = 250 ms, echo time (TE) = 5.7 ms, flip angle (FA) = 90°, 2 averages, field of view (FOV) = 22 cm, matrix = 256 × 256, slice thickness = 3 mm, 40 axial slices to cover the whole brain], which was identical to the prescription of the functional acquisitions. Functional images were acquired with a T2\*-weighted, reverse spiral acquisition sequence (gradient recall echo, TR = 2000 ms, TE = 30 ms, FA = 90°, FOV = 22 cm, matrix = 64 × 64, slice thickness = 3 mm with no gap, 40 axial slices to cover the whole brain, acquisition voxel size = 3 × 3 × 3 mm) which has been shown to minimize signal drop-out in regions such as ventral striatum and orbitofrontal cortex that are vulnerable to susceptibility artifact (Glover and Law, 2001). The intermediate template and fMRI images were acquired using a GE Quadrature sending and receiving head coil. Four initial volumes were discarded from each run to allow for equilibration of the scanner signal. A high-quality T1-weighted structural image was obtained with a 3-D volume inversion recovery fast spoiled gradient recalled echo (IR-FSPGR) protocol (TR = 12.3 ms, TE = 5.2 ms, FA = 9°, TI = 650 ms, FOV = 26 cm, matrix = 256 × 256 for in-plane resolution of 1 mm; slice thickness = 1 mm with no gap, 160 contiguous axial slices to cover the whole brain), using an eight-channel GE phase array receiving head coil. E-prime was used to present stimuli (Psychology Software Tools, Pittsburgh, PA, USA). Participants wore glasses with built-in mirrors (NordicNeuro Labs) in order to view the projected stimuli inside the scanner.

### PREPROCESSING OF fMRI DATA

An initial series of preprocessing steps was carried out. First, we removed k-space outliers in raw data that were two standard deviations away from the mean and substituted them with the average value from neighboring voxels. Next, a  $B_0$  field map was used in the reconstruction of the images to remove the distortions that resulted from magnetic field inhomogeneity [IEEE-TIME, 10:629-637, 1991]. The variance due to physiological responses (i.e., cardiac and respiratory sources) was removed using regression (Glover et al., 2000). Additional preprocessing and image analysis was performed in SPM5 (Wellcome Department of Cognitive Neurology, London, UK; <http://www.fil.ion.ucl.ac.uk>). The T2 overlay was co-registered to the functional images, and then the high-resolution T1 image was co-registered to overlay. T1 images were normalized to the scalped T1 template and the functional volumes were normalized to the Montreal Neurological Institute (MNI) template using the previously computed transformation matrix. Images were smoothed using an isotropic 8 mm full-width-half maximum (FWHM) Gaussian kernel.

### ANALYSIS

fMRI comparisons of interest were implemented as linear contrasts. Realignment parameters were added as covariates of no interest at the first level. Z-score images from individual analyses were entered into second-level random-effects analyses (one-sample and two-sample *t*-tests) implemented in SPM5. Second-level maps were thresholded at  $p < 0.001$ , cluster-level

corrected for multiple comparisons via family wise error correction. Regions of interests (ROIs) were selected from a systematic review of fMRI fear conditioning studies (Sehlmeyer et al., 2009), and defined using the automated anatomical labeling atlas (AAL). They included amygdala, hippocampus, vmPFC (bilateral medial orbital frontal gyrus), dmPFC (bilateral superior medial frontal gyrus), and insula. Only the clusters within the regions of interest that survived family wise error small volume correction were extracted and used for further analysis. Symptom severity was assessed via the CAPS, which consists of three subscales (reexperiencing symptoms, avoidance symptoms, and hyperarousal symptoms) that are summed for a total score. Bivariate correlations were computed between CAPS score and the extracted BOLD signal during fear acquisition and fear extinction.

## RESULTS

Subjects were successfully fear conditioned, as reported elsewhere (Garfinkel et al., unpublished). As predicted, fear conditioning activated a network of fear processing regions in response to CS+ presentation in both PTSD patients and Combat-exposed Controls. There were no between-group differences in brain activation patterns or skin conductance responses during the conditioning phase (Garfinkel et al., unpublished).

### CORRELATIONS WITH AVOIDANCE SYMPTOMS

#### Fear acquisition

To explore whether avoidance symptoms were associated with differential neural activation patterns during conditioning, the CAPS avoidance subscale was entered as a regressor in a whole brain analysis during the fear acquisition phase. During CS+ (as compared to CS-) greater CAPS avoidance was associated with greater activity in right hippocampus ([33, -27, -9],  $k = 9$ ,  $z = 3.62$ ,  $p = 0.025$ , SVC; see **Table 1**). This correlation remained significant after controlling for levels of reexperiencing and hyperarousal symptoms. No other significant associations were observed.

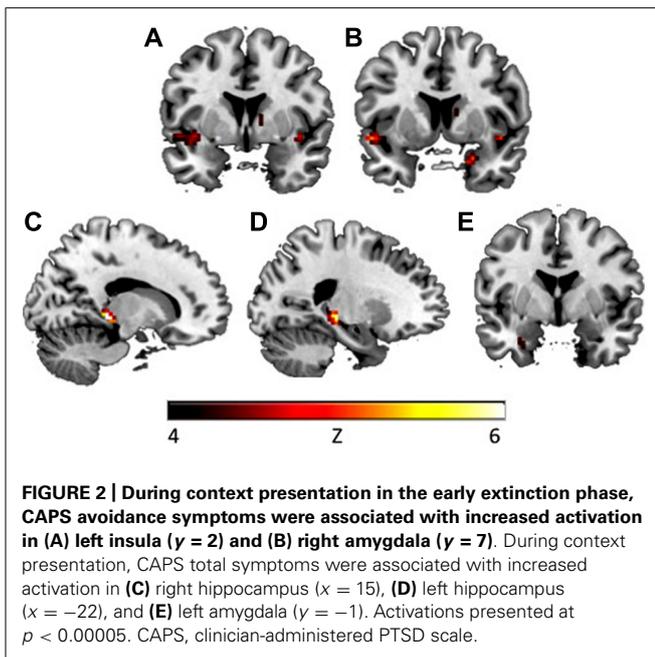
#### Fear extinction

To explore whether avoidance symptoms were associated with differential activation patterns during extinction learning, the CAPS avoidance subscale was entered as a regressor in a whole brain analysis during the fear extinction phase. During context presentation prior to CS presentation (as compared to fixation), avoidance was associated with greater activity in left hippocampus ([-21, -30, -3],  $k = 3$ ,  $z = 3.42$ ,  $p = 0.047$ , SVC), left insula ([-48, 6, -9],  $k = 21$ ,  $z = 3.52$ ,  $p = 0.05$ , SVC), and right amygdala ([24, 3, -24],  $k = 1$ ,  $z = 3.22$ ,  $p = 0.025$ , SVC; see **Figure 2**). Correlations with insula and amygdala remained significant after controlling for other PTSD symptom clusters. During CS+E (as compared to CS-), greater avoidance was associated with greater activity in right amygdala ([24, 3, -24],  $k = 5$ ,  $z = 3.89$ ,  $p = 0.002$ , SVC), vmPFC ([3, 45, -6],  $k = 14$ ,  $z = 3.35$ ,  $p = 0.05$ , SVC), dmPFC ([-6, 30, 57],  $k = 36$ ,  $z = 3.77$ ,  $p = 0.049$ , SVC), left insula ([-39, -18, -3],  $k = 2$ ,  $z = 3.45$ ,  $p = 0.05$ , SVC), and left hippocampus ([-24, -21, -12],  $k = 17$ ,  $z = 3.79$ ,  $p = 0.011$ , SVC; see **Figure 3**). The correlation with left hippocampus remained significant after controlling for other PTSD symptom clusters.

**Table 1 | Correlations with CAPS avoidance symptoms.**

Contrast map and brain region	Cluster size	MNI coordinates (x, y, z)	Analysis (z)
Fear acquisition			
CS+ > CS-			
<b>Right hippocampus</b>	<b>9</b>	<b>33, -27, -9</b>	<b>3.62</b>
Fear extinction			
Context presentation > fixation			
Right superior temporal gyrus/insula	80	45, -12, -27	3.77
Left superior temporal gyrus	67	-51, 6, -9	3.64
<b>Left insula</b>	<b>21</b>	<b>-48, 6, -9</b>	<b>3.52</b>
Right middle temporal gyrus	61	42, -60, -6	3.94
Left middle temporal gyrus	61	-63, -63, 6	3.72
Left insula/thalamus	76	-27, -30, 12	3.83
<b>Left hippocampus</b>	<b>3</b>	<b>-21, -30, -3</b>	<b>3.42</b>
Right caudate	79	15, 21, 3	3.73
<b>Right amygdala</b>	<b>1</b>	<b>24, 3, -24</b>	<b>3.22</b>
Fear extinction			
CS+ > CS-			
Left inferior/middle temporal gyrus	185	-51, -6, -36	4.07
<b>Left insula</b>	<b>2</b>	<b>-39, -18, -3</b>	<b>3.45</b>
Left cerebellum	232	-9, -87, -33	4.17
Left inferior temporal gyrus	154	-72, -30, -12	4.67
Right middle temporal gyrus	333	60, -54, -6	5.11
Left superior temporal gyrus	162	-63, -51, 21	4.15
Right precuneus	184	12, -75, 48	3.72
<b>Right amygdala</b>	<b>5</b>	<b>24, 3, -24</b>	<b>3.89</b>
<b>Medial orbital frontal gyrus</b>	<b>14</b>	<b>3, 45, -6</b>	<b>3.35</b>
<b>Superior medial frontal gyrus</b>	<b>36</b>	<b>-6, 30, 57</b>	<b>3.77</b>
<b>Left hippocampus</b>	<b>17</b>	<b>-24, -21, -12</b>	<b>3.79</b>
Fear Extinction			
Fixation after CS+ > fixation after CS-			
Right fusiform/parahippocampal gyrus	318	36, -12, -30	5.03
<b>Right amygdala</b>	<b>9</b>	<b>27, 3, -27</b>	<b>3.62</b>
Left inferior orbital frontal gyrus	532	-45, 21, -12	4.47
<b>Left insula (anterior)</b>	<b>70</b>	<b>-39, 15, -12</b>	<b>4.35</b>
<b>Medial orbital frontal gyrus</b>	<b>11</b>	<b>-6, 21, -15</b>	<b>3.40</b>
Left insula/middle temporal gyrus	230	-69, -27, -12	3.74
<b>Left insula (posterior)</b>	<b>52</b>	<b>-36, -18, 6</b>	<b>3.63</b>
Right superior temporal pole	77	54, 18, -15	3.68
<b>Right insula</b>	<b>11</b>	<b>42, 6, -6</b>	<b>3.66</b>
Right hippocampus/parahippocampal gyrus	134	36, -30, -12	3.90
<b>Right hippocampus</b>	<b>26</b>	<b>36, -30, -12</b>	<b>3.90</b>
Supplementary motor area	255	-3, 9, 72	3.74
Right precentral gyrus	102	27, -24, 75	3.53

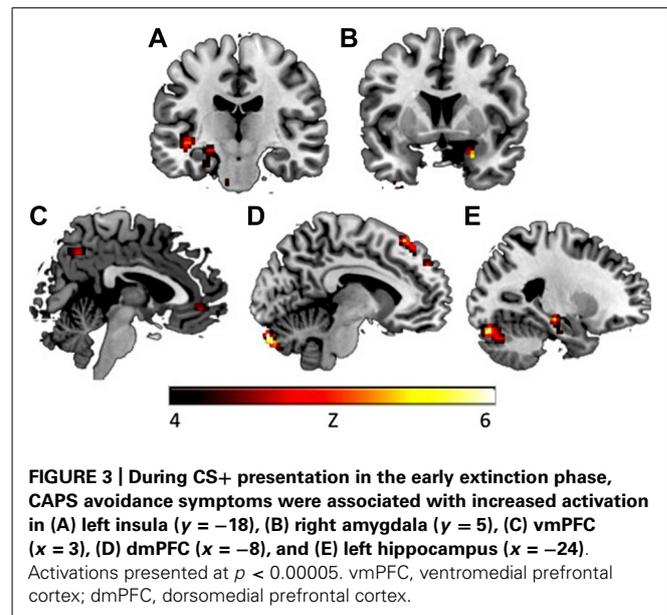
\*Regions of interest (ROIs) in **bold**; significant at  $p < 0.05$ , family wise error corrected for multiple comparisons across the ROI. All other activations are presented at  $p < 0.001$ , cluster-level corrected for multiple comparisons via family wise error correction. MNI, Montreal Neurologic Institute.



Immediately following CS+E, during the period that involved shock administration while in the acquisition phase, greater avoidance was associated with greater activity in right amygdala ([27, 3, -27],  $k = 9$ ,  $z = 3.62$ ,  $p = 0.007$ , SVC), right insula ([42, 6, -6],  $k = 11$ ,  $z = 3.66$ ,  $p = 0.038$ , SVC), right hippocampus ([36, -30, -12],  $k = 26$ ,  $z = 3.9$ ,  $p = 0.009$ , SVC), left anterior insula ([-39, 15, -12],  $k = 70$ ,  $z = 4.35$ ,  $p = 0.004$ , SVC), left posterior insula ([-36, -18, 6],  $k = 52$ ,  $z = 3.63$ ,  $p = 0.042$ , SVC), and a trend for vmPFC ([33, -27, -9],  $k = 11$ ,  $z = 3.4$ ,  $p = 0.06$ , SVC; see **Figure 4**). Correlations with left and right insula remained significant after controlling for other PTSD symptom clusters. Whole brain activations are reported in **Table 1**.

#### CORRELATIONS WITH CAPS TOTAL

To explore whether total PTSD symptoms were associated with differential activation patterns during fear acquisition and extinction learning, correlations were also computed between BOLD signal and CAPS total score. During extinction, in response to context presentation prior to CS, CAPS total was associated with greater activity in right hippocampus (2 clusters: [15, -36, 0],  $k = 4$ ,  $z = 3.71$ ,  $p = 0.017$ , SVC; [18, -33, -2],  $k = 2$ ,  $z = 3.42$ ,  $p = 0.039$ , SVC), left hippocampus (2 clusters: [-24, -30, -3],  $k = 23$ ,  $z = 3.93$ ,  $p = 0.011$ , SVC; [-27, -15, -12],  $k = 2$ ,  $z = 3.45$ ,  $p = 0.042$ , SVC), and left amygdala ([-30, 0, -27],  $k = 1$ ,  $z = 3.21$ ,  $p = 0.021$ , SVC; see **Table 2**; **Figure 2**). Left ( $p = 0.015$ , SVC) and right ( $p = 0.032$ , SVC) hippocampal activity was also significantly correlated with the sum of reexperiencing and hyperarousal symptom clusters (i.e., total CAPS score minus avoidance subscale). Immediately following CS+E, CAPS total was associated with greater activity in right amygdala (2 clusters: [33, -3, -27],  $k = 4$ ,  $z = 3.35$ ,  $p = 0.015$ , SVC; [27, 3, -27],  $k = 1$ ,  $z = 3.10$ ,  $p = 0.029$ , SVC), right hippocampus (2 clusters: [36, -9, -27],  $k = 37$ ,  $z = 3.83$ ,  $p = 0.012$ , SVC; [21, -36, 6],



$k = 6$ ,  $z = 3.31$ ,  $p = 0.05$ , SVC), and right insula (2 clusters: [45, 0, -6],  $k = 32$ ,  $z = 4.39$ ,  $p = 0.003$ , SVC; [33, -15, 6],  $k = 4$ ,  $z = 3.59$ ,  $p = 0.044$ , SVC; see **Figure 4**). Whole brain activations are reported in **Table 2**.

#### CORRELATIONS WITH OTHER SYMPTOM CLUSTERS

For completeness, we also conducted an exploratory analysis to investigate whether reexperiencing or hyperarousal symptoms were associated with unique activation patterns during fear acquisition and extinction learning. During extinction, in response to context presentation prior to CS, CAPS hyperarousal was associated with greater activity in left hippocampus ([-24, -24, -12],  $k = 27$ ,  $z = 4.17$ ,  $p = 0.004$ , SVC). Reexperiencing symptoms were not associated with differential activation in any ROI.

#### TIME COURSE OF ROI ACTIVATION

We implemented a series of repeated-measures ANOVAs to investigate the degree to which the different stages of extinction (context presentation, CS+E presentation, and immediate aftermath of CS+E presentation) activated key limbic and prefrontal regions. This analysis revealed that dmPFC activation varied by stage ( $F(26,2) = 3.82$ ,  $p = 0.035$ ). *Post-hoc* LSD tests showed that dmPFC exhibited greater activation during context processing ( $M = 0.26$ ,  $SD = 0.67$ ) than during CS+E presentation ( $M = -0.47$ ,  $SD = 1.12$ ;  $p = 0.048$ ) or during the immediate aftermath of CS+E presentation ( $M = -0.16$ ,  $SD = 0.72$ ;  $p = 0.045$ ). No other ROI demonstrated significantly different activation across stages.

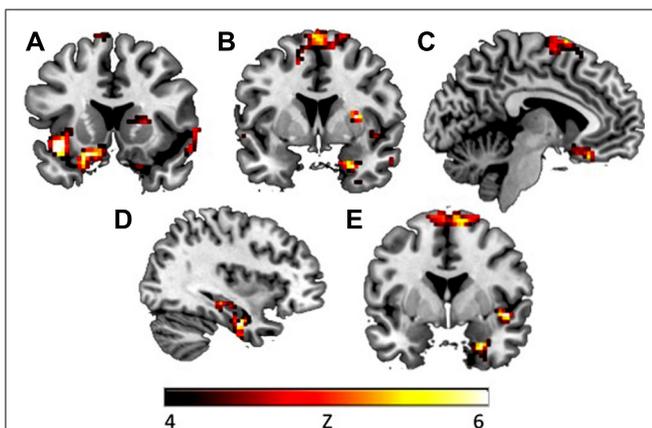
#### DISCUSSION

In this fMRI study, we investigated the neural underpinnings of the link between fear extinction learning and avoidant symptoms in PTSD. We found that amongst individuals with PTSD, greater avoidance symptomatology was associated with greater activation in emotion processing circuits in response to conditioned cues and contextual information, both during fear acquisition and fear

**Table 2 | Correlations with CAPS total symptoms.**

Contrast map and brain region	Cluster size	MNI coordinates (x, y, z)	Analysis (z)
Fear extinction			
Context presentation > fixation			
<b>Right hippocampus (posterior)</b>	<b>4</b>	<b>15, -36, 0</b>	<b>3.71</b>
<b>Right hippocampus (anterior)</b>	<b>2</b>	<b>18, -33, -2</b>	<b>3.42</b>
<b>Left hippocampus (posterior)</b>	<b>23</b>	<b>-24, -30, -3</b>	<b>3.93</b>
<b>Left hippocampus (anterior)</b>	<b>2</b>	<b>-27, -15, -12</b>	<b>3.45</b>
<b>Left amygdala</b>	<b>1</b>	<b>-30, 0, -27</b>	<b>3.21</b>
Fear extinction			
Fixation after CS+ > fixation after CS-			
Right fusiform/parahippocampal gyrus	168	36, -9, -30	4.4
<b>Right hippocampus (anterior)</b>	<b>37</b>	<b>36, -9, -27</b>	<b>3.83</b>
<b>Right amygdala (lateral)</b>	<b>4</b>	<b>33, -3, -27</b>	<b>3.35</b>
<b>Right amygdala (medial)</b>	<b>1</b>	<b>27, 3, -27</b>	<b>3.10</b>
Left cerebellum/fusiform gyrus	97	-24, -42, -15	3.47
Supplementary motor area	256	9, -3, 69	3.90
<b>Right hippocampus (posterior)</b>	<b>6</b>	<b>21, -36, 6</b>	<b>3.31</b>
<b>Right insula (anterior)</b>	<b>32</b>	<b>45, 0, -6</b>	<b>4.39</b>
<b>Right insula (posterior)</b>	<b>4</b>	<b>33, -15, 6</b>	<b>3.59</b>

\*Regions of interest (ROIs) in **bold**; significant at  $p < 0.05$ , family wise error corrected for multiple comparisons across the ROI. All other activations are presented at  $p < 0.001$ , cluster-level corrected for multiple comparisons via family wise error correction. MNI, Montreal Neurologic Institute.



**FIGURE 4 |** During fixation after CS+ in the early extinction phase, CAPS avoidance symptoms were associated with increased activation in (A) left insula ( $y = -15$ ), (B) right amygdala ( $y = 3$ ), and (C) vmPFC ( $x = -6$ ). During fixation after CS+, CAPS total symptoms were associated with increased activation in (D) right hippocampus ( $x = 36$ ), (E) right insula and right amygdala ( $y = 0$ ). Activations presented at  $p < 0.00005$ .

extinction. This pattern was observed during presentation of context immediately prior to the CS (i.e., “context alone”), during the presentation of the CS+E, and immediately following the CS+E. Correlations with insula, amygdala, and hippocampus survived after controlling for other PTSD symptom clusters. Heightened responses to previously conditioned stimuli in individuals with

more avoidant symptoms or more severe PTSD could indicate a deficiency in safety learning, consistent with PTSD symptomatology. The close link between avoidance symptoms and fear circuit activation suggests that this symptom cluster may be a key component of fear extinction deficits in PTSD and/or may be particularly amenable to change through extinction-based therapies.

The multimodal nature of this experiment enhanced its applicability to PTSD. Research demonstrates that multimodal trauma experiences may exacerbate PTSD symptoms, but also that multimodal treatment can enhance efficacy. For instance, extending the duration of context presentation (including tactile, visual, and olfactory cues) before foot-shock administration (in mice) increases generalization and avoidance (Sauerhofer et al., 2012). This finding was interpreted to suggest that fostering multimodal learning enhances conditioning (Sauerhofer et al., 2012). Conversely, emerging evidence suggests that individuals with high avoidance may benefit more from treatment when it incorporates multisensory trauma cues, which provides less opportunity for further avoidance (Rizzo et al., 2009; Norrholm and Jovanovic, 2010). For instance, in virtual reality exposure therapy, patients are immersed in simulations of trauma-relevant environments that allow for precise control of stimulus conditions. Directly delivering these multimodal cues can help circumvent clinical avoidance. Changes in PTSD symptom clusters, and particularly symptoms in the avoidant cluster, may be the key mechanism of PTSD treatment efficacy (Monson et al., 2012). Our paradigm is inherently multimodal because it involves manipulation of context (visual) and conditioned cues (tactile). Thus, it may be a more effective

PTSD probe than unimodal paradigms. That avoidance symptoms were extensively correlated with brain activation while the other symptom clusters were not suggests that the multimodal nature of the paradigm was particularly effective at drawing out avoidant tendencies.

The current study suggests that avoidance symptoms are associated with hyperactivity in a variety of regions key to emotion processing and extinction learning (Sehlmeyer et al., 2009), including hippocampus, amygdala, insula, and medial prefrontal regions. In our data, avoidance symptoms were associated with greater hippocampal activity across both fear acquisition and fear extinction phases. Correlations between avoidance and hippocampal activity were observed during the presentation of context alone, during the presentation of previously conditioned cues, and immediately following the presentation of previously conditioned cues (during the period in which participants were shocked while in the acquisition phase). These findings are consistent with the role of hippocampus in contextual information processing (Maren et al., 2013) and with its role in “binding” contextual information with fear cues (Fanselow, 2000; Maren, 2001). Hippocampal activation was also associated with greater overall symptom severity, even after controlling for avoidance symptoms. This finding could help explain previous reports that PTSD patients exhibit greater hippocampal activity than healthy controls during fear acquisition and extinction learning (Bremner et al., 2005) or during non-fear related encoding (Werner et al., 2009). Interestingly, Milad et al. (2009) report *reduced* hippocampal activity in PTSD patients during extinction recall, which typically occurs the day after the conditioning and extinction phases. Conversely, in the present study, we found associations with avoidance during extinction *learning*, during the first 16 trials of extinction. This may reflect enhanced encoding or processing of conditioned associations formed during the acquisition phase in individuals with higher avoidance. This would suggest that higher avoidance is not only related to the expression of acquired fear, but also to fear learning. It is also possible that higher hippocampal activation in high avoidance patients reflects emotional rather than memory processing. Indeed, it has been demonstrated that anterior hippocampal regions in humans (which are analogous to ventral hippocampal regions in rodents; Moser and Moser, 1998), are involved in affect processing (Fanselow and Dong, 2010).

Avoidance symptom severity was also positively associated with amygdala and insula activity. These associations were present during context alone, during the presentation of previously conditioned cues, and immediately following conditioned cues. This too is consistent with previous animal and human findings. During extinction, high-anxious rats show hyperactivation (increased c-Fos expression) of the central nucleus of the amygdala (Muigg et al., 2008). Similarly, neuroimaging studies of individuals with PTSD report amygdala hyperactivity (Milad et al., 2009) and insula hyperactivity (Bremner et al., 2005) during extinction learning. Greater amygdala and insula activation during extinction is also correlated with trait anxiety (Barrett and Armony, 2009; Sehlmeyer et al., 2011). Insula and amygdala are key regions in salience detection and anticipation of negative events (Armony and LeDoux, 1997; Paulus and Stein, 2006). These regions are also

associated with negative emotion production in PTSD, more generally (Shin and Liberzon, 2010). Greater activity in these emotion generation regions could thus reflect hyperactive fear responding to signals previously paired with negative outcomes. It could also reflect a failure to encode safety signals, or failure to adapt to or integrate new contextual information into previously learned contingencies (Liberzon and Sripada, 2008; Garfinkel and Liberzon, 2009).

We also found that avoidance symptom severity correlated with greater dmPFC and vmPFC activation. Greater dmPFC activity is consistent with previous findings of dmPFC/dACC hyperactivity in PTSD. For instance, Milad et al. (2009) report greater dACC activity during extinction recall in patients with PTSD, and Rougemont-Buckling et al. (2011) report exaggerated dACC activation in response to context presentation during late conditioning and early extinction. Other studies support that dmPFC/dACC hyperactivity in PTSD is also present during cognitive interference tasks such as oddball tasks (Bryant et al., 2005; Felmingham et al., 2009), Stroop tasks (Shin et al., 2007) and the Multi-Source Interference Task (Shin et al., 2011). The greater vmPFC activity found in highly avoidant patients, on the other hand, seemingly diverges from some previous reports of hypoactive vmPFC in PTSD (Bremner et al., 2005; Milad et al., 2009; Rougemont-Buckling et al., 2011). However, Barrett and Armony (2009) report that greater trait anxiety is associated with greater vmPFC activity during extinction. In our data, it is possible that vmPFC hyperactivity could represent a compensatory response to down-regulate amygdala activity. More broadly, the relationship between greater symptom severity and widespread hyperactivity across fear and emotion circuitry suggests that PTSD symptomatology is associated with greater neural reactivity during extinction learning. This hyperactivity may give rise to aberrant extinction retention.

Our findings suggest that PTSD symptoms and avoidance symptoms in particular are associated with exaggerated fear circuit activity. Previous fear conditioning studies in PTSD have largely focused on reexperiencing and hyperarousal symptom clusters. Reexperiencing symptoms have been demonstrated to be associated with greater fear-potentiated startle during fear acquisition and extinction (Glover et al., 2011; Norrholm et al., 2011). Hyperarousal symptoms, too, are associated with exaggerated fear responding (Jovanovic et al., 2010; Glover et al., 2011; Norrholm et al., 2011). One study reported that in response to script-driven imagery, avoidance symptoms were negatively correlated with vmPFC/rACC activation (Hopper et al., 2007). To our knowledge, however, ours is the first study to demonstrate a link between avoidance symptoms and greater reactivity to cue and context processing. Studies using animal models can provide important insight into the link between avoidance and fear extinction. Some investigators have suggested that the construct of avoidance involves both non-associative novelty fear, which is ameliorated by habituation, and stimulus-specific associative fear, which is ameliorated by extinction training (Pamplona et al., 2011). Generalized avoidance behavior in mice is reduced through both habituation and through extinction training (Costanzi et al., 2011; Pamplona et al., 2011), suggesting that both novelty fear and stimulus-specific fear contribute to avoidance behavior. Our data is consistent with this dual conceptualization, since it

demonstrates that avoidance is associated with exaggerated limbic responding to both context (novel) and cues (conditioned). Individuals with greater avoidance symptoms may be more sensitive to both types of stimuli.

The relationship between fear extinction deficits and avoidance symptoms in PTSD might be bidirectional. Previous studies suggest that extinction deficits can lead to the development of avoidance symptoms, and conversely that pre-existent “higher avoidance” can be a contributor to extinction deficits. In support of the first hypothesis, greater fear in response to aversive stimuli is associated with greater levels of subsequent avoidance in rats (Chen et al., 2012). Additionally, pre-trauma deficits in extinction learning are associated with greater risk for developing PTSD after trauma in Dutch soldiers (Lommen et al., 2013). Thus, our findings of greater activity in fear circuits could reflect a mechanism by which individuals develop greater avoidance symptomatology, and provide additional support for the notion of harnessing fear extinction for the purpose of effective avoidance reduction. Alternatively, avoidance could precede fear extinction deficits, in that avoidance symptoms could result in greater fear responding (or amygdala hyperactivity) when confronted with fear-related stimuli that are usually avoided. Theoretical models of PTSD suggest that chronic avoidance leads to greater intensity of avoided cognitions and emotions (Hayes et al., 1999). Additionally, Aderka et al. (2013) recently reported that fear and avoidance predict each other during cognitive-behavioral therapy for social anxiety disorder. Our findings could also reflect resistance to extinction in individuals with greater avoidance. There is evidence to suggest that individuals with greater avoidance have poorer response to CBT (Taylor et al., 2001) and greater rates of attrition (Glynn et al., 1999), though other studies have found that avoidant coping predicts better response to exposure therapy (Leiner et al., 2012). Furthermore, avoidance symptoms may not be as responsive to trauma-focused treatment as other PTSD symptom clusters (Glynn et al., 1999). Longitudinal studies are needed to determine whether fear circuit hyperactivity during extinction is better understood as a risk factor for avoidance symptoms or as a consequence of these symptoms.

Our study had several limitations. First, our paradigm is multimodal in that it involves both visual and tactile cues. However, the conditioned stimuli were primarily visual in nature. Thus,

future studies on the relationship between PTSD symptoms and fear conditioning abnormalities could use additional modalities, such as olfactory cues, to further probe the multimodal nature of the link between avoidance symptoms and fear extinction. Second, the design used in this study, i.e., fear conditioning followed by fear extinction, does not allow us to clearly disambiguate the effects of extinction learning from the potential effects of differential recall of CS+ memory trace. For example, the amygdala hyperactivity we observed during extinction may reflect either extinction learning or recall of the CS+ conditioning (see Quirk and Mueller, 2008). Similarly, greater hippocampal activity during both conditioning and extinction learning in avoidant individuals could indicate overconsolidation of fear during conditioning, greater recall of conditioning in the extinction phase, or overgeneralization of fear expression into a neutral context. Future studies could use on-line expectancy ratings to help distinguish between these alternatives. A full factorial design utilizing both new and previously viewed contexts in conjunction with new and previously viewed cues would also be helpful in distinguishing novelty fear from the effects of conditioning. From a clinical perspective, however, PTSD symptoms could be similarly exacerbated by either deficient extinction learning or excessive acquisition-related fear. As such, this phase of extinction may provide a valuable target for research on treatment-enhancing approaches.

In conclusion, our results demonstrate that individuals with greater levels of avoidance exhibit hyperactivation in brain regions involved in fear expression during the presentation of previously conditioned cues and contextual information. This represents a potential brain-based mechanism contributing to the maintenance of fear memories in PTSD patients. Our findings suggest that ameliorating impaired inhibition of fear is an important treatment target for PTSD, in particular for PTSD patients with high levels of avoidance.

## ACKNOWLEDGMENTS

The research reported in this article was supported by grants from the Michigan Institute for Clinical and Health Research (U028028) to SG, from the National Institute of Mental Health (R24 MH075999) to IL, and from the Telemedicine and Advanced Technology Research Center (W81XWH-08-2-0208) to IL.

## REFERENCES

- Aderka, I. M., Mclean, C. P., Huppert, J. D., Davidson, J. R., and Foa, E. B. (2013). Fear, avoidance and physiological symptoms during cognitive-behavioral therapy for social anxiety disorder. *Behav. Res. Ther.* 51, 352–358. doi: 10.1016/j.brat.2013.03.007
- APA. (2000). *Diagnostic and Statistical Manual of Mental Disorders*, 4th Edn. Washington, DC: American Psychiatric Press.
- Armony, J. L., and LeDoux, J. E. (1997). How the brain processes emotional information. *Ann. N. Y. Acad. Sci.* 821, 259–270. doi: 10.1111/j.1749-6632.1997.tb48285.x
- Barrett, J., and Armony, J. L. (2009). Influence of trait anxiety on brain activity during the acquisition and extinction of aversive conditioning. *Psychol. Med.* 39, 255–265. doi: 10.1017/S0033291708003516
- Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Gusman, F. D., Charney, D. S., et al. (1995). The development of a Clinician-Administered PTSD Scale. *J. Trauma. Stress* 8, 75–90. doi: 10.1002/jts.2490080106
- Bremner, J. D., Vermetten, E., Schmahl, C., Vaccarino, V., Vythilingam, M., Afzal, N., et al. (2005). Positron emission tomographic imaging of neural correlates of a fear acquisition and extinction paradigm in women with childhood sexual-abuse-related post-traumatic stress disorder. *Psychol. Med.* 35, 791–806. doi: 10.1017/S0033291704003290
- Bryant, R. A., Felmingham, K. L., Kemp, A. H., Barton, M., Peduto, A. S., Rennie, C., et al. (2005). Neural networks of information processing in posttraumatic stress disorder: a functional magnetic resonance imaging study. *Biol. Psychiatry* 58, 111–118. doi: 10.1016/j.biopsych.2005.03.021
- Bryant, R. A., Marosszeky, J. E., Crooks, J., Baguley, I., and Gurka, J. (2000). Coping style and post-traumatic stress disorder following severe traumatic brain injury. *Brain Inj.* 14, 175–180. doi: 10.1080/026990500120826
- Chen, X., Li, Y., Li, S., and Kirouac, G. J. (2012). Early fear as a predictor of avoidance in a rat model of post-traumatic stress disorder. *Behav. Brain Res.* 226, 112–117. doi: 10.1016/j.bbr.2011.09.004
- Costanzi, M., Cannas, S., Saraulli, D., Rossi-Arnaud, C., and Cestari, V. (2011). Extinction after retrieval: effects on the associative and nonassociative components of remote contextual fear memory. *Learn. Mem.* 18, 508–518. doi: 10.1101/lm.2175811

- Fanselow, M. S. (2000). Contextual fear, gestalt memories, and the hippocampus. *Behav. Brain Res.* 110, 73–81. doi: 10.1016/S0166-4328(99)00186-2
- Fanselow, M. S., and Dong, H. W. (2010). Are the dorsal and ventral hippocampus functionally distinct structures? *Neuron* 65, 7–19. doi: 10.1016/j.neuron.2009.11.031
- Felmingham, K. L., Williams, L. M., Kemp, A. H., Rennie, C., Gordon, E., and Bryant, R. A. (2009). Anterior cingulate activity to salient stimuli is modulated by autonomic arousal in posttraumatic stress disorder. *Psychiatry Res.* 173, 59–62. doi: 10.1016/j.psychres.2008.12.005
- Garfinkel, S. N., and Liberzon, I. (2009). Neurobiology of PTSD: a review of neuroimaging findings. *Psychiatr. Ann.* 39, 370–381. doi: 10.3928/00485713-20090527-01
- Glover, E. M., Phifer, J. E., Crain, D. F., Norrholm, S. D., Davis, M., Bradley, B., et al. (2011). Tools for translational neuroscience: PTSD is associated with heightened fear responses using acoustic startle but not skin conductance measures. *Depress. Anxiety* 28, 1058–1066. doi: 10.1002/da.20880
- Glover, G. H., and Law, C. S. (2001). Spiral-in/out BOLD fMRI for increased SNR and reduced susceptibility artifacts. *Magn. Reson. Med.* 46, 515–522. doi: 10.1002/mrm.1222
- Glover, G. H., Li, T. Q., and Ress, D. (2000). Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn. Reson. Med.* 44, 162–167. doi: 10.1002/1522-2594(200007)44:1<162::AID-MRM23>3.0.CO;2-E
- Glynn, S. M., Eth, S., Randolph, E. T., Foy, D. W., Urbaitis, M., Boxer, L., et al. (1999). A test of behavioral family therapy to augment exposure for combat-related posttraumatic stress disorder. *J. Consult. Clin. Psychol.* 67, 243–251. doi: 10.1037/0022-006X.67.2.243
- Hayes, S. C., Strosahl, K. D., and Wilson, K. G. (1999). *Acceptance and Commitment Therapy: An Experiential Approach to Behavior Change*. New York, USA: Guilford Press.
- Hendrix, C. C., Erdmann, M. A., and Briggs, K. (1998). Impact of Vietnam veterans' arousal and avoidance on spouses' perceptions of family life. *Am. J. Fam. Ther.* 26, 115–128. doi: 10.1080/01926189808251092
- Hopper, J. W., Frewen, P. A., Van Der Kolk, B. A., and Lanius, R. A. (2007). Neural correlates of reexperiencing, avoidance, and dissociation in PTSD: symptom dimensions and emotion dysregulation in responses to script-driven trauma imagery. *J. Trauma. Stress* 20, 713–725. doi: 10.1002/jts.20284
- Jovanovic, T., Norrholm, S. D., Blanding, N. Q., Davis, M., Duncan, E., Bradley, B., et al. (2010). Impaired fear inhibition is a biomarker of PTSD but not depression. *Depress. Anxiety* 27, 244–251. doi: 10.1002/da.20663
- Jovanovic, T., Norrholm, S. D., Fennell, J. E., Keyes, M., Fiallos, A. M., Myers, K. M., et al. (2009). Posttraumatic stress disorder may be associated with impaired fear inhibition: relation to symptom severity. *Psychiatry Res.* 167, 151–160. doi: 10.1016/j.psychres.2007.12.014
- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., and Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch. Gen. Psychiatry* 62, 593–602. doi: 10.1001/archpsyc.62.6.593
- Kuhn, E., Blanchard, E. B., and Hickling, E. J. (2003). Posttraumatic stress disorder and psychosocial functioning within two samples of MVA survivors. *Behav. Res. Ther.* 41, 1105–1112. doi: 10.1016/S0005-7967(03)00071-8
- Lauterbach, D., Bak, C., Reiland, S., Mason, S., Lute, M. R., and Earls, L. (2007). Quality of parental relationships among persons with a lifetime history of posttraumatic stress disorder. *J. Trauma. Stress* 20, 161–172. doi: 10.1002/jts.20194
- Leiner, A. S., Kearns, M. C., Jackson, J. L., Astin, M. C., and Rothbaum, B. O. (2012). Avoidant coping and treatment outcome in rape-related posttraumatic stress disorder. *J. Consult. Clin. Psychol.* 80, 317–321. doi: 10.1037/a0026814
- Liberzon, I., and Sripada, C. S. (2008). The functional neuroanatomy of PTSD: a critical review. *Prog. Brain Res.* 167, 151–169. doi: 10.1016/S0079-6123(07)67011-3
- Lissek, S., Powers, A. S., McClure, E. B., Phelps, E. A., Woldehawariat, G., Grillon, C., et al. (2005). Classical fear conditioning in the anxiety disorders: a meta-analysis. *Behav. Res. Ther.* 43, 1391–1424. doi: 10.1016/j.brat.2004.10.007
- Lommen, M. J., Engelhard, I. M., Sijbrandij, M., Van Den Hout, M. A., and Hermans, D. (2013). Pre-trauma individual differences in extinction learning predict posttraumatic stress. *Behav. Res. Ther.* 51, 63–67. doi: 10.1016/j.brat.2012.11.004
- Lunney, C. A., and Schnurr, P. P. (2007). Domains of quality of life and symptoms in male veterans treated for posttraumatic stress disorder. *J. Trauma. Stress* 20, 955–964. doi: 10.1002/jts.20269
- Mahan, A. L., and Ressler, K. J. (2012). Fear conditioning, synaptic plasticity and the amygdala: implications for posttraumatic stress disorder. *Trends Neurosci.* 35, 24–35. doi: 10.1016/j.tins.2011.06.007
- Malta, L. S., Levitt, J. T., Martin, A., Davis, L., and Cloitre, M. (2009). Correlates of functional impairment in treatment-seeking survivors of mass terrorism. *Behav. Ther.* 40, 39–49. doi: 10.1016/j.beth.2007.12.007
- Maren, S. (2001). Neurobiology of Pavlovian fear conditioning. *Annu. Rev. Neurosci.* 24, 897–931. doi: 10.1146/annurev.neuro.24.1.897
- Maren, S., Phan, K. L., and Liberzon, I. (2013). The contextual brain: implications for fear conditioning, extinction and psychopathology. *Nat. Rev. Neurosci.* 14, 417–428. doi: 10.1038/nrn3492
- Milad, M. R., Orr, S. P., Lasko, N. B., Chang, Y., Rauch, S. L., and Pitman, R. K. (2008). Presence and acquired origin of reduced recall for fear extinction in PTSD: results of a twin study. *J. Psychiatr. Res.* 42, 515–520. doi: 10.1016/j.jpsyres.2008.01.017
- Milad, M. R., Pitman, R. K., Ellis, C. B., Gold, A. L., Shin, L. M., Lasko, N. B., et al. (2009). Neurobiological basis of failure to recall extinction memory in posttraumatic stress disorder. *Biol. Psychiatry* 66, 1075–1082. doi: 10.1016/j.biopsych.2009.06.026
- Milad, M. R., Quinn, B. T., Pitman, R. K., Orr, S. P., Fischl, B., and Rauch, S. L. (2005). Thickness of ventromedial prefrontal cortex in humans is correlated with extinction memory. *Proc. Natl. Acad. Sci. U.S.A.* 102, 10706–10711. doi: 10.1073/pnas.0502441102
- Milad, M. R., Wright, C. I., Orr, S. P., Pitman, R. K., Quirk, G. J., and Rauch, S. L. (2007). Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. *Biol. Psychiatry* 62, 446–454. doi: 10.1016/j.biopsych.2006.10.011
- Monson, C. M., Macdonald, A., Vorstenbosch, V., Shnaider, P., Goldstein, E. S., Ferrier-Auerbach, A. G., et al. (2012). Changes in social adjustment with cognitive processing therapy: effects of treatment and association with PTSD symptom change. *J. Trauma. Stress* 25, 519–526. doi: 10.1002/jts.21735
- Moser, M. B., and Moser, E. I. (1998). Functional differentiation in the hippocampus. *Hippocampus* 8, 608–619. doi: 10.1002/(SICI)1098-1063(1998)8:6<608::AID-HIPO3>3.0.CO;2-7
- Muigg, P., Hetzenauer, A., Hauer, G., Hauschild, M., Gaburro, S., Frank, E., et al. (2008). Impaired extinction of learned fear in rats selectively bred for high anxiety – evidence of altered neuronal processing in prefrontal-amygdala pathways. *Eur. J. Neurosci.* 28, 2299–2309. doi: 10.1111/j.1460-9568.2008.06511.x
- Norrholm, S. D., and Jovanovic, T. (2010). Tailoring therapeutic strategies for treating posttraumatic stress disorder symptom clusters. *Neuropsychiatr. Dis. Treat.* 6, 517–532. doi: 10.2147/NDT.S10951
- Norrholm, S. D., Jovanovic, T., Olin, I. W., Sands, L. A., Karapanou, I., Bradley, B., et al. (2011). Fear extinction in traumatized civilians with posttraumatic stress disorder: relation to symptom severity. *Biol. Psychiatry* 69, 556–563. doi: 10.1016/j.biopsych.2010.09.013
- North, C. S., Nixon, S. J., Shariat, S., Mallonee, S., Mcmillen, J. C., Spitznagel, E. L., et al. (1999). Psychiatric disorders among survivors of the Oklahoma City bombing. *JAMA* 282, 755–762. doi: 10.1001/jama.282.8.755
- North, C. S., Oliver, J., and Pandya, A. (2012). Examining a comprehensive model of disaster-related posttraumatic stress disorder in systematically studied survivors of 10 disasters. *Am. J. Public Health* 102, e40–e48. doi: 10.2105/AJPH.2012.300689
- Orr, S. P., Metzger, L. J., Lasko, N. B., Macklin, M. L., Peri, T., and Pitman, R. K. (2000). De novo conditioning in trauma-exposed individuals with and without posttraumatic stress disorder. *J. Abnorm. Psychol.* 109, 290–298. doi: 10.1037/0021-843X.109.2.290
- Pamplona, F. A., Henes, K., Micale, V., Mauch, C. P., Takahashi, R. N., and Wotjak, C. T. (2011). Prolonged fear incubation leads to generalized avoidance behavior in mice. *J. Psychiatr. Res.* 45, 354–360. doi: 10.1016/j.jpsyres.2010.06.015
- Paulus, M. P., and Stein, M. B. (2006). An insular view of anxiety. *Biol. Psychiatry* 60, 383–387. doi: 10.1016/j.biopsych.2006.03.042
- Quirk, G. J., and Mueller, D. (2008). Neural mechanisms of extinction learning and retrieval. *Neuropsychopharmacology* 33, 56–72. doi: 10.1038/sj.npp.1301555

- Riggs, D. S., Byrne, C. A., Weathers, F. W., and Litz, B. T. (1998). The quality of the intimate relationships of male Vietnam veterans: problems associated with posttraumatic stress disorder. *J. Trauma. Stress* 11, 87–101. doi: 10.1023/A:1024409200155
- Rizzo, A., Reger, G., Gahm, G., Difede, J., and Rothbaum, B. O. (2009). "Virtual reality exposure therapy for combat-related PTSD," in *Post-traumatic Stress Disorder: Basic Science and Clinical Practice*, eds P. J. Shiromani, T. M. Keane, and J. E. Ledoux (Totowa, NJ: Humana Press), 375–399. doi: 10.1007/978-1-60327-329-9\_18
- Rougemont-Bucking, A., Linnman, C., Zeffiro, T. A., Zeidan, M. A., Lebron-Milad, K., Rodriguez-Romaguera, J., et al. (2011). Altered processing of contextual information during fear extinction in PTSD: an fMRI study. *CNS Neurosci. Ther.* 17, 227–236. doi: 10.1111/j.1755-5949.2010.00152.x
- Ruscio, A. M., Weathers, F. W., King, L. A., and King, D. W. (2002). Male war-zone veterans' perceived relationships with their children: the importance of emotional numbing. *J. Trauma. Stress* 15, 351–357. doi: 10.1023/A:1020125006371
- Samper, R. E., Taft, C. T., King, D. W., and King, L. A. (2004). Posttraumatic stress disorder symptoms and parenting satisfaction among a national sample of male vietnam veterans. *J. Trauma. Stress* 17, 311–315. doi: 10.1023/B:JOTS.0000038479.30903.ed
- Sauerhofer, E., Pamplona, F. A., Bedenk, B., Moll, G. H., Dawirs, R. R., Von Horsten, S., et al. (2012). Generalization of contextual fear depends on associative rather than non-associative memory components. *Behav. Brain Res.* 233, 483–493. doi: 10.1016/j.bbr.2012.05.016
- Schnurr, P. P., and Lunney, C. A. (2008). Exploration of gender differences in how quality of life relates to posttraumatic stress disorder in male and female veterans. *J. Rehabil. Res. Dev.* 45, 383–393. doi: 10.1682/JRRD.2007.06.0099
- Sehlmeyer, C., Dannlowski, U., Schoning, S., Kugel, H., Pyka, M., Pfeleiderer, B., et al. (2011). Neural correlates of trait anxiety in fear extinction. *Psychol. Med.* 41, 789–798. doi: 10.1017/S0033291710001248
- Sehlmeyer, C., Schoning, S., Zwitserlood, P., Pfeleiderer, B., Kircher, T., Arolt, V., et al. (2009). Human fear conditioning and extinction in neuroimaging: a systematic review. *PLoS ONE* 4:e5865. doi: 10.1371/journal.pone.0005865
- Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., et al. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J. Clin. Psychiatry* 59(Suppl. 20), S22–S33; quiz 34–57.
- Shin, L. M., Bush, G., Milad, M. R., Lasko, N. B., Brohawn, K. H., Hughes, K. C., et al. (2011). Exaggerated activation of dorsal anterior cingulate cortex during cognitive interference: a monozygotic twin study of posttraumatic stress disorder. *Am. J. Psychiatry* 168, 979–985. doi: 10.1176/appi.ajp.2011.09121812
- Shin, L. M., Bush, G., Whalen, P. J., Handwerker, K., Cannistraro, P. A., Wright, C. I., et al. (2007). Dorsal anterior cingulate function in posttraumatic stress disorder. *J. Trauma. Stress* 20, 701–712. doi: 10.1002/jts.20231
- Shin, L. M., and Liberzon, I. (2010). The neurocircuitry of fear, stress, and anxiety disorders. *Neuropsychopharmacology* 35, 169–191. doi: 10.1038/npp.2009.83
- Solomon, Z., and Mikulincer, M. (2007). Posttraumatic intrusion, avoidance, and social functioning: a 20-year longitudinal study. *J. Consult. Clin. Psychol.* 75, 316–324. doi: 10.1037/0022-006X.75.2.316
- Sripada, R. K., King, A. P., Garfinkel, S. N., Wang, X., Sripada, C. S., Welsh, R. C., et al. (2012a). Altered resting-state amygdala functional connectivity in men with posttraumatic stress disorder. *J. Psychiatry Neurosci.* 37, 241–249. doi: 10.1503/jpn.110069
- Sripada, R. K., King, A. P., Welsh, R. C., Garfinkel, S. N., Wang, X., Sripada, C. S., et al. (2012b). Neural dysregulation in posttraumatic stress disorder: evidence for disrupted equilibrium between salience and default mode brain networks. *Psychosom. Med.* 74, 904–911. doi: 10.1097/PSY.0b013e318273bf33
- Taylor, S., Fedoroff, I. C., Koch, W. J., Thordarson, D. S., Fecteau, G., and Nicki, R. M. (2001). Posttraumatic stress disorder arising after road traffic collisions: patterns of response to cognitive-behavior therapy. *J. Consult. Clin. Psychol.* 69, 541–551. doi: 10.1037/0022-006X.69.3.541
- Werner, N. S., Meindl, T., Engel, R. R., Rosner, R., Riedel, M., Reiser, M., et al. (2009). Hippocampal function during associative learning in patients with posttraumatic stress disorder. *J. Psychiatr. Res.* 43, 309–318. doi: 10.1016/j.jpsychires.2008.03.011

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 June 2013; accepted: 26 September 2013; published online: 17 October 2013.

Citation: Sripada RK, Garfinkel SN and Liberzon I (2013) Avoidant symptoms in PTSD predict fear circuit activation during multimodal fear extinction. *Front. Hum. Neurosci.* 7:672. doi: 10.3389/fnhum.2013.00672

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2013 Sripada, Garfinkel and Liberzon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Non-verbal emotion communication training induces specific changes in brain function and structure

Benjamin Kreifelts<sup>1\*</sup>, Heike Jacob<sup>1</sup>, Carolin Brück<sup>1</sup>, Michael Erb<sup>2</sup>, Thomas Ethofer<sup>1,2</sup> and Dirk Wildgruber<sup>1</sup>

<sup>1</sup> Department of Psychiatry and Psychotherapy, Eberhard Karls University of Tübingen, Tübingen, Germany

<sup>2</sup> Department of Biomedical Magnetic Resonance, Eberhard Karls University of Tübingen, Tübingen, Germany

## Edited by:

Martin Klasen, RWTH Aachen University, Germany

## Reviewed by:

Liliana R. R. Demenescu, RWTH Aachen, Germany

Eliza M. Alawi, RWTH Aachen University, Germany

## \*Correspondence:

Benjamin Kreifelts, Department of Psychiatry and Psychotherapy, Eberhard Karls University of Tübingen, Calwerstr. 14, 72076 Tübingen, Germany  
e-mail: benjamin.kreifelts@med.uni-tuebingen.de

The perception of emotional cues from voice and face is essential for social interaction. However, this process is altered in various psychiatric conditions along with impaired social functioning. Emotion communication trainings have been demonstrated to improve social interaction in healthy individuals and to reduce emotional communication deficits in psychiatric patients. Here, we investigated the impact of a non-verbal emotion communication training (NECT) on cerebral activation and brain structure in a controlled and combined functional magnetic resonance imaging (fMRI) and voxel-based morphometry study. NECT-specific reductions in brain activity occurred in a distributed set of brain regions including face and voice processing regions as well as emotion processing- and motor-related regions presumably reflecting training-induced familiarization with the evaluation of face/voice stimuli. Training-induced changes in non-verbal emotion sensitivity at the behavioral level and the respective cerebral activation patterns were correlated in the face-selective cortical areas in the posterior superior temporal sulcus and fusiform gyrus for valence ratings and in the temporal pole, lateral prefrontal cortex and midbrain/thalamus for the response times. A NECT-induced increase in gray matter (GM) volume was observed in the fusiform face area. Thus, NECT induces both functional and structural plasticity in the face processing system as well as functional plasticity in the emotion perception and evaluation system. We propose that functional alterations are presumably related to changes in sensory tuning in the decoding of emotional expressions. Taken together, these findings highlight that the present experimental design may serve as a valuable tool to investigate the altered behavioral and neuronal processing of emotional cues in psychiatric disorders as well as the impact of therapeutic interventions on brain function and structure.

**Keywords: fMRI, VBM, superior temporal sulcus, fusiform face area, neuroticism, emotion sensitivity**

## INTRODUCTION

Perception and correct interpretation of non-verbal emotional cues from voice and face is essential for intact social interaction and social functioning. Typically, this ability is acquired during childhood and youth in a rather implicit manner as a part of our upbringing. Nevertheless, these perceptive skills can be trained explicitly. It has been shown that such training in healthy individuals improves both their confidence and their decoding performance (Constanzo, 1992), their social and interpersonal skills (Matsumoto and Hwang, 2011) and heightens their non-verbal perceptiveness and sensitivity (Klinzing and Jackson, 1987).

The processing of non-verbal emotional signals has been found to be altered in different psychiatric conditions including depression, anxiety disorders, bipolar disorder, schizophrenia, psychopathy, and borderline personality disorder (e.g., Domes et al., 2009; Demenescu et al., 2010; Kohler et al., 2010, 2011; Dawel et al., 2012; Samame et al., 2012). Consequentially, emotion perception trainings may also appear as a means to reduce this deficit and improve social communication and functioning in psychiatric patients. In this regard, structured behavioral

trainings have been found to be effective in ameliorating facial emotion recognition and social functioning in schizophrenia (Kurtz and Richardson, 2012). However, much less is known about the neural bases of such emotional communication trainings, both in healthy individuals and psychiatric patients. While a recent functional magnetic resonance imaging (fMRI) study in schizophrenic patients described alterations in cerebral activation patterns following a facial affect recognition training (Habel et al., 2010), in healthy individuals, to our knowledge, no study describing the neuronal effects of an emotion communication training has been published to date.

Using fMRI in healthy individuals, it was demonstrated that the processing of non-verbal emotional facial and vocal cues (i.e., facial expressions and tone of voice) is associated with increased activation of sensory cortices specialized for the processing of human (emotional) voices (e.g., Belin et al., 2000; Wildgruber et al., 2006; Ethofer et al., 2012) and faces (e.g., Kanwisher et al., 1997; Posamentier and Abdi, 2003), and limbic brain areas (e.g., Posamentier and Abdi, 2003; Wildgruber et al., 2006; Brück et al., 2011b). The combined perception and integration of these signals

is associated with a further increase in activation in the posterior temporal sulcus (pSTS), thalamus, the face processing area in the fusiform gyrus (FFA) and the amygdala (e.g., Pourtois et al., 2005; Kreifelts et al., 2007, 2009, 2010; Robins et al., 2009; Klasen et al., 2011, for a review see Klasen et al., 2012). Irrespective of the sensory modality of non-verbal cue presentation, lateral prefrontal and supplementary motor cortices are engaged in the evaluation of these signals and response selection (e.g., Schirmer and Kotz, 2006; Brück et al., 2011b; Ethofer et al., 2013).

In the present study, it was our goal to determine the effects of a non-verbal emotional communication training (NECT) on the neuronal processing of non-verbal cues from voice and face in healthy individuals as a feasibility study for further investigations in patients with psychiatric conditions.

To this end, healthy participants either received a four-week NECT or an equally intensive but essentially non-communicative training in Sudoku as control condition. The NECT took place as group training with the central training elements revolving around a board game complemented by non-verbal exercises. Sudoku, in contrast, is a popular Japanese pastime in form of number riddles to be solved by pure logic. Before and after the training interval, the participants took part in three fMRI experiments: During the first experiment the participants had to judge the valence of short video sequences portraying faces speaking short sentence with varying (happy, neutral, angry) vocal and facial expressions. Valence ratings and cerebral responses were treated as outcome variables. The additional two fMRI experiments were standard localizer experiments to identify voice- and face-selective brain regions.

First, behavioral and cerebral data were analyzed within the framework of an analysis of variance (ANOVA) with the factors time point [before (T0) vs. after (T1) training], training type (NECT vs. SUDOKU) and non-verbal emotion (emotional vs. neutral cues). Both general and emotion-specific effects were assessed.

With regard to the training outcome, we expected that cerebral effects of NECT would occur foremost in those brain areas involved in the processing of non-verbal emotional signals as well as in face- and voice-selective cortices.

Secondly, as it has been shown that certain personality factors such as neuroticism, on the one hand, influence the processing of non-verbal emotional signals (e.g., Stein et al., 2007; Cremers et al., 2010; Suslow et al., 2010; Brück et al., 2011a; Kehoe et al., 2012), and on the other hand, are associated with an increased risk to develop a psychiatric condition (e.g., the association of neuroticism and depression and anxiety; Brandes and Bienvenu, 2006; Klein et al., 2011), we investigated if personality factors modulate the behavioral or cerebral effects of NECT. To this end, the NEO-Five Factor Inventory (NEO-FFI; Borkenau and Ostendorf, 2008) capturing the personality factors neuroticism, extraversion, openness to experience, agreeableness and conscientiousness was completed by all participants. The individual personality ratings were then used as explanatory covariates for observed NECT effects.

Thirdly, NECT-associated changes in the behavioral and cerebral sensitivity to non-verbal emotional cues were correlated to identify behaviorally relevant cerebral correlates of NECT.

Finally, cerebral correlates of training have not only been observed on the functional but also on the structural level (e.g., Draganski et al., 2004). Therefore, we further tested if NECT increased gray matter (GM) volume in face- and voice-selective cortices and other brain regions with NECT-induced changes in their neuronal responses.

## MATERIALS AND METHODS

### PARTICIPANTS

Sixteen healthy and right-handed (Edinburgh Handedness Inventory; Oldfield, 1971) individuals were initially included into the study. None of the participants reported any current or past substance abuse problems, neurological or psychiatric illnesses, nor indicated any hearing difficulties or uncorrected vision impairments. Moreover, none of the participants reported to be taking any medication. Of the 16 participants, eight were randomized to the NECT group [4 females; mean age = 24.88 years, standard deviation (*SD*) =  $\pm 1.89$  years] and eight were randomized to the SUDOKU training group. Of these eight, two participants had to be excluded from the study (i.e., one participant did not attend the training and one had to be excluded due to a structural cerebral anomaly). Thus, six individuals remained in the Sudoku training group (2 females; mean age = 25.33 years, *SD* =  $\pm 1.51$  years) and 14 individuals (6 females; mean age = 25.07 years, *SD* =  $\pm 1.69$  years) were included in the analyses.

At the beginning of the study all participants completed the German version of the NEO-FFI (Borkenau and Ostendorf, 2008) based on the works of Costa and McCrae. The NEO-FFI is a multidimensional self-report personality inventory with 60 items which assesses the following five personality factors: neuroticism (N), extraversion (E), openness (O) to experience, agreeableness (A), and conscientiousness (C). The NECT group scored as follows (mean  $\pm$  *SD*): N:  $1.21 \pm 0.33$ ; E:  $2.57 \pm 0.31$ ; O:  $2.68 \pm 0.22$ ; A:  $2.55 \pm 0.71$ ; C:  $2.47 \pm 0.5$ . The results of the SUDOKU group were: N:  $1.74 \pm 0.71$ ; E:  $2.33 \pm 0.43$ ; O:  $2.82 \pm 0.39$ ; A:  $2.67 \pm 0.4$ ; C:  $2.44 \pm 0.32$ .

### ETHICS STATEMENT

The study was performed according to the principles of the Code of Ethics of the World Medical Association (Declaration of Helsinki), and with the approval of the ethics review board of the University of Tübingen. Before inclusion in the study, all participants gave written informed consent. The participants were monetarily compensated for the study participation.

### STIMULUS MATERIAL, TASKS, AND PROCEDURE

The stimulus material comprised a set of 120 video films of 10 professional actors (5 females) speaking short sentences. The stimuli contained verbal (i.e., six short German sentences with self-referential content) as well as non-verbal information (i.e., facial expression and tone of voice) about the emotional states of the respective speakers. The non-verbal expressions differed in their valence: One third (40) of the video sequences portrayed an angry expression, one third a happy expression and one third an emotionally neutral expression. The stimulus material was balanced with respect to the valence of verbal stimulus content which

included negative, positive and neutral sentences. The stimuli had a mean duration of 1459 ms ( $SD = 317$  ms). For further details on stimulus production, editing, validation and selection please refer to Jacob et al. (2013).

The software Presentation (Neurobehavioral Systems Inc., Albany, CA, USA) was used for the experiment. Video films were back-projected onto a screen in the scanner bore about 50 cm behind the participant's head. The participants viewed the screen through a mirror system mounted on the head coil. Magnetic resonance compatible headphones (Sennheiser electronic GmbH & Co. KG, Wedemark-Wennebostel, Germany; in-house modified) were used for sound transmission.

The 120 video films used for the study were divided into two equal blocks balanced for non-verbal and verbal stimulus valence as well as the gender of the speaker. The order of stimulus presentation was randomized within blocks, and the sequence of the two stimulus blocks was balanced across participants. In each block 10 null events each with a duration of 10 s were randomly inserted in the stimulus sequence. The stimulus blocks were presented within separate imaging runs. Stimulus onset was jittered relative to scan onset in steps of  $1/4$  repetition time (TR). It was the participant's task to judge the emotional state of the speaker on a 4-point valence scale (“--” = highly negative, “-” = negative, “+” = positive, “++” = highly positive) as precisely and as fast as possible. A horizontally flipped scale was used for half of the participants. The participants' responses were transmitted through button presses on a four-buttons fiber optic response pad (LUMItouch, Photon Control Inc., Burnaby, BC, Canada). The response window was set to 5 s in duration and was time-locked to the onset of the videos. After the offset of the videos, the 4-point valence scale was presented. The participants were acquainted with the experimental setting through a short training session outside the scanner. The stimuli employed in the training session were not part of the stimulus set used in the main experiment.

After the main experiment two standard functional localizer experiments were run to identify face- and voice-selective brain regions. The face localizer was adapted from previous studies on face processing (Kanwisher et al., 1997; Epstein et al., 1999) and included pictures from four different categories (faces, houses, objects, and natural scenes) presented using a block-design. The voice localizer data were acquired using a block design experiment validated in previous research (Belin et al., 2000; Kreifelts et al., 2009). The employed stimuli included human voices (e.g., speech, sighs, laughs), animal sounds (cries of various animals), and environmental sounds (e.g., doors, telephones, cars). For details on the localizer experiments see Kreifelts et al. (2010).

All participants took part in the fMRI experiments once before the training (T0) and once after training (T1). Randomization to the different types of training took place directly after the first measurement session before analyzing any data.

#### **NON-VERBAL EMOTION COMMUNICATION TRAINING (NECT)**

NECT is a four-week group training program (18 days, 1 h per day) consisting of a game, theoretical discussions and

supplementary exercises held in a group setting with eight participants.

#### **Non-verbal communication game**

The central part of the training was a non-verbal communication game in the form of a board game. Playing the game involves extensive practice expressing emotions as well as understanding emotions as the players take turns expressing emotions and perceiving emotions in the other players. The basic rules can be summarized as follows: Sentence cards (i.e., cards, each displaying a single sentence) are put face down on their allotted space on the board. Emotion cards (i.e., cards, each displaying an emotion label) are placed in the middle of the board. Six Emotion cards are then put face up on their allotted spaces on the board numbered from 1 to 6. Each player team chooses one of six tokens to represent their position on the board. Each team also receives a set of Number cards ranging from 1 to 6. The team that was chosen to start draws one of their six Number cards referring to one of the six Emotion cards. Each of the two players within a team has to draw a Sentence card and read the sentence in the respective emotional tone of voice and with an appropriate facial expression. The other teams have to guess which emotion was conveyed by the tone of voice and the facial expression. In the first round, each team is allowed to discuss their decision in a non-verbal manner (i.e., by pointing to the respective number cards referring to the different emotions). Once all quiz teams have come to a decision, they show their cards simultaneously. In case of a correct answer, the respective quiz team is allowed to move their token one step in the direction of the goal. The team which performed the emotional expressions is allowed to move their token in the direction of the goal for the number of spaces indicated by the number of correct answers by the other teams based on their performance. In the second round, the procedure is similar, but each team is given two sets of Number cards, one for each team player. Now, each quiz team is allowed to silently discuss their decision, however not by pointing to the cards but by using means of non-verbal communication (e.g., by using facial expressions or gestures of the respective emotion). The quiz teams are allowed to move their token in the direction of the goal one step in case of correct AND matching answers. In the third round, the quiz teams are not allowed to communicate within the teams. Each team player has to make her/his own decision. The quiz team is allowed to move their token in the direction of the goal one step in case of correct AND matching answers. To diversify the game and increase its training effect, some modifications were made during the course of the training (e.g., to avoid imitations and to broaden the range of expressions, one of the two acting players had to wait outside the room until his/her teammate has finished his/her performance).

#### **Non-verbal communication—supplementary exercises**

The supplementary exercises were based on a playful approach to improve non-verbal communication skills. Exercises were provided by a book written by Funcke (2006) describing different exercises training the ability to express and perceive non-verbal emotional signals. The following exercises were part of the training:

- a. Körpersprachenspiel (i.e., game of body language; pages 203–204)
- b. Der Ton macht die Musik (i.e., it is not what you say, it is the way you say it; pages 139–140)
- c. Mimische Kette (i.e., chain of facial expressions; pages 169–170)
- d. Audienz beim Papst (i.e., audience with the pope; pages 187–188)
- e. Menschen-Memory (i.e., human memory game; pages 59–60)
- f. Aus der Zeitung lesen (i.e., reading from the newspaper; pages 137–138)
- g. Briefe lesen (i.e., reading letters; pages 189–190)
- h. Texte rezitieren (i.e., reciting texts; pages 127–130)

### Non-verbal communication—theory

Short theory units aimed at sensitizing the participants for non-verbal cues in daily life were an additional part of the training. Participants were encouraged to discuss their own attention to non-verbal signals in daily life (e.g., facial expressions, gestures, tone of voice) depending on the social context.

### SUDOKU TRAINING

Sudoku (Japanese *sūdoku*, short for *suji wa dokushin ni kagiru*, meaning “the numerals must remain single”) is a Japanese pastime where numbers have to be filled into a grid following certain rules. Typically, the grid has the width and the height of nine fields each resulting in an overall number of 81 fields. Some of the fields are prefilled with numbers ranging from 1 to 9. The player has to fill in the remaining fields adhering to the following rules: Every number between 1 and 9 has to be filled in exactly once in each row and each column of the grid. Moreover, each of these numbers has to appear exactly once in each of the nine 3-by-3 sub-grids of the main grid. The solution of each Sudoku puzzle can be found by logically applying the rules of the game. The difficulty is defined by the amount of prefilled fields and the numbers already given.

The SUDOKU training consisted of a four-week training program (19 days, 1 h per day). Participants were seated at single desks oriented in the same direction of view. Each participant received a pencil, several sheets of scratch papers and a copy of a Sudoku book (Rossa, 2009). This book includes different types of Sudoku puzzles and the degree of difficulty ranges from very easy (first chapter) to difficult (last chapter). Participants were instructed to solve the Sudoku puzzles one at a time. Participants were allowed to use the scratch papers to make notes, but the use of any other aids as well as talking to each other or copying from each other was prohibited. The training was monitored continuously by the examiner. Subsequent to each training session, the solved Sudoku puzzles were checked and errors were marked. In the following training session, Sudoku puzzles with errors had to be repeated before a new one could be started.

### IMAGE ACQUISITION

High resolution structural T1-weighted images [176 slices, slice thickness 1 mm, no gap, TR = 2300 ms, echo time (TE) = 2.96 ms, time to inversion (TI) = 1100 ms, voxel size:  $1 \times 1 \times 1 \text{ mm}^3$ , field of view (FoV) =  $256 \times 256 \text{ mm}^2$ , magnetization

prepared rapid acquisition gradient echo (MPRAGE) sequence] and functional images [30 axial slices acquired in an interleaved descending order, slice thickness 4 mm + 1 mm gap, TR = 1.7 s, TE = 30 ms, voxel size:  $3 \times 3 \times 5 \text{ mm}^3$ , FoV =  $192 \times 192 \text{ mm}^2$ , echo-planar imaging (EPI) sequence] were recorded with a 3 T scanner (Siemens TIM TRIO, Erlangen, Germany). A field map [36 slices, slice thickness 3 mm + 1 mm gap, TR = 400 ms, TE(1) = 5.19 ms, TE(2) = 7.65 ms, voxel size:  $3 \times 3 \times 4 \text{ mm}^3$ ] was acquired to correct for image distortions.

### ANALYSIS OF BEHAVIORAL DATA

Valence ratings and response times were treated as behavioral outcome variables. First, the valence ratings were transformed from symbolic to numerical values ( $-- = 1$ ,  $- = 2$ ,  $+ = 3$ ,  $++ = 4$ ). Then mean absolute valence ratings for neutral and emotional (i.e., positive and negative) non-verbal cues were calculated on the above arbitrary scale where a value of 2.5 indicates “neutral” valence. These absolute valence ratings and the response times were then analyzed using IBM SPSS Statistics Version 19 (IBM Corp., Armonk, NY, USA) within the framework of a three-factorial ANOVA for repeated measure with non-verbal emotion (emotional, neutral) and time point (T0, T1) as within-subject factors and training type (NECT, SUDOKU) as between-subject factor. To clarify potential interactions between the participants’ personality and the experimental factors, additional ANOVAs with the separate NEO-FFI personality factors as covariates were performed. For the correlation of individual training-associated changes in the behavioral sensitivity to non-verbal emotional cues with the respective cerebral activation patterns, the interaction term  $T0[\text{EMO} - \text{NEU}] - T1[\text{EMO} - \text{NEU}]$  was calculated from the valence ratings as well as from the response times of each participant.

### ANALYSIS OF fMRI DATA

Imaging data were analyzed with statistical parametric mapping software (SPM5, Wellcome Department of Imaging Neuroscience, London, UK). Preprocessing of the images comprised realignment, unwarping to correct for field distortions and to remove residual movement-related variance due to interactions between motion and field distortions (Andersson et al., 2001), coregistration with the anatomical data, normalization into MNI space (Montreal Neurological Institute, resampled voxel size:  $3 \times 3 \times 3 \text{ mm}^3$ ), and smoothing with a Gaussian filter [8 mm full width at half maximum (FWHM)]. The first five EPI images were discarded to exclude measurements preceding T1 equilibrium.

Responses to the stimuli of the main experiment were modeled separately for each trial as event-related responses employing a stick function time-locked to stimulus onset convolved with the hemodynamic responses function (HRF). For the face and voice localizer experiments, responses to the single categories (faces, houses, objects, and scenes in the face localizer and human voices, animal sounds, and environmental sounds in the voice localizer) were separately modeled using a box-car function corresponding to the duration of the respective blocks of stimuli convolved with the HRF. A high-pass filter with a cut-off frequency of 1/128 Hz was applied to reduce low-frequency components in the data. Serial autocorrelations within the data were accounted for by

modeling the error term as an autoregressive process (Friston et al., 2002).

For the main experiment and the localizer experiments, data from the individual first-level general linear models were employed to create contrast images for each subject. These were then submitted to a second-level random-effect analysis to enable population inference.

In the main experiment, brain regions showing stronger responses during task performance than during rest were identified via the main contrast where all events were contrasted against the implicit resting baseline. Brain regions exhibiting stronger responses to non-verbal emotional stimuli (i.e., positively and negatively valenced expressions) than to non-verbal neutral stimuli were identified via the contrast EMO > NEU.

### **General and emotion-specific effects of NECT**

To allow inference on general training effects for the processing of non-verbal expressions and differential training effects for the processing of non-verbal emotional and neutral expressions respectively, both of the above mentioned individual contrasts of interest were submitted to separate two-factorial whole-brain ANOVAs with time point (T0, T1) as within-subject factor and training type (NECT, SUDOKU) as between-subject factor. Results of the whole-brain analyses are reported at a height threshold of  $p < 0.001$ , uncorrected, and an extent threshold of  $k = 50$  voxels corresponding to  $p < 0.05$ , family wise error (FWE) corrected for multiple comparisons across the whole-brain at the cluster level.

The source of any observed interaction effects was determined by applying one-sample  $t$ -tests to the mean parameter estimates extracted from the clusters with significant interaction effects for the contrast T0–T1 for both training types (NECT, SUDOKU).

### **Correlations of cerebral responses with personality (NEO-FFI)**

In order to further investigate the relationship between the participants' personality traits and observed cerebral correlates of NECT, the mean contrast estimates were extracted from all brain regions with significant effects of NECT and correlated with the five NEO-FFI factors in the NECT group. To ascertain the specificity of potential correlations, the equivalent analysis was also performed in the SUDOKU group. All resulting  $p$  values are reported two-tailed and were Bonferroni-corrected for the number of NEO-FFI factors ( $n = 5$ ).

### **Correlations of cerebral responses with NECT-associated behavioral changes in sensitivity to non-verbal emotional cues**

As final add-on analyses, NECT-specific linear relationship of changes in the behavioral correlates of sensitivity to non-verbal emotional cues [i.e., the following contrast in valence ratings and response times: NECT<sub>T0[EMO – NEU] – T1[EMO – NEU]</sub> vs. SUDOKU<sub>T0[EMO – NEU] – T1[EMO – NEU]</sub>] were tested by performing a correlation analysis between the respective behavioral and cerebral activation contrasts at the population level. Statistical analyses were performed with a threshold of  $p < 0.001$ , uncorrected, at voxel level for whole-brain analyses. Results were FWE-corrected at cluster level ( $p < 0.05$ ). A more sensitive threshold of  $p < 0.01$ , uncorrected at voxel level, was used for regions of

interest (ROI) analyses within face- and voice-selective regions as well as within regions with general NECT-associated changes in cerebral activation. Here, small volume correction (SVC; Worsley et al., 1996) for the size of the respective ROI together with FWE correction ( $p < 0.05$ ) at cluster level was applied. Observed effects were post-hoc further investigated by extracting the mean contrast estimates from significant clusters for both training groups (NECT, SUDOKU) and performing separate bivariate correlation analyses (Pearson).

### **Functional localizer experiments for voice- and face-selective brain regions**

For the face localizer experiment, the responses to faces were contrasted against the responses to houses, objects, and natural scenes, while for the voice localizer experiment responses to human voices were contrasted against animal and environmental sounds.

Results of the whole-brain group analyses are reported at a height threshold of  $p < 0.0001$ , uncorrected, and an extent threshold of  $k = 10$  voxels corresponding to  $p < 0.05$ , FWE corrected for multiple comparisons across the whole-brain at the cluster level. The strict statistical threshold at the voxel level allows a non-overlapping localization of face- and voice-selective brain areas. Face- and voice-selective regions were then used as ROI for additional analyses of NECT effects with the contrasts described above using SVC and FWE-correction.

The Automated Anatomic Labeling (AAL) toolbox implemented in SPM (Tzourio-Mazoyer et al., 2002) was used for the anatomic cluster labeling.

### **ANALYSIS OF STRUCTURAL MRI DATA**

The goal of the voxel based morphometry analysis was to identify brain regions with changes in GM volume associated with NECT. The voxel based morphometry toolbox (VBM8) implemented in the SPM8 software environment was used for the preprocessing of the high resolution structural T1 images. Preprocessing was performed using the VBM8 default pipeline recommended for longitudinal data and included realignment, bias correction, segmentation into GM, white matter and cerebrospinal fluid and normalization. The voxels were resized to  $1.5 \times 1.5 \times 1.5 \text{ mm}^3$  during preprocessing. VBM8 by default produces modulated images. This means that the voxel values of the GM images are multiplied by the non-linear components of the normalization procedure thus correcting the data for individual brain sizes before the setup of a statistical model and enabling the analysis of relative differences in regional GM volume. The GM images were smoothed with an 8-mm FWHM Gaussian kernel and then analyzed within the framework of a two factorial ANOVA with time point as within-subject factor and training type as between-subject factor. The interaction term between time point and training type was defined as contrast of interest. Primary ROIs for the VBM analysis were those regions with a significant effect of NECT at the level of cerebral activation (fMRI) and face- as well as voice-selective brain regions. These analyses were complemented with a whole-brain analysis to rule out unspecific NECT associated structural alterations. Results are reported at a height threshold of  $p < 0.001$ , uncorrected, and  $p < 0.05$  with

SVC for the respective ROI using FWE-correction at cluster level. For the whole-brain analysis FWE-correction at cluster level was used in the same fashion. For areas with a significant interaction between time point and training type, mean GM values were extracted and results validated by testing the time point by training type interaction term after correcting for differences in GM volume at T0.

## RESULTS

### BEHAVIORAL RESPONSES

#### Valence ratings

The  $2 \times 2 \times 2$ -factorial ANOVA revealed that emotional stimuli received more extreme valence ratings than neutral stimuli [ $F_{(1, 12)} = 35.6, p < 0.001$ ; see **Figure 1**]. However, no significant effects were observed for the factors time point, training type, or the interactions between the three experimental factors [all  $F_{(1, 12)} \leq 3.0$ , all  $p \geq 0.11$ ]. Additional ANOVAs with the separate NEO-FFI personality factors as covariates did not yield any significant interactions with any of the experimental factors [all  $F_{(1, 12)} = 1.5$ , all  $p = 0.24$ ].

#### Response times

Responses to emotional stimuli were faster than to neutral stimuli [ $F_{(1, 12)} = 38.8, p < 0.001$ ; emotional stimuli: mean  $\pm$  SEM:  $1639 \pm 32$  ms, neutral stimuli:  $1800 \pm 42$  ms]. For the remaining factors time point, training type, or the interactions between all three experimental factors no significant effects emerged [all  $F_{(1, 12)} \leq 3.4$ , all  $p \geq 0.09$ ]. Furthermore, no significant interactions between the NEO-FFI factors and any of the experimental factors were observed [all  $F_{(1, 12)} \leq 4.3$ , all  $p \geq 0.06$ ].

### CEREBRAL RESPONSES

#### General and emotion-specific effects of NECT

NECT-specific reductions in brain activity were revealed as an interaction between time point and training type within the framework of a whole-brain ANOVA in a distributed set of brain

regions including face- and voice-selective areas of the visual and auditory cortices, STS, inferior frontal cortex, motor-related regions and cerebellum (see **Figures 2A–C**; **Tables 1, 2**). Post-hoc analysis of the mean parameter estimates from the clusters with significant interaction effects revealed that the interaction was uniformly due to a NECT-associated decrease in cerebral activation (all  $T \geq 5.5$ , all  $p \leq 0.001$ ) while no significant change in the cerebral activation patterns occurred in the SUDOKU group (all  $T \leq 2.2, p \geq 0.08$ ). An additional investigation of NECT effects specific for emotional non-verbal signals as compared to neutral non-verbal stimuli, framed as a second order interaction between time point, training type and non-verbal emotional stimulus content (i.e., emotional vs. neutral), did not yield any whole-brain significant results. Applying a more sensitive threshold of  $p < 0.01$  uncorrected, however, demonstrated several clusters in inferolateral and dorsolateral regions of the prefrontal cortex as well as the anterior and middle cingulum extending into the posterior cingulum and precuneus (see **Table 3**). Please note that these results are only reported for the purpose of completeness and are purely descriptive.

#### Correlations of cerebral responses with personality (NEO-FFI)

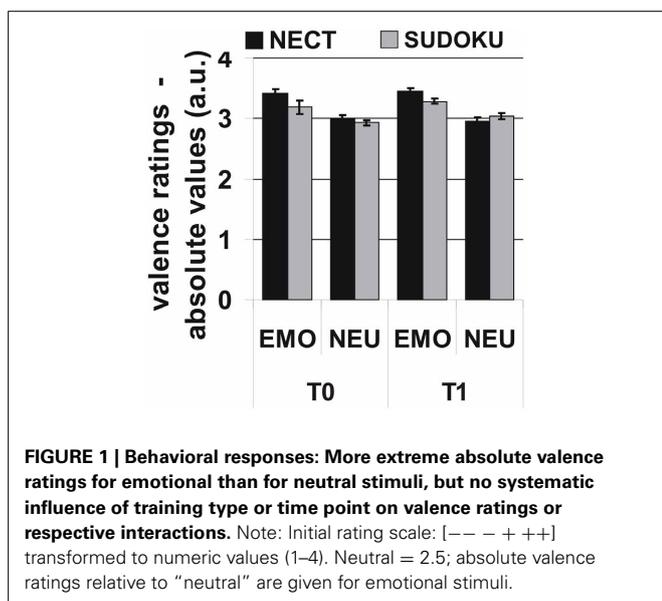
There were no significant differences with regard to the separate NEO-FFI personality factors or age between the training groups [all  $\text{abs}(t_{(12)}) \leq 1.9$ , all  $p \geq 0.09$ ].

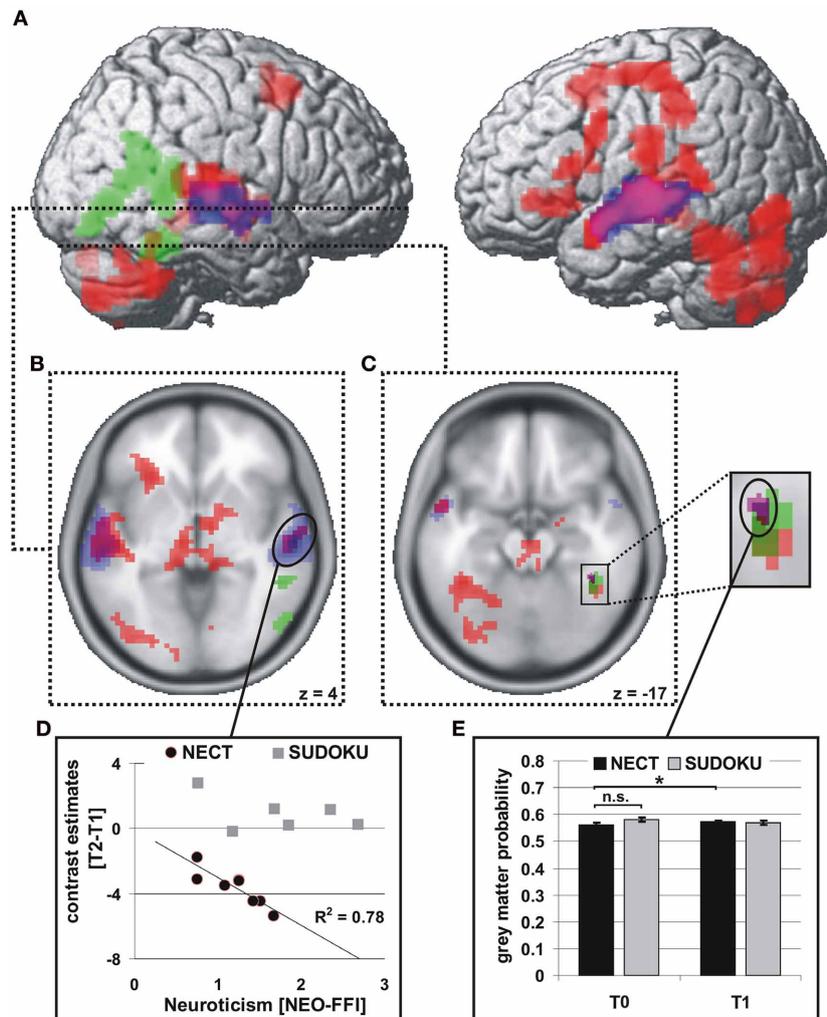
Of the five personality factors investigated, solely neuroticism exhibited a significant linear relationship with the size of NECT-associated decreases in cerebral activation during processing of non-verbal cues from voice and face in three regions, namely the right middle/posterior superior temporal gyrus (STG;  $r = -0.88, p = 0.004$ ; see **Figures 2B,D**), the left STS ( $r = -0.78, p = 0.024$ ; see **Figure 2B**) and the midbrain/thalamus ( $r = -0.71, p = 0.048$ ; see **Figure 2B**). Only the linear relationship in the right middle/posterior STG survived correction for the number of personality factors tested ( $p = 0.018$ ). No equivalent significant linear relationships between the effects of SUDOKU training and any of the personality factors were observed.

#### Correlations of cerebral responses with NECT-associated behavioral changes in sensitivity to non-verbal emotional cues

**Valence ratings.** The ROI analysis of the voice- and face-selective brain regions and brain regions with significant NECT-effects revealed that the right STS face area (STS-FA; coordinates of peak voxel:  $48x, -57y, 15z$ ;  $Z$  value of peak voxel: 3.7; cluster size: 34 voxels; see **Figures 3A,B**) and the right FFA (coordinates of peak voxel:  $42x, -42y, -24z$ ;  $Z$  value of peak voxel: 3.0; cluster size: 2 voxels; see **Figures 3A,C,D**) exhibited a positive linear relationship between NECT-associated changes in sensitivity to emotional non-verbal cues and NECT-associated increases in sensitivity to such cues. At the whole-brain level no such relationship was observed.

**Response times.** At the whole-brain level significant training specific (i.e., NECT vs. SUDOKU) associations between changes in the behavioral emotion-sensitivity and cerebral responses were observed in the left temporal pole and the right inferior frontal gyrus (see **Table 4** and **Figure 4A**). In a ROI analysis of the voice- and face-selective brain regions and brain regions with significant





**FIGURE 2 | Brain regions with significant functional and structural effects of NECT.** fMRI: interaction of training type with time point of measurement (red; **A–C**),  $p < 0.001$ , uncorrected,  $k > 47$  voxels, corresponding to  $p < 0.05$ , FWE corrected at cluster level. Voice-selective areas (blue) and face-selective areas (green),  $p < 0.0001$ , uncorrected,  $p < 0.05$  FWE whole-brain-corrected at

cluster level. **(D)** Linear relationship between training-associated changes in neural activity and neuroticism in the NECT-group in the right mSTG. VBM: **(E)** interaction of training type with time point of measurement in the right FFA (purple; **C**),  $p < 0.001$ , uncorrected,  $p < 0.05$  small-volume FWE corrected (right FFA) with a significant increase in NECT group. \* $p < 0.05$ .

NECT-effects with a more sensitive voxel-wise threshold of  $p < 0.01$ , this effect was also found in the midbrain/thalamus ROI (coordinates of peak voxel:  $-6x, -24y, -6z$ ; Z value of peak voxel: 3.9; cluster size: 65 voxels; see **Figure 4B**). Post-hoc analyses revealed that the observed effects were uniformly driven by the difference between a significant negative linear relationship between the changes in the response time correlate of emotion sensitivity  $T0[EMO - NEU] - T1[EMO - NEU]$  and its cerebral counterpart in the SUDOKU group [ $r_{(6)} \geq 0.89, p \leq 0.02$ ], on the one hand, and the reversed (i.e., positive) linear relationship in the NECT group (see **Figure 4C**). This relationship in the NECT group was significant in the left temporal pole [ $r_{(4)} = 0.72; p = 0.046$ ] and non-significant in the right inferior frontal gyrus [ $r_{(4)} = 0.58, p = 0.13$ ] and the midbrain/thalamus ROI [ $r_{(4)} = 0.45, p = 0.26$ ]. In other words, an increase in the response time

difference between non-verbal emotional and neutral cues was accompanied by an increase in activation differences between these types of cues in the NECT group. In the SUDOKU group, the same behavioral pattern was associated with decreased activation differences between emotional and neutral cues.

#### NECT-ASSOCIATED CHANGES IN GRAY MATTER VOLUME

Only in the right FFA an interaction between time point and training type was observed with regard to GM volume (coordinates of peak voxel:  $40x, -42y, -18z$ ; Z value of peak voxel: 3.5; cluster size: 24 voxels; see **Figures 2C,E**). Neither in the other face- or voice-selective areas or in the brain regions with NECT-associated activation decreases nor at the whole-brain level such an effect could be observed. The interaction effect in the right FFA was driven by a significant increase in GM volume

**Table 1 | Brain areas with training (NECT) specific changes in their responses to non-verbal cues from voice and face as determined by the whole-brain interaction analysis between time point (T0, T1) and training type (NECT, SUDOKU).**

	x	y	z	Z-score (peak voxel)	Cluster size (voxel)
L superior and middle temporal gyri/ L postcentral gyrus/ L Rolandic operculum/ L supramarginal gyrus/ L Heschl's gyrus <b>[left STS]</b>	-48	-18	6	5.3	486
R+L cerebellum/ L fusiform gyrus/ L inferior and middle occipital gyri/ L inferior temporal gyrus	9	-78	-45	4.6	833
Midbrain/ L+R thalamus/ R+L lingual gyri/ R cerebellum/ L+R parahippocampal gyrus/ R hippocampus <b>[midbrain/thalamus]</b>	6	-15	-12	4.5	359
L+R supplementary motor area/ L precentral gyrus/ L postcentral gyrus/ L inferior parietal gyrus/ L superior and middlefrontal gyri	-3	9	48	4.5	294
R superior and middle temporal gyri/ R Heschl's gyrus <b>[right middle/posterior STG]</b>	57	-24	3	4.3	146
L insula/ L inferior frontal gyrus, partes orbitalis and triangularis	-27	27	-3	4.1	68
L inferior frontal gyrus, partes opercularis and triangularis/ L precentral gyrus	-39	15	15	4.0	72
R+L cerebellum/ R fusiform gyrus	33	-48	-39	3.8	74

Activations thresholded at  $p < 0.001$ , uncorrected with a cluster size  $k \geq 50$ , corresponding to  $p < 0.05$ , FWE corrected for multiple comparisons across the whole-brain. Voxel size  $3 \times 3 \times 3 \text{ mm}^3$ . The anatomical descriptions in brackets represent short descriptions of the respective clusters. The clusters are referred to using these descriptions in the text.

**Table 2 | Voice- and face-selective brain areas with training (NECT) specific changes in their responses to non-verbal cues from voice and face.**

	x	y	z	Z-score (peak voxel)	Cluster size (voxel)
ROI: R temporal voice area (TVA)	60	-18	-3	6.5	237
Effect of NECT within R TVA: R superior temporal gyrus	57	-24	3	4.3	41
ROI: L temporal voice area	-54	-24	0	5.4	346
Effect of NECT within L TVA: L superior and middle temporal gyrus	-51	-21	-3	5.0	177
ROI: R fusiform face area	42	-45	-21	5.0	45
Effect of NECT within R FFA: R fusiform gyrus	42	-51	-21	3.4	8
ROI: R STS face area	48	-39	6	5.1	272
Effect of NECT within R STS-FA: R superior and middle temporal gyrus	45	-36	12	3.8	11

Localizer experiment activations are thresholded at  $p < 0.0001$ , uncorrected, with a cluster size  $k \geq 10$  (face localizer) corresponding to  $p < 0.05$ , FWE corrected for multiple comparisons across the whole-brain. Within-ROI NECT effects are thresholded at  $p < 0.001$ , uncorrected. All within-ROI cluster NECT-effects are significant with  $p < 0.05$ , FWE corrected for multiple comparisons across the respective ROI (SVC). Voxel size  $3 \times 3 \times 3 \text{ mm}^3$ .

in the NECT group [ $t_{(7)} = 2.2$ ,  $p = 0.03$ ] while there were no differences between the two training groups at T0 [ $t_{(12)} = 1.5$ ,  $p = 0.148$ ]. After removal of variance in GM volume at T0, the NECT-associated increase in GM volume remained significant [ $t_{(7)} = 3.0$ ,  $p = 0.005$ ].

## DISCUSSION

This is, to our knowledge, the first combined fMRI and structural cerebral imaging study, reporting the specific effects of a four-week-long NECT in healthy individuals. Our results afford a first view on the neuronal underpinnings of non-verbal

trainings which—implemented in various forms—have been demonstrated to be effective in enhancing non-verbal decoding skills and sensitivity (e.g., Klinzing and Jackson, 1987; Constanzo, 1992; Matsumoto and Hwang, 2011).

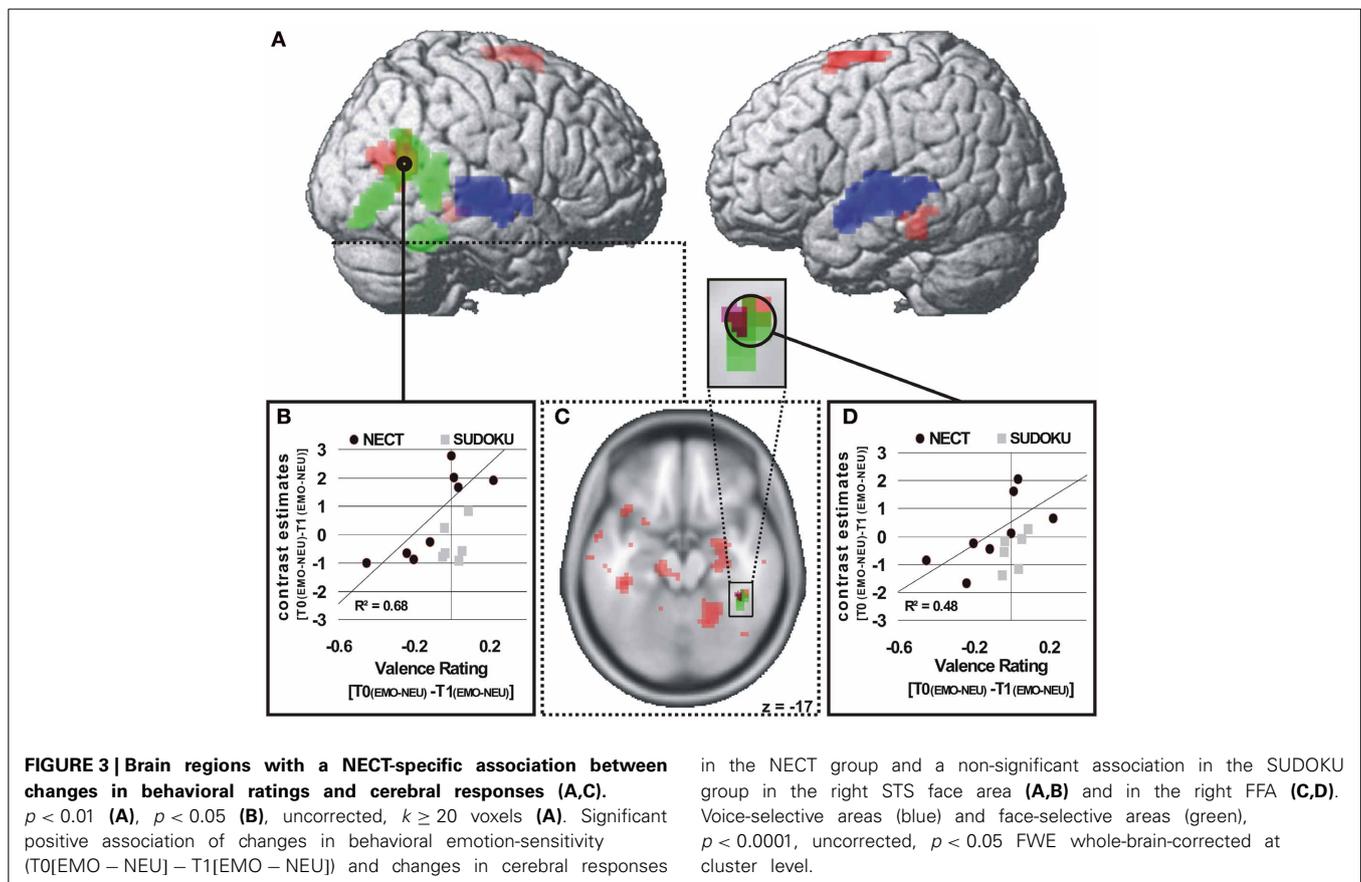
## GENERAL EFFECTS OF NECT ON CEREBRAL ACTIVATION PATTERNS

In accordance with our expectations, NECT-specific alterations in brain activity were observed in a distributed set of brain regions including face- and voice-selective areas of the visual and auditory cortex, the STS, inferior frontal cortex, insula and thalamus which have all been demonstrated to be involved in the processing

**Table 3 | Brain areas with differential training (NECT) specific changes in their responses to emotional and neutral non-verbal cues from voice and face.**

	x	y	z	Z-score (peak voxel)	Cluster size (voxel)
R inferior frontal gyrus, partes triangularis, orbitalis and opercularis	57	24	9	3.7	92
R middle frontal gyrus/ R precentral gyrus	45	9	42	3.5	56
L calcarine gyrus/ L precuneus	-6	-48	6	3.4	41
R inferior and middle frontal gyri, partes orbitals	39	36	-9	3.3	45
L+R anterior cingulum/ L+R superior frontal gyrus, pars medialis	3	45	12	3.3	41
R+L middle cingulum/ R precuneus/ L posterior cingulum	6	-27	36	3.2	121
R+L thalamus	6	-18	15	3.2	53

Activations thresholded at  $p < 0.01$ , uncorrected with a cluster size  $k \geq 40$ . Voxel size  $3 \times 3 \times 3 \text{ mm}^3$ . Reported for descriptive purposes only.



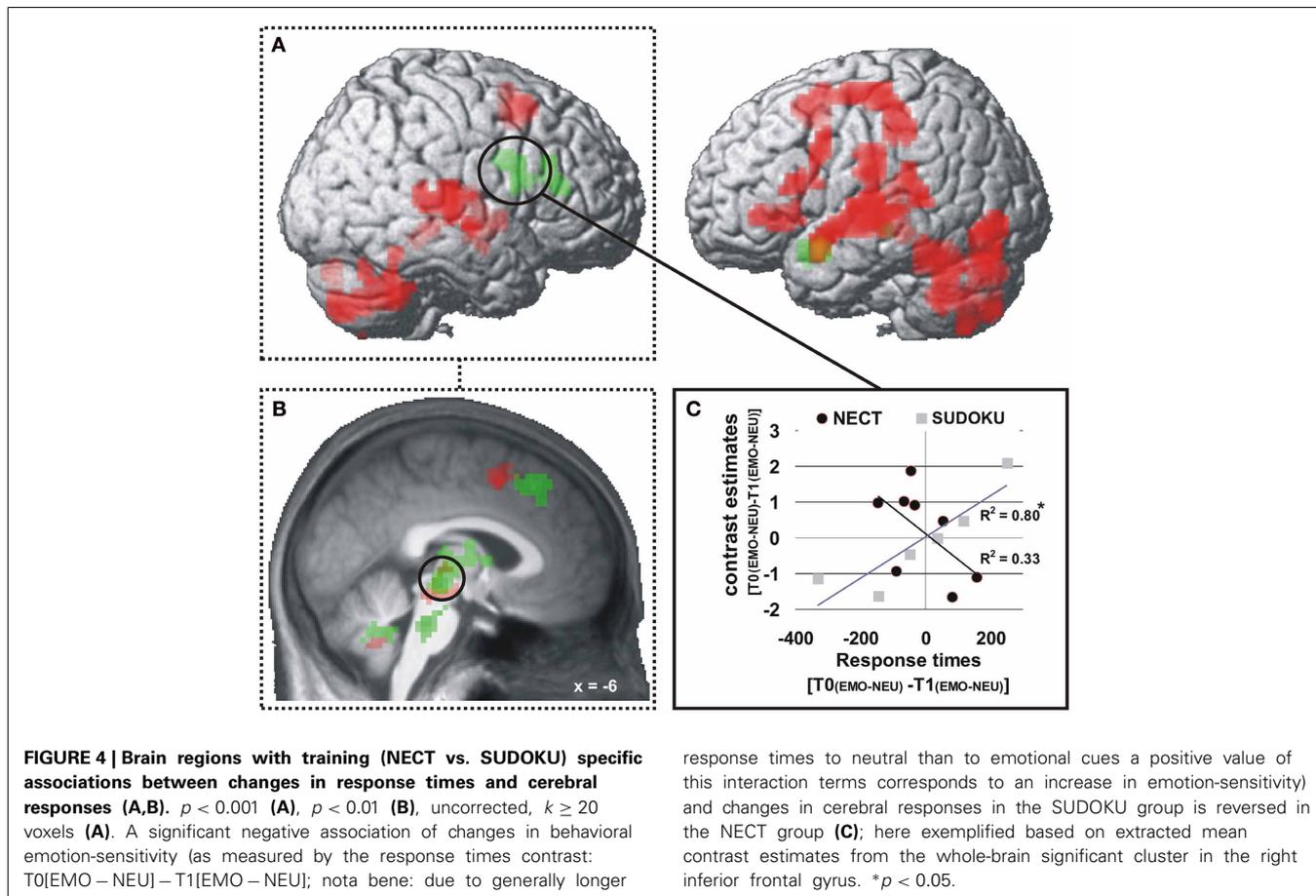
of audiovisual non-verbal emotional signals (for reviews see Campanella and Belin, 2007; Brück et al., 2011b; Kreifelts et al., 2013). Moreover, NECT induced altered activation patterns in motor-related regions, the cerebellum and the parietal cortex. All brain regions with NECT-specific activation alterations showed a uniform pattern with decreased activation during the perception and evaluation of non-verbal signals from voice and face which fit in well with earlier reports of decreased cerebral activation after procedural (Friston et al., 1992; Steele and Penhune, 2010) and perceptual (Schiltz et al., 1999) learning. These decreases may constitute a correlate of training-induced familiarization with the evaluation of face/voice stimuli including less effortful perceptual

processing as well as a facilitation of stimulus-directed attention, decision making and response selection. It is worth noting, that the observed decreases in activation appear similar to the effect of so called “repetition suppression,” a phenomenon occurring after the repeated presentation of identical stimuli. Repetition suppression effects presumably reflect bottom-up sharpening of neural responses (Larsson and Smith, 2012) but also top-down mediated perceptual expectations (Summerfield et al., 2008). Although, in the present study, “simple” repetition suppression can be ruled out as source of the observed effects due to the fact that each stimulus was shown exactly twice during the course of the study in both training groups, nevertheless NECT might tap

**Table 4 | Brain areas with training (NECT vs. SUDOKU) specific associations between training-induced changes in the behavioral sensitivity to non-verbal emotional cues as measured by response times and the respective cerebral activation patterns (contrast of interest:  $T0[EMO - NEU] - T1[EMO - NEU]$ ).**

	x	y	z	Z-score (peak voxel)	Cluster size (voxel)
L superior temporal pole	-39	9	-18	4.28	47*
Midbrain	-3	-24	-9	4.25	24
R inferior frontal gyrus, partes opercularis and triangularis	36	9	24	3.95	49*
R inferior frontal gyrus, pars triangularis	48	27	27	3.77	20

Activations thresholded at  $p < 0.001$ , uncorrected, with a minimal cluster size of  $k \geq 20$  voxels. Activations with a cluster size  $k \geq 46$ , correspond to a significance level of  $p < 0.05$ , FWE corrected for multiple comparisons across the whole-brain and are marked with an asterisk. Voxel size  $3 \times 3 \times 3 \text{ mm}^3$ .



into similar bottom-up and top-down neuronal tuning processes, not via the repetition of identical stimuli but via similarities between the tasks and stimuli of the training and during the fMRI experiment.

Decreases of cerebral responses through NECT in healthy individuals also stand in contrast to the observation of cerebral activation increases observed in schizophrenic patients after an emotion recognition training (Habel et al., 2010). Given the paucity of data with regard to the neural correlates of such non-verbal trainings, further studies are needed which directly compare training effects in psychiatric populations and healthy individuals as an active control condition to determine if there

are indeed fundamental differences in the neural correlates of non-verbal emotional trainings between these groups.

With regard to differential NECT training effects for emotional and neutral non-verbal signals, the results of our study were negative: There was a non-significant tendency toward greater absolute valence rating differences between emotional and neutral stimuli after NECT at the behavioral level (see Figure 1) and several brain areas in the right orbitofrontal cortex, right lateral prefrontal cortex as well as areas in the anterior and middle cingulate cortex extending into the precuneus where differential NECT training effects for emotional and neutral non-verbal stimuli failed whole-brain significance. Two potential reasons for

this negative finding should be considered: Firstly, the negative finding might be due to a lack of power. This notion is supported by the suggestive correspondence between the behavioral tendency toward a more sensitive discrimination of emotional and neutral stimuli and sub-threshold cerebral activation patterns occurring in brain areas previously associated with the evaluation of emotional stimuli, mentalizing (i.e., inferences on the mental states and intentions), and supramodal emotion processing (Amodio and Frith, 2006; Brück et al., 2011b; Klasen et al., 2011, 2012).

A second reason might be that in the context of a valence rating task “neutral” and “emotional” stimuli are processed similarly as both types of stimuli are evaluated with respect to their emotional valence. While offering a direct behavioral correlate of emotional evaluation, this experimental context may have decreased activation differences for “emotional” and “neutral” stimuli relative to, for example, an implicit emotion processing design with a gender discrimination task as demonstrated in previous studies (Hariri et al., 2000; Lange et al., 2003).

#### **PERSONALITY-DEPENDENT NEURONAL EFFECTS OF NECT**

The middle and posterior aspects of the right STG, and to a lesser degree also the left STS and posterior thalamus/midbrain exhibited a specific characteristic in that neuronal NECT effects here were strongly and positively correlated with the personality trait of neuroticism. In light of the known association between neuroticism and depression as well as anxiety disorders (Brandes and Bienvu, 2006; Klein et al., 2011) with high levels of neuroticism in anxious and depressed patient samples, this association may appear relevant for future studies with respect to the strength of observable neuronal NECT effects in clinical samples of anxious and depressed patients. Here, the present data set suggests that the present experimental design may be even more sensitive in detecting cerebral correlates of NECT in clinical groups with strong neurotic personality traits.

#### **NECT INDUCED CHANGES IN SENSITIVITY TO EMOTIONAL SIGNALS: CORRELATIONS BETWEEN BEHAVIOR AND CEREBRAL ACTIVATION**

For both, the valence rating correlate as well as the response time correlate of changes in emotion sensitivity, corresponding cerebral activation patterns were observed:

##### ***Valence ratings***

Exclusively within the face-selective areas in the right pSTS and the right fusiform gyrus a NECT-specific correlation between changes in the behavioral sensitivity to non-verbal emotional signals from voice and face and corresponding activation patterns was observed. The degree to which individuals after NECT perceived greater absolute valence differences between neutral and emotional stimuli was correlated with an increase in cerebral sensitivity to emotional as compared to neutral non-verbal stimuli. The exclusive observation of such associations in face-selective cortices supports the hypothesis that sensory tuning in the decoding of facial expressions lies at the neural basis of NECT-induced behavioral alterations in sensitivity to non-verbal emotional cues. It remains an open question if the lack

of such a correlation in the voice-selective sensory cortices is the result of an innate preference of humans to base the evaluation of social signals on visual stimulus components (e.g., DePaulo et al., 1978), or if face-selective sensory cortices are more training sensitive in the plasticity of their neuronal responses than voice-selective areas. Thirdly, despite the audiovisual nature of NECT, a potential tendency of the participants toward the use of visual non-verbal signals during NECT might explain the observed patterns of association between behavioral data and cerebral activation. Moreover, the observation of correlations between emotion sensitivity and cerebral activation patterns both in the STS-FA and FFA, justifies to very cautiously argue against the view of a strict functional segregation of these two face-processing modules (Haxby et al., 2000) where the FFA processes invariant aspects of faces (e.g., gender or identity) and the STS-FA processes dynamic aspects of faces (e.g., facial expressions or gaze).

Finally, the largest cortical cluster with a correlation between training induced increases in facial emotion recognition in schizophrenic patients and corresponding increases in cerebral responses (Habel et al., 2010) was observed in the pSTS in a very similar location as the association between behavioral and cerebral correlates of NECT. This points to the STS-FA as a central cortical module linking behavioral and neuronal effects of non-verbal emotional recognition and communication trainings.

##### ***Response times***

For changes in the response time-based behavioral measure of emotion sensitivity, neural activation correlates were revealed in the left temporal pole, the right inferior frontal gyrus and midbrain/thalamus. While individually increased behavioral sensitivity to emotional cues was correlated with a decrease in cerebral sensitivity to these cues in the SUDOKU group, NECT led to the reversed pattern with parallel changes in behavioral and cerebral correlates of emotion sensitivity.

Parallel to the association of valence rating changes and cerebral activation patterns, interpretations of the above findings in a rather small study like the present one have to be phrased with caution and need to be treated as preliminary.

The most striking feature of the analyses presented here is that NECT induced a qualitative change in the association of behavioral and cerebral correlates of emotion-sensitivity. This finding points to the left temporal pole, the right inferior frontal gyrus and the midbrain/thalamus as cerebral structures which instantiate NECT-driven training effects as interfaces between the evaluation of non-verbal signals and the selection of responses to these signals. The assumption that the right inferior frontal gyrus is implicated in this process is in line with findings that the lateral prefrontal cortex is activated during the evaluation of non-verbal emotional stimulus content (e.g., Schirmer and Kotz, 2006; Brück et al., 2011b; Ethofer et al., 2013) and structurally connected to voice- and face-selective cortices, the face/voice integration region of the STS as well as the supplementary motor area (Ethofer et al., 2013). Also, the temporal pole has been implicated in emotion processing (Olson et al., 2007) and multimodal integration of audiovisual emotional signal (Kreifelts et al., 2007)

although its specific functional properties remain at least partially unresolved. Olson et al. (2007) hypothesized that the temporal pole binds complex, perceptual inputs to visceral emotional responses. Present results suggest an additional role of the temporal pole in binding complex emotional percepts to voluntary motor responses. As for the midbrain/thalamus cluster, a comparison with the activation coordinates of previous studies on audiovisual integration and supramodal representation of emotional signals (Kreifelts et al., 2007, 2010; Klasen et al., 2011) supports the hypothesis that the location of this cluster overlaps with a brain region in the posterior thalamus which has been demonstrated not only to be involved in the audiovisual integration of emotional signals (Kreifelts et al., 2007, 2010; Klasen et al., 2011) but also to exhibit a linear relationship of hemodynamic with behavioral responses (i.e., emotion recognition rates; Kreifelts et al., 2007).

### NECT-INDUCED STRUCTURAL PLASTICITY

The increase in GM volume after NECT in the right FFA represents, to our knowledge, the first demonstration of specific structural plasticity induced by a complex emotional communication training. These results fit in well with the growing body of studies investigating the structural effects of various motor-related trainings (Driemeyer et al., 2008; Taubert et al., 2010; Granert et al., 2011) but also visual perceptual training (Ditye et al., 2013). These concordantly indicate dynamic increases in GM volume in areas functionally associated with the trained tasks. Regarded from this perspective, the present structural findings argue for a pivotal position of the right FFA within the cerebral network exhibiting dynamic functional and structural alterations as neuronal correlates of NECT. Nevertheless, one may ask if the small size of the FFA as a functional cortical module does not somewhat boost the sensitivity to detect even smaller effects.

### LIMITATIONS AND PERSPECTIVES

A limitation on the implications of the present study is its small sample size which allows only large (i.e., with regard to effect

size) effects of NECT to be discovered. Therefore, a training study with a larger sample size may depict a more complex and fine-grained pattern of cerebral alterations induced by NECT. Nevertheless, the considerable strength of the observed effects affords a first impression of NECT-induced functional and structural cerebral alterations and may serve as a good starting point for future studies including larger samples of healthy individuals and patients with psychiatric conditions. A second important point pertains specifically to the neurofunctional correlates of NECT: In the present study these are inherently tied to the cognitive context of a valence rating task. Depending on the psychiatric conditions investigated in future studies, it would be sensible to accommodate disorder-specific neuropsychological and behavioral alterations in the processing of non-verbal emotional signals within the experimental setup (i.e., both stimulus material and task) in order to capture an optimized estimate of disease-relevant behavioral alterations and their neuronal correlates.

### CONCLUSION

Here, we demonstrated in healthy participants, an association of NECT-induced changes in brain function and structure with changes in the evaluation of non-verbal emotional stimulus content and neuroticism. Based on these findings we conclude that the present experimental design may be a very valuable neuroimaging probe not only for the investigation of the neural bases of altered processing of non-verbal emotional cues in psychiatric disorders but also for the assessment of the influence of therapeutic interventions on brain function and structure.

### ACKNOWLEDGMENTS

We would like to thank Hyeri Lee for developing the layout of the non-verbal communication game. Furthermore, we acknowledge support by the Deutsche Forschungsgemeinschaft and the Open Access Publishing Fund of Tübingen University.

### REFERENCES

- Amodio, D. M., and Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* 7, 268–277. doi: 10.1038/nrn1884
- Andersson, J. L., Hutton, C., Ashburner, J., Turner, R., and Friston, K. (2001). Modeling geometric deformations in EPI time series. *Neuroimage* 13, 903–919. doi: 10.1006/nimg.2001.0746
- Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., and Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature* 403, 309–312. doi: 10.1038/35002078
- Borkenau, P., and Ostendorf, F. (2008). *NEO-FFI NEO-Fünf-Faktoren-Inventar Nach Costa und McCrae. 2. Neu Normierte und Vollständige Überarbeitete Auflage.* Göttingen: Hogrefe.
- Brandes, M., and Bienvenu, O. J. (2006). Personality and anxiety disorders. *Curr. Psychiatry Rep.* 8, 263–269. doi: 10.1007/s11920-006-0061-8
- Brück, C., Kreifelts, B., Kaza, E., Lotze, M., and Wildgruber, D. (2011a). Impact of personality on the cerebral processing of emotional prosody. *Neuroimage* 58, 259–268. doi: 10.1016/j.neuroimage.2011.06.005
- Brück, C., Kreifelts, B., and Wildgruber, D. (2011b). Emotional voices in context: a neurobiological model of multimodal affective information processing. *Phys. Life Rev.* 8, 383–403. doi: 10.1016/j.plrev.2011.10.002
- Campanella, S., and Belin, P. (2007). Integrating face and voice in person perception. *Trends Cogn. Sci.* 11, 535–543. doi: 10.1016/j.tics.2007.10.001
- Constanzo, M. (1992). Training students to decode verbal and non-verbal cues: effects on confidence and performance. *J. Educ. Psychol.* 84, 308–313. doi: 10.1037/0022-0663.84.3.308
- Creemers, H. R., Demenescu, L. R., Aleman, A., Renken, R., Van Tol, M. J., Van Der Wee, N. J., et al. (2010). Neuroticism modulates amygdala-prefrontal connectivity in response to negative emotional facial expressions. *Neuroimage* 49, 963–970. doi: 10.1016/j.neuroimage.2009.08.023
- Dawel, A., O'Kearney, R., McKone, E., and Palermo, R. (2012). Not just fear and sadness: meta-analytic evidence of pervasive emotion recognition deficits for facial and vocal expressions in psychopathy. *Neurosci. Biobehav. Rev.* 36, 2288–2304. doi: 10.1016/j.neubiorev.2012.08.006
- Demenescu, L. R., Kortekaas, R., Den Boer, J. A., and Aleman, A. (2010). Impaired attribution of emotion to facial expressions in anxiety and major depression. *PLoS ONE* 5:e15058. doi: 10.1371/journal.pone.0015058
- DePaulo, B. M., Rosenthal, R., Eisenstat, R. A., Rogers, P. L., and Finkelstein, S. (1978). Decoding discrepant nonverbal cues. *J. Pers. Soc. Psychol.* 36, 313–323. doi: 10.1037/0022-3514.36.3.313

- Ditye, T., Kanai, R., Bahrami, B., Muggleton, N. G., Rees, G., and Walsh, V. (2013). Rapid changes in brain structure predict improvements induced by perceptual learning. *Neuroimage* 81, 205–212. doi: 10.1016/j.neuroimage.2013.05.058
- Domes, G., Schulze, L., and Herpertz, S. C. (2009). Emotion recognition in borderline personality disorder—a review of the literature. *J. Pers. Disord.* 23, 6–19. doi: 10.1521/pedi.2009.23.1.6
- Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., and May, A. (2004). Neuroplasticity: changes in grey matter induced by training. *Nature* 427, 311–312. doi: 10.1038/427311a
- Driemeyer, J., Boyke, J., Gaser, C., Buchel, C., and May, A. (2008). Changes in gray matter induced by learning—revisited. *PLoS ONE* 3:e2669. doi: 10.1371/journal.pone.0002669
- Epstein, R., Harris, A., Stanley, D., and Kanwisher, N. (1999). The parahippocampal place area: recognition, navigation, or encoding. *Neuron* 23, 115–125. doi: 10.1016/S0896-6273(00)80758-8
- Ethofer, T., Bretscher, J., Gschwind, M., Kreifelts, B., Wildgruber, D., and Vuilleumier, P. (2012). Emotional voice areas: anatomic location, functional properties, and structural connections revealed by combined fMRI/DTI. *Cereb. Cortex* 22, 191–200. doi: 10.1093/cercor/bhr113
- Ethofer, T., Bretscher, J., Wiethoff, S., Bisch, J., Schlipf, S., Wildgruber, D., et al. (2013). Functional responses and structural connections of cortical areas for processing faces and voices in the superior temporal sulcus. *Neuroimage* 76C, 45–56. doi: 10.1016/j.neuroimage.2013.02.064
- Friston, K. J., Frith, C. D., Passingham, R. E., Liddle, P. F., and Frackowiak, R. S. (1992). Motor practice and neurophysiological adaptation in the cerebellum: a positron tomography study. *Proc. Biol. Sci.* 248, 223–228. doi: 10.1098/rspb.1992.0065
- Friston, K. J., Glaser, D. E., Henson, R. N., Kiebel, S., Phillips, C., and Ashburner, J. (2002). Classical and Bayesian inference in neuroimaging: applications. *Neuroimage* 16, 484–512. doi: 10.1006/nimg.2002.1091
- Funcke, A. (2006). *Vorstellbar: Methoden von Schauspielern und Regisseuren für den ganz normalen Trainer*. Bonn: managerSeminar Verlags GmbH.
- Granert, O., Peller, M., Gaser, C., Groppa, S., Hallett, M., Knutzen, A., et al. (2011). Manual activity shapes structure and function in contralateral human motor hand area. *Neuroimage* 54, 32–41. doi: 10.1016/j.neuroimage.2010.08.013
- Habel, U., Koch, K., Kellermann, T., Reske, M., Frommann, N., Wolwer, W., et al. (2010). Training of affect recognition in schizophrenia: neurobiological correlates. *Soc. Neurosci.* 5, 92–104. doi: 10.1080/17470910903170269
- Hariri, A. R., Bookheimer, S. Y., and Mazziotta, J. C. (2000). Modulating emotional responses: effects of a neocortical network on the limbic system. *Neuroreport* 11, 43–48. doi: 10.1097/00001756-200001170-00009
- Haxby, J. V., Hoffman, E. A., and Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends Cogn. Sci.* 4, 223–233. doi: 10.1016/S1364-6613(00)01482-0
- Jacob, H., Kreifelts, B., Brück, C., Nizielski, S., Schütz, A., and Wildgruber, D. (2013). Nonverbal signals speak up: association between perceptual nonverbal dominance and emotional intelligence. *Cogn. Emot.* 27, 783–799. doi: 10.1080/02699931.2012.739999
- Kanwisher, N., McDermott, J., and Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302–4311.
- Kehoe, E. G., Toomey, J. M., Balsters, J. H., and Bokde, A. L. (2012). Personality modulates the effects of emotional arousal and valence on brain activation. *Soc. Cogn. Affect. Neurosci.* 7, 858–870. doi: 10.1093/scan/nsr059
- Klases, M., Kenworthy, C. A., Mathiak, K. A., Kircher, T. T., and Mathiak, K. (2011). Supramodal representation of emotions. *J. Neurosci.* 31, 13635–13643. doi: 10.1523/JNEUROSCI.2833-11.2011
- Klases, M., Chen, Y. H., and Mathiak, K. (2012). Multisensory emotions: perception, combination and underlying neural processes. *Rev. Neurosci.* 23, 381–392. doi: 10.1515/revneuro-2012-0040
- Klein, D. N., Kotov, R., and Bufferd, S. J. (2011). Personality and depression: explanatory models and review of the evidence. *Annu. Rev. Clin. Psychol.* 7, 269–295. doi: 10.1146/annurev-clinpsy-03210-104540
- Klinzing, H. G., and Jackson, I. (1987). Training teachers in nonverbal sensitivity and nonverbal behavior. *Int. J. Educ. Res.* 11, 589–600. doi: 10.1016/0883-0355(87)90018-8
- Kohler, C. G., Hoffman, L. J., Eastman, L. B., Healey, K., and Moberg, P. J. (2011). Facial emotion perception in depression and bipolar disorder: a quantitative review. *Psychiatry Res.* 188, 303–309. doi: 10.1016/j.psychres.2011.04.019
- Kohler, C. G., Walker, J. B., Martin, E. A., Healey, K. M., and Moberg, P. J. (2010). Facial emotion perception in schizophrenia: a meta-analytic review. *Schizophr. Bull.* 36, 1009–1019. doi: 10.1093/schbul/sbn192
- Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., and Wildgruber, D. (2007). Audiovisual integration of emotional signals in voice and face: an event-related fMRI study. *Neuroimage* 37, 1445–1456. doi: 10.1016/j.neuroimage.2007.06.020
- Kreifelts, B., Ethofer, T., Huberle, E., Grodd, W., and Wildgruber, D. (2010). Association of trait emotional intelligence and individual fMRI-activation patterns during the perception of social signals from voice and face. *Hum. Brain Mapp.* 31, 979–991. doi: 10.1002/hbm.20913
- Kreifelts, B., Ethofer, T., Shiozawa, T., Grodd, W., and Wildgruber, D. (2009). Cerebral representation of non-verbal emotional perception: fMRI reveals audiovisual integration area between voice- and face-sensitive regions in the superior temporal sulcus. *Neuropsychologia* 47, 3059–3066. doi: 10.1016/j.neuropsychologia.2009.07.001
- Kreifelts, B., Wildgruber, D., and Ethofer, T. (2013). “Audiovisual integration of emotional information from voice and face,” in *Integrating Face and Voice in Person Perception*, eds P. Belin, S. Campanella, and T. Ethofer (New York; Heidelberg; Dordrecht; London: Springer), 225–252.
- Kurtz, M. M., and Richardson, C. L. (2012). Social cognitive training for schizophrenia: a meta-analytic investigation of controlled research. *Schizophr. Bull.* 38, 1092–1104. doi: 10.1093/schbul/sbr036
- Lange, K., Williams, L. M., Young, A. W., Bullmore, E. T., Brammer, M. J., Williams, S. C., et al. (2003). Task instructions modulate neural responses to fearful facial expressions. *Biol. Psychiatry* 53, 226–232. doi: 10.1016/S0006-3223(02)01455-5
- Larsson, J., and Smith, A. T. (2012). fMRI repetition suppression: neuronal adaptation or stimulus expectation. *Cereb. Cortex* 22, 567–576. doi: 10.1093/cercor/bhr119
- Matsumoto, D., and Hwang, H. S. (2011). Evidence for training the ability to read microexpressions of emotion. *Motiv. Emot.* 35, 181–191. doi: 10.1007/s11031-011-9212-2
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- Olson, I. R., Plotzker, A., and Ezzyat, Y. (2007). The Enigmatic temporal pole: a review of findings on social and emotional processing. *Brain* 130, 1718–1731. doi: 10.1093/brain/awm052
- Posamentier, M. T., and Abdi, H. (2003). Processing faces and facial expressions. *Neuropsychol. Rev.* 13, 113–143. doi: 10.1023/A:1025519712569
- Pourtois, G., De Gelder, B., Bol, A., and Crommelinck, M. (2005). Perception of facial expressions and voices and of their combination in the human brain. *Cortex* 41, 49–59. doi: 10.1016/S0010-9452(08)70177-1
- Robins, D. L., Hunyadi, E., and Schultz, R. T. (2009). Superior temporal activation in response to dynamic audio-visual emotional cues. *Brain Cogn.* 69, 269–278. doi: 10.1016/j.bandc.2008.08.007
- Rossa, W. (2009). *Sudoku für die Ferien: 400 Zahlenrätsel von leicht bis schwer*. München: Goldmann Verlag.
- Samame, C., Martino, D. J., and Strejilevich, S. A. (2012). Social cognition in euthymic bipolar disorder: systematic review and meta-analytic approach. *Acta Psychiatr. Scand.* 125, 266–280. doi: 10.1111/j.1600-0447.2011.01808.x
- Schiltz, C., Bodart, J. M., Dubois, S., DeJardin, S., Michel, C., Roucoux, A., et al. (1999). Neuronal mechanisms of perceptual learning: changes in human brain activity with training in orientation discrimination. *Neuroimage* 9, 46–62. doi: 10.1006/nimg.1998.0394
- Schirmer, A., and Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends Cogn. Sci.* 10, 24–30. doi: 10.1016/j.tics.2005.11.009
- Steele, C. J., and Penhune, V. B. (2010). Specific increases within global decreases: a functional magnetic resonance imaging investigation of

- five days of motor sequence learning. *J. Neurosci.* 30, 8332–8341. doi: 10.1523/JNEUROSCI.5569-09.2010
- Stein, M. B., Simmons, A. N., Feinstein, J. S., and Paulus, M. P. (2007). Increased amygdala and insula activation during emotion processing in anxiety-prone subjects. *Am. J. Psychiatry* 164, 318–327. doi: 10.1176/appi.ajp.164.2.318
- Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M. M., and Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nat. Neurosci.* 11, 1004–1006. doi: 10.1038/nn.2163
- Suslow, T., Kugel, H., Reber, H., Bauer, J., Dannlowski, U., Kersting, A., et al. (2010). Automatic brain response to facial emotion as a function of implicitly and explicitly measured extraversion. *Neuroscience* 167, 111–123. doi: 10.1016/j.neuroscience.2010.01.038
- Taubert, M., Lohmann, G., Margulies, D. S., Villringer, A., and Ragert, P. (2010). Long-term effects of motor training on resting-state networks and underlying brain structure. *Neuroimage* 57, 1492–1498. doi: 10.1016/j.neuroimage.2011.05.078
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978
- Wildgruber, D., Ackermann, H., Kreifelts, B., and Ethofer, T. (2006). Cerebral processing of linguistic and emotional prosody: fMRI studies. *Prog. Brain Res.* 156, 249–268. doi: 10.1016/S0079-6123(06)56013-3
- Worsley, K., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., and Evans, A. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Hum. Brain Mapp.* 4, 74–90.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 01 June 2013; accepted: 18 September 2013; published online: 17 October 2013.
- Citation:* Kreifelts B, Jacob H, Brück C, Erb M, Ethofer T and Wildgruber D (2013) Non-verbal emotion communication training induces specific changes in brain function and structure. *Front. Hum. Neurosci.* 7:648. doi: 10.3389/fnhum.2013.00648
- This article was submitted to the journal *Frontiers in Human Neuroscience*. Copyright © 2013 Kreifelts, Jacob, Brück, Erb, Ethofer and Wildgruber. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Emotional sounds modulate early neural processing of emotional pictures

Antje B. M. Gerdes<sup>1\*</sup>, Matthias J. Wieser<sup>2</sup>, Florian Bublitzky<sup>1</sup>, Anita Kusay<sup>1</sup>, Michael M. Plichta<sup>3</sup> and Georg W. Alpers<sup>1,4</sup>

<sup>1</sup> Department of Psychology, School of Social Sciences, University of Mannheim, Mannheim, Germany

<sup>2</sup> Department of Psychology, University of Würzburg, Würzburg, Germany

<sup>3</sup> Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim/Heidelberg University, Mannheim, Germany

<sup>4</sup> Otto-Selz Institute, University of Mannheim, Mannheim, Germany

## Edited by:

Martin Klasen, RWTH Aachen University, Germany

## Reviewed by:

Ewald Naumann, University of Trier, Germany

Rebecca Watson, University of Glasgow, UK

Sarah Jessen, Max Planck Institute for Human Cognitive and Brain Sciences, Germany

## \*Correspondence:

Antje B. M. Gerdes, Chair of Clinical and Biological Psychology, Department of Psychology, School of Social Sciences, University of Mannheim, L 13, 17, D-68131 Mannheim, Germany  
e-mail: gerdes@uni-mannheim.de

In our natural environment, emotional information is conveyed by converging visual and auditory information; multimodal integration is of utmost importance. In the laboratory, however, emotion researchers have mostly focused on the examination of unimodal stimuli. Few existing studies on multimodal emotion processing have focused on human communication such as the integration of facial and vocal expressions. Extending the concept of multimodality, the current study examines how the neural processing of emotional pictures is influenced by simultaneously presented sounds. Twenty pleasant, unpleasant, and neutral pictures of complex scenes were presented to 22 healthy participants. On the critical trials these pictures were paired with pleasant, unpleasant, and neutral sounds. Sound presentation started 500 ms before picture onset and each stimulus presentation lasted for 2 s. EEG was recorded from 64 channels and ERP analyses focused on the picture onset. In addition, valence and arousal ratings were obtained. Previous findings for the neural processing of emotional pictures were replicated. Specifically, unpleasant compared to neutral pictures were associated with an increased parietal P200 and a more pronounced centroparietal late positive potential (LPP), independent of the accompanying sound valence. For audiovisual stimulation, increased parietal P100 and P200 were found in response to all pictures which were accompanied by unpleasant or pleasant sounds compared to pictures with neutral sounds. Most importantly, incongruent audiovisual pairs of unpleasant pictures and pleasant sounds enhanced parietal P100 and P200 compared to pairings with congruent sounds. Taken together, the present findings indicate that emotional sounds modulate early stages of visual processing and, therefore, provide an avenue by which multimodal experience may enhance perception.

**Keywords:** emotional pictures, emotional sounds, audiovisual stimuli, ERPs, P100, P200, LPP

## INTRODUCTION

In everyday life people are confronted with an abundance of different emotional stimuli from the environment. Typically, these cues are transmitted through multiple sensory channels and especially audiovisual stimuli (e.g., information from face and voice in the social interaction context) are highly prevalent. Only a fraction of this endless stream of information however is consciously recognized, is attended to and more elaborately processed (Schupp et al., 2006). To cope with limited processing capacities, emotionally relevant cues have been suggested to benefit from prioritized information processing (Vuilleumier, 2005). Despite the high relevance of multimodal emotional processing, emotion research has mainly focused on investigating unimodal stimuli (Campanella et al., 2010). Furthermore, existing studies on multimodal stimuli predominantly investigated how emotional faces and emotional voices are integrated (for a recent review see Klasen et al., 2012). As expected, most of the studies generally indicate that behavioral outcome is based on interactive integration

of multimodal emotional information (de Gelder and Bertelson, 2003; Mothes-Lasch et al., 2012). For example, emotion recognition is improved in response to redundant multimodal compared to unimodal stimuli (Vroomen et al., 2001; Kreifelts et al., 2007; Paulmann and Pell, 2011). Furthermore, the identification and evaluation of an emotional facial expression is biased toward the valence of simultaneously presented affective prosodic stimuli and vice versa (de Gelder and Vroomen, 2000; de Gelder and Bertelson, 2003; Focker et al., 2011; Rigoulot and Pell, 2012). Such interactions between emotional face and voice processing even occur when subjects were asked to ignore concurrent sensory information (Collignon et al., 2008) and were shown to be independent of attentional resources (Vroomen et al., 2001; Focker et al., 2011). In addition, the processing of emotional cues can even alter responses to non-related events coming from a different sensory modality which may indicate that an emotional context can modulate the excitability of sensory regions (Dominguez-Borrás et al., 2009).

Regarding cortical stimulus processing, event-related potentials (ERP) to picture cues are well-suited to investigate the time course of attentional and emotional processes (Schupp et al., 2006). Already early in the visual processing stream, differences have been shown for emotional as compared to neutral pictures for the P100, P200, and the early posterior negativity (EPN). These early components may relate to facilitated sensory processing fostering detection and categorization processes. Later processing stages have been associated with detailed evaluation of emotional visual cues (e.g., the late positive potential, LPP). The P100 component indexes early sensory processing within the visual cortex, which is modulated by spatial attention and may reflect a sensory gain control mechanisms to attended stimuli (Luck et al., 2000). Studies on emotion processing have reported enhanced P100 amplitudes for unpleasant pictures and threatening conditions—but also for pleasant stimuli which has been interpreted as an early attentional orientation toward emotional cues (see e.g., Pourtois et al., 2004; Brosch et al., 2008; Bublatzky and Schupp, 2012). Further, as an indicator of early selective stimulus encoding the EPN has been related to stimulus arousal for both pleasant and unpleasant picture materials (Schupp et al., 2004). In addition, the P200 has been considered as an index of affective picture processing (Carretie et al., 2001a, 2004). Enhanced P200 amplitudes in response to unpleasant and pleasant cues suggest that emotional cues mobilize automatic attention resources (Carretie et al., 2004; Delplanque et al., 2004; Olofsson and Polich, 2007). In addition to affective scenes, enhanced P200 amplitudes were also reported for emotional words (e.g., Kanske and Kotz, 2007) and facial expressions (Eimer et al., 2003). Subsequent in the visual processing stream, the LPP over centro-parietal sensors (developing around 300 ms after stimulus onset) is sensitive for emotional intensity (Cuthbert et al., 2000; Schupp et al., 2000; Bradley et al., 2001). Further, the LPP has been associated to working memory and competing tasks indicating the operation of capacity-limited processing (for a review see Schupp et al., 2006). Taken together, affect-modulation of visual ERPs can be identified at both early and later processing stages.

Research on multimodal integration of emotional faces and voices has also reported an early modulation of ERP components (i.e., around 100 ms poststimulus). These effects have been interpreted as evidence for an early influence of one modality on the other (de Gelder et al., 1999; Pourtois et al., 2000; Liu et al., 2012). Comparing unimodal and multimodal presentations of human communication, Stekelenburg and Vroomen (2007) observed an effect of multimodality on the N100 and the P200 component time-locked to the sound onset. They report a decrease in amplitude and latency for the presentation of congruent auditory and visual human stimuli compared to unimodally presented sounds. Likewise, Paulmann et al. (2009) suggested that an advantage of congruent multimodal human communication cues compared to unimodal auditory perception is reflected by a systematic decrease of P200 and N300 components. In a recent study, videos of facial expressions and body language with and without emotionally congruent human sounds were investigated (Jessen and Kotz, 2011). Focusing on auditory processing, the N100 amplitude was strongly reduced in the audiovisual compared to the auditory condition, indicating a significant

impact of visual information on early auditory processing. Further, simultaneously presented congruent emotional face-voice combinations elicited enhanced P200 and P300 amplitudes for emotional relative to neutral audiovisual stimuli, irrespective of valence (Liu et al., 2012). Taken together, these studies support the notion that audiovisual compared to unimodal stimulation is characterized by reduced and speeded processing effort.

Regarding the match or mismatch of emotional information from different sensory channels, differences in ERPs to congruent and incongruent information have been reported. De Gelder et al. (1999) presented angry voices with congruent (angry) or incongruent (sad) faces and observed a mismatch negativity effect (MMN) around 180 ms after stimulus onset for incongruent compared to congruent combinations. Likewise, Pourtois et al. (2000) investigated multimodal integration with congruent and incongruent pairings of emotional facial expression and emotional prosody. They reported delayed auditory processing for the incongruent condition as indexed by a delayed posterior P2b component in response to incongruent compared to congruent face-voice-trials (Pourtois et al., 2002).

Beyond face-voice integration, there are only very few studies, which investigated interactions of emotional picture and sound stimuli. On the one hand, there are some studies which included bodily gestures to investigate multimodal interactions—see above (Stekelenburg and Vroomen, 2007; Jessen and Kotz, 2011; Jessen et al., 2012), on the other side, there are studies investigating interactions between musical and visual stimuli (Baumgartner et al., 2006a,b; Logeswaran and Bhattacharya, 2009; Marin et al., 2012). For instance, music can enhance the emotional experience of emotional pictures (Baumgartner et al., 2006a). Combined (congruent) presentation of pictures and music enhanced peripheral physiological responses and evoked stronger cortical activation (alpha density) in comparison to unimodal presentations. Similarly, presenting congruent or incongruent pairs of complex affective pictures and affective human sounds led to an increased P200 as well as an enhanced LPP in response to congruent compared to incongruent stimulus pairs (Spreckelmeyer et al., 2006). Thus, multimodal simultaneity is not limited to human communication.

Building upon these findings, the present study examines how picture processing is influenced by simultaneously presented complex emotional sounds (e.g., sounds of a car crash, laughing children). We did not aim at optimizing mutual influences by semantic matches of related audiovisual stimulus pairs (such as the picture and the sound of an accident), instead, we wanted to examine the interaction of valence-specific pairs (such as the sight of a child and the sound of a crash). Overall, based on previous findings we expect that emotional information of one modality modulate the EEG components in response to the other modality. Specifically, we expect that the presentation of emotional sounds modulate early as well as later processing stages of visual processing. It is expected that picture processing is generally affected by a concurrent sound compared to pictures only. Furthermore, emotional sounds should differentially modulate visual processing according to their congruence or incongruence to the emotional content of the pictures.

## MATERIALS AND METHODS

### PARTICIPANTS

Participants were recruited from the University of Mannheim as well as via personal inquiry and advertisements in local newspapers. The group consisted of 22 participants<sup>1</sup> (11 female) with a mean age of  $M = 21.32$ ,  $SD = 2.85$ . Participation in the study was voluntary and students received class credits for participation. External participants received a small gift, but no financial reimbursement. The study protocol was approved by the ethics committee of the University of Mannheim.

Exclusion criteria included any severe physical illness as well as current psychiatric or neurological disorder and depression as indicated by a score of 39 or higher on the German version of the Self-Rating Depression Scale [SDS, CIPS (1986)]. Also participants reported normal or corrected-to-normal vision and audition and no use of psychopharmaca. In addition, the following questionnaires were completed: a personal data form, the German version of the SDS ( $M = 31.48$ ,  $SD = 4.05$ ), the German version of the Positive and Negative Affect Schedule (Positive affect:  $M = 30.90$ ,  $SD = 5.66$ , Negative affect:  $M = 11.14$ ,  $SD = 1.11$ , Krohne et al., 1996), as well as the German Version of the State-Trait-Anxiety Inventory (Trait version:  $M = 33.95$ ,  $SD = 6.90$ , State:  $M = 30.62$ ,  $SD = 3.94$ , Laux et al., 1981)<sup>2</sup>.

### STIMULUS MATERIALS

The stimulus material consisted of 20 pleasant, 20 unpleasant, and 20 neutral pictures selected from the International Affective Picture System (Lang et al., 2008) as well as the same amount of pleasant, unpleasant and neutral sounds selected from the International Affective Digitalized Sounds database (Bradley and Lang, 2007)<sup>3</sup>. Stimuli were selected for comparable valence and arousal ratings between pleasant and unpleasant stimuli and between pictures and sounds. Furthermore, different content categories (human, animals, inanimate) were represented in the most balanced way possible between the valence categories as well as between sound and pictures. The original sound stimuli of the IADS were cut to a duration of 2 s and used in this edited version<sup>4</sup> (see also Noulhiane et al., 2007; Mella et al., 2011).

<sup>1</sup>From originally 27 participants,  $N = 5$  were excluded due to technical problems or extensive artifacts.

<sup>2</sup>Between male and female participants there were no differences except for age: male participants,  $M = 22.55$ ,  $SD = 3.50$ , were slightly older than female participants,  $M = 20.09$ ,  $SD = 1.22$ ,  $t(21) = 2.19$ ;  $p = 0.04$ .

<sup>3</sup>Nos. of the selected pictures from the IAPS: *pleasant*: 2071, 2165, 2224, 2344, 2501, 4250, 4599, 4607, 4659, 4681, 8030, 8461, 8540, 1812, 5831, 5551, 5910, 7280, 8170, 8502; *unpleasant*: 3000, 3005.1, 3010, 3053, 3080, 3150, 3170, 3350, 6230, 6350, 6360, 6510, 9250, 9252, 9902, 9921, 6415, 9570, 6300; *neutral*: 2372, 2385, 2512, 2514, 2516, 2595, 2635, 2830, 7493, 7640, 1675, 5395, 5920, 7037, 7043, 7170, 7207, 7211, 7242, 7487. Nos. of selected sounds from the IADS: *pleasant*: 110, 112, 200, 202, 220, 226, 230, 351, 815, 816, 150, 151, 172, 717, 726, 809, 810, 813, 817, 820; *unpleasant*: 241, 242, 255, 260, 276, 277, 278, 284, 285, 286, 290, 292, 296, 105, 422, 501, 600, 703, 711, 713; *neutral*: 246, 262, 361, 368, 705, 720, 722, 723, 113, 152, 171, 322, 358, 373, 374, 376, 382, 698, 700, 701.

<sup>4</sup>The edited sounds were preliminary tested for valence and arousal in a separate pilot study and this unpublished pretest showed that a presentation duration of 2 s is adequate to elicit emotional reactions comparable to the original sounds.

### EXPERIMENTAL PROCEDURE

Upon arrival in the laboratory the location and procedure were introduced and participants read and signed the informed consent form. The electrode cap and electrodes were then attached. Afterwards, participants were seated on a chair approximately 100 cm away from the monitor (resolution:  $1280 \times 960$  pixel) in the separate EEG booth and were asked to fill in the questionnaires. Upon finishing the preparation phase, participants were informed about the procedure and instructed to view the pictures presented on the computer monitor and listen to the sounds presented through headphones (AKG K77). Also they were told to move as little as possible. Practice trials were presented in order to customize participants to the procedure before the main experiment was started. Overall, the experimental part consisted of 60 visual (pictures only) and 180 audiovisual trials<sup>5</sup>. Visual and audiovisual trials were presented in randomized order.

During visual trials, 20 pleasant, 20 neutral, and 20 unpleasant pictures were displayed for 2 s each. After 50% of the trials 9-point-scales of the Self-Assessment-Manikin (Bradley and Lang, 1994) were presented for ratings of valence and arousal. To shorten the experimental procedure, the participants rated only 50% of all stimulus presentations. The selection of the stimuli was counterbalanced across participants so that all stimulus presentations were rated by 50% of the participants. In cases of no rating, an interval of 2000 ms followed.

For the audiovisual condition, sounds were presented for 2 s with pictures being presented 500 ms after sound onset with a total duration of also 2 s resulting in an overall trial length of 2.5 s. Again stimuli had to be rated in 50% of the trials and the task was to rate valence and arousal elicited by the combination of both, picture and sound. The sound and picture onset were asynchronous as the grasp of the emotional meaning of a sound is not as precise and clearly defined with the onset as compared to a picture. To ensure that the emotional meaning of the sound was present when the picture was presented, we decided to present the picture after a delay of 500 ms.

Overall, the audiovisual condition consisted of 180 trials. Every picture condition (pleasant, neutral, and unpleasant) was paired with every sound condition (pleasant, neutral and unpleasant). This results in nine different conditions with 20 trials with pleasant pictures and pleasant sounds, 20 trials with unpleasant pictures and unpleasant sounds (congruent), 20 trials with pleasant pictures paired with unpleasant sounds and 20 trials with unpleasant pictures with pleasant sounds (incongruent). Additionally, pleasant, unpleasant and neutral pictures were paired each with neutral sounds (60 trials) as well as pleasant and unpleasant sounds with neutral pictures (40 trials).

Ratings were completed using the corresponding keyboard button. Overall, the experimental session lasted about 45 min.

### DATA ACQUISITION AND PREPROCESSING

Electrophysiological data were collected with a 64-channel recording system (actiCAP, Brain Products GmbH, Munich) with

<sup>5</sup>Originally, the experimental part also comprised 60 unimodal trials with unpleasant, neutral and pleasant sounds. As the analysis focused on visual ERPs only, these trials were not considered for further analysis.

a sampling rate of 1 kHz. Electrodes were recorded according to the international 10–20-system. FCz served as the reference electrode and AFz as the ground electrode. Scalp impedance was kept below 10 k $\Omega$ . Data was recorded with an EEG-amplifier Brain-Amp-MR Amplifier (Brain Products GmbH, Munich, Germany).

EEG-data were offline re-referenced to an average reference and filtered (Notch filter of 50 Hz; IIR filter: high cut-off 30 Hz; low cut-off 0.1 Hz) using Brainvision Analyzer 2 (by Brain Products GmbH). Ocular correction was conducted via a semi-automatic Independent Component Analysis (ICA)-based correction process. For data reduction stimulus-synchronized segments with a total length of 1600 ms lasting from 100 ms before and 1500 ms after picture onset were extracted. These segments were then passed through an automatic Artifact Rejection algorithm also provided by Brainvision Analyzer 2. Artifacts were defined with the following criteria: a voltage step of more than 50.0  $\mu$ V/ms, a voltage difference of 200  $\mu$ V within the segments, amplitudes of less than  $-100 \mu$ V or more than 100  $\mu$ V and a maximum voltage difference of more than 0.50 V within 100-ms intervals.

Afterwards all remaining segments (97.5%) for each condition, sensor and participant were baseline corrected (100 ms before stimulus onset) and averaged to calculate the ERPs from the spontaneous EEG.

## STATISTICAL ANALYSIS

### Self-report data

The affective ratings for valence and arousal were analyzed by separate repeated measure analyses of variance (ANOVAs).

**Visual vs. audiovisual condition.** Within-subject variables were *Modality* (visual vs. audiovisual trials), and *Stimulus Category* (congruent pleasant vs. congruent unpleasant vs. congruent neutral). In terms of comparableness of the visual and audiovisual trials for valence, we only considered congruent audiovisual trials for this analysis.

**Audiovisual condition.** Separate repeated measures ANOVAs for audiovisual trials only were conducted with the within-subject variables *Sound Category* (pleasant vs. unpleasant vs. neutral) and *Picture Category* (pleasant vs. unpleasant vs. neutral).

**Congruency.** To test specific differences between congruent and incongruent trials separately for pleasant and unpleasant pictures, planned *t*-tests were conducted at *p*-value < 0.05.

In order to correct for violations of sphericity the Greenhouse-Geisser corrected *p*-value was used to test for significance. Separate ANOVAs as well as *post-hoc t*-tests (bonferroni-corrected) were used for follow up analyses.

### Electrophysiological data

As sound stimuli develop their emotional meaning over time and thus, the emotional onset is not clearly defined, ERPs were locked to picture onsets only. Based on visual inspection and previous research, three time windows and sensor areas were identified: for the P100 component, the mean activity in a time window from 90 to 120 ms was averaged over parietal and occipital electrodes (left: P3, O1; right: P4, O2); for the P200, mean activity between

170 and 230 ms was averaged over parietal and central electrodes (left: P3, C3, right P4, C4—see Stekelenburg and Vroomen, 2007) and the LPP was scored at CP1 and CP2 in a time interval ranging from 400 to 600 ms (see Schupp et al., 2000, 2007)<sup>6</sup>.

**Visual vs. audiovisual condition.** To investigate the general influence of the sound presentation on picture processing, mean amplitudes for P100, P200, and LPP were subjected to separate repeated measures analyses of variances (ANOVAs). Within-subject variables were *Modality* (visual vs. audiovisual trials), *Stimulus Category* (congruent pleasant vs. congruent unpleasant vs. congruent neutral), and *Electrode Site*<sup>7</sup>. In terms of comparableness of the visual and audiovisual trials for valence, we only considered congruent audiovisual trials for this analysis.

**Audiovisual condition.** To further examine the influence of the emotional content of the sounds on picture processing and possible interactions of the emotional contents, for the P100, P200, and the LPP separate repeated measures ANOVAs for audiovisual trials only were conducted with the within-subject variables *Sound Category* (pleasant, unpleasant, neutral) and *Picture Category* (pleasant, unpleasant, neutral) and *Electrode Site*.

**Congruency.** To test specific differences between congruent and incongruent trials separately for pleasant and unpleasant pictures, planned *t*-tests were conducted at *p*-value < 0.05.

In order to correct for violations of sphericity the Greenhouse-Geisser corrected *p*-value was used to test for significance (according to Picton et al., 2000). Effects of *Electrode Site* were only considered if they interact with one of the other variables. Separate ANOVAs as well as *post-hoc t*-tests (bonferroni-corrected) were used for follow up analyses.

## RESULTS

### SELF-REPORT DATA

#### Valence

**Visual vs. audiovisual condition.** For the valence ratings a significant main effect of *Stimulus Category*,  $F_{(2, 42)} = 353.61$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.94$ , was observed, as well as a significant interaction of *Modality* and *Stimulus Category*,  $F_{(2, 42)} = 7.01$ ,  $p = 0.003$ ,  $\eta_p^2 = 0.25$ , but no significant main effect of *Modality*. As expected, unpleasant stimuli were rated as more unpleasant than neutral or pleasant stimuli and pleasant stimuli were rated as most pleasant [unpleasant vs. neutral:  $t_{(21)} = 19.91$ ,  $p < 0.01$ ; pleasant vs. neutral  $t_{(21)} = 13.03$ ,  $p < 0.01$ ; pleasant vs. unpleasant:  $t_{(21)} = 20.41$ ,  $p < 0.01$ ]. Following the interaction, audiovisual pairs with pleasant sounds and pictures were rated as more pleasant than pleasant pictures only,  $t_{(21)} = 3.47$ ,  $p < 0.01$ , whereas unpleasant sounds with unpleasant pictures were rated as marginally more unpleasant than unpleasant pictures only,  $t_{(21)} = 1.89$ ,  $p < 0.10$ —see **Table 1**.

<sup>6</sup>No processing differences were observed at PO9/10 within the EPN time window.

<sup>7</sup>For the P100, four individual electrodes were entered into the ANOVA (P3, O1, P4, O2), for the P200 the electrodes P3, C3, P4, and C4 and for the LPP, CP1, and CP2 were entered.

**Audiovisual condition.** Focusing on audiovisual trials only, the ANOVA with the within-subject Factor *Sound Category* and *Picture Category* revealed a significant main effect of *Sound Category*,  $F_{(2, 42)} = 161.45, p < 0.001, \eta_p^2 = 0.89$ , a significant main effect of *Picture Category*,  $F_{(2, 42)} = 270.07, p < 0.001, \eta_p^2 = 0.93$ , as well as a significant interaction of *Sound* and *Picture Category*,  $F_{(4, 84)} = 26.53, p < 0.001, \eta_p^2 = 0.56$ . Overall, audiovisual presentations with unpleasant pictures were rated as more unpleasant than presentations with neutral or pleasant pictures. Presentations with pleasant pictures were rated as most pleasant, for all comparisons  $p < 0.01$ . Similarly, audiovisual presentations with unpleasant sounds were rated as more unpleasant than presentations with neutral or pleasant sounds and presentations with pleasant sounds were rated more pleasant than presentations with other sounds, for all comparisons  $p < 0.01$ .

Following the interaction, audiovisual pairs with pleasant pictures were rated as most pleasant if they were accompanied with

a pleasant sound and most unpleasant if they were paired with an unpleasant sound, for all comparisons  $p < 0.01$ .

Similarly, presentation with neutral pictures were rated as most pleasant if combined with a pleasant and as most unpleasant if they were combined with unpleasant sounds, for all comparisons  $p < 0.01$ . Presentation with unpleasant pictures were also rated as more unpleasant in combination with an unpleasant sound, for all comparisons  $p < 0.01$ , but there was no significant difference between unpleasant pictures with neutral or pleasant sounds,  $t_{(21)} = 0.789; ns$ —see **Figure 1**.

**Congruency.** Comparing the valence ratings of congruent and incongruent audiovisual trials, valence ratings to pleasant pictures with congruent sounds were significantly more pleasant than pleasant pictures with incongruent sounds,  $t_{(21)} = 12.87, p < 0.01$ . Furthermore, valence ratings of unpleasant pictures with congruent sounds were significantly more unpleasant than unpleasant pictures with incongruent sounds,  $t_{(21)} = 7.27, p < 0.01$ .

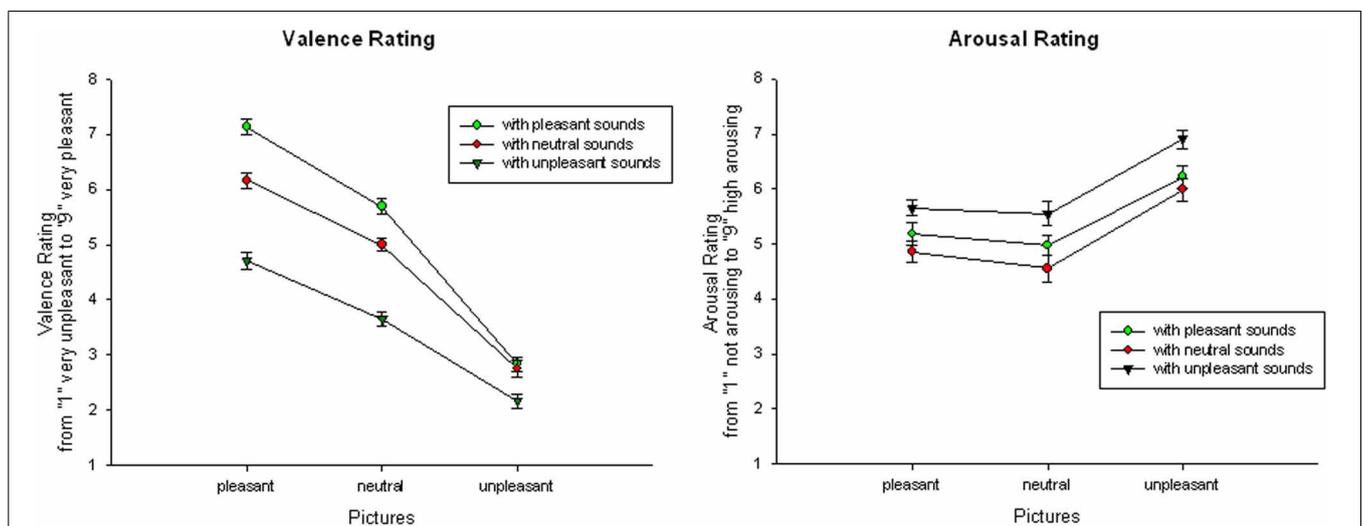
**Arousal**

**Visual vs. audiovisual condition.** For the arousal ratings we found a significant main effect of *Modality*,  $F_{(1, 21)} = 18.87, p < 0.001, \eta_p^2 = 0.47$ , and a significant main effect of *Stimulus Category*,  $F_{(2, 42)} = 47.13, p < 0.001, \eta_p^2 = 0.69$ , but no significant interaction. Overall, audiovisual presentations were rated as more arousing than pictures only,  $t_{(21)} = 4.34, p < 0.01$ . As expected, unpleasant stimuli were rated as more arousing than neutral stimuli [unpleasant vs. neutral:  $t_{(21)} = 10.36, p < 0.01$ ; pleasant vs. neutral:  $t_{(21)} = 2.15, ns$ ]. Furthermore, unpleasant stimuli were significant rated as more arousing than pleasant stimuli,  $t_{(21)} = 6.90, p < 0.01$ —see **Table 1**.

**Audiovisual condition.** For the arousal ratings, a significant main effect of *Picture Category*,  $F_{(2, 42)} = 43.54, p < 0.001, \eta_p^2 = 0.68$ ,

**Table 1 | Mean and standard deviation for valence and arousal ratings of pleasant, neutral and unpleasant visual and congruent audiovisual presentations.**

Rating	Emotion Category	Visual M (SD)	Audiovisual M (SD)
Valence	Pleasant	6.87 (0.76)	7.13 (0.65)
	Neutral	5.06 (0.39)	5.00 (0.54)
	Unpleasant	2.36 (0.63)	2.17 (0.59)
Arousal	Pleasant	4.81 (1.12)	5.18 (0.95)
	Neutral	4.30 (1.16)	4.55 (1.14)
	Unpleasant	6.50 (0.92)	6.90 (0.81)



**FIGURE 1 | Valence (left) and arousal (right) ratings for audiovisual presentations: Mean and SEMs of valence and arousal ratings for pleasant, neutral, and unpleasant pictures in combination with pleasant, neutral, and unpleasant sounds.**

and a significant main effect of *Sound Category*,  $F_{(2, 42)} = 37.06$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.64$ , occurred, but no significant interaction. Overall, stimulus presentations with unpleasant pictures were rated as more arousing than presentations with neutral or pleasant pictures and presentations with pleasant pictures were rated as more arousing than presentations with neutral pictures, for all comparisons  $p < 0.01$ . Similarly, stimulus presentations with unpleasant sounds were rated as more arousing than presentations with neutral or pleasant sounds, for all comparisons  $p < 0.01$ , but presentations with pleasant sounds were not rated as significantly more arousing than presentations with neutral sounds,  $t_{(21)} = 1.39$ ,  $n_s$ —see **Figure 1**.

**Congruency.** Specifically comparing congruent and incongruent stimulus pairs, arousal ratings to pleasant pictures with incongruent sounds were significantly more arousing than with congruent sounds,  $t_{(21)} = 12.46$ ,  $p < 0.01$ . In contrast, arousal ratings to unpleasant pictures with congruent sounds were significantly more arousing than with incongruent sounds,  $t_{(21)} = 8.39$ ,  $p < 0.01$ .

## ELECTROPHYSIOLOGICAL DATA

### P100 component

**Visual vs. audiovisual condition.** For the P100 amplitudes, we found a significant main effect of *Picture Category*,  $F_{(2, 42)} = 3.70$ ,  $p = 0.041$ ,  $\eta_p^2 = 0.15$ , and a significant main effect of *Electrode Site*,  $F_{(3, 63)} = 33.47$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.61$ , but no other significant main effect or interaction. P100 amplitudes in response to pleasant trials were significant higher than in response to unpleasant trials and there was no significant difference between the visual and audiovisual condition—see **Table 2**.

**Audiovisual condition.** For the P100 amplitudes, we found a significant main effect of *Sound Category*,  $F_{(2, 42)} = 4.803$ ,  $p = 0.014$ ,  $\eta_p^2 = 0.19$ , and a significant main effect of *Electrode Site*,  $F_{(3, 63)} =$

$25.06$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.54$ , as well as a significant interaction of *Sound Category* and *Electrode Site*,  $F_{(6, 126)} = 4.04$ ,  $p = .006$ ,  $\eta_p^2 = 0.16$ . No other main effect or interaction was significant.

Following the interaction, P100 amplitudes on parietal electrodes (P3, P4) were enhanced when pictures were accompanied by pleasant sounds [P3:  $F_{(2, 42)} = 4.86$ ,  $p < 0.05$ ,  $\eta_p^2 = 0.19$ ; P4:  $F_{(2, 42)} = 7.27$ ,  $p < 0.01$ ,  $\eta_p^2 = 0.26$ ] compared to pictures with neutral sounds, whereas this effect was not significant on central electrodes. Additionally, P100 amplitudes to pictures with unpleasant sounds compared to neutral sounds were enhanced on P4 [ $t_{(21)} = 3.23$ ,  $p < 0.01$ ]—see **Figure 2**.

**Congruency.** Specifically comparing congruent and incongruent audiovisual pairs, parietal P100 (P4) was enhanced in response to unpleasant pictures with incongruent (pleasant) compared to unpleasant pictures with congruent sounds,  $t_{(21)} = 2.93$ ,  $p < 0.01$ —see **Figure 3**.

### P200 component

**Visual vs. audiovisual condition.** For the P200 amplitudes, we found a significant main effect of *Modality*,  $F_{(1, 21)} = 4.44$ ,  $p = 0.047$ ,  $\eta_p^2 = 0.18$ , a significant main effect of *Stimulus Category*,  $F_{(2, 42)} = 3.80$ ,  $p = 0.034$ ,  $\eta_p^2 = 0.15$ , and a significant main effect of *Electrode Site*,  $F_{(3, 63)} = 69.07$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.77$ , but no other significant main effect or interaction. P200 amplitudes in response to audiovisual trials were significantly enhanced compared to unimodal picture trials,  $t_{(21)} = 2.11$ ,  $p < 0.05$ . Furthermore, independent of *Modality*, unpleasant stimulus presentations elicited stronger P200 amplitudes than neutral presentations,  $t_{(21)} = 2.77$ ,  $p < 0.05$ —see **Table 3**.

**Audiovisual condition.** For the P200 amplitudes, we found a significant main effect of *Sound Category*,  $F_{(2, 42)} = 6.752$ ,  $p = 0.004$ ,  $\eta_p^2 = 0.24$ , a significant main effect of *Electrode Site*,  $F_{(3, 63)} = 57.11$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.73$ , as well as a significant interaction of *Sound Category* and *Electrode Site*,  $F_{(6, 126)} = 11.31$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.35$ . No other main effect or interaction was significant.

Following the interaction, P200 amplitudes on parietal electrodes (P3, P4) were enhanced when pictures were accompanied by emotional sounds [P3:  $F_{(2, 42)} = 15.52$ ,  $p < 0.01$ ,  $\eta_p^2 = 0.43$ ; P4:  $F_{(2, 42)} = 12.36$ ,  $p < 0.01$ ,  $\eta_p^2 = 0.37$ ], whereas this effect was not significant on central electrodes—see **Figure 2**.

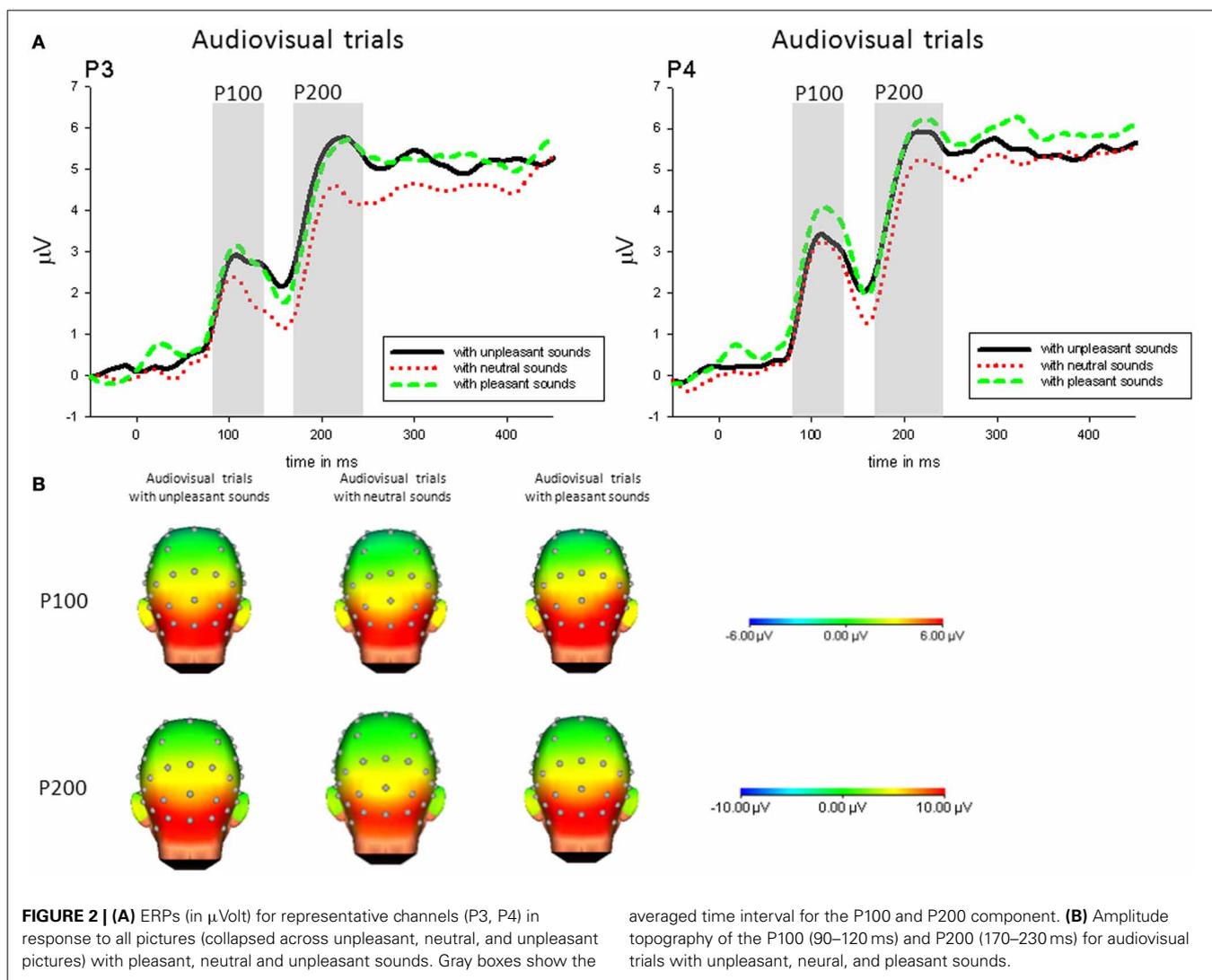
**Congruency.** Specifically comparing congruent and incongruent stimulus pairs, parietal P200 (P4) was enhanced in response to unpleasant pictures with incongruent (pleasant) compared to unpleasant pictures with congruent sounds,  $t_{(21)} = 2.32$ ,  $p < 0.05$ —see **Figure 3**.

### Late positive potential (LPP)

**Visual vs. audiovisual condition.** For the LPP, we found a significant main effect of *Stimulus Category*,  $F_{(2, 42)} = 7.50$ ,  $p = 0.002$ ,  $\eta_p^2 = 0.263$ . No other main effect or interaction was significant. The LPP in response to unpleasant trials was significantly enhanced compared to neutral,  $t_{(21)} = 2.64$ ,  $p < 0.05$ , or pleasant presentations,  $t_{(21)} = 2.95$ ,  $p < 0.05$ —see **Table 4**.

**Table 2 | Mean and standard deviation for the P100 amplitude on parietal (P3,P4) and occipital electrodes (O1,O2) in response to visual and congruent audiovisual presentations.**

P100	Emotion Category	Visual M (SD)	Audiovisual M (SD)
P3	Pleasant	2.04 (1.67)	3.02 (2.43)
	Neutral	2.16 (2.00)	2.16 (2.71)
	Unpleasant	2.05 (1.91)	2.57 (2.84)
P4	Pleasant	2.86 (1.66)	4.12 (2.38)
	Neutral	2.36 (1.61)	2.90 (2.23)
	Unpleasant	2.89 (1.98)	3.24 (2.28)
O1	Pleasant	6.24 (3.38)	7.29 (4.71)
	Neutral	6.23 (4.31)	6.24 (4.56)
	Unpleasant	6.06 (4.60)	5.77 (4.66)
O2	Pleasant	7.25 (3.40)	8.28 (4.92)
	Neutral	6.76 (4.18)	6.89 (4.49)
	Unpleasant	6.94 (4.51)	6.40 (4.38)



**Audiovisual condition.** For audiovisual trials, there was a significant main effect of *Picture Category*,  $F(2, 42) = 13.95$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.399$ . No other main effect or interaction was significant. The LPP in response to trials with unpleasant pictures was significantly enhanced compared to trials with neutral,  $t_{(21)} = 3.99$ ,  $p < 0.01$ , or pleasant pictures,  $t_{(21)} = 3.70$ ,  $p < 0.01$ . Furthermore, in response to presentations containing pleasant pictures compared to neutral pictures an enhanced LPP was found,  $t_{(21)} = 2.91$ ,  $p < 0.05$ —see **Figure 4**.

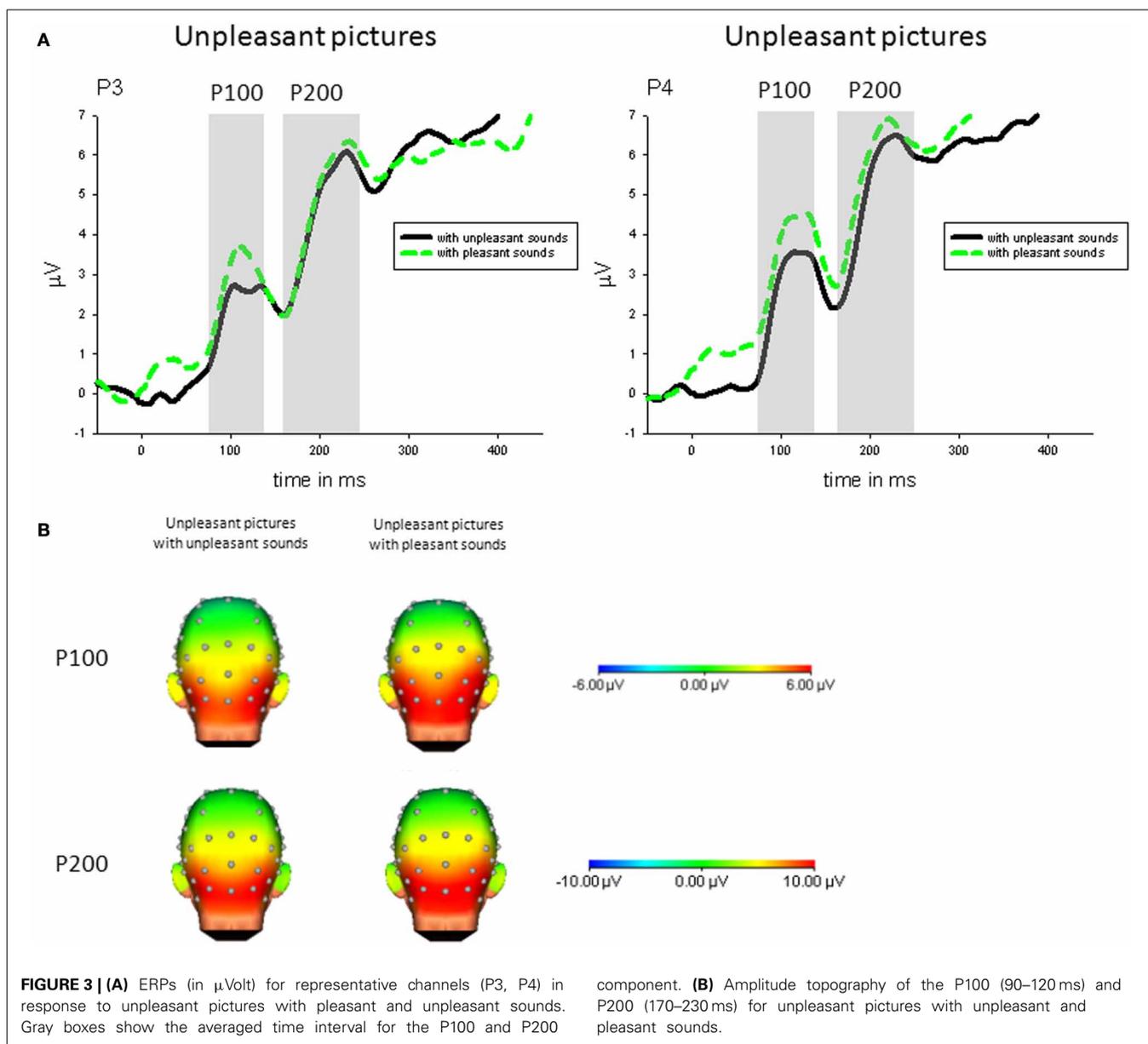
**Congruency.** For the LPP, there was no significant difference between congruent and incongruent trials, all  $ps > 0.19$ .

## DISCUSSION

The present study investigated the impact of concurrent emotional sounds on picture processing. Extending previous research on emotional face-voice pairings, the utilized stimulus material (pictures and sounds) covered a wide range of semantic contents (Bradley and Lang, 2000; Lang et al., 2008). Results showed that high arousing unpleasant compared to neutral pictures were

associated with an increased parietal P200 and a more pronounced centro-parietal LPP regardless of the accompanying sound. For audiovisual stimulation, increased parietal P100 and P200 amplitudes were found in response to all pictures which were accompanied by unpleasant or pleasant sounds compared to pictures with neutral sounds. Most importantly, parietal P100 and P200 were enhanced in response to unpleasant pictures with incongruent (pleasant) compared to congruent sounds. Additionally, subjective ratings clearly showed that both emotional information—sounds and pictures—revealed a significant impact on valence and arousal ratings.

Regarding the neural processing, indicators of selective processing of emotional compared to neutral pictures were replicated. Independent of the accompanying sound, unpleasant compared to neutral pictures were associated with an increased P200 and a more pronounced LPP. These findings are in line with studies reporting that unpleasant stimuli were associated with an enhanced P200 which is thought to originate in the visual association cortex and reflect enhanced attention toward unpleasant picture cues (Carrette et al., 2001a,b, 2004). Similarly, the LPP



was more pronounced in response to unpleasant pictures compared to neutral indicating sustained processing and enhanced perception of high arousing material (Schupp et al., 2000; Brown et al., 2012). Most recent research reported enhanced LPP amplitudes to both, high arousing pleasant and unpleasant stimuli (Cuthbert et al., 2000; Schupp et al., 2000). In the current study, the lack of enhanced LPP amplitudes for pleasant pictures might be explained in terms of emotional intensity. Thus, pleasant pictures (and audiovisual pairs containing pleasant pictures) were rated as less arousing than unpleasant pictures (and audiovisual pairs containing unpleasant pictures).

Comparing visual and audiovisual stimulation, pictures with preceding congruent sounds were associated with enhanced P200 amplitudes regardless of picture and sound valence compared to pictures without sounds. This may be interpreted as an enhanced attentional allocation to the pictures when they were

accompanied by congruent sounds. Similarly, rating data revealed that audiovisual pairs were perceived as more arousing and more emotional intense than visual stimuli alone. Thus, the enhanced P200 might reflect an increased salience of a picture when it is accompanied by a (congruent) sound. Consequently, pictures with sounds seem to receive a higher salience in contrast to pictures without sounds. Generally, the finding of altered P200 amplitude is in line with previous studies on multimodal information (see also Jessen and Kotz, 2011). However, in contrast to the present finding of enhanced P200 for multimodal information, several studies reported reduced P200 amplitudes to multimodal compared to unimodal stimulation in multimodal human communication (Stekelenburg and Vroomen, 2007; Paulmann et al., 2009). This has been interpreted as an indicator of facilitated processing of multimodal redundant information and state that multimodal emotion processing is less effortful than unimodal

**Table 3 | Mean and standard deviation for the P200 amplitude on parietal (P3,P4) and occipital electrodes (O1,O2) in response to visual and congruent audiovisual presentations.**

P200	Emotion Category	Visual <i>M (SD)</i>	Audiovisual <i>M (SD)</i>
P3	Pleasant	3.96 (2.44)	4.52 (1.95)
	Neutral	4.43 (2.84)	3.79 (2.47)
	Unpleasant	4.79 (2.64)	4.81 (2.97)
P4	Pleasant	4.71 (2.65)	5.05 (2.99)
	Neutral	4.54 (2.44)	4.27 (2.77)
	Unpleasant	5.01 (2.90)	5.13 (2.81)
O1	Pleasant	-2.08 (2.30)	-1.49 (2.30)
	Neutral	-1.86 (1.91)	-1.55 (2.02)
	Unpleasant	-1.82 (2.26)	-1.22 (2.00)
O2	Pleasant	-2.26 (2.63)	-1.89 (1.94)
	Neutral	-2.69 (2.73)	-1.12 (2.04)
	Unpleasant	-2.15 (2.41)	-1.51 (1.79)

**Table 4 | Mean and standard deviation for the late positive potential (LPP) on CP1 and CP2 in response to visual and audiovisual presentations separately for pleasant, neutral and unpleasant presentations.**

LPP	Picture Category	Visual <i>M (SD)</i>	Audiovisual <i>M (SD)</i>
CP1	Pleasant	2.39 (1.18)	2.57 (1.30)
	Neutral	1.96 (1.05)	1.99 (0.96)
	Unpleasant	2.78 (1.83)	3.45 (2.15)
CP2	Pleasant	2.36 (1.34)	2.67 (1.33)
	Neutral	1.86 (0.98)	1.90 (0.91)
	Unpleasant	3.00 (1.73)	3.65 (1.99)

processing. However, variant findings may relate to methodological differences regarding the stimulus material (faces and voices vs. more complex stimuli), focus of analyses (auditory or visual evoked potentials) and order and timing of the presentation (simultaneous vs. shifted presentation of sound and pictures). As (congruent) sound and picture stimuli did not transport redundant but additional information in the current study (cf. face-voice pairings), the present findings of generally enhanced responses to multimodal stimuli may rather reflect intensified salience detection than a facilitated processing.

Regarding the specific findings for audiovisual stimulation, an increased parietal P100 and an increased P200 was observed in response to all pictures which were accompanied by unpleasant or pleasant sounds compared to pictures with neutral sounds. The modulation of early visual components as the P100 by emotional sounds may be interpreted as evidence that emotional sounds may unspecifically increase sensory sensitivity or selective attention to consequently improve perceptual processing of all incoming visual stimuli (Mangun, 1995; Hillyard et al., 1998; Kolassa et al., 2006; Brosch et al., 2009). Likewise, the increased P200 amplitude to all pictures which came along with emotional sounds could be interpreted as an unspecific enhancement of

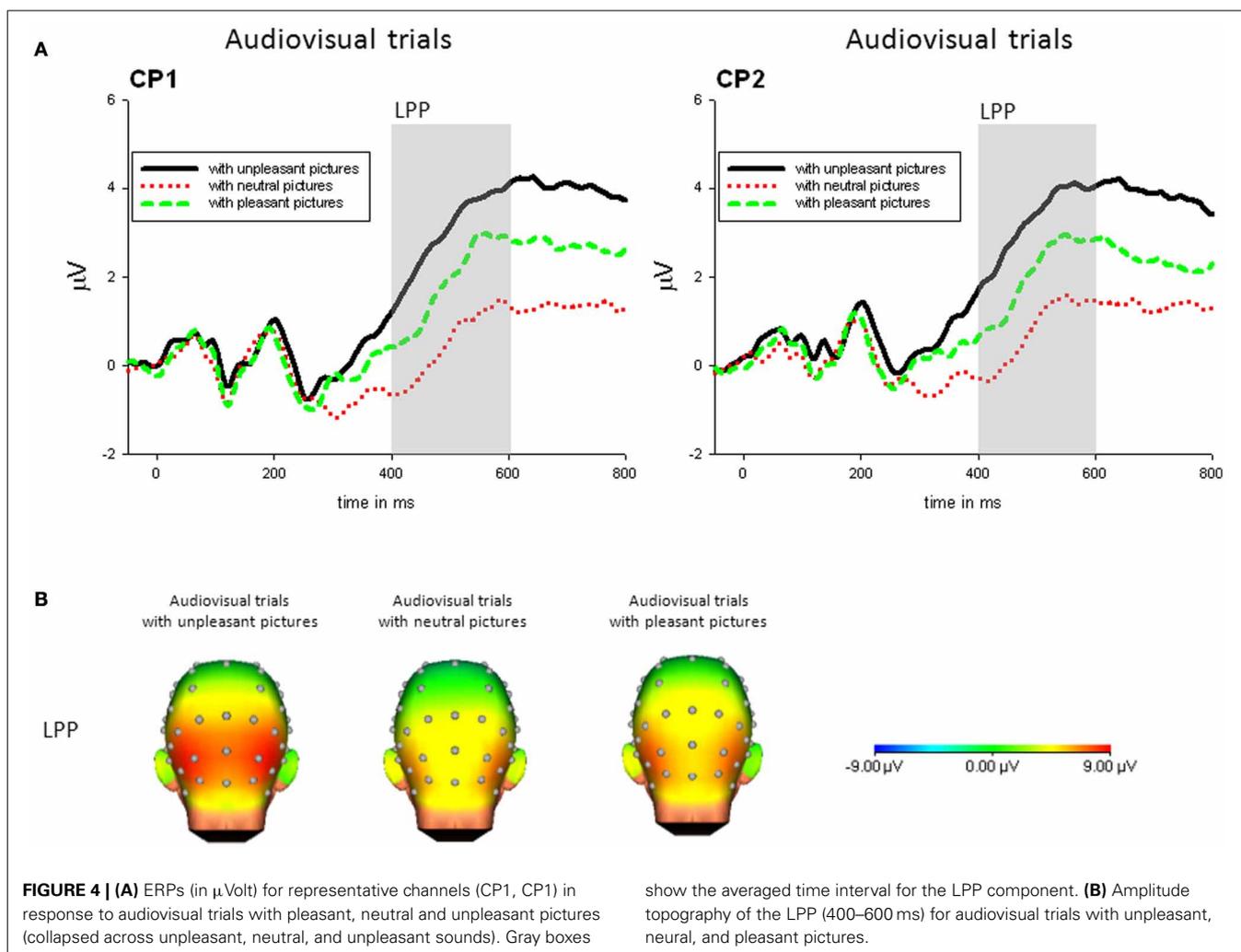
attentional resources toward the visual stimuli if any emotional information was conveyed by the sounds. Both P100 and P200 may reflect an important mechanism to support fast discrimination between relevant and irrelevant information (in all sensory channels) and thus to prepare all senses for following relevant information in order to facilitate rapid and accurate behavioral responses (Öhman et al., 2001, 2000).

Of particular interest, the emotional mismatch of visual and auditory stimuli revealed a pronounced impact on picture processing. Specifically, a reduction of P100 and P200 amplitudes was observed for unpleasant pictures with congruent (unpleasant) compared to incongruent (pleasant) sounds. This finding indicates that unpleasant pictures processing is facilitated when they were preceded by congruent unpleasant sounds. In contrast, the incongruent combination (unpleasant picture and pleasant sounds), may require more attentional resources as indicated by enhanced P100 and P200 responses. This finding is in line with previous research on emotional perceptual integration suggesting facilitated processing for emotional congruent information (de Gelder et al., 1999; Pourtois et al., 2002; Meeren et al., 2005). Regarding the question why an incongruency effect was only found for unpleasant pictures paired with pleasant sounds, we can only speculate that this mismatch is much more behaviorally relevant as the opposite one (pleasant picture with unpleasant sound). The sudden onset of an aversive visual event after pleasant sounds might indicate that immediate change of behavior is needed to avoid potential surprising harm. However, when there is an aversive sound present but then the visual signal provides information which is non-threatening, this is not as arousing and relevant for the organism to change behavior at the onset of the visual event. All the more, this finding also warrants further research on the timing and order of multi-modal affective stimulation.

Subsequent processing stages of the pictures were not modulated by concurrent emotional sounds. Specifically, LPPs to unpleasant picture did not vary as a function of picture-sound congruency in the present study. These findings contrast with a recent study reporting later visual processing modulated by congruent auditory information (Spreckelmeyer et al., 2006). However, future studies will need to integrate crossmodal resource competition (cf. Schupp et al., 2007, 2008).

Regarding the underlying brain structures, our results are in line with functional imaging data suggesting that multisensory interaction takes place in posterior superior temporal cortices (Pourtois et al., 2005; Ethofer et al., 2006a). Furthermore, recent fMRI studies suggested that emotional incongruence is accompanied with higher BOLD-responses (e.g., in a cingulate-frontoparietal network) compared to congruent information (Müller et al., 2011, 2012b). However, further studies reported enhanced neural activation in response to congruent compared to incongruent information (Spreckelmeyer et al., 2006; Klasen et al., 2011; Liu et al., 2012). Thus, future studies are needed to clarify whether congruent information is processed in a facilitated or intensified fashion and which brain regions are significantly involved in these processes.

Complementary findings are provided by verbal report data. Similar to the ERP findings, a congruency effect specifically pronounced for unpleasant picture materials with unpleasant sounds



was revealed for arousal ratings. Specifically, more pronounced arousal was reported for unpleasant pictures with congruent as compared to incongruent sounds. Further, pleasant picture ratings were generally lower in arousal. In addition, valence congruence revealed lower arousal ratings in comparison to pleasant pictures with unpleasant sounds. Accounting for that difference between unpleasant and pleasant pictures, an evolutionary perspective may be of particular relevance. From a survival point of view, the detection of possibly threatening visual information is much more relevant (Öhman and Wiens, 2003) when the auditory domain prompts the anticipation of unpleasant stimulation. Conversely, the violation of anticipated pleasant visual information triggered by unpleasant sounds appears behaviorally less momentous.

### LIMITATIONS

Several limitations of the present study need to be acknowledged. Regarding congruency effects, the present study focused on emotional rather than on semantic mis/match. Accordingly, picture and sound stimuli were not specifically balanced with regards to their semantic content. For example, pictures depicting animals could be accompanied by human or environmental

sounds and vice versa. Consequently, a systematic differentiation between emotional and semantic (in)-congruency cannot be inspected in the present study. Further, as for other studies, the question occurs whether the present findings actually reflect multimodal integration of emotional information (Ethofer et al., 2006b) or rather enhancement effects due to increased (emotional) intensity of audiovisual compared to unimodal stimuli. To elucidate this question in detail, future studies will need to systematically vary emotional intensity during unimodal and multimodal presentations.

Furthermore, it is important to mention that our comparison of visual and audiovisual stimuli is to be seen with caution. In line and to be comparable with several existing studies on multimodal emotion processing (e.g., Pourtois et al., 2000, 2002; Müller et al., 2012a), we defined the baseline to 100 ms preceding the multimodal stimulation (picture onset) which is favorable because (1) it is as close as possible to the relevant time epoch and therefore corrects for relevant potential level shifts and (2) it subtracts audio-evoked brain activity and therefore multimodal effects are less confounded. However, for comparison of multimodal vs. visual only, this baseline definition corrects for a pure double-stimulation effect in the multimodal condition but the different

stimulation during the baseline might lead to incommensurable effects. Future studies could investigate this with adequate experimental paradigms.

## CONCLUSION

The present study support the notion of multimodal impact of emotional sounds on affective picture processing. Early components of visual processing (P100, P200) were modulated by the concurrent presentation of emotional sounds. Further, the congruence of sound and picture materials was important, especially

for unpleasant picture processing. In contrast, later indices of facilitated processing of emotional pictures (LPPs) remained relatively unaffected by the sound stimuli. Taken together, further evidence is provided for early interactions of multimodal emotional information beyond human communication.

## ACKNOWLEDGMENTS

This work was supported by the Research Group “Emotion and Behavior” which is sponsored by the German Research Society (DFG; FOR 605; GE 1913/3-1).

## REFERENCES

- Baumgartner, T., Esslen, M., and Jancke, L. (2006a). From emotion perception to emotion experience: emotions evoked by pictures and classical music. *Int. J. Psychophysiol.* 60, 34–43. doi: 10.1016/j.ijpsycho.2005.04.007
- Baumgartner, T., Lutz, K., Schmidt, C. F., and Jancke, L. (2006b). The emotional power of music: how music enhances the feeling of affective pictures. *Brain Res.* 1075, 151–164. doi: 10.1016/j.brainres.2005.12.065
- Bradley, B. P., and Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology* 37, 204–215. doi: 10.1111/1469-8986.3720204
- Bradley, M. M., Codispoti, M., Cuthbert, B. N., and Lang, P. J. (2001). Emotion and motivation I: defensive and appetitive reactions in picture processing. *Emotion* 1, 276–298. doi: 10.1037/1528-3542.1.3.276
- Bradley, M. M., and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* 25, 49–59. doi: 10.1016/0005-7916(94)90063-9
- Bradley, M. M., and Lang, P. J. (2007). *The International Affective Digitized Sounds (2nd Edn. IADS-2): Affective ratings of sounds and instruction manual*. Technical report B-3. (Gainesville, FL: University of Florida).
- Brosch, T., Grandjean, D., Sander, D., and Scherer, K. R. (2009). Cross-modal emotional attention: emotional voices modulate early stages of visual processing. *J. Cogn. Neurosci.* 21, 1670–1679. doi: 10.1162/jocn.2009.21110
- Brosch, T., Sander, D., Pourtois, G., and Scherer, K. R. (2008). Beyond fear: rapid spatial orienting toward positive emotional stimuli. *Psychol. Sci.* 19, 362–370. doi: 10.1111/j.1467-9280.2008.02094.x
- Brown, S. B., van Steenbergen, H., Band, G. P., de Rover, M., and Nieuwenhuis, S. (2012). Functional significance of the emotion-related late positive potential. *Front. Hum. Neurosci.* 6:33. doi: 10.3389/fnhum.2012.00033
- Bublitzky, F., and Schupp, H. (2012). Pictures cueing threat: brain dynamics in viewing explicitly instructed danger cues. *Soc. Cogn. Affect. Neurosci.* 7, 611–622. doi: 10.1093/scan/nsr032
- Campanella, S., Bruyer, R., Froidbise, S., Rossignol, M., Joassin, F., Kornreich, C., et al. (2010). Is two better than one? A cross-modal oddball paradigm reveals greater sensitivity of the P300 to emotional face-voice associations. *Clin. Neurophysiol.* 121, 1855–1862. doi: 10.1016/j.clinph.2010.04.004
- Carretie, L., Hinojosa, J. A., Martin-Loeches, M., Mercado, F., and Tapia, M. (2004). Automatic attention to emotional stimuli: neural correlates. *Hum. Brain Mapp.* 22, 290–299. doi: 10.1002/hbm.20037
- Carretie, L., Martin-Loeches, M., Hinojosa, J. A., and Mercado, F. (2001a). Emotion and attention interaction studied through event-related potentials. *J. Cogn. Neurosci.* 13, 1109–1128. doi: 10.1162/089892901753294400
- Carretie, L., Mercado, F., Tapia, M., and Hinojosa, J. A. (2001b). Emotion, attention, and the ‘negativity bias’, studied through event-related potentials. *Int. J. Psychophysiol.* 41, 75–85. doi: 10.1016/S0167-8760(00)00195-1
- Collegium Internationale Psychiatricae Scalarum (CIPS), (1986). *Selbstbeurteilungs-Depressions-Skala (SDS) nach Zung*, 3rd Edn. Weinheim: Beltz.
- Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., et al. (2008). Audio-visual integration of emotion expression. *Brain Res.* 1242, 126–135. doi: 10.1016/j.brainres.2008.04.023
- Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., and Lang, P. J. (2000). Brain potentials in affective picture processing: covariation with autonomic arousal and affective report. *Biol. Psychol.* 52, 95–111. doi: 10.1016/S0301-0511(99)00044-7
- de Gelder, B., and Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends Cogn. Sci.* 7, 460–467. doi: 10.1016/j.tics.2003.08.014
- de Gelder, B., Bocker, K. B., Tuomainen, J., Hensen, M., and Vroomen, J. (1999). The combined perception of emotion from voice and face: early interaction revealed by human electric brain responses. *Neurosci. Lett.* 260, 133–136. doi: 10.1016/S0304-3940(98)00963-X
- de Gelder, B., and Vroomen, J. (2000). The perception of emotions by ear and eye. *Cogn. Emot.* 14, 289–311. doi: 10.1080/026999300378824
- Delplanque, S., Lavoie, M. E., Hot, P., Silvert, L., and Sequeira, H. (2004). Modulation of cognitive processing by emotional valence studied through event-related potentials in humans. *Neurosci. Lett.* 356, 1–4. doi: 10.1016/j.neulet.2003.10.014
- Dominguez-Borras, J., Trautmann, S. A., Erhard, P., Fehr, T., Herrmann, M., and Escera, C. (2009). Emotional context enhances auditory novelty processing in superior temporal gyrus. *Cereb. Cortex* 19, 1521–1529. doi: 10.1093/cercor/bhn188
- Eimer, M., Holmes, A., and McGlone, F. P. (2003). The role of spatial attention in the processing of facial expression: an ERP study of rapid brain responses to six basic emotions. *Cogn. Affect. Behav. Neurosci.* 3, 97–110. doi: 10.3758/CABN.3.2.97
- Ethofer, T., Anders, S., Erb, M., Droll, C., Royen, L., Saur, R., et al. (2006a). Impact of voice on emotional judgment of faces: an event-related fMRI study. *Hum. Brain Mapp.* 27, 707–714. doi: 10.1002/hbm.20212
- Ethofer, T., Pourtois, G., and Wildgruber, D. (2006b). Investigating audiovisual integration of emotional signals in the human brain. *Prog. Brain Res.* 156, 345–361. doi: 10.1016/S0079-6123(06)56019-4
- Focker, J., Gondan, M., and Roder, B. (2011). Preattentive processing of audio-visual emotional signals. *Acta Psychol. (Amst.)* 137, 36–47. doi: 10.1016/j.actpsy.2011.02.004
- Hillyard, S. A., Vogel, E. K., and Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 353, 1257–1270. doi: 10.1098/rstb.1998.0281
- Jessen, S., and Kotz, S. A. (2011). The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. *Neuroimage* 58, 665–674. doi: 10.1016/j.neuroimage.2011.06.035
- Jessen, S., Obleser, J., and Kotz, S. A. (2012). How bodies and voices interact in early emotion perception. *PLoS ONE* 7:e36070. doi: 10.1371/journal.pone.0036070
- Kanske, P., and Kotz, S. A. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Res.* 1148, 138–148. doi: 10.1016/j.brainres.2007.02.044
- Klasen, M., Chen, Y.-H., and Mathiak, K. (2012). Multisensory emotions: perception, combination and underlying neural processes. *Rev. Neurosci.* 23, 381–392. doi: 10.1515/revneuro-2012-0040
- Klasen, M., Kenworthy, C. A., Mathiak, K. A., Kircher, T. T., and Mathiak, K. (2011). Supramodal representation of emotions. *J. Neurosci.* 31, 13635–13643. doi: 10.1523/JNEUROSCI.2833-11.2011
- Kolassa, I.-T., Musial, F., Kolassa, S., and Miltner, W. H. (2006). Event-related potentials when identifying or color-naming threatening schematic stimuli in spider phobic and non-phobic individuals. *BMC Psychiatry* 6:38. doi: 10.1186/1471-244X-6-38
- Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., and Wildgruber, D. (2007). Audiovisual integration of emotional signals in voice and face: an event-related fMRI study. *Neuroimage* 37, 1445–1456. doi: 10.1016/j.neuroimage.2007.06.020

- Krohne, H. W., Egloff, B., Kohlmann, C.-W., and Tausch, A. (1996). Untersuchungen mit einer deutschen Version der "Positive and Negative Affect Schedule" (PANAS) [Investigations with a German version of the Positive and Negative Affect Schedule (PANAS)]. *Diagnostica* 42, 139–156.
- Lang, P. J., Bradley, M. M., and Cuthbert, B. (2008). "International affective picture system (IAPS): instruction manual and affective ratings," in *Technical Report A-7*, (Gainesville, FL: University of Florida).
- Laux, L., Glanzmann, P., Schaffner, P., and Spielberger, C. D. (1981). *Das State-Trait-Angstinventar (STAI) [State-Trait Anxiety Inventory (STAI)]*. Weinheim: Beltz Test.
- Liu, T., Pinheiro, A., Zhao, Z., Nestor, P. G., McCarley, R. W., and Niznikiewicz, M. A. (2012). Emotional cues during simultaneous face and voice processing: electrophysiological insights. *PLoS ONE* 7:e31001. doi: 10.1371/journal.pone.0031001
- Logeswaran, N., and Bhattacharya, J. (2009). Crossmodal transfer of emotion by music. *Neurosci. Lett.* 455, 129–133. doi: 10.1016/j.neulet.2009.03.044
- Luck, S. J., Woodman, G. F., and Vogel, E. K. (2000). Event-related potential studies of attention. *Trends Cogn. Sci.* 4, 432–440. doi: 10.1016/S1364-6613(00)01545-X
- Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology* 32, 4–18. doi: 10.1111/j.1469-8986.1995.tb03400.x
- Marin, A., Gingras, B., and Bhattacharya, J. (2012). Crossmodal transfer of arousal, but not pleasantness, from the musical to the visual domain. *Emotion* 12, 618–631. doi: 10.1037/a0025020
- Meeren, H. K. M., van Heijnsbergen, C. C. R. J., and de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proc. Natl. Acad. Sci. U.S.A.* 102, 16518–16523. doi: 10.1073/pnas.0507650102
- Mella, N., Conty, L., and Pouthas, J. (2011). The role of physiological arousal in time perception: psychophysiological evidence from an emotion regulation paradigm. *Brain Cogn.* 75, 182–187. doi: 10.1016/j.bandc.2010.11.012
- Mothes-Lasch, M., Miltner, W. H., and Straube, T. (2012). Processing of angry voices is modulated by visual load. *Neuroimage* 63, 485–490. doi: 10.1016/j.neuroimage.2012.07.005
- Müller, V., Kellermann, T. S., Seligman, S. C., Turetsky, B. I., and Eickhoff, S. B. (2012a). Modulation of affective face processing deficits in schizophrenia by congruent emotional sounds. *Soc. Cogn. Affect. Neurosci.* doi: 10.1093/scan/nss107. [Epub ahead of print].
- Müller, V. I., Cieslik, E. C., Turetsky, B. I., and Eickhoff, S. B. (2012b). Crossmodal interactions in audiovisual emotion processing. *Neuroimage* 60, 553–561. doi: 10.1016/j.neuroimage.2011.12.007
- Müller, V. I., Habel, U., Derntl, B., Schneider, F., Zilles, K., Turetsky, B. I., et al. (2011). Incongruence effects in crossmodal emotional integration. *Neuroimage* 54, 2257–2266. doi: 10.1016/j.neuroimage.2010.10.047
- Noulhiane, M., Mella, N., Samson, S., Ragot, R., and Pouthas, V. (2007). How emotional auditory stimuli modulate time perception. *Emotion* 7, 697–704. doi: 10.1037/1528-3542.7.4.697
- Öhman, A., Flykt, A., and Esteves, F. (2001). Emotion drives attention: detecting the snake in the grass. *J. Exp. Psychol. Gen.* 130, 466–478. doi: 10.1037/0096-3445.130.3.466
- Öhman, A., Flykt, A., and Lundqvist, D. (2000). "Unconscious emotion: evolutionary perspectives, psychophysiological data and neuropsychological mechanisms," in *Cognitive Neuroscience of Emotion*, eds R. D. R. Lane and L. Nadel (New York, NY: Oxford University Press), 296–327.
- Öhman, A., and Wiens, S. (2003). "On the automaticity of autonomic responses in emotion: an evolutionary perspective," in *Handbook of Affective Sciences*, eds R. J. Davidson, K. R. Scherer, and H. H. Goldsmith (New York, NY: Oxford University Press), 256–275.
- Olofsson, J. K., and Polich, J. (2007). Affective visual event-related potentials: arousal, repetition, and time-on-task. [Journal; Peer Reviewed Journal]. *Biol. Psychol.* 75, 101–108. doi: 10.1016/j.biopsycho.2006.12.006
- Paulmann, S., Jessen, S., and Kotz, S. A. (2009). Investigating the multi-modal nature of human communication: insights from ERPs. *J. Psychophysiol.* 23, 63–76. doi: 10.1027/0269-8803.23.2.63
- Paulmann, S., and Pell, M. D. (2011). Is there an advantage for recognizing multi-modal emotional stimuli? *Motiv. Emot.* 35, 192–201. doi: 10.1007/s11031-011-9206-0
- Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R. Jr., et al. (2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology* 37, 127–152. doi: 10.1111/1469-8986.3720127
- Pourtois, G., Debatte, D., Despland, P. A., and de Gelder, B. (2002). Facial expressions modulate the time course of long latency auditory brain potentials. *Cogn. Brain Res.* 14, 99–105. doi: 10.1016/S0926-6410(02)00064-2
- Pourtois, G., de Gelder, B., Bol, A., and Crommelinck, M. (2005). Perception of facial expressions and voices and of their combination in the human brain. *Cortex* 41, 49–59. doi: 10.1016/S0010-9452(08)70177-1
- Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B., and Crommelinck, M. (2000). The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport* 11, 1329–1333. doi: 10.1097/00001756-200004270-00036
- Pourtois, G., Grandjean, D., Sander, D., and Vuilleumier, P. (2004). Electrophysiological correlates of rapid spatial orienting towards fearful faces. *Cereb. Cortex* 14, 619–633. doi: 10.1093/cercor/bhh023
- Rigoulot, S., and Pell, M. D. (2012). Seeing emotion with your ears: emotional prosody implicitly guides visual attention to faces. *PLoS ONE* 7:e30740. doi: 10.1371/journal.pone.0030740
- Schupp, H., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Ito, T., and Lang, P. J. (2000). Affective picture processing: the late positive potential is modulated by motivational relevance. *Psychophysiology* 37, 257–261. doi: 10.1111/1469-8986.3720257
- Schupp, H. T., Flaisch, T., Stockburger, J., and Junghöfer, M. (2006). Emotion and attention: event-related brain potential studies. *Prog. Brain Res.* 156, 31–51. doi: 10.1016/S0079-6123(06)56002-9
- Schupp, H. T., Junghöfer, M., Weike, A. I., and Hamm, A. O. (2004). The selective processing of briefly presented affective pictures: an ERP analysis. *Psychophysiology* 41, 441–449. doi: 10.1111/j.1469-8986.2004.00174.x
- Schupp, H. T., Stockburger, J., Bublatzky, E., Junghöfer, M., Weike, A. I., and Hamm, A. O. (2008). The selective processing of emotional visual stimuli while detecting auditory targets: an ERP analysis. *Brain Res.* 1230, 168–176. doi: 10.1016/j.brainres.2008.07.024
- Schupp, H. T., Stockburger, J., Codispot, M., Junghöfer, M., Weike, A. I., and Hamm, A. O. (2007). Selective visual attention to emotion. *J. Neurosci.* 27, 1082–1089. doi: 10.1523/JNEUROSCI.3223-06.2007
- Spreckelmeyer, K. N., Kutas, M., Urbach, T. P., Altenmüller, E., and Munte, T. F. (2006). Combined perception of emotion in pictures and musical sounds. *Brain Res.* 1070, 160–170. doi: 10.1016/j.brainres.2005.11.075
- Stekelenburg, J. J., and Vroomen, J. (2007). Neural correlates of multi-sensory integration of ecologically valid audiovisual events. *J. Cogn. Neurosci.* 19, 1964–1973. doi: 10.1162/jocn.2007.19.12.1964
- Vroomen, J., Driver, J., and de Gelder, B. (2001). Is cross-modal integration of emotional expressions independent of attentional resources? *Cogn. Affect. Behav. Neurosci.* 1, 382–387. doi: 10.3758/CABN.1.4.382
- Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends Cogn. Sci.* 9, 585–594. doi: 10.1016/j.tics.2005.10.011

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 01 June 2013; accepted: 24 September 2013; published online: 18 October 2013.

Citation: Gerdes ABM, Wieser MJ, Bublatzky F, Kusay A, Plichta MM and Alpers GW (2013) Emotional sounds modulate early neural processing of emotional pictures. *Front. Psychol.* 4:741. doi: 10.3389/fpsyg.2013.00741

This article was submitted to *Emotion Science*, a section of the journal *Frontiers in Psychology*.

Copyright © 2013 Gerdes, Wieser, Bublatzky, Kusay, Plichta and Alpers. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Attention and multisensory integration of emotions in schizophrenia

Mikhail Zvyagintsev<sup>1,2,3\*</sup>, Carmen Parisi<sup>1</sup>, Natalia Chechko<sup>1,3</sup>, Andrey R. Nikolaev<sup>4</sup> and Klaus Mathiak<sup>1,3</sup>

<sup>1</sup> Department of Psychiatry, Psychotherapy and Psychosomatics, Medical School, RWTH Aachen University, Aachen, Germany

<sup>2</sup> IZKF Aachen, RWTH Aachen University, Aachen, Germany

<sup>3</sup> JARA – Translational Brain Medicine, Aachen, Germany

<sup>4</sup> Laboratory for Perceptual Dynamics, University of Leuven, Leuven, Belgium

## Edited by:

Benjamin Kreifelts, University of Tübingen, Germany

## Reviewed by:

Gregor R. Szycik, Hannover Medical School, Germany

Diana Robins, Georgia State University, USA

## \*Correspondence:

Mikhail Zvyagintsev, Department of Psychiatry, Psychotherapy and Psychosomatics, RWTH Aachen University, Pauwelsstr. 30, 52074 Aachen, Germany  
e-mail: mzyvagitsev@ukaachen.de

The impairment of multisensory integration in schizophrenia is often explained by deficits of attentional selection. Emotion perception, however, does not always depend on attention because affective stimuli can capture attention automatically. In our study, we specify the role of attention in the multisensory perception of emotional stimuli in schizophrenia. We evaluated attention by interference between conflicting auditory and visual information in two multisensory paradigms in patients with schizophrenia and healthy participants. In the first paradigm, interference occurred between physical features of the dynamic auditory and visual stimuli. In the second paradigm, interference occurred between the emotional content of the auditory and visual stimuli, namely fearful and sad emotions. In patients with schizophrenia, the interference effect was observed in both paradigms. In contrast, in healthy participants, the interference occurred in the emotional paradigm only. These findings indicate that the information leakage between different modalities in patients with schizophrenia occurs at the perceptual level, which is intact in healthy participants. However, healthy participants can have problems with the separation of fearful and sad emotions similar to those of patients with schizophrenia.

**Keywords:** schizophrenia, attention, multisensory integration, emotions, interference

## INTRODUCTION

Schizophrenia is a severe mental disorder characterized by impairments in a wide spectrum of psychological functions. Eight separable cognitive domains represent essential deficits in schizophrenia: speed of processing, attention/vigilance, working memory, verbal learning and memory, visual learning and memory, reasoning and problem solving, verbal comprehension, and social cognition (Nuechterlein et al., 2004). Among these domains, the deficit of attention stands out because the selection of relevant information is crucial for any perceptual or cognitive function. Correspondingly, the attentional deficit itself may cause impairment in other domains. This view has been held from the beginning of twentieth century, when Eugen Bleuler (1911) proposed that most schizophrenia deficits originate from the fundamental deficit of attention.

Contrary to the well-established role of attention in the perception of physical stimuli (Posner and Peterson, 1990), findings on the effect of attention on perception of emotions are inconsistent. Some works showed that the perception of emotions is automatic (Vuilleumier et al., 2001; Pessoa and Ungerleider, 2004; Vuilleumier, 2005), whereas other studies have demonstrated that attention contributes to the selection of emotional stimuli (Pessoa et al., 2002, 2005; Erthal et al., 2005; Mitchell et al., 2007; Lim and Pessoa, 2008). Emotions are severely affected in schizophrenia. Unpredictable or inappropriate emotional responses and anhedonia are typical clinical features of the disease. In fact, all

aspects of emotional behavior, such as emotion expression, experience, and recognition, are impaired in schizophrenia (Trémeau, 2006). Specifying the effect of attention on emotional behavior in patients with schizophrenia is important for the development of diagnostics and treatment.

The aim of our study is to specify the role of attention in the perception of emotions in schizophrenia. Because of the limited processing resources, the attentional selection of relevant information is crucial for perception. The role of attention in perception can be estimated by measuring a conflict that occurs between incongruent features of the same stimulus. Consequently, the most common approach to test selective attention is based on the interference effect. For example, in the classical Stroop task (Stroop, 1935), participants are presented with words written with inks of different colors and their task is either to read a word ignoring a color or to name a color ignoring a word. When the name of the color corresponds to the ink color, the participants' responses are facilitated. The Stroop interference has been widely used as a measure of attention for studying perceptual encoding, processing, and decision in healthy people (reviewed in MacLeod, 1991). Patients with schizophrenia show significantly stronger Stroop interference compared to healthy participants, suggesting the presence of attentional deficit (Perlstein et al., 1998; Barch et al., 1999, see Henik and Salo, 2004 for review).

In our study, we were interested in cases in which emotional stimuli come from different sensory modalities. Impairment of

multisensory integration is a well-known problem of schizophrenia (Ross et al., 2007; Szycik et al., 2009; Seubert et al., 2010; Williams et al., 2010), which can be explained by an attention deficit (de Jong et al., 2010). Early studies in healthy people did not find any effect of attention on multisensory integration (reviewed in De Gelder and Bertelson, 2003), but recent works indicate the specific role of attention in multisensory perception (Talsma et al., 2010; Zvyagintsev et al., 2011; Roudaia et al., 2013). Consequently, we will focus on the role of attention in the multisensory perception of emotional stimuli in patients with schizophrenia.

To study attention in multisensory integration, the Stroop task can be modified so that the interfering stimuli arrive from different modalities (Zvyagintsev et al., 2009; Klasen et al., 2011). If the stimuli from different modalities interfere in their emotional content, the interference effect can be used as a tool for studying the multisensory integration of emotional information. For example, de Gelder et al. (2005) investigated the categorization of happy and sad facial expressions presented together with voices with happy and sad prosody. The interference effect from *auditory* incongruent stimuli was weaker in patients with schizophrenia than in healthy participants. The authors explained this finding by impaired cross-modal integration of the emotional stimuli in schizophrenia. In the second experiment, the authors investigated the categorization of the voices with happy and sad prosody in the presence of emotionally congruent and incongruent faces. It was found that the interference effect from *visual* incongruent stimuli was higher in patients with schizophrenia than in healthy participants. To explain the inconsistency between the first and second experiments, the authors assumed that cross-modal interference in schizophrenia depends on the target modality: hypo-integration of the auditory incongruent stimuli and hyper-integration of the visual incongruent stimuli may occur because of general visual dominance in audiovisual perception. However, in another work, the same authors found weaker interference from the *visual* incongruent stimuli in patients with schizophrenia than in healthy participants in the categorization of voices with fear and happy prosody (de Jong et al., 2009).

The inconsistency of these results conceals the factors that contribute to impairment of multisensory integration in schizophrenia, particularly the role of attention. In the present study, we used the multimodal interference effect as a measure of attention in the perception of emotional stimuli coming from different modalities. Similar to de Gelder et al. (2005), we examined the emotional interference of auditory and visual stimuli in patients with schizophrenia. However, instead of happy and sad emotions, we used fearful and sad emotions. We chose these emotions because the categorization of the fearful facial expression suffers most in patients with schizophrenia compared with healthy participants (Kohler et al., 2003; Johnston et al., 2006; Schneider et al., 2006; Habel et al., 2010, see also Morris et al., 2009 for a review). In addition, the categorization of sad faces is also impaired in schizophrenia (Johnston et al., 2006; Habel et al., 2010), and the misattribution of sad and fearful faces is the highest among other facial expressions for healthy participants (Johnston et al., 2006; Habel et al., 2010). These observations suggest that fearful and sad emotions may be more likely to be confused than happy

and sad emotions, making the task demanding even for healthy participants.

We used two multisensory paradigms in which congruency of the auditory and visual information was manipulated when participants were categorizing the visual stimuli. In one paradigm, interference occurred in the spatiotemporal properties of the dynamic auditory and visual stimuli. In another paradigm, interference occurred in the emotional content of the auditory and visual stimuli. Here, participants were asked to categorize sad and fearful facial expressions while listening to the pseudowords with sad and fearful prosody. We assessed the interference effects in two paradigms for patients with schizophrenia compared to healthy participants. We hypothesized that because of attentional deficit in schizophrenia, the interference effects in patients should occur in both paradigms. However, healthy participants may be able to avoid the information leakage between modalities and accurately select the target stimuli.

## METHODS

### PARTICIPANTS

Twenty patients with schizophrenia and twenty healthy participants took part in the study. Healthy participants were recruited via public advertisement. They had normal or corrected to normal vision, normal hearing, and no history of neurological comorbidity, psychiatric illness, and psychopharmacological therapy. Patients were recruited among inpatients in the Clinic for Psychiatry, Psychotherapy and Psychosomatics, RWTH Aachen University Hospital, Germany. The diagnosis of schizophrenia was made by the treating physician according to the ICD-10. Symptoms were assessed with the Positive and Negative Symptom Scale (PANSS) by an experienced neuropsychologist. All patients received the second-generation antipsychotic medication with  $66 \pm 28\%$  of the daily defined maximal dose (DDD, WHO Collaborating Centre for Drug Statistics Methodology, 2012). Five patients were additionally taking anti-depressive medication ( $DDD = 39 \pm 37\%$ ). The groups were matched for gender, age and parental education. The sociodemographic and illness-related characteristics of participants are listed in the **Table 1**. Participants of both groups received 10 € for their participation. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of the RWTH Aachen University. Written informed consent was obtained from all participants following a complete description of the study and all experimental procedures.

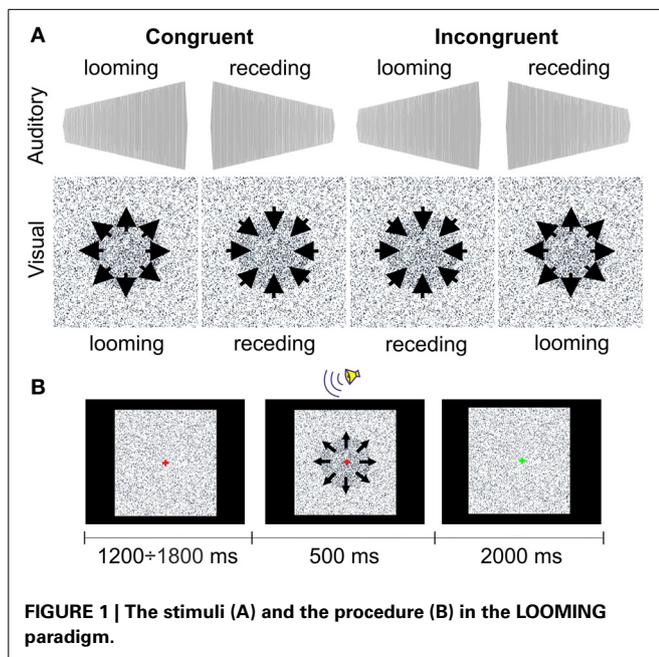
### STIMULI FOR THE LOOMING PARADIGM

We used two paradigms that will be referred to as LOOMING and FACE.

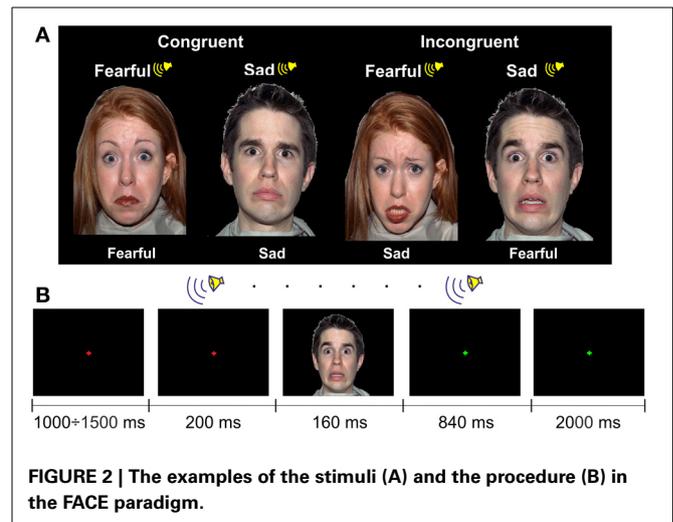
The LOOMING paradigm included two audiovisually congruent and two incongruent conditions (**Figure 1A**). The looming and receding sounds were the 500-ms sine waves that linearly rose or fell in intensity with initial and terminal intensities of 42(57) and 57(42) dB, respectively. A recent study performed by Bach et al. (2011) showed that at these intensities, patients with schizophrenia do not differ from healthy participants in accuracy of the sound direction detection (e.g., receding or looming). Both sounds had an initial falling/raising time of

**Table 1 | Sociodemographic characteristics of the groups and the illness-related data for the patients.**

	Patients, <i>n</i> = 15 (mean ± SD)	Controls, <i>n</i> = 18 (mean ± SD)	Two-sample <i>t</i> -test, <i>df</i> = 32 ( <i>p</i> )
Age, years	42.9 ± 7.7	42.5 ± 7.1	0.1 (0.9)
Gender	13 m, 7 f	13 m, 7 f	0 (1)
Education, years	12.0 ± 1.4	11.4 ± 1.8	1.3 (0.2)
Parental Education, years	9.9 ± 1.7	10.0 ± 1.7	0.1 (0.9)
Disease duration, years	15.7 ± 8.7	—	—
PANSS positive, score	11.4 ± 3.9	—	—
PANSS negative, score	13.1 ± 4.0	—	—
PANSS general, score	26.2 ± 5.6	—	—
PANSS total, score	50.5 ± 12.0	—	—



10 ms. The auditory stimuli were prepared using Csounds 5.09 Software ([www.csounds.com](http://www.csounds.com)). They were delivered binaurally via Sennheiser HD 600 headphones (Sennheiser Electronics Corp., CT). The visual stimuli were prepared using Presentation 7.0 Software (Neurobehavioral Systems, Inc., Albany, California; [www.neurobs.com](http://www.neurobs.com)). Therefore, we created a matrix of 200 × 200 quadrants on the laptop screen. Each quadrant represented a 4 × 4 matrix of the display's pixels using a standard resolution of 1280 × 1024 pixels. The color of each quadrant was randomly assigned to the standard RGB gray scale values in the range from 0 (white) to 255 (black). During the entire session, the quadrants



randomly changed their color intensity within this range with an interval of 16.7 ms. This was perceived as a twinkling of the background. The purpose of this twinkling was to make detection of the circle size changes (see below) more difficult. In the center of the matrix, we created a circle by increasing the color intensity of the quadrants, forming a filled circular shape by 10%. This circle either increased or decreased in size from 8 to 17° (looming condition) or from 17 to 8° (receding condition). The onsets of the circle appearance and the sound stimulus were synchronized with Presentation 7.0 software.

#### STIMULI FOR THE FACE PARADIGM

The FACE paradigm included two audiovisual congruent and two audiovisual incongruent conditions (**Figure 2A**). However, here the congruency was related to the emotional content of the stimuli. For the visual stimuli, we selected 60 faces of 13 male and 13 female actors presenting an equal number of fearful and sad facial expressions from the NimStim Face Stimulus Set (Tottenham et al., 2009). The NimStim stimuli included faces with open and closed mouths. We chose an equal number of open- and closed-mouth faces and counterbalanced them between the conditions.

One face subtended ~12 × 16° of a visual angle (width × height) with a viewing distance of ~60 cm. The faces were presented on a black background.

Auditory stimuli were emotional pseudowords from the in-house dataset. They consisted of 84 sound files that contained 7 pseudowords pronounced by 3 male and 3 female actors with fear and sad prosodies (see examples of the stimuli in Supplementary Material).

The visual and auditory stimuli were presented on a laptop using Presentation 7.0 Software. The auditory stimuli were delivered binaurally via Sennheiser HD 600 headphones at 55 dB SPL.

#### EXPERIMENTAL PROCEDURE

To evaluate attentional impairment, before the experiment, all participants underwent neuropsychological testing, which

included the *Trail Making Test A* and *B* (TMT-A, TMT-B), the vocabulary test [*Wortschatztest* (WST)], and the *forward* and *backward digit span* tests [WST and digit span were taken from the German version of the *Wechsler Adult Intelligence Scale* (WAIS-R), (Tewes, 1991); **Table 2**]. The TMT-A, -B aimed to test attention and processing speed, WST aimed to test verbal intelligence, and the digit span test aimed to test attention and working memory (Mesholam-Gately et al., 2009). In addition, symptoms of patients with schizophrenia were assessed with the positive and negative syndrome scale (PANSS; **Table 1**).

In the LOOMING paradigm, 60 congruent trials with either audiovisual looming or receding stimuli and 60 incongruent trials were presented to each participant. In a looming congruent trial, the auditory stimulus increased in loudness and the visual stimulus increased in size. In a receding congruent trial, the auditory stimulus decreased in loudness and the visual stimulus decreased in size. In incongruent trials, the directions of the changes for the auditory and visual stimuli were opposite. Altogether, 120 trials were presented in random order to each participant.

Participants were instructed to look at the fixation cross of  $0.1^\circ$  of the visual angle centered at the screen. A trial started with a change of the color of the cross from green to red. Then, after a delay (which varied randomly between 1200 and 1800 ms), an audiovisual stimulus was presented for 500 ms (**Figure 1B**). After offset of the audiovisual stimulus, a green fixation cross was shown for 2000 ms. The task for the participants was to recognize the type of the visual stimulus. After the onset of the audiovisual stimulus, the participants had to press one of two buttons, which indicated the looming or receding type of the *visual* stimulus. Participants were instructed to answer as precisely and as quickly as possible. The response time was limited to 2000 ms after the offset of the audiovisual stimulus. The response buttons were counterbalanced between participants. The inter-trial interval varied randomly between 3700 to 4300 ms. The experiment lasted  $\sim 8$  min.

In the FACE paradigm, 60 congruent trials with either sad or fearful emotional content and 60 incongruent trials were presented to each participant. In the congruent trials, the emotional content of the auditory and visual stimuli was matched, and in the incongruent trials, the emotional content was different. The

face and the voice were always matched in gender, but a matched auditory stimulus was chosen randomly for a particular visual stimulus. The sequence of 120 trials was randomized for each participant.

Participants were instructed to look at the fixation cross of  $0.1^\circ$  of the visual angle centered at the screen. A trial started with a change of the color of the cross from green to red. Then, after a delay (which varied randomly between 1000 and 1500 ms), a pseudoword with a duration of 1000 ms was presented. In 200 ms after the pseudoword onset, a face was presented for 160 ms (**Figure 2**). After the offset of the visual stimulus, a green fixation cross was presented for 2840 ms, indicating to participants the response interval. The task was to categorize the facial expression: the participants had to press one of two buttons, which indicated sad or fearful emotion. Participants were instructed to answer as precisely and as quickly as possible. The response time was limited to 3000 ms after the onset of the visual stimulus, i.e., until the fixation cross changed the color. The type of the response button was counterbalanced between participants. The inter-trial interval varied randomly between 4200 to 4700 ms. The experiment lasted  $\sim 9$  min.

The order of application of the LOOMING and FACE paradigms was counterbalanced between participants.

## STATISTICAL ANALYSIS

For each participant and each paradigm, we considered the following variables:

1. the percentage of the trials with responses, regardless of condition and response correctness, as a ratio of the trials with responses to all trials;
2. the percentage of the trials with correct responses, regardless of condition, as a ratio of correct responses to all responses;
3. the average response time for all trials with responses, regardless of correctness of the response;
4. the percentage of the trials with correct responses for each condition as a ratio of correct responses to all responses for this condition (further referred as an accuracy rate).

To verify that both groups of participants attended to each paradigm and followed the instructions, we compared the percentage of the trials with responses between groups using a two-sample *t*-test for each paradigm separately.

We then tested the overall correctness of responses between groups by comparing the percentage of the trials with correct responses regardless of condition using a two-sample *t*-test for each paradigm separately.

Next, we compared the average response time for all trials with a response (regardless of its correctness) between the groups using a two-sample *t*-test for each paradigm separately.

The accuracy rates were averaged across stimulus conditions for congruent and incongruent trials in each paradigm. The accuracy rates were submitted to a repeated-measures ANOVA for each paradigm separately. In ANOVA, we used the within-factor of Congruency (congruent vs. incongruent) and the between-factor of Group (patients with schizophrenia vs. healthy

**Table 2 | Neuropsychological assessment and comparison of healthy participants and patients with schizophrenia.**

Test	Patients (mean $\pm$ SD)	Controls (mean $\pm$ SD)	Two-sample <i>t</i> -test, <i>df</i> = 32
TMTA, seconds	31.6 $\pm$ 11.5	23.1 $\pm$ 8.3	<b>2.7 (0.01)*</b>
TMTB, seconds	79.4 $\pm$ 38.1	50.0 $\pm$ 13.2	<b>3.3 (0.01)*</b>
WST, items	29.2 $\pm$ 6.5	34.1 $\pm$ 3.6	<b>2.9 (0.01)*</b>
Digit span forward, items	7.5 $\pm$ 2.2	8.4 $\pm$ 2.5	1.2
Digit span backward, items	5.8 $\pm$ 2.1	7.0 $\pm$ 2.3	0.7

\**p* < 0.05.

participants). Whenever ANOVA revealed an interaction, we proceeded with the *post-hoc* LSD test.

The statistical analysis was performed with STATISTICA 10.0 software (StatSoft, Inc., Tulsa, OK).

## RESULTS

One patient did not finish the study, and two patients and two healthy participants responded in less than 70% of trials in one of the paradigms. Two patients responded at the chance level in both congruent conditions (53 and 57% of correct answers) in one of the paradigms. These participants were excluded from further analyses. The remaining participants in both groups were the same for both paradigms.

The results of neuropsychological assessment and comparison between the groups are reported in the **Table 2**. Patients were significantly slower than healthy participants in the TMT-A and TMT-B and made significantly more errors in the WST. Although the digit span tests revealed the lower performance in the schizophrenia group than in controls, the differences between the groups in these tests were insignificant. The results of neuropsychological testing suggest that in our study, patients with schizophrenia had a moderate impairment of attention and a lower verbal IQ level than healthy participants.

We did not observe any difference between groups in the total number of responses in both paradigms: LOOMING [ $t_{(31)} = 0.9, p = 0.4$ ] and FACE [ $t_{(31)} = 0.7, p = 0.5$ ; **Table 3**]. This suggests that participants of both groups followed the instructions and were attentive. Further, we did not find any difference between groups in the average response time for the trials with responses: LOOMING [ $t_{(31)} = 0.6, p = 0.7$ ] and FACE [ $t_{(31)} = 0.8, p = 0.4$ ].

However, the difference between groups was observed in the accuracy rate in both paradigms (**Table 3**). Therefore, we submitted this measure to ANOVAs for each paradigm separately.

In the LOOMING paradigm, the ANOVA revealed a significant effect of Congruency  $F_{(1, 31)} = 6.3, p = 0.01$  and a Congruency  $\times$  Group interaction:  $F_{(1, 31)} = 4.7, p = 0.03$ , but no effect of Group [ $F_{(1, 31)} = 3.5, p = 0.08$ ]. The *post-hoc* test showed that the effect of Congruency was significant only in the patients'

group ( $p < 0.01$ ), but not in the healthy participants' group ( $p = 0.8$ ).

In the FACE paradigm, the ANOVA revealed a significant effect of Congruency:  $F_{(1, 31)} = 24.3, p < 0.001$  and Group:  $F_{(1, 31)} = 23.2, p < 0.001$  and no Congruency  $\times$  Group interaction [ $F_{(1, 31)} = 3.1, p = 0.1$ ] (**Figure 3**).

## DISCUSSION

The present study examined the effect of attention on the categorization of visual stimuli in two multisensory paradigms in patients with schizophrenia and healthy controls. Attention was measured as an interference effect between conflicting auditory and visual information. In the first (LOOMING) paradigm, the interference occurred between *dynamic physical features* of the auditory and visual stimuli, i.e., it occurred at the perceptual level. In the second (FACE) paradigm, the interference occurred between the *emotional content* of the auditory and visual stimuli.

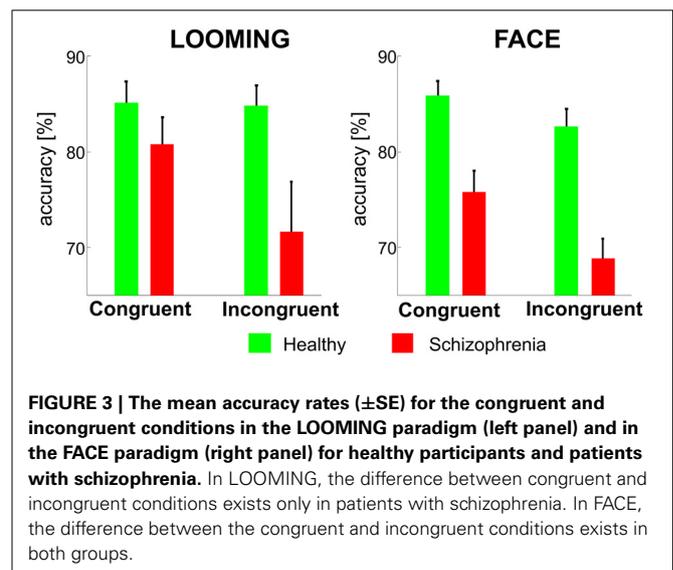
In the perceptual (LOOMING) paradigm, the interference was observed in patients with schizophrenia, but not in healthy participants. In other words, the incongruent auditory input had an impact on visual stimuli categorization only in patients. Because interference in this paradigm occurred between physical features of the auditory and visual stimuli, this finding indicates a leakage of sensory information between the auditory and visual modalities. This suggests insufficiency of the attentional mechanism, which is responsible for the separation of relevant and irrelevant information flow in schizophrenia.

In the emotional (FACE) paradigm, the interference of auditory and visual stimuli occurred in both groups. Because interference in this paradigm occurred between the emotional contents of the audiovisual stimuli, this finding indicates facilitated the fusion of emotional information coming from the auditory and visual modalities, supporting the view that emotion perception is automatic (Vuilleumier et al., 2001; Pessoa and Ungerleider, 2004; Vuilleumier, 2005).

**Table 3 | Overall performance of participants in two paradigms.**

Paradigm	Parameter	Patients (mean $\pm$ SE)	Controls (mean $\pm$ SE)	Two-sample t-test, df = 32
LOOMING	Total, %	94.2 $\pm$ 1.6	96.4 $\pm$ 1.2	0.9
	Correct, %	76.2 $\pm$ 3.8	85.0 $\pm$ 2.0	<b>2.1 (0.04)*</b>
	Response time, ms	1208 $\pm$ 36	1235 $\pm$ 40	0.6
FACE	Total, %	99.3 $\pm$ 0.3	99.7 $\pm$ 0.1	0.7
	Correct, %	72.4 $\pm$ 1.9	84.8 $\pm$ 1.5	<b>5.1 (0.001)**</b>
	Response time, ms	1250 $\pm$ 56	1280 $\pm$ 44	0.8

\* $p < 0.05$ , \*\* $p < 0.01$ .



The results obtained in two paradigms cannot be compared directly because of several differences between them, e.g., different durations of the stimuli, response time windows and inter-trial intervals; different synchrony of the stimulus onsets; and different physical properties of the stimuli. However, we can compare the paradigms qualitatively and consider their consistency with the previous findings.

The results from the perceptual paradigm corroborate the previous studies based on the Stroop task, which showed higher interference between task-relevant and irrelevant information in patients with schizophrenia compared to healthy controls (Perlstein et al., 1998; Barch et al., 1999; Boucart et al., 1999, see Henik and Salo, 2004 for review). Our study extends this observation to multisensory perception: patients with schizophrenia have difficulties in concentrating on the task-relevant information, which comes not only from the same modality but also from different modalities.

In the emotional paradigm, the interference effects were similar for the healthy and schizophrenia groups. This observation is inconsistent with the view on the general impairment of the categorization of emotions in schizophrenia (Kohler et al., 2003; Johnston et al., 2006; Habel et al., 2010). It is also distinct from the observation that interference between auditory and visual emotional information in face categorization is lower in patients with schizophrenia compared to healthy participants (de Gelder et al., 2005). The latter finding was interpreted as evidence for the impairment of multisensory integration in schizophrenia. Our result can be explained by the types of emotional stimuli used. de Gelder et al. (2005) used happy and sad emotions, whereas we used fearful and sad emotions. Indeed, the common finding in schizophrenia studies was the impaired categorization of fearful facial expressions (Kohler et al., 2003; Johnston et al., 2006) and the mis-categorization of sad and fearful faces (Johnston et al., 2006). The significant effect of Group observed in our study indicates that patients had difficulties in the categorization of emotional faces even in the congruent condition, whereas de Gelder et al. (2005) did not observe differences in the categorization of the congruent stimuli between patients and healthy controls. Taken together, these observations suggest that the categorization task in our study was more difficult than in the study by de Gelder et al. (2005). This is a possible consequence of the larger similarity between fearful and sad emotions compared to happy and sad emotions. The similar emotional information can be easily confused because it competes for the same brain resources of

recognition and categorization. Future research should include a wider range of tested emotions, perhaps including mixtures of emotion, to determine a difference in emotion confusion between patients and controls.

A possible limitation of our experimental design is the absence of the unimodal stimuli as a control condition. Although the multisensory interference is repeatedly reported (Talsma et al., 2010), without such a control, we cannot completely exclude the possibility that the patients had impairments of multisensory binding. Moreover, despite previous research suggesting that the second-generation antipsychotics have little influence on attention or even improve it in the long term (Bilder et al., 2002), we cannot completely exclude the effect of medication on the results of our experiment.

In sum, our results indicate that the deficit of attention in schizophrenia results in a mixture of the multimodal stimuli, which can be separated by healthy participants. This indicates that the fusion of task-relevant and irrelevant information that comes via the auditory and visual channels occurs at a relatively low perceptual level. This finding demonstrates how the fundamental deficit of attention in schizophrenia (Bleuler, 1911; Nuechterlein et al., 2004) may affect multisensory integration.

## AUTHOR CONTRIBUTIONS

Mikhail Zvyagintsev, Carmen Parisi, and Klaus Mathiak designed the study; Mikhail Zvyagintsev and Carmen Parisi prepared the protocol for the study; Carmen Parisi performed the data collection; Mikhail Zvyagintsev, Andrey R. Nikolaev and Carmen Parisi analyzed the data; Mikhail Zvyagintsev and Andrey R. Nikolaev wrote the manuscript; all authors contributed to and approved the final manuscript.

## ACKNOWLEDGMENTS

This research was supported by the START AG (121/11 and 143/13) and DFG MA 2631/4-1 (IRTG 1328, JARA-BRAIN). The authors thank Björn Kutzner (RWTH Aachen University) and the Brain Imaging Facility of IZKF Aachen (RWTH Aachen University) for technical support.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fnhum.2013.00674/abstract>

## REFERENCES

- Bach, D. R., Buxtorf, K., Strik, W. K., Neuhoff, J. G., and Seifritz, E. (2011). Evidence for impaired sound intensity processing in schizophrenia. *Schizophr. Bull.* 37, 426–431. doi: 10.1093/schbul/sbp092
- Barch, D. M., Carter, C. S., Hachten, P. C., Usher, M., and Cohen, J. D. (1999). The “benefits” of distractibility: mechanisms underlying increased Stroop effects in schizophrenia. *Schizophr. Bull.* 25, 749–762. doi: 10.1093/oxfordjournals.schbul.a033416
- Bleuler, E. (1911). *Dementia Praecox or the Group of Schizophrenias*. Translated by Zinkin J. New York, NY: International Universities Press, (1950).
- Bilder, R. M., Goldman, R. S., Volavka, J., Czobor, P., Hoptman, M., Sheitman, B., et al. (2002). Neurocognitive effects of clozapine, olanzapine, risperidone, and haloperidol in patients with chronic schizophrenia or schizoaffective disorder. *Am. J. Psychiatry* 159, 1018–1028. doi: 10.1176/appi.ajp.159.6.1018
- Boucart, M., Mobarek, N., Cuervo, C., Danion, J. M. (1999). What is the nature of increased Stroop interference in schizophrenia? *Acta Psychol. (Amst.)* 101, 3–25. doi: 10.1016/S0001-6918(98)00037-7
- De Gelder, B., and Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends Cogn. Sci.* 7, 460–467. doi: 10.1016/j.tics.2003.08.014
- de Gelder, B., Vroomen, J., de Jong, S. J., Masthoff, E. D., Trompenaars, F. J., and Hodiament, P. (2005). Multisensory integration of emotional faces and voices in schizophrenics. *Schizophr. Res.* 72, 195–203. doi: 10.1016/j.schres.2004.02.013
- de Jong, J. J., Hodiament, P. P., Van den Stock, J., and de Gelder, B. (2009). Audiovisual emotion recognition in schizophrenia: reduced integration of facial and vocal affect. *Schizophr. Res.* 107,

- 286–293. doi: 10.1016/j.schres.2008.10.001
- de Jong, J. J., Hodiomont, P. P., and de Gelder, B. (2010). Modality-specific attention and multisensory integration of emotions in schizophrenia: reduced regulatory effects. *Schizophr. Res.* 122, 136–143. doi: 10.1016/j.schres.2010.04.010
- Erthal, F. S., de Oliveira, L., Mocaiber, I., Pereira, M. G., Machado-Pinheiro, W., Volchan, E., et al. (2005). Load-dependent modulation of affective picture processing. *Cogn. Affect. Behav. Neurosci.* 5, 388–395. doi: 10.3758/CABN.5.4.388
- Habel, U., Chechko, N., Pauly, K., Koch, K., Backes, V., Seiferth, N., et al. (2010). Neural correlates of emotion recognition in schizophrenia. *Schizophr. Res.* 122, 113–123. doi: 10.1016/j.schres.2010.06.009
- Henik, A., and Salo, R. (2004). Schizophrenia and the Stroop effect. *Behav. Cogn. Neurosci. Rev.* 42–59. doi: 10.1177/1534582304263252
- Johnston, P. J., Dvir, H., and Karayanidis, F. (2006). Facial emotion processing in schizophrenia: no evidence for a deficit specific to negative emotions in a differential deficit design. *Psychiatry Res.* 143, 51–61. doi: 10.1016/j.psychres.2005.08.006
- Klasen, M., Kenworthy, C. A., Mathiak, K. A., Kircher, T. T., and Mathiak, K. (2011). Supramodal representation of emotions. *J. Neurosci.* 31, 13635–13643. doi: 10.1523/JNEUROSCI.2833-11.2011
- Kohler, C. G., Turner, T. H., Bilker, W. B., Brensinger, C. M., Siegel, S. J., Kanes, S. J., et al. (2003). Facial emotion recognition in schizophrenia: intensity effects and error pattern. *Am. J. Psychiatry* 160, 1768–1774. doi: 10.1176/appi.ajp.160.10.1768
- Lim, S. L., and Pessoa, L. (2008). Affective learning increases sensitivity to graded emotional faces. *Emotion* 8, 96–103. doi: 10.1037/1528-3542.8.1.96
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychol. Bull.* 109, 163–203. doi: 10.1037/0033-2909.109.2.163
- Mesholam-Gately, R. I., Giuliano, A. J., Goff, K. P., Faraone, S. V., and Seidman, L. J. (2009). Neurocognition in first-episode schizophrenia: a meta-analytic review. *Neuropsychology* 23, 315–336. doi: 10.1037/a0014708
- Mitchell, D. G., Nakic, M., Fridberg, D., Kamel, N., Pine, D. S., and Blair, R. J. (2007). The impact of processing load on emotion. *Neuroimage* 34, 1299–1309. doi: 10.1016/j.neuroimage.2006.10.012
- Morris, R. W., Weickert, C. S., and Loughland, C. M. (2009). Emotional face processing in schizophrenia. *Curr. Opin. Psychiatry* 22, 140–146. doi: 10.1097/YCO.0b013e328324f895
- Nuechterlein, K. H., Barch, D. M., Gold, J. M., Goldberg, T. E., Green, M. F., and Heaton, R. K. (2004). Identification of separable cognitive factors in schizophrenia. *Schizophr. Res.* 15;72, 29–39. doi: 10.1016/j.schres.2004.09.007
- Perlstein, W. M., Carter, C. S., Barch, D. M., and Baird, J. W. (1998). The Stroop task and attention deficits in schizophrenia: a critical evaluation of card and single-trial Stroop methodologies. *Neuropsychology* 12, 414–425. doi: 10.1037/0894-4105.12.3.414
- Pessoa, L., McKenna, M., Gutierrez, E., and Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *Proc. Natl. Acad. Sci. U.S.A.* 99, 11458–11463. doi: 10.1073/pnas.172403899
- Pessoa, L., Padmala, S., and Morland, T. (2005). Fate of unattended fearful faces in the amygdala is determined by both attentional resources and cognitive modulation. *Neuroimage* 28, 249–255. doi: 10.1016/j.neuroimage.2005.05.048
- Pessoa, L., and Ungerleider, L. G. (2004). Neuroimaging studies of attention and the processing of emotion-laden stimuli. *Prog. Brain Res.* 144, 171–182. doi: 10.1016/S0079-6123(03)14412-3
- Posner, M. I., and Peterson, S. E. (1990). The attention system of the human brain. *Annu. Rev. Neurosci.* 13, 25–42. doi: 10.1146/annurev.ne.13.030190.000325
- Ross, L. A., Saint-Amour, D., Leavitt, V. M., Molholm, S., Javitt, D. C., and Foxe, J. J. (2007). Impaired multisensory processing in schizophrenia: deficits in the visual enhancement of speech comprehension under noisy environmental conditions. *Schizophr. Res.* 97, 173–183. doi: 10.1016/j.schres.2007.08.008
- Roudaia, E., Sekuler, A. B., Bennett, P. J., and Sekuler, R. (2013). Aging and audio-visual and multi-cue integration in motion. *Front. Psychol.* 4:267 doi: 10.3389/fpsyg.2013.00267
- Schneider, F., Gur, R. C., Koch, K., Backes, V., Amunts, K., Shah, N. J., et al. (2006). Impairment in the specificity of emotion processing in schizophrenia. *Am. J. Psychiatry* 163, 442–447. doi: 10.1176/appi.ajp.163.3.442
- Seubert, J., Loughead, J., Kellermann, T., Boers, F., Brensinger, C. M., and Habel, U. (2010). Multisensory integration of emotionally valenced olfactory-visual information in patients with schizophrenia and healthy controls. *J. Psychiatry Neurosci.* 35, 185–194. doi: 10.1503/jpn.090094
- Stroop, J. R. (1935). “Studies of interference in serial verbal reactions”. *J. Exp. Psychol.* 18, 643–662. doi: 10.1037/h0054651
- Szycik, G. R., Münte, T. F., Dillo, W., Mohammadi, B., Samii, A., Emrich, H. M., et al. (2009). Audiovisual integration of speech is disturbed in schizophrenia: an fMRI study. *Schizophr. Res.* 110, 111–118. doi: 10.1016/j.schres.2009.03.003
- Talsma, D., Senkowski, D., Soto-Faraco, S., and Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends Cogn. Sci.* 14, 400–410. doi: 10.1016/j.tics.2010.06.008
- Tewes, U. (1991). *HAWIE-R: Hamburg-Wechsler Intelligenztest für Erwachsene; Handbuch und Testanweisung*. Bern: Huber-Verlag.
- Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., et al. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Res.* 168, 242–249. doi: 10.1016/j.psychres.2008.05.006
- Trémeau, F. (2006). A review of emotion deficits in schizophrenia. *Dialogues Clin. Neurosci.* 8, 59–70. Available online at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3181757>
- Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends Cogn. Sci.* 9, 585–594. doi: 10.1016/j.tics.2005.10.011
- Vuilleumier, P., Armony, J. L., Driver, J., and Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron* 30, 829–841. doi: 10.1016/S0896-6273(01)00328-2
- WHO Collaborating Centre for Drug Statistics Methodology. (2012). *Guidelines for ATC Classification and DDD Assignment, Oslo, 2011*. Oslo: WHO Collaborating Centre for Drug Statistics Methodology.
- Williams, L. E., Light, G. A., Braff, D. L., and Ramachandran, V. S. (2010). Reduced multisensory integration in patients with schizophrenia on a target detection task. *Neuropsychologia* 48, 3128–3136. doi: 10.1016/j.neuropsychologia.2010.06.028
- Zvyagintsev, M., Nikolaev, A. R., Sachs, O., and Mathiak, K. (2011). Early attention modulates perceptual interpretation of multisensory stimuli. *Neuroreport* 22, 586–591. doi: 10.1097/WNR.0b013e32834901f3
- Zvyagintsev, M., Nikolaev, A. R., Thönnessen, H., Sachs, O., Dammers, J., and Mathiak, K. (2009). Spatially congruent visual motion modulates activity of the primary auditory cortex. *Exp. Brain Res.* 198, 391–402. doi: 10.1007/s00221-009-1830-5

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 06 June 2013; accepted: 26 September 2013; published online: 18 October 2013.

Citation: Zvyagintsev M, Parisi C, Chechko N, Nikolaev AR and Mathiak K (2013) Attention and multisensory integration of emotions in schizophrenia. *Front. Hum. Neurosci.* 7:674. doi: 10.3389/fnhum.2013.00674

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2013 Zvyagintsev, Parisi, Chechko, Nikolaev and Mathiak. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Connecting multimodality in human communication

Christina Regenbogen<sup>1,2</sup>, Ute Habel<sup>1,2</sup> and Thilo Kellermann<sup>1,2\*</sup>

<sup>1</sup> Department of Psychiatry, Psychotherapy, and Psychosomatics, Medical School, RWTH Aachen University, Aachen, Germany

<sup>2</sup> JARA Translational Brain Medicine, Jülich/Aachen, Germany

## Edited by:

Yu-Han Chen, University of New Mexico, USA

## Reviewed by:

Claus Lamm, University of Vienna, Austria

Sebastian Korb, University of Wisconsin Madison, USA

## \*Correspondence:

Thilo Kellermann, Department of Psychiatry, Psychotherapy, and Psychosomatics, Medical School, RWTH Aachen University, Pauwelsstr. 30, 52074 Aachen, Germany  
e-mail: tkellermann@ukaachen.de

A successful reciprocal evaluation of social signals serves as a prerequisite for social coherence and empathy. In a previous fMRI study we studied naturalistic communication situations by presenting video clips to our participants and recording their behavioral responses regarding empathy and its components. In two conditions, all three channels transported congruent emotional or neutral information, respectively. Three conditions selectively presented two emotional channels and one neutral channel and were thus bimodally emotional. We reported channel-specific emotional contributions in modality-related areas, elicited by dynamic video clips with varying combinations of emotionality in facial expressions, prosody, and speech content. However, to better understand the underlying mechanisms accompanying a naturalistically displayed human social interaction in some key regions that presumably serve as specific processing hubs for facial expressions, prosody, and speech content, we pursued a reanalysis of the data. Here, we focused on two different descriptions of temporal characteristics within these three modality-related regions [right fusiform gyrus (FFG), left auditory cortex (AC), left angular gyrus (AG) and left dorsomedial prefrontal cortex (dmPFC)]. By means of a finite impulse response (FIR) analysis within each of the three regions we examined the post-stimulus time-courses as a description of the temporal characteristics of the BOLD response during the video clips. Second, effective connectivity between these areas and the left dmPFC was analyzed using dynamic causal modeling (DCM) in order to describe condition-related modulatory influences on the coupling between these regions. The FIR analysis showed initially diminished activation in bimodally emotional conditions but stronger activation than that observed in neutral videos toward the end of the stimuli, possibly by bottom-up processes in order to compensate for a lack of emotional information. The DCM analysis instead showed a pronounced top-down control. Remarkably, all connections from the dmPFC to the three other regions were modulated by the experimental conditions. This observation is in line with the presumed role of the dmPFC in the allocation of attention. In contrary, all incoming connections to the AG were modulated, indicating its key role in integrating multimodal information and supporting comprehension. Notably, the input from the FFG to the AG was enhanced when facial expressions conveyed emotional information. These findings serve as preliminary results in understanding network dynamics in human emotional communication and empathy.

**Keywords: dynamic causal modeling, finite impulse response, empathy, emotion, prosody, facial expressions, speech content, multimodality**

## INTRODUCTION

Human communication relies on a dynamic information exchange of several communication channels. A successful reciprocal evaluation of social signals serves as a prerequisite for social coherence and empathy. During this perceptive process, information from multiple channels is integrated. For example, when we engage in a conversation with someone and see their worried expression, notice a cracking tone of voice and finally understand by verbal information that something tragic has happened, we base our social inference on the available information—here, from facial expressions, prosody, and speech content. In empathy, according to the contextual appraisal process (De Vignemont and Singer, 2006) emotional cues that are presented can either lead

to an immediate empathic reaction (“early appraisal model”) or a slightly delayed reaction. This deferred response is assumed to be due to intermediate steps of modulating an earlier, automatically elicited response (“late appraisal model”). Emotional cues therefore benefit from a congruent contextual embedding so they can be interpreted correctly and initiate an empathic reaction. Underlying are two different concepts that help us create an inner picture of the outside world. By means of “bottom-up” processing, sensations come in from the external world. We hear, see, touch, taste our environment and these sensations are cues to infer phenomenal reality and are automatically processed, even if we do not pay attention to them (Vuilleumier et al., 2001). Parallel to this, “top-down” processing of incoming information takes

place which restricts the amount of information and works as a selection filter, but also amplifies certain information, depending on the necessity and appropriateness of that specific piece of information (for an introduction see Desimone and Duncan, 1995). Well-researched in the cognitive domain such as visual attention (Desimone, 1996) top-down efforts can be applied to emotional processes, reappraising and modulating hues of emotions in the information we get (Wright et al., 2008; Lee et al., 2010; Mühlberger et al., 2012).

On a neural level, when faced with emotional multimodal information such as the presented example, bottom-up mechanisms include the stimuli's perception. This takes place in unimodal cortices (Kanwisher et al., 1997) as well as in multi- and supramodal areas which combine inputs and perceive emotionality irrespective of the modality it is presented in (Buckner et al., 2000; Campanella and Belin, 2007; Peelen et al., 2010; Seubert et al., 2010). Top-down influences guide this process and further integrate information.

When we reconsider the anecdotal social communication example stated above against the background of our everyday experiences it seems, however, quite unnatural to always have the luxury of perceiving as many as three unequivocal sources of information when inferring someone's mental state. In contrary, quite often, one or several channels are completely non-existent, as in email conversation (no facial expressions, and prosody), or on the telephone (no facial expressions). However, based on previous experiences and given the established effort to build a "good gestalt" (Wertheimer, 1923), our brains fill in information. Over the course of development, we learn how someone usually looks when they have a sad-sounding voice, we can also imagine how someone would sound that we see with tears running down their face but a muted speech.

Recent theoretical advances in the formulation of basic principles stating how the brain might function may be resorted to in order to unite descriptions of top-down and bottom-up processes on a cognitive level and hierarchical feed-back and feed-forward connections on a neural level—albeit maintaining a distinction between the two. In this article, we will refer to these concept(s) as the "Bayesian brain hypothesis" of cortical functioning. In short, Bayesian descriptions of cortical functioning rest upon the hierarchical organization of the neocortex (Maunsell and Van Essen, 1983) where Bayesian surprise is propagated from a lower level to the next higher one. Conversely, predictions about the causes of sensory inputs (inferences) are passed from higher levels to lower levels (Mumford, 1992). Deviations from these predictions have been referred to as (Bayesian) surprise (Itti and Baldi, 2009), prediction error (Rao and Ballard, 1999) or free energy (Friston, 2010). Against this background, the brain can be regarded as a Bayesian machine which tries to infer the causes of sensory inputs. In doing so, actual sensory input is tried to be explained by a kind of prior beliefs, expectations or predictions, whilst discrepancies between these expectations and actual inputs are reported from lower levels to higher ones. At each hierarchical stage higher levels try to explain away the prediction errors that are passed by the level below (Friston, 2005). On a cognitive or behavioral level of description the salience or novelty of stimuli can be regarded as examples of such a kind of

surprise or prediction error that are assumed to underlie e.g., the startle response or the mismatch negativity (Feldman and Friston, 2010). In a similar vein, effects of incongruence, like those observed in the Stroop task (Stroop, 1935), the Posner paradigm (Posner and Petersen, 1990) or in Simon tasks (Simon, 1969), may have a neural basis which can be characterized as Bayesian, with a hierarchical processing of stimuli and subsequent (motor) responding. When we now return to the topic of multimodal emotional communication, many aspects of the Bayesian brain hypothesis may be considered relevant in experimental situations where different modalities contain different or at least ambiguous emotional connotation. This may also be again linked to the "appraisal model of empathy" (De Vignemont and Singer, 2006) where incongruence may lead to a failure of sufficient embedding of an emotional cue in order to establish successful multimodal integration and/or an unequivocal empathic response.

In a functional magnetic resonance imaging (fMRI) study investigating empathy and its components emotion recognition and affective responses we presented short video clips to our participants (Regenbogen et al., 2012a). These clips presented actors who told short stories and expressed emotions on different communication channels. In two conditions, all channels uniformly transported either only emotional (condition E) or neutral (condition N) information. Three other conditions selectively presented two emotional channels and one neutral channel (conditions "neutral prosody," E/nP; "neutral facial expression," E/nF; and "neutral speech content," E/nC) while a last condition presented unintelligible speech content (condition E/iC). Subjects indicated the actors' emotional valence and their own while fMRI was recorded.

Our findings confirmed the speculation formulated above: multimodal emotionality showed a facilitative effect for empathy, assessed by behavioral ratings of self- and other emotion, and was associated with multimodal integration in thalamus and precuneus. While behavioral empathy clearly decreased when one out of three channels presented non-emotional information neural network activation was still (although weaker) present in regions responsible for inferring someone's mental state and understanding their actions (e.g., temporoparietal junction) and in the dorsomedial prefrontal cortex (dmPFC), a region which serves self-referential functioning (Wolf et al., 2010) such that it decouples one's own from other people's perspectives on the self (D'Argembeau et al., 2007; Lamm et al., 2010) and is part of an anterior mentalizing network (Schnell et al., 2011). Further, while neutral facial expressions made it more difficult to recognize the emotion presented in the video clip, neutral speech content mainly influenced participants' ratings of their own affective state (Regenbogen et al., 2012b). This latter finding supports the notion proposed in the late empathic appraisal model (De Vignemont and Singer, 2006). We suggest that emotion-congruent speech content may be necessary for not only recognizing someone else's emotion, but also sharing it, and thus reveal empathy.

Further, emotions displayed via facial expressions, prosody, and speech content were associated with bilateral activations in fusiform gyri (FFG), auditory cortices (AC), and left angular

gyrus (AG), respectively (Figure 1). The identification of these four different network-nodes associated with the perception of human interactions constituted a prerequisite for understanding the dynamics that underlie multimodal integration and at the same time being partly able to explain participants' behavioral empathy decline to incomplete information.

In this reanalysis of the same data set, we pursued two different descriptions of temporal characteristics within these three modality-related regions (FFG, AC, AG) and the dmPFC: First, we examined the post-stimulus time-courses within each of the six conditions because of the prolonged stimulus durations (~10 s) during which different temporal patterns regarding activation levels are likely to occur. This exploratory description of the temporal characteristics of the BOLD response during the videos was accomplished by means of a finite impulse response (FIR) analysis within the FFG, AC, and AG, respectively. Second, the network dynamics in terms of effective connectivity between these areas and the dmPFC were analyzed using dynamic causal modeling (DCM) in order to describe stimulus condition-related modulatory influences on the coupling between these regions.

## METHODS

### PREVIOUS fMRI STUDY

#### Participants and task

The final sample of the previous fMRI study (Regenbogen et al., 2012a) consisted of 27 right-handed healthy participants (13 females,  $M$  age = 34.07 years,  $SD$  = 9.82 years, no history of psychiatric disorder, neurological illness, current substance abuse or lifetime substance addiction) with normal or corrected-to-normal vision and positive MR scanning inclusion criteria.

Subjects were informed about the study protocol, familiarized with the stimulus presentation environment and gave written informed consent. During the fMRI experiment, they were presented with 96 short video clips depicting actors telling a self-related story. After the clip, they were asked to rate the emotional valence of the presented actor, as well as their own (each scaled  $-3$  very negative,  $+3$  very positive) and to indicate this via button press. If this resulted in a matching of other's and own (correct) emotion, the answer was defined as empathic. The experimental set-up was designed according to the Declaration of Helsinki and the study was approved by the local institutional review board.

#### Stimuli, fMRI design and data analysis

Stimuli consisted of 96 thoroughly evaluated video clips (for details please refer to Regenbogen et al., 2012b) with an average duration of 11.8 s ( $SD$  = 1 s). Clips depicted either a male or a female conversational partner who told self-related stories of different emotional valence (disgust, fear, happiness, sadness, or neutral). Six conditions with 16 videos each (all emotions collapsed) displayed different combinations of prosody, facial expression, and speech content. "All emotional" (E) included emotional stories with congruent emotional facial expression and prosody. "All neutral" (N) contained neutral stories with neutral facial expression and prosody. "Neutral prosody" (E/nP), "neutral face" (E/nF), or "neutral speech content" (E/nC) had

the respective channel presented neutral while the two other channels transferred an emotion. "Incomprehensive speech content" (E/iC) consisted of emotional prosody and facial expressions, yet incomprehensive foreign speech content (Polish, Russian, Croatian, Yugoslavian, or Serbian).

Functional images were obtained on a 3 T Trio® MR scanner (Siemens Medical Systems, Erlangen, Germany) during a single session using T2\* weighted echo-planar imaging (EPI) sensitive to blood oxygenation level dependent (BOLD) changes (voxel size:  $3.125 \times 3.125 \times 3.1$  mm, matrix size:  $64 \times 64$ , field of view (FoV):  $200 \times 200$  mm<sup>2</sup>, 36 axial (AC-PC) slices, gap 0.356 mm, TR/TE = 2000/30 ms, flip angle: 76°, 1180 volumes, total duration: 39.33 min).

Data analysis was performed with SPM8 (Wellcome Department of Cognitive Neurology, London). Data pre-processing included realignment, coregistration of the mean functional image into Montreal Neurologic Institute (MNI) space which delivered the priors for a segmentation process in which the mean image was non-linearly segmented into gray matter, white matter, and cerebrospinal fluid (unified segmentation, Ashburner and Friston, 2005). Fitting the mean functional image's gray matter, white matter, and cerebrospinal fluid segments with the tissue probability maps yielded the normalization parameters. These were applied to the time series, including resampling to a voxel size of  $1.5 \times 1.5 \times 1.5$  mm. Spatial smoothing on normalized images was carried out with an isotropic 8 mm FWHM (full width at half maximum) Gaussian kernel.

On a single-subject level, one regressor for each condition and one modeling the rating period were created and six realignment parameters were included as covariates of no interest. Serial autocorrelations were accounted for by including a first order autoregressive model. A mixed-effects general linear model (GLM) was used for group-level inference with subjects as random effects and conditions as fixed effects (for both factors heteroscedasticity assumed).

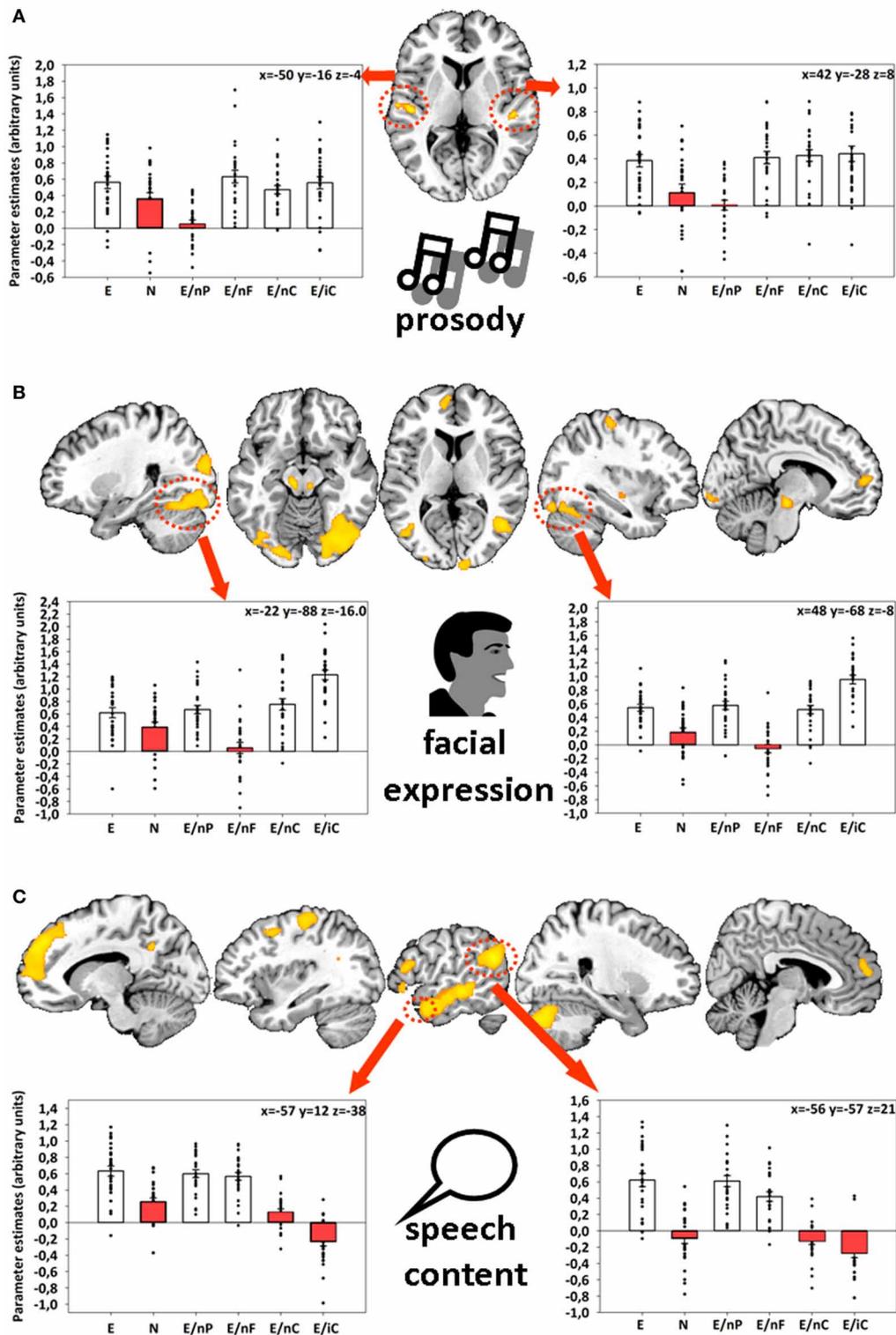
Three channel-specific contrasts were calculated in order to extract the emotional information of one communication channel each:  $[(E > E/nP) \cap (E > N)]$  for emotional prosody,  $[(E > E/nF) \cap (E > N)]$  for emotional facial expression, and  $[(E > E/nC) \cap (E > N)]$  for emotional speech content. To identify brain regions that are activated when emotions are presented through at least two of the three channels a four-fold conjunction analysis (Nichols et al., 2005) analyzed all three conditions with emotion present in 2/3 of the channels compared to the fully natural one [bimodal:  $(E/nP > N) \cap (E/nF > N) \cap (E/nC > N) \cap (E > N)$ ]. Further details regarding the fMRI experiment can be found in Regenbogen et al. (2012a).

### FINITE IMPULSE RESPONSE (FIR) ANALYSIS

The FIR analysis was constrained to regions of interest (ROI) and was performed using the SPM toolbox MarsBaR (<http://marsbar.sourceforge.net/>). The ROIs were specified based on

<sup>1</sup>Because contrasts did not only differ with respect to their emotionality, but also regarding ambiguity, all contrasts were analyzed in conjunction with  $E > N$ .

## Channel-specific emotional contribution of ...



**FIGURE 1 | Sliced MNI brain templates and overlays comprising the group activations of (A) emotional prosody, (B) emotional facial expression, and (C) emotional speech content. Within several activated clusters (indicated by red dotted shapes), individuals' parameter estimates**

were plotted onto cluster mean voxel estimates ( $\pm SEM$ ). MNI coordinates indicate the location of the maximum statistics. Contrast resulted from a random-effects GLM. Re-printed with kind permission of Elsevier (Regenbogen et al., 2012a).

the peak activation in the respective contrast (“emotional prosody,” “emotional facial expression,” “emotional speech content,” “bimodal”). Time-series were extracted from the left AC ( $x = -50, y = -17, z = -5$ ), right FFG ( $x = 48, y = -68, z = -8$ ), left AG ( $x = -56, y = -57, z = 21$ ) and left dmPFC ( $x = -2, y = 21, z = 54$ ). Data per ROI were averaged within a sphere of 10 mm radius around these peak coordinates, except for the AC, where the whole supra-threshold cluster was averaged. A window length (time bin) of 2 s was used, which corresponded to the repetition time (TR), and the order was set to 9 (corresponding to 18 s post-stimulus time) so that responses to the whole videos were captured including the expected hemodynamic delay.

The extracted mean values (one for each time-bin, totaling to 9 per subject and condition) were analyzed separately per ROI in IBM® SPSS® (version 20) by means of a two-way repeated measures ANOVA. The two factors were TIME with nine levels referring to the time bins and COND with three levels referring to the type of video. Two levels of the latter factor were identical for all ROIs, namely “all emotional” (E) and “all neutral” (N), while the last level was region-specific, sending neutral information to the respective modality-responsive region: E/nP for the AC, E/nF for the FFG and E/nC for the AG, respectively.

We tested the interaction term TIME  $\times$  COND for significance and performed two-tailed *post-hoc* *t*-tests for contrasts which focused on the comparison between the three conditions at each time-point. Because this analysis is rather exploratory and the Bonferroni correction is quite conservative, we decided to correct the threshold for each ANOVA individually [i.e., with  $(9 - 1) \times (3 - 1) = 16$  independent comparisons] instead of including the number of ROIs (which would have resulted in a correction factor of  $16 \times 3 = 48$ ). It must be emphasized that this procedure must be regarded as exploratory, because it does not adhere to the formally correct adjustment. Nevertheless, this approach is more conservative than not correcting for multiple comparisons at all.

### DYNAMIC CAUSAL MODELING (DCM) ANALYSIS

The regions for the DCM analysis were by name the same as in the FIR analysis. Because DCM infers on the causes of the MR-signal—namely on the hidden neuronal states or their activity induced by the stimuli—only those voxel of a region were included that survived an uncorrected threshold of  $p < 0.01$  in an *F*-contrast spanning the columns of interest in the first-level design of each subject. Before subjected to the DCM analysis, the data were adjusted in the sense that they were high-pass filtered (as described in the usual first-level analysis) and the effects of the covariates (intercept and realignment parameters) were removed. In order to address effective connectivity between these regions we used *post-hoc* Bayesian model selection (BMS) as introduced by Friston and Penny (Friston and Penny, 2011). Since we could hardly restrict the potentially huge model space a priori we decided to explore a large model space in a *post-hoc* fashion which is unfeasible for a standard BMS procedure because of its computational burden. Therefore, we applied *post-hoc* BMS by specifying a superordinate, “full” DCM whose subspace was searched under the Laplace approximation to determine an optimized DCM (Friston and Penny, 2011). We

assumed an almost full average connectivity structure between the four regions, where just the reciprocal connections between AC and FFG were omitted. Corresponding modulatory connectivities were included for the six different inputs or conditions (see Regenbogen et al., 2012a). These six conditions also served as driving inputs on the two nodes AC and FFG. The ensuing optimized DCMs for each subject were then averaged by means of Bayesian parameter averaging (BPA) as implemented in SPM8.

## RESULTS

### FIR ANALYSIS

The ANOVA comparing each nine time-bins (TIME) and three conditions (COND) yielded the following results in each region (Figure 2).

#### Auditory cortex

The ANOVA showed significant main effects of COND [ $F_{(2, 52)} = 27.65, p < 0.001$ ], TIME [ $F_{(2.16, 56.10)} = 117.64, p < 0.001$ ], and a significant interaction of COND  $\times$  TIME [ $F_{(7.70, 200.16)} = 9.24, p < 0.001$ ].

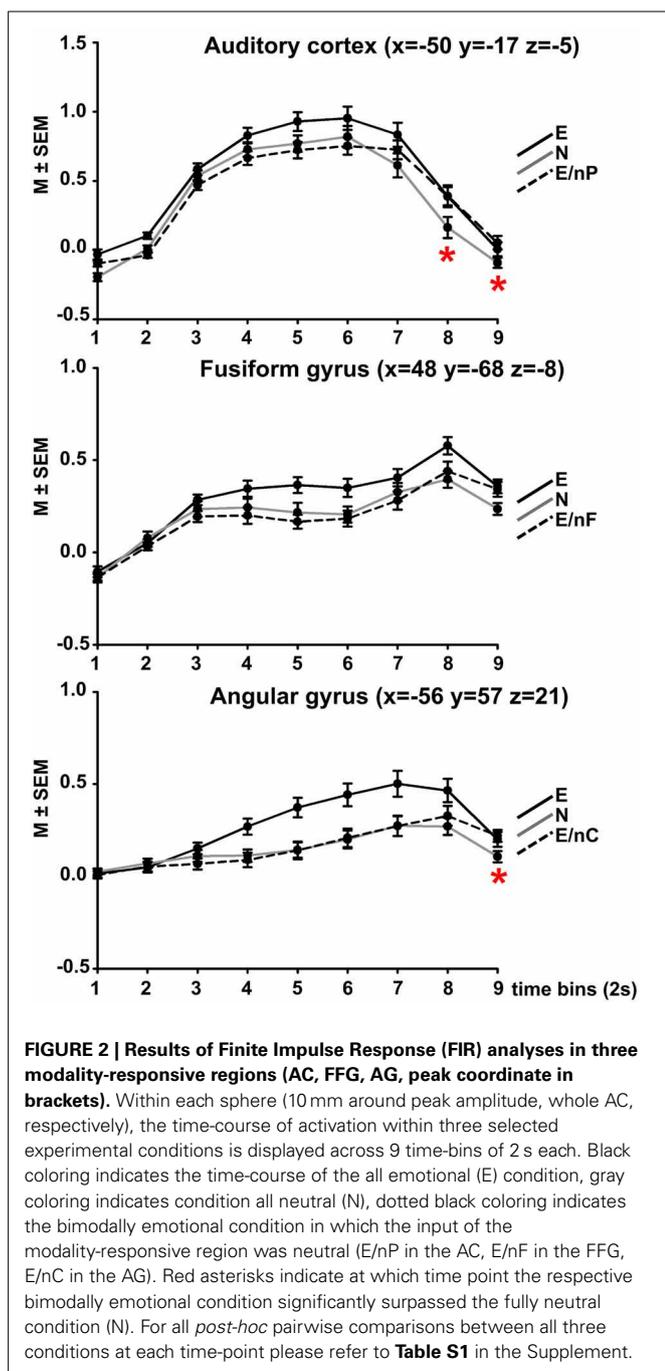
A complete listing of *post-hoc* comparisons of interest are listed in Table S1 in the Supplement and are Bonferroni-corrected for 16 independent comparisons (corrected threshold  $p = 0.001581$ ). Figure 2 displays the time-course plotted at intervals of 1 TR = 2 s. Condition E surpassed condition N at each time point (significantly at time-points 5, 7, and 9) and also condition E/nP (significantly at time-points 2–6), condition E/nP was actually lowest of the three across the first 6 time-points, even lower than N (n.s.). However, at time-point 7, condition E/nP surpassed condition N and this was significant at time-points 8 and 9.

#### Fusiform gyrus

The ANOVA showed significant main effects of COND [ $F_{(1.38, 35.90)} = 9.92, p = 0.001$ ], TIME [ $F_{(2.71, 70.44)} = 45.66, p < 0.001$ ], and a significant interaction of COND  $\times$  TIME [ $F_{(8.17, 212.36)} = 5.36, p = 0.002$ ]. For *post-hoc* comparisons please refer to Table S1. Condition E again surpassed condition N at each time point (significantly at time-points 5, 6, 8, and 9) and also condition E/nF (significantly at time-points 4, 5, 6, and 8), condition E/nP was actually lowest of the three across the first 7 time-points, even lower than N (n.s.). However, at time-point 7, condition E/nP surpassed condition N (n.s.).

#### Angular gyrus

The ANOVA showed significant main effects of COND [ $F_{(2, 52)} = 25.31, p < 0.001$ ], a significant main effect of TIME [ $F_{(1.57, 40.85)} = 27.65, p < 0.001$ ] and a significant interaction of COND  $\times$  TIME [ $F_{(5.64, 146.72)} = 8.69, p < 0.001$ ]. Condition E surpassed condition N at each time point (significantly at time-points 4–8) and also condition E/nC (significantly at time-points 4–8), condition E/nC was lowest of the three across the first 7 time-points (exception for time-point 6 where it was slightly above), even lower than N (n.s.). However, at time-point 8, condition E/nP surpassed condition N (significantly at time-point 9). (Figure 2; Table S1).



## DCM RESULTS

The *post-hoc* BMS showed a clear peak on one model which had a posterior probability of more than 0.99 for which reason we constrain the descriptions to this winning model. The context-dependent changes according to the six experimental conditions in the coupling between regions are illustrated in **Figure 3**. From visual inspection of the modulatory inputs it is quite striking that always the same connections were affected by one of the six conditions—apart from one exception, namely condition E/nP during which the AC→dmPFC connection was affected in addition to the other ones. Moreover, the general

pattern of modulated connections was of the form that all connections *from* dmPFC and all connections *to* AG were affected, albeit the size and even the sign of those modulations differed between conditions. The specific modulatory effects of each of the six conditions on the connections are briefly described in the following paragraphs.

### All emotional—E

While the dmPFC exerted a negative force (to AC and AG) or minor negative one (to FFG), the input from AC to AG showed a positive input in this (consistently emotional) experimental condition. This positive input was found in all other conditions with emotional speech content, but manifested strongest here. In both conditions with neutral speech content the AC exerted a slightly negative (N) or moderately negative (E/nC) influence. Further, the FFG exerted a strongly positive connection to the AG.

### All neutral—N

The fully neutral condition lead to an increase of inputs from dmPFC to both, FFG and AC. The input from dmPFC to AG stayed negative as in E but was less pronounced. Condition N (as well as E/nC) resulted in a decrease of connectivity from AC to the AG as it was present in all other conditions with emotional speech content. Further, the increasing modulatory inputs from FFG to AG were smallest in this condition (and negative in condition E/nF) while this connectivity was large in any other condition with an emotional facial expression.

### Neutral prosody—E/nP

Like in condition E, the dmPFC exerted negative inputs to AC, AG and FFG, although these were small to AC and AG. Further, uniquely present here was a negative input back from AC to the dmPFC. The input from AC to AG was slightly positive and again, the connection from FFG to AG was (strongly) positive.

### Neutral face—E/nF

When the facial expression was selectively switched to neutral, connectivity from dmPFC to the FFG and to the AG was slightly enhanced. Further, inputs from dmPFC to the AC were slightly decreased. Like in all other conditions with emotional speech content (except N and E/nC) connectivity from AC to AG was positive. This condition was also the only condition in which the input from FFG to AG was not increased but significantly decreased.

### Neutral speech content—E/nC

When speech content was neutral, connectivity from dmPFC to the AG was increased. This was also the case in condition E/nF. Both other connections originating in the dmPFC were moderately negative (AC and FFG). This condition also resulted in a decrease of connectivity from AC to the AG and a strong increase in connectivity from FFG to AG.

### Incomprehensive content—E/iC

In the foreign language condition, connectivity from dmPFC was mildly negative to all other regions (AC, FFG, AG). Further, both the AC (weak) and the FFG (strong) exerted a positive connection on the AG.

**DISCUSSION**

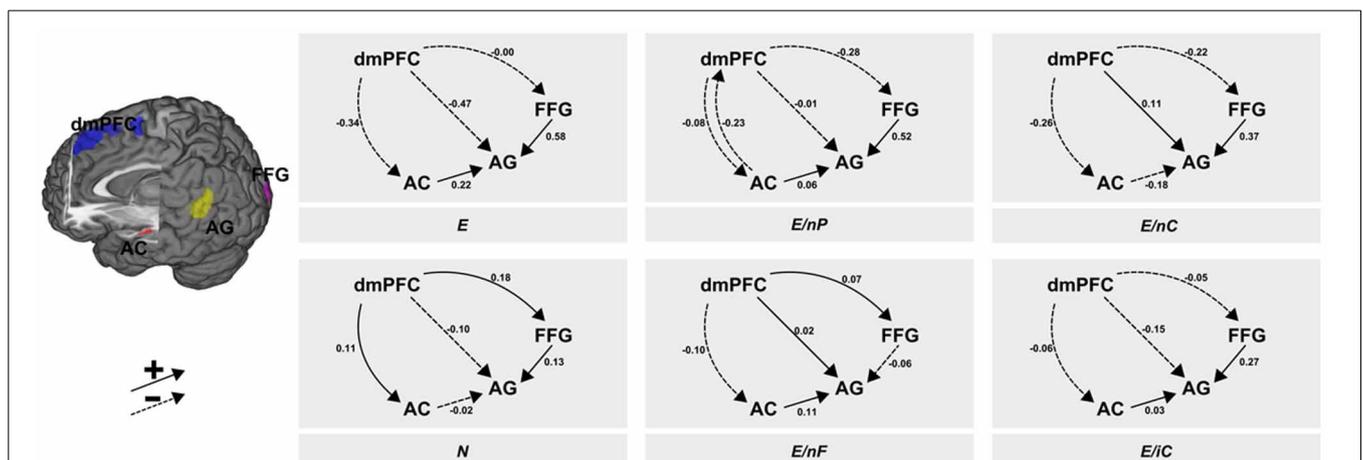
We aimed at clarifying several aspects related to the temporal development of neural activation in regions that are responsible for processing emotionality in three communication channels in an experiment studying empathy with naturalistic stimulus material. Based on a reanalysis of a data set, these findings led way to an effective connectivity analysis in which we further sought to disentangle the relationships between these regions and an area responsible for processing multi-channel information and serving self-referential functions in the dorsomedial prefrontal cortex.

The FIR analysis of the three experimental conditions in which two channels were emotional and one was neutral resulted in significantly lower activations of the respective channel-related area compared to the fully emotional condition. The time-course of the neutral condition was in between the other conditions for the first 5–6 time-bins. After about 2/3 of the video clip duration (time-bin 7 for the AC, time-bin 8 for the FFG and for the AG) the activation in the respective bimodal condition surpassed that of condition N. Although this was statistically significant for the AC at time-points 8 and 9, and for the AG at time-point 9, it was observable that the bimodal condition changed from “below neutral” to “above neutral.”

This initially diminished activation (time-bin 1 to approx. 6 or 7) might have been the result of a suppression by other emotionally informed regions in order to inhibit the incongruent communication channel—or in other words—to reduce the prediction error reported by that region. As evident in **Figure 1** while referring to the original analysis of this data-set (Regenbogen et al., 2012a), the whole-brain fMRI analysis showed that on average, bimodal conditions resulted—at least descriptively—in even less activation in the respective area (bimodal AC for E > E/nP (uncorr.  $p = 0.001$ ), bimodal FFG for E > E/nF (uncorr.  $p = 0.005$ ), and left AG for E/nC (n.s.,  $p = 0.667$ )) than the fully neutral condition. In this reanalysis, we were able to disentangle

the underlying time-course of the actual process to some degree. A prolonged presentation of the respective incongruent information may have then emphasized bottom-up processes which resulted in a stronger activation than the response observed in neutral videos toward the end of the stimuli. On a conceptual level, this could mean that although presented with incongruent information and thereby ambiguous (emotional and neutral) cues, contextual embedding still takes place to some degree, giving way to a successful late appraisal (De Vignemont and Singer, 2006) and empathy.

The general pattern of the *post-hoc* BMS, however, did not corroborate our hypotheses regarding this just-stated emphasis of bottom-up processes. One striking aspect of the graphical illustration of the DCM results (**Figure 3**) is the control that the dmPFC exerts on all other regions irrespective of the experimental condition. This pattern rather suggests a top-down control of the higher region dmPFC on the two lower regions (AC and FFG) as well as on AG. The fact that most of the coupling parameters (13 out of 18) have a negative sign may indicate a reduction of prediction errors reported by these lower regions. Remarkably, the condition transporting consistently emotional information (E) shows a unidirectional suppression of all three regions by the dmPFC, which might underscore the suppression of prediction error in the non-ambiguous emotional condition. The other likewise non-ambiguous but neutral condition (N), however, shows a synaptic gain in the connectivity from dmPFC to the two lower regions. As a rather speculative interpretation of this finding we may consider the neutral condition as being less informative so that this synaptic gain possibly reflects some cognitive effort in order to maintain attention in light of the rather boring content of the received message. Taking into account its documented role in self-referential processing (Wolf et al., 2010) and differentiation into self and other (Lamm et al., 2010), the dmPFC may—on a speculative note—not necessarily play a central role in fully neutral



**FIGURE 3 | Results of the DCM analysis.** The six panels show the modulatory inputs of the six different conditions on the connections between the regions. The stimuli conveyed neutral or emotional information via facial expressions, prosody and speech content. These channels were either consistent [all neutral (N) or all emotional (E)] or one of the three channels

was neutral whereas the other two transmitted emotional information (E/nP, neutral prosody; E/nF, neutral facial expression; E/nC, neutral speech content). The last condition contained incomprehensible speech (E/iC). AC, auditory cortex; FFG, fusiform gyrus; AG, angular gyrus; dmPFC, dorsomedial prefrontal cortex.

clips but rather in clips requiring some mental reappraisal of agency attribution.

Regarding the Bayesian brain hypothesis outlined in the introduction our results do not substantiate the idea that incongruent neutral information displayed by one out of three communication channels leads to a prediction error in the corresponding region at the lowest level (e.g., in FFG for E/nF or AC for E/nP) which then would be passed on to higher levels (dmPFC or AG). One reason for this lack of support is due to the rather bold assumption regarding the hierarchical ranking of the selected cortical regions: the AC and FFG at the lowest level, the dmPFC at the highest, and the AG somewhere in between. Another related argument pertains to the hierarchical distance between these regions which is assumed to be quite large, including polysynaptic pathways probably via several regions. This possibility raises the question if there are too many pathways and (cognitive) processes involved challenging the idea of a uniquely recordable hierarchy among the selected regions.

A recent meta-analytical review of vigilant attention proposed that the medial prefrontal cortex [including the dorsal anterior cingulate cortex (dACC)] and the anterior insula also belong to a network for “energizing” and performance monitoring (Langner and Eickhoff, 2012). While the dACC and the anterior insula are assumed to reactivate vigilant attention in cases of mind wandering when attention cannot be sustained, the medial prefrontal cortex (roughly corresponding to the dmPFC in the present study) withholds preplanned responses and maintains the intention and preparation to respond. The present results of the task-dependent connectivity of the dmPFC may point toward its role as an “energizer” of sensory modalities, allocating attention to functionally specialized, sensory regions. At least the ambiguous conditions lacking emotionality in facial expressions (E/nF) and speech content (E/nC) lead to an enhancement in the connectivity from the dmPFC to the corresponding region (FFG and AG, respectively). The consistently neutral condition resulted in a similar increase in the strength of connectivity from dmPFC to FFG and AC. This phenomenon may be explained by “inverse effectiveness,” where obscured, ambiguous or otherwise degraded information in a specific modality leads to an enhanced and thereby compensating allocation of resources on that modality (Senkowski et al., 2011; Stevenson et al., 2012). In the present study, the less clear or unique a source of information is, the greater is the need to take into account more available pieces of information and rely on those. Here, additional information is increasingly attended and used if information is obscured or ambiguous. Given the dmPFC’s role in processing at least two channels (Regenbogen et al., 2012a), i.e., multimodal information, it seems plausible that it was involved in redistributing information flow in cases of ambiguity as well as increased effort to carry out agency distribution while it may have been harder to mentalize or empathize with someone who was displaying an ambiguous message. Although this interpretation might bear some plausibility, it reveals some limitations of the present study which will be discussed below.

Assuming that a crucial part of semantic processing is accomplished in AG, which roughly corresponds to Wernicke’s area in our study, it is quite intriguing that the FFG → AG connection is

always strongly enhanced by stimuli containing emotional facial expressions, while this connection is only moderately enhanced in the consistently neutral condition and even suppressed in condition E/nF. Although this might underscore the importance of visual stimuli and corroborate the Colavita visual dominance effect (Colavita, 1974), our results should be regarded as preliminary because of our assumptions about semantic processing and all restrictions that pertain to implicit or explicit assumptions in every DCM study, i.e., inclusion of brain regions or restrictions of connectivity structure.

## LIMITATIONS

As some of the interpretations already hinted to weaknesses and limitations of the present study these should be elaborated on in this section. Regarding the FIR time-course analyses it must be emphasized again that these should be regarded as exploratory. The aim of this analysis was to look for some differences in the time-course of selected modality-responsive regions of interest in respective conditions. Although the time-by-condition interaction was significant for these regions, one has to keep in mind that due to the relatively large degrees of freedom particularly for the factor TIME quite subtle differences in the time-courses may yield the interaction significant. Moreover, the *post-hoc* *t*-tests have been corrected only for the number of independent comparisons within each of the three regions, yielding these tests more liberal and therefore exploratory.

With respect to the analyses of effective connectivity it must be emphasized that model selection can be regarded as both boon and bane of neuroscientific progress. On the one hand, the model-based approach, where explicitly devised models compete against each other for describing and explaining the same data set, brings forth our understanding of the neuronal implementations of cognitive processes. Maybe model selection is better suited in comparing competing explanations than conventional null-hypothesis significance tests where one single model is tested against a usually uninteresting chance explanation. On the other hand, model selection procedures heavily rely on the validity of (the definition of) the model space. In other words, one has to assure that at least one reasonable good model is included in the model space where it is hard to decide if a specific space is broad enough and suitable models are included.

In a similar vein, the selection of brain regions in a neuroscientific approach using BMS (for fMRI) is crucial as model selection requires the same data to be explained by all models and a different choice of regions changes the data set. Additionally, more regions make the potential model space grow exponentially rendering an exhaustive model space impossible to deal with. In a recent study, the models to be compared included only four regions and three different inputs and though the searched space was based on constraints it nevertheless included more than 4000 models which was managed by a high-performance computing cluster (Kellermann et al., 2012).

Because of this computational burden we decided to apply a *post-hoc* BMS which searches a subspace of one “full” model without the need to invert each single model within the search space (Friston and Penny, 2011). Although such a greedy approach makes the search of a huge space feasible within several seconds it

has been criticized because of combinatory explosion and therefore inclusion of many invalid models bearing the hazard of selecting such an invalid model purely by chance (Lohmann et al., 2012). Because of this risk, a greedy search without constraints should be characterized as exploratory and results as well as conclusions must be retained with care.

Contrary to the objections above one might also argue that the constraints we put on the model space (omitting connections between AC and FFG) excludes too many (plausible?) models *a priori*. Another argument refers to the number of regions being too little in order to adequately describe neuronal processing of multimodal emotional stimuli. For example, the “hierarchical gap” induced by restraining the data to a few regions may have resulted in a critical disregard of potentially available and valuable data. On the one hand, including intermediate levels may bridge that gap, but on the other hand it may severely impede putting reasonable constraints on plausible models and/or extremely complicates the problem of model space complexity.

Challenges for future studies therefore ask for both, inclusion of more regions (and thus more available data) which makes potential models more complex and at the same time putting reasonable constraints on the competing models permitting hypothesis-driven model selection. The antagonism between hypotheses-driven and exploratory studies remains, of course, but it might be regarded as a continuum wherein an individual study may be placed. As mentioned above, the present study should be regarded as an exploratory analysis of the effective connectivity among selected regions during multimodal processing of emotional stimuli.

## CONCLUSION

The best model, according to a greedy *post-hoc* Bayesian model selection, revealed a dynamic top-down connectivity from the dorsomedial prefrontal cortex (dmPFC) to at least two supposedly lower regions specialized for processing facial (fusiform gyrus, FFG) and auditory (auditory cortex, AC) stimuli as well as to a third region putatively involved in speech comprehension (angular gyrus, AG). The multimodal stimuli modulated these connectivities often in a sense that lacking emotional information in one of the three experimentally manipulated channels caused an increased input from the dmPFC to the respective channel-sensitive region. This pattern strengthens the suggested role of the dmPFC in executive functions like the allocation of cognitive resources and attention, as well as its role in mentalizing, empathy, and self-referential processing. While the experimental stimuli modulated all connections stemming from the dmPFC they also changed all connections toward the AG. Its property as a main receiving region might underscore its importance in the integration of information from different modalities in order to support comprehension. Notably in this regard are the couplings directed from the FFG to the AG as these were quite large in conditions in which the facial expression conveyed emotional information. These results indicate that the AG is possibly strongly supplied with emotional visual information which might be a hint toward another example of visual dominance.

Summarized, the results underlying the processing of complex dynamic stimulus with social relevance could be characterized in

the temporal domain and show compensation effects driven by bottom-up effects in channel-sensitive regions potentially giving way to late appraisal of empathy-eliciting information. Further, the analysis of the dynamics between those regions and an important hub in social cognition and executive functions suggests a differential role in allocation attention toward incoming information depending on the emotional content and relevance of the stimulus, as well as the integration of multiple emotional and neutral modalities.

## ACKNOWLEDGMENTS

This study was supported by the Deutsche Forschungsgemeinschaft (DFG: IRTG1328, Ha3202/2-2), JARA-BRAIN, the Interdisciplinary Center for Clinical Research of the Medical Faculty of the RWTH Aachen University (N2-6, N4-4).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fnhum.2013.00754/abstract>

**Table S1 | *Post-hoc* comparisons for each repeated measures ANOVA for the regions of interest.**

## REFERENCES

- Ashburner, J., and Friston, K. J. (2005). Unified segmentation. *Neuroimage* 26, 839–851. doi: 10.1016/j.neuroimage.2005.02.018
- Buckner, R. L., Koutstaal, W., Schacter, D. L., and Rosen, B. R. (2000). Functional MRI evidence for a role of frontal and inferior temporal cortex in amodal components of priming. *Brain* 123, 620–640. doi: 10.1093/brain/123.3.620
- Campanella, S., and Belin, P. (2007). Integrating face and voice in person perception. *Trends Cogn.Sci.* 11, 535–543. doi: 10.1016/j.tics.2007.10.001
- Colavita, F. B. (1974). Human sensory dominance. *Percept. Psychophys.* 16, 409–412. doi: 10.3758/BF03203962
- D’Argembeau, A., Ruby, P., Collette, F., Degueldre, C., Baeteau, E., Luxen, A., et al. (2007). Distinct regions of the medial prefrontal cortex are associated with self-referential processing and perspective taking. *J. Cogn. Neurosci.* 19, 935–944. doi: 10.1162/jocn.2007.19.6.935
- De Vignemont, F., and Singer, T. (2006). The empathic brain: how, when and why? *Trends Cogn.Sci.* 10, 435–441. doi: 10.1016/j.tics.2006.08.008
- Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proc. Natl. Acad. Sci. U.S.A.* 93, 13494–13499. doi: 10.1073/pnas.93.24.13494
- Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222. doi: 10.1146/annurev.ne.18.030195.001205
- Feldman, H., and Friston, K. J. (2010). Attention, uncertainty and free-energy. *Front. Hum. Neurosci.* 4:215. doi: 10.3389/fnhum.2010.00215
- Friston, K. (2005). A theory of cortical responses. *Phil. Trans. R. Soc. B* 360, 815–836. doi: 10.1098/rstb.2005.1622
- Friston, K. (2010). The free energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787
- Friston, K., and Penny, W. (2011). *Post hoc* Bayesian model selection. *Neuroimage* 56, 2089–2099. doi: 10.1016/j.neuroimage.2011.03.062
- Itti, L., and Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Res.* 49, 1295–1306. doi: 10.1016/j.visres.2008.09.007
- Kanwisher, N., McDermott, J., and Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302–4311.
- Kellermann, T., Regenbogen, C., De Vos, M., Mößnang, C., Finkelmeyer, A., and Habel, U. (2012). Effective connectivity of the human cerebellum during visual attention. *J. Neurosci.* 32, 11453–11460. doi: 10.1523/JNEUROSCI.0678-12.2012

- Lamm, C., Meltzoff, A. N., and Decety, J. (2010). How do we empathize with someone who is not like us? A functional magnetic resonance imaging study. *J. Cogn. Neurosci.* 22, 362–376. doi: 10.1162/jocn.2009.21186
- Langner, R., and Eickhoff, S. B. (2012). Sustaining attention to simple tasks: a meta-analytic review of the neural mechanisms of vigilant attention. *Psychol. Bull.* 139, 870–900. doi: 10.1037/a0030694
- Lee, K. Y., Lee, T. H., Yoon, S. J., Cho, Y. S., Choi, J. S., and Kim, H. T. (2010). Neural correlates of top-down processing in emotion perception: an ERP study of emotional faces in white noise versus noise-alone stimuli. *Brain Res.* 1337, 56–63. doi: 10.1016/j.brainres.2010.03.094
- Lohmann, G., Erfurth, K., Müller, K., and Turner, R. (2012). Critical comments on dynamic causal modelling. *Neuroimage* 59, 2322–2329. doi: 10.1016/j.neuroimage.2011.09.025
- Maunsell, J. H., and Van Essen, D. C. (1983). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *J. Neurosci.* 3, 2563–2586.
- Mühlberger, A., Neumann, R., Lozo, L., Müller, M., and Hetteringer, M. (2012). Bottom-up and top-down influences of beliefs on emotional responses: fear of heights in a virtual environment. *Stud. Health Technol. Inform.* 181, 133–137. doi: 10.3233/978-1-61499-121-2-133
- Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biol. Cybern.* 66, 241–251. doi: 10.1007/BF00198477
- Nichols, T., Brett, M., Andersson, J., Wager, T., and Poline, J. B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage* 25, 653–660. doi: 10.1016/j.neuroimage.2004.12.005
- Peelen, M. V., Atkinson, A. P., and Vuilleumier, P. (2010). Supramodal representations of perceived emotions in the human brain. *J. Neurosci.* 30, 10127–10134. doi: 10.1523/JNEUROSCI.2161-10.2010
- Posner, M. I., and Petersen, S. E. (1990). The attention system of the human brain. *Ann. Rev. Neurosci.* 13, 25–42. doi: 10.1146/annurev.ne.13.030190.000325
- Rao, R. P., and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive field effects. *Nat. Neurosci.* 2, 79–87. doi: 10.1038/4580
- Regenbogen, C., Schneider, D. A., Gur, R. E., Schneider, F., Habel, U., and Kellermann, T. (2012a). Multimodal human communication – targeting facial expressions, speech content and prosody. *Neuroimage* 60, 2346–2356. doi: 10.1016/j.neuroimage.2012.02.043
- Regenbogen, C., Schneider, D. A., Finkelmeyer, A., Kohn, N., Derntl, B., Kellermann, T., et al. (2012b). The differential contribution of facial expressions, prosody, and speech content to empathy. *Cogn. Emot.* 26, 995–1014. doi: 10.1080/02699931.2011.631296
- Schnell, K., Bluschke, S., Konradt, B., and Walter, H. (2011). Functional relations of empathy and mentalizing: an fMRI study on the neural basis of cognitive empathy. *Neuroimage* 54, 1743–1754. doi: 10.1016/j.neuroimage.2010.08.024
- Senkowski, D., Saint-Amour, D., Hofle, M., and Foxe, J. J. (2011). Multisensory interactions in early evoked brain activity follow the principle of inverse effectiveness. *Neuroimage* 56, 2200–2208. doi: 10.1016/j.neuroimage.2011.03.075
- Seubert, J., Kellermann, T., Loughead, J., Boers, F., Brensinger, C., Schneider, F., et al. (2010). Processing of disgusted faces is facilitated by odor primes: a functional MRI study. *Neuroimage* 53, 746–756. doi: 10.1016/j.neuroimage.2010.07.012
- Simon, J. R. (1969). Reaction towards the source of stimulation. *J. Exp. Psychol.* 81, 174–176. doi: 10.1037/h0027448
- Stevenson, R. A., Bushmakin, M., Kim, S., Wallace, M. T., Puce, A., and James, T. W. (2012). Inverse effectiveness and multisensory interactions in visual event-related potentials with audiovisual speech. *Brain Topogr.* 25, 308–326. doi: 10.1007/s10548-012-0220-7
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662. doi: 10.1037/h0054651
- Vuilleumier, P., Armony, J. L., Driver, J., and Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron* 30, 829–841. doi: 10.1016/S0896-6273(01)00328-2
- Wertheimer, M. (1923). “Untersuchungen zur Lehre von der Gestalt II,” in *Psychologische Forschung* 4, 301–350. Translation published in Ellis, W. (1938). *A Source Book of Gestalt Psychology* (London: Routledge and Kegan Paul), 71–88.
- Wolf, I., Dziobek, I., and Heekeren, H. R. (2010). Neural correlates of social cognition in naturalistic settings: a model-free analysis approach. *Neuroimage* 49, 894–904. doi: 10.1016/j.neuroimage.2009.08.060
- Wright, P., Albarracín, D., Brown, R. D., Li, H., He, G., and Liu, Y. (2008). Dissociated responses in the amygdala and orbitofrontal cortex to bottom-up and top-down components of emotional evaluation. *Neuroimage* 39, 894–902. doi: 10.1016/j.neuroimage.2007.09.014

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 April 2013; accepted: 21 October 2013; published online: 08 November 2013.

Citation: Regenbogen C, Habel U and Kellermann T (2013) Connecting multimodality in human communication. *Front. Hum. Neurosci.* 7:754. doi: 10.3389/fnhum.2013.00754

This article was submitted to the journal *Frontiers in Human Neuroscience*. Copyright © 2013 Regenbogen, Habel and Kellermann. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Dissociating task difficulty from incongruence in face-voice emotion integration

Rebecca Watson<sup>1,2\*</sup>, Marianne Latinus<sup>2,3</sup>, Takao Noguchi<sup>2,4</sup>, Oliver Garrod<sup>2</sup>, Frances Crabbe<sup>2</sup> and Pascal Belin<sup>2,3,5</sup>

<sup>1</sup> Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, Netherlands

<sup>2</sup> Centre for Cognitive Neuroimaging, Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, UK

<sup>3</sup> Institut de Neurosciences de la Timone and Aix-Marseille Université, Marseille, France

<sup>4</sup> Department of Psychology, University of Warwick, Coventry, UK

<sup>5</sup> International Laboratories for Brain, Music and Sound, Université de Montréal and McGill University, Montreal, QC, Canada

## Edited by:

Martin Klasen, RWTH Aachen University, Germany

## Reviewed by:

Anne-Marie Brouwer, Max Planck Institute for Biological Cybernetics, Germany

Antje B. M. Gerdes, University of Würzburg, Germany

Mariateresa Sestito, Università di Parma, Italy

## \*Correspondence:

Rebecca Watson, Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Oxfordlaan 55, Maastricht 6229 EV, Netherlands  
e-mail: rebecca.watson@maastrichtuniversity.nl

In the everyday environment, affective information is conveyed by both the face and the voice. Studies have demonstrated that a concurrently presented voice can alter the way that an emotional face expression is perceived, and vice versa, leading to emotional conflict if the information in the two modalities is mismatched. Additionally, evidence suggests that incongruence of emotional valence activates cerebral networks involved in conflict monitoring and resolution. However, it is currently unclear whether this is due to task difficulty—that incongruent stimuli are harder to categorize—or simply to the detection of mismatching information in the two modalities. The aim of the present fMRI study was to examine the neurophysiological correlates of processing incongruent emotional information, independent of task difficulty. Subjects were scanned while judging the emotion of face-voice affective stimuli. Both the face and voice were parametrically morphed between anger and happiness and then paired in all audiovisual combinations, resulting in stimuli each defined by two separate values: the degree of incongruence between the face and voice, and the degree of clarity of the combined face-voice information. Due to the specific morphing procedure utilized, we hypothesized that the clarity value, rather than incongruence value, would better reflect task difficulty. Behavioral data revealed that participants integrated face and voice affective information, and that the clarity, as opposed to incongruence value correlated with categorization difficulty. Cerebrally, incongruence was more associated with activity in the superior temporal region, which emerged after task difficulty had been accounted for. Overall, our results suggest that activation in the superior temporal region in response to incongruent information cannot be explained simply by task difficulty, and may rather be due to detection of mismatching information between the two modalities.

**Keywords:** multisensory integration, emotion perception, functional magnetic resonance adaptation, incongruence, affective conflict

## INTRODUCTION

The recognition and understanding of emotion from the face and voice is a crucial part of social cognition and inter-personal relationships. In everyday life, however, the evaluation of emotion is rarely based on the expression of either of these modalities alone. Rather, we usually see a face whilst simultaneously hearing a voice, and combine this information to create a coherent, unified percept.

Neuroimaging studies of audiovisual emotion perception have typically compared the response to congruent audiovisual stimuli with purely auditory or visual ones. Regions that respond more to both, or each of the unimodal sources alone are assumed to play a part in integrating information from the two modalities. Studies using this approach have particularly emphasized the integrative role of the superior temporal gyrus (STG)/middle temporal gyrus (MTG) as well as the posterior STS (pSTS; Pourtois et al., 2005;

Ethofer et al., 2006a; Kreifelts et al., 2007, 2009, 2010; Robins et al., 2009), amygdala (Dolan et al., 2001; Ethofer et al., 2006a,b) and insula (Ethofer et al., 2006a); but also regions presumed to be part of the “visual” or “auditory” systems, such as the fusiform gyrus (Kreifelts et al., 2010) and anterior superior temporal gyrus (STG; Robins et al., 2009).

Another, less utilized approach has been to employ a congruence design where a condition with emotionally congruent bimodal stimulation is compared to emotionally incongruent bimodal cues (e.g., Dolan et al., 2001; Klasen et al., 2011; Müller et al., 2011). This comparison follows the assumption that only in a congruent condition can the unimodal inputs be truly integrated into a viable percept, and thus at the neural level regions responding more to congruent as opposed to incongruent information would be presumed to be involved in integrating multisensory information. Furthermore, a congruence design also

allows the researcher to focus on the effects of affective conflict. Affective conflict can occur when information conveyed by the two modalities is not congruent, and we concurrently receive two or more different emotional inputs. Brain regions activated by affective conflict can be isolated by employing the reverse contrast; that is, incongruent vs. congruent information.

Non-emotional conflict has been studied extensively using both behavioral and neuroimaging experiments (e.g., Carter et al., 1998; MacDonald et al., 2000; Durston et al., 2003; Weissman et al., 2004). In contrast, very few studies have focused on the effects of affective conflict. Behaviorally, emotion conflict has been indicated by decreases in categorization accuracy and increased reaction times in incongruent compared to congruent conditions (de Gelder and Vroomen, 2000; Dolan et al., 2001; Collignon et al., 2008). Congruence effects have rarely been explored at the cerebral level, and studies which have done so have mainly focussed on isolating regions which integrate affective information, as opposed to those responding to emotional conflict.

For example, in an early study Dolan et al. (2001) compared activation in response to congruent and incongruent affective face-voice stimuli. They observed that there was an enhanced response in the left amygdala to congruent fearful stimuli (fearful voice + fearful face) compared with incongruent ones (happy voice + fearful face), suggesting that the amygdala is important for emotional crossmodal sensory convergence, specifically during fear processing. More recently, Klasen et al. (2011) investigated the multimodal representation of emotional information with dynamic stimuli expressing facial and vocal emotions congruently and incongruently. The authors observed that both congruent and incongruent audiovisual stimuli evoked larger responses in thalamus and superior temporal regions, compared with unimodal conditions, but that congruent emotions (compared to incongruent) elicited higher activation in the amygdala, insula, ventral posterior cingulate (vPCC), temporo-occipital, and auditory cortices. The vPCC exhibited differential reactions to congruency and incongruency for all emotion categories, and the amygdala for all but one, leading the authors to conclude that these may be regions specifically involved in integrating affective information from the face and the voice.

Recently, Müller et al. (2011) conducted a study which focussed on the neural correlates of audiovisual emotional incongruence processing. The authors scanned subjects as they judged emotional expressions in static faces while concurrently being exposed to emotional (scream, laughter) or neutral (yawning) sounds. The imaging data revealed that incongruence of emotional valence between faces and sounds led to increased activation in the middle cingulate cortex, right superior frontal cortex, right supplementary motor area as well as the right temporoparietal junction. These results correspond to those of Klasen et al. (2011), who observed that incongruent emotions (compared to congruent) activated a similar frontoparietal network and also the bilateral caudate nucleus. However, in contrast to the findings of Dolan et al. (2001), Klasen et al. (2011), Müller et al. (2011) reported that congruent compared to incongruent conditions did not evoke significantly increased activation in any brain region.

The limited, and on occasions conflicting evidence means that the effects of emotion incongruence still remain a relatively open question. Importantly, it should also be noted that the described studies confounded aspects of task difficulty with stimulus incongruence. Typically, the judgment of emotion is more difficult in an incongruent condition. Generally aspects of task difficulty are inherent to the task: emotional congruency facilitates emotion recognition, which is the major benefit of multimodal emotions. As such, congruent and incongruent trials are usually by definition characterized by differences in difficulty levels. However, this means that the neural correlates of task difficulty have still not been fully disentangled from the pure effects of emotional incongruency.

The purpose of the present study was to examine the neurophysiological correlates of processing incongruent emotional information, independent of task difficulty. A secondary aim was to search for regions specifically processing congruent stimuli, which could be presumed to be involved in multisensory integration. We parametrically morphed dynamic faces and voices between anger and happiness, and paired the resultant visual and auditory morphs to create a set of audiovisual stimuli that varied in not only the degree of incongruence between the face and voice, but also their presumed difficulty to classify. We assigned the audiovisual stimuli two values: one corresponding to the degree of incongruence between the face and the voice, and another corresponding to the degree of clarity in the combined face-voice information. Due to our use of morphing techniques, we hypothesized that the clarity value of a stimulus would be more related to task difficulty than its incongruence value, and that this distinction would allow us to separately examine the effects of incongruence and task difficulty. Participants were scanned in a rapid, efficiency-optimized event-related design [specifically, the continuous carry-over design (Aguirre, 2007)] while viewing the parametrically morphed audiovisual movies, and performing a 2-alternative forced choice emotion classification task. On the basis of previous results, we hypothesized that if activation in those networks associated with conflict monitoring was due to task difficulty, unclear as opposed to incongruent stimuli would provoke a response in these areas. We also hypothesized that once task difficulty was accounted for, stimulus incongruence might instead evoke responses in regions more associated with audiovisual (specifically, audiovisual affective) processing.

## MATERIALS AND METHODS

### PARTICIPANTS

Ten English-speaking participants [4 males and 6 females; mean age 27 years (SD  $\pm$  13 years)] took part in a pre-test of stimuli, in order to ensure there was appropriate categorization of unimodal emotion (see Appendix), and a new group of eighteen participants [10 males, 8 females, mean age: 25 years (SD  $\pm$  3.7 years)] were scanned in the main fMRI experiment. All had self-reported normal or corrected vision and hearing. The study was approved by University of Glasgow Ethics Committee and conformed to the Declaration of Helsinki. All volunteers provided informed written consent before, and received payment at the rate of £6 p/h for participation.

## STIMULI

### **Video recording**

Two actors (one male, one female) participated in the video recording sessions. Both had studied drama at University level. The actors were paid for their participation at the rate of £6 p/h. Each actor sat in a recording booth, and was given instructions through an outside microphone connected to speakers within the booth. The actor wore a head cap, in order to hide the hair, and a marked head panel was fitted to the cap, which was used to determine head position. A Di3d capture system (see Winder et al., 2008) was used for the video recording. The actor sat between two camera pods, at a distance of 143 cm away from them both. Thus, each camera captured a slight side-view of the face, as opposed to a directly frontal view. Each pod consisted of a vertical arrangement of 3 different cameras. The top and bottom cameras were black and white, and were used to capture general shape information. The middle camera in each pod was a color camera, used to capture texture and color information. A lamp was placed behind each camera, and luminance kept constant at 21 amps. Video information was recorded by Di3D software on this PC as a series of jpegs at high resolution (2 megapixels). Vocal sound-information was transmitted via a Microtech Geffell GMBH UMT 800 microphone—positioned above the actor—to a second PC outside the booth, and was recorded at 44100 Hz using Adobe Audition (Adobe Systems Incorporated, San Jose, CA).

The actors were instructed to express anger and happiness in both the face and the voice. The sound “ah” was chosen as it contains no linguistic information. They were asked to sit as still as possible, in order to keep head movement to a minimum. Audiovisual expressions were produced a number of times, with a pause of three seconds between each repetition. The actor clapped in front of their face before they produced each set of expressions, which provided markers when later matching the audio recording to the video.

### **Video processing**

Video output was split into a number of different sequences, where each sequence was made up of a number of jpegs (frames) and each repetition of each emotional expression formed one sequence. Two final sequences were chosen for each actor. Using the Di3D software, 43 landmarks were placed around the face and facial features in the first and last frame of the sequence, forming a landmark-mesh. An existing generic mesh was applied to the beginning and the end of the sequence (i.e., first and last jpeg), which was then warped to fit the landmark-mesh. The first mesh was then used to estimate the mesh position in the second jpeg, which was then used to estimate the position in the third and so on. This forward tracking/mesh estimation was then carried out in the opposite direction (i.e., the last mesh was used to estimate the mesh position in the jpeg before it). The two side-views of the actor, one from each camera pod, were merged together, forming one directly frontal view of the face. We smoothed the converging line, which ran from the forehead to the chin down the middle of the face, using average facial texture information. Any head movement was removed by tracking and aligning the eight marked points on the head panel, so that they were always in the same position throughout the sequence.

### **Audio processing**

In addition to the original sound recording, a duplicate reduced-noise version was also produced. A recording made in the empty booth provided a “noise-baseline,” which was used to remove noise using a Fourier transform. The entire reduced-noise audio recording for each actor was then edited in Adobe Premiere (Adobe Systems Incorporated, San Jose, CA). Using the actor claps as markers for the start of each emotional expression, the audio sequences corresponding to the correct video sequence frames (at a frame rate of 25 frames per second) were identified and split into separate clips. The separate audio samples were then normalized for mean amplitude using Adobe Audition.

### **Video morphing**

The video morphing was performed independently on the texture and shape components of the 4D models. The texture was warped onto a common template shape using the piecewise-affine warp and the morph was then performed as a weighted linear sum on the RGB pixel values; the shape was normalized for rigid head position (i.e., rotation, translation) using a combination of the ICP (Besl and McKay, 1992) and the RANSAC (Bolles and Fischler, 1981) methods and the morph was then performed as a weighted linear sum on the vertex coordinates. To account for timing differences between two expressions, pairs of matching anchor frames were selected in the two sequences corresponding to similar movement stages (for example, “mouth first opens,” “maximum mouth opening,” etc.). The sequence pairs were broken up into segment pairs between the anchor points and the lengths of the pixel and vertex time courses for the segment pairs were rescaled to be equal for the pair using linear interpolation. The new length was chosen as the average length of the segment over the pair. Finally, the segment pairs were reassembled into the full sequence pair and the morph was performed at each frame of the sequence. Five morph levels were chosen—ranging from 10 to 90% of one expression, in 20% steps—and the same morph level was used at each frame of the sequence, producing a total of five morph sequences which were rendered to video using 3DS Max.

### **Audio morphing**

Auditory stimuli were edited using Adobe Audition 2.0. In order to generate the auditory components to the “morph-videos” three temporal and three frequency points were identified and landmarks corresponding to these set in the MATLAB-based morphing algorithm STRAIGHT (Kawahara, 2003), which were then used to generate a morph continuum between the two affective vocalizations equivalent to the faces. Two continua of voices—one for each actor, and consisting of five different voices ranging from 90% angry to 90% happy in 20% steps—were then generated by resynthesis based on a logarithmic interpolation of the angry and happy voices temporal and frequency anchor templates to a 50% average.

### **Audiovisual movie production**

The auditory and visual morphing procedures produced five dynamic face videos and five audio samples for each actor. Within actor, these stimuli were all equal length. In order to ensure all stimuli were of equal length, we edited video and audio clips

between actors. In all video clips, seven important temporal landmarks that best characterized the facial movements related to the vocal production were determined, and the frames at which they occurred were identified. These landmarks were the first movement of the chin, first opening of lips, maximum opening of the mouth, first movement of the lips inwards, time point at which the teeth met, closing of the lips, and the last movement of the chin. The theoretical average frames for these landmarks were then calculated, and the videos edited so the occurrence of these landmarks matched in all clips. Editing consisted of inserting or deleting video frames during fairly motionless periods. The editing produced ten adjusted video clips, each 18 frames (720 ms) long. The audio samples were then also adjusted in accordance with the temporal landmarks identified in the video clips, in order to create 10 vocalizations (5 for each actor) of equal length. Within actor, the five visual and five auditory clips were then paired together in all possible combinations. This resulted in a total of 25 audiovisual stimuli for each actor, parametrically varying in congruence between face and voice affective information (see **Figure 1**).

**Definition of stimulus clarity and incongruence**

Each stimulus was assigned “clarity” and “incongruence” values, which took into account the emotion morph of both the face and the voice. Incongruence was defined as the absolute (abs) value of face morph level minus voice morph level:

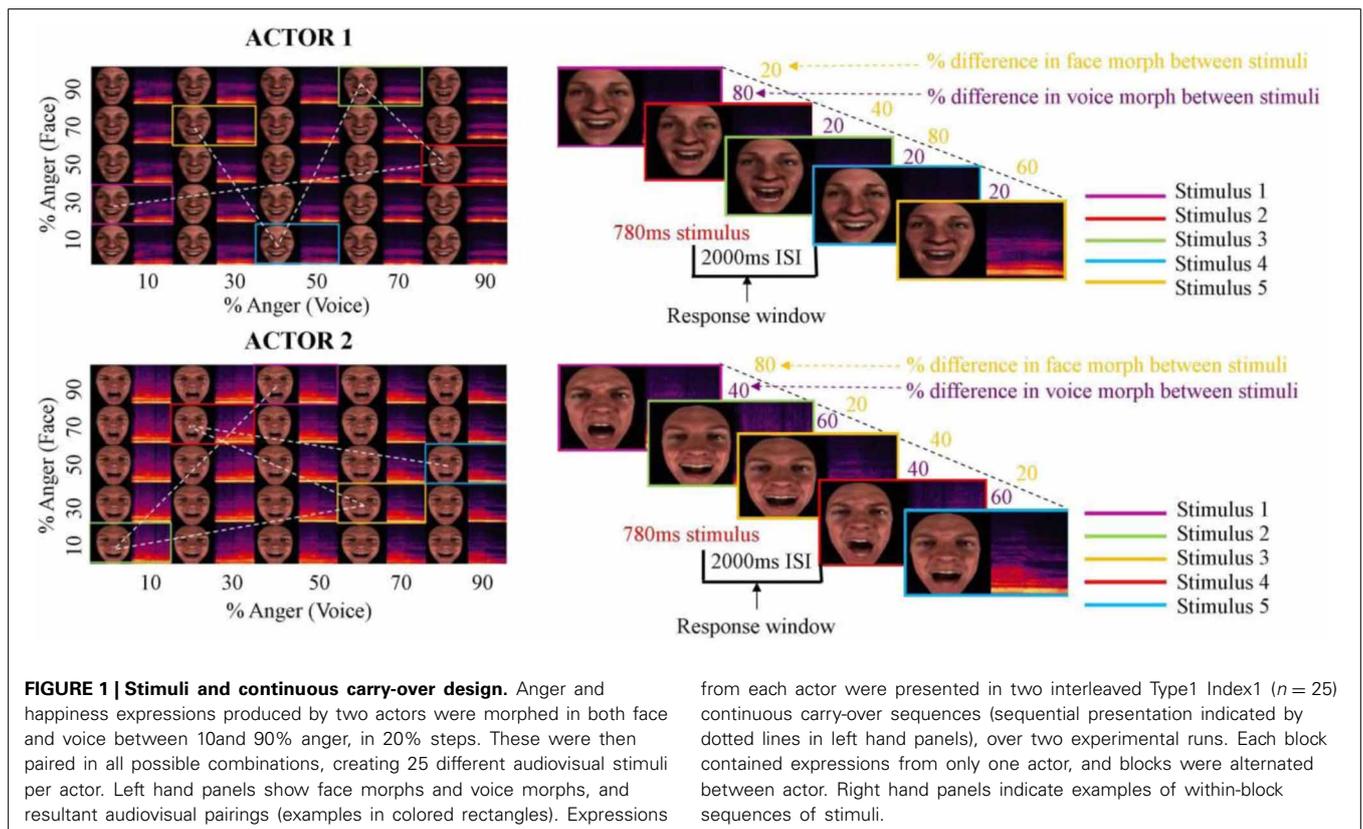
$$\text{Incongruence} = \text{abs}(\% \text{ anger in face} - \% \text{ anger in voice})$$

Therefore, the higher values indicated the highest degree of incongruence.

However, we recognized that although completely congruent stimuli were all assigned the same value, some would presumably be easier to categorize than others (e.g., 90% angry face-90% angry voice as compared to 50% angry face-50% angry voice). Therefore, we took into account the clarity of the combined information of the face and the voice. To calculate a clarity value, we determined the average percentage of “anger” information contained in the stimulus. For example, the 90% angry face-90% angry voice stimulus contained 90% anger informativeness, and the 10% angry face-90% angry voice contained 50% anger informativeness—as did the 50% angry face-50% angry voice stimulus. Clarity was thus calculated:

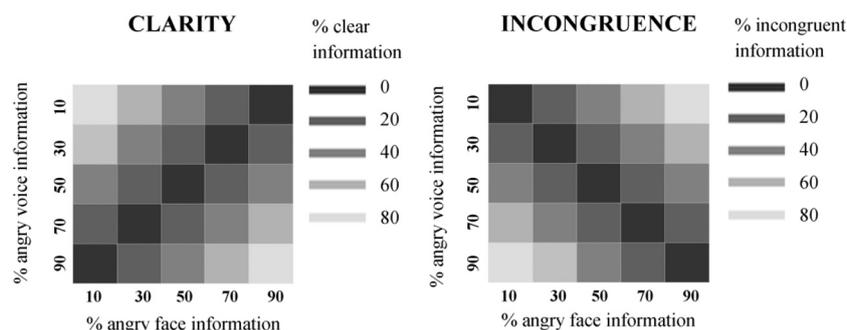
$$\text{Clarity} = \text{abs}[50\% - (\text{average \% anger information})]*2$$

This resulted in clarity values which were a 90° rotation of our incongruence values in the 2D face-voice space (see **Figure 2**), where the values indicated the level of clear affective information contained within the stimulus as a whole. The higher values indicated a clearer combined emotion representation and lower values indicated an unclear combined emotion representation. For clarity and incongruence values assigned to each stimulus, refer to **Figure 2**. It should also be noted that there was a significant negative correlation between clarity and incongruency values ( $r = -0.556, p < 0.0001$ ).



**FIGURE 1 | Stimuli and continuous carry-over design.** Anger and happiness expressions produced by two actors were morphed in both face and voice between 10 and 90% anger, in 20% steps. These were then paired in all possible combinations, creating 25 different audiovisual stimuli per actor. Left hand panels show face morphs and voice morphs, and resultant audiovisual pairings (examples in colored rectangles). Expressions

from each actor were presented in two interleaved Type1 Index1 ( $n = 25$ ) continuous carry-over sequences (sequential presentation indicated by dotted lines in left hand panels), over two experimental runs. Each block contained expressions from only one actor, and blocks were alternated between actor. Right hand panels indicate examples of within-block sequences of stimuli.



**FIGURE 2 | Clarity and incongruence values assigned to face-voice stimuli. Left:** Clarity values. Each square in the grid represents one audiovisual stimulus. The horizontal value indicates the percentage of angry face information contained in the stimulus, and the vertical value represents the percentage of angry voice information contained in the stimulus. The color represents the clarity value of the stimulus (darkest = lowest clarity value = unclear combined information in the stimulus; lightest = highest clarity value = clear combined information in the stimulus). **Right:**

Incongruence values. Each square in the grid represents one audiovisual stimulus. The horizontal value indicates the percentage of angry face information contained in the stimulus, and the vertical value represents the percentage of angry voice information contained in the stimulus. The color represents the incongruence value of the stimulus (darkest = lowest incongruence value = no difference in affect level of face and voice; lightest = highest = highest incongruence value = maximal difference in affect level of face and voice).

## DESIGN AND PROCEDURE

### Continuous carry-over experiment

In the main experiment, stimuli were presented by way of a continuous carry over design (Aguirre, 2007). In summary, these are efficiency-optimized, serially balanced fMRI sequences in which every stimulus precedes and follows every other stimulus (i.e., the “Type 1 Index 1” sequence), which account for stimulus counterbalancing.

Carry-over designs are particularly useful when there is more complex variation in a set of stimuli, whose differences can be expressed as parametric changes along a number of different axes (e.g., face and voice emotion morph). A full description of the continuous carry over design can be found in Aguirre (2007). Stimuli were presented using the Psychtoolbox in Matlab, via electrostatic headphones (NordicNeuroLab, Norway) at a sound pressure level of 80 dB as measured using a Lutron SL-4010 sound level meter. Before they were scanned, subjects were presented with sound samples to verify that the sound pressure level was comfortable and loud enough considering the scanner noise. Audiovisual movies were presented in two scanning runs (over two different days) while blood oxygenation-level dependent (BOLD) signal was measured in the fMRI scanner.

The stimulus order followed two interleaved  $N = 25$  Type1 Index 1 sequences (one for each of the speaker continua; ISI: 2 s; Nonyane and Theobald, 2007), which shuffles stimuli within the continuum so that each stimulus is preceded by itself and every other within-continuum in a balanced manner. The sequence was interrupted by seven 20 s silent periods, which acted as a baseline, and at the end of a silent period the last 5 stimuli of the sequence preceding the silence were repeated before the sequence continued. These stimuli were removed in our later analysis. Participants were instructed to perform a 2 alternative forced choice emotion classification task (responses: Angry or Happy) using 2 buttons of an MR compatible response pad (NordicNeuroLab, Norway). They were also instructed to pay attention to both the face and voice, but could use the information presented in whatever way they wished to make their decision on emotion. Reaction times

(relative to stimulus onset) were collected using Matlab with a response window limited to 2 s.

### Imaging parameters

Functional images covering the whole brain (slices = 32, field of view =  $210 \times 210$  mm, voxel size =  $3 \times 3 \times 3$  mm) were acquired on a 3T Tim Trio Scanner (Siemens) with a 12 channel head coil, using an echoplanar imaging (EPI) sequence (interleaved,  $TR = 2$  s,  $TE = 30$  ms, Flip Angle =  $80^\circ$ ) were acquired in both the carry-over and localizer experiments. In total, we acquired 1560 EPI image volumes for the carry-over experiment, split into two scanning sessions consisting of 780 EPI volumes. The first 4 s of the functional run consisted of “dummy” gradient and radio frequency pulses to allow for steady state magnetization during which no stimuli were presented and no fMRI data collected. MRI was performed at the Centre for Cognitive Neuroimaging (CCNi) in Glasgow, UK.

At the end of each fMRI session, high-resolution T1-weighted structural images were collected in 192 axial slices and isotropic voxels ( $1 \text{ mm}^3$ ; field of view:  $256 \times 256 \text{ mm}^2$ ,  $TR = 1900$  ms,  $TE = 2.92$  ms, time to inversion = 900 ms,  $FA = 9^\circ$ ).

### Imaging analysis

SPM8 software (Wellcome Department of Imaging Neuroscience, London, UK) was used to pre-process and analyse the imaging data. First the anatomical scan was AC-PC centered, and this correction applied to all EPI volumes.

Functional data were motion corrected using a spatial transformation which realigned all functional volumes to the first volume of the run and subsequently realigned the volumes to the mean volumes. The anatomical scan was co-registered to the mean volume and segmented. The anatomical and functional images were then normalized to the Montreal Neurological Institute (MNI) template using the parameters issued from the segmentation keeping the voxel resolution of the original scans ( $1 \times 1 \times 1$  and  $3 \times 3 \times 3$ , respectively). Functional images were then smoothed with a Gaussian function (8 mm FWHM).

EPI time series were analyzed using the general linear model as implemented in SPM8. Functional data was further analyzed in two separate two-level random effects designs:

**Clarity.** Brain activity time-locked to stimulus onset and duration was modeled against the 1st (linear) expansion of two parametric modulators: incongruence, then clarity. The second parametric modulator (clarity) was automatically orthogonalized with respect to the first (incongruence), meaning that any variance associated with incongruence was removed. The linear expansion allowed us to search for regions which showed a stepped, linear increase or decrease in signal in line with the linear increase/decrease in clarity of the audiovisual stimuli. This analysis is illustrated in (Figure 3A). The contrast for the effect of the second parametric modulator—clarity—was entered into separate second-level, group RFX analysis. We then further examined both positive and negative correlations of BOLD signal with clarity.

**Incongruence.** Brain activity time-locked to stimulus onset was modeled against the 1st (linear) expansion of two parametric modulators: clarity, then incongruence. In contrast to the previous model, any variance associated with clarity was regressed out, isolating any effects due only to the degree of incongruence between the face and voice. This analysis is illustrated in (Figure 3B). The contrast for the effect of the second parametric modulator—incongruence—was entered into separate second-level, group RFX analyses. As in our clarity analysis, we then examined both positive and negative effects.

Reported results from the experimental run are from whole-brain analyses, masked by an experimental audiovisual vs. base-line contrast thresholded at  $p < 0.001$  (voxel-level uncorrected), and are reported descriptively at a threshold of  $p < 0.05$  (FWE voxel-level corrected).

## RESULTS

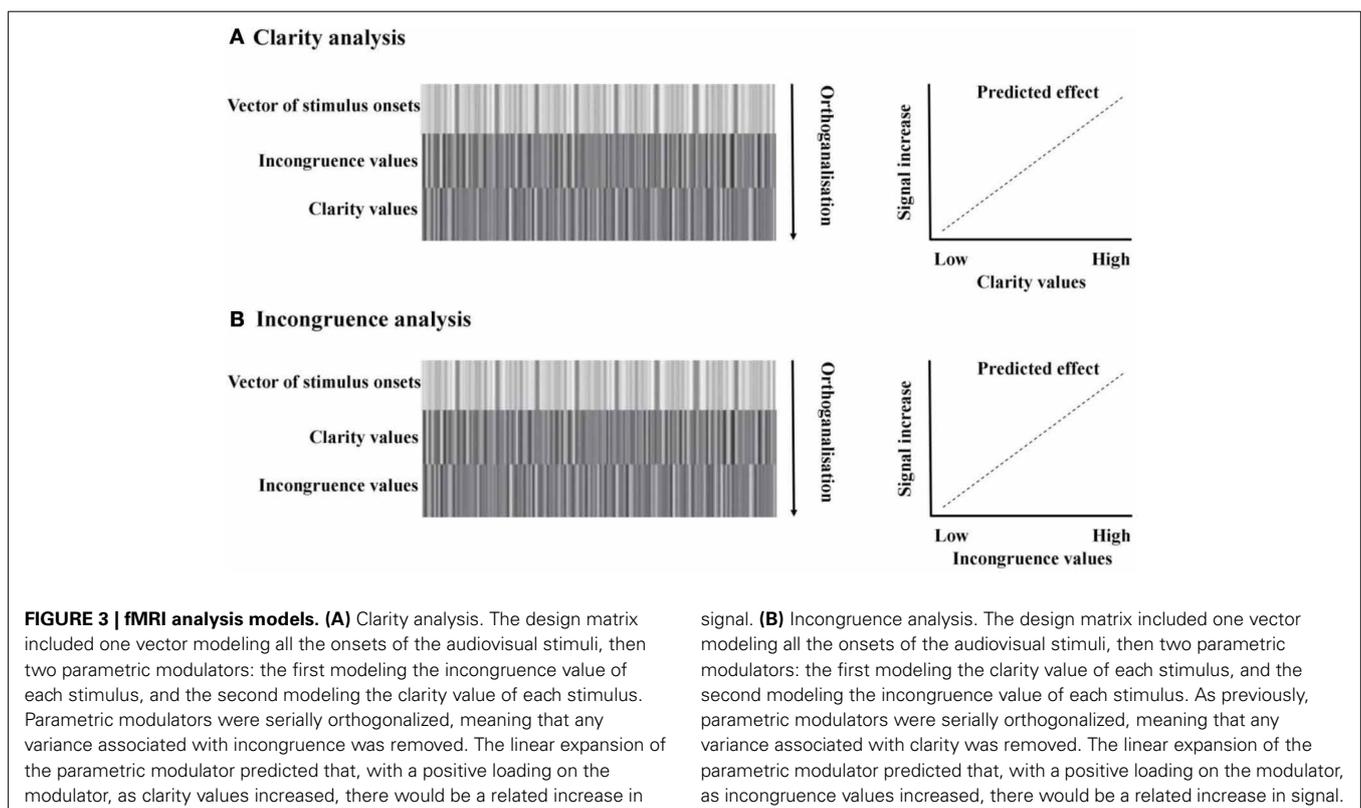
### BEHAVIORAL DATA

#### Effects of face and voice morph

**Categorical data.** Each participant's mean categorization values for each audiovisual emotion morph stimulus (collapsed across actor) was submitted to a two factor (face morph and voice morph), fully within subjects repeated measures ANOVA, with 5 levels per factor (percentage of "anger" information in the morph). This was in order to assess the overall contributions of face and voice emotion morph on categorical response.

The percentages of anger identification were of 96.3% ( $\pm 4.7\%$ ) for the 90% angry face-90% angry voice stimulus and 2.78% ( $\pm 3.59\%$ ) for the 90% happy face-90% happy voice stimulus. The percentage of anger identification for the 50% ambiguous angry-happy stimulus was 49.4% ( $\pm 16.9\%$ ). The repeated measures ANOVA highlighted a main effect of voice morph [ $F_{(1.14, 19.4)} = 15.3$ ,  $p < 0.002$ ,  $\eta_p^2 = 0.473$ ] and of face morph [ $F_{(2.02, 34.3)} = 348$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.953$ ], and also a significant voice  $\times$  face interaction [ $F_{(5.78, 98.1)} = 6.78$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.285$ ].

*Post-hoc*, we compared categorization values between each of our completely congruent stimuli (i.e., 10% angry face-10% angry voice; 30% angry face-30% angry voice; 50% angry



face-50% angry voice; 70% angry face-70% angry voice; 90% angry face-90% angry voice) in five paired *t*-tests. Each stimulus was compared to the next one (i.e., 10% angry face-10% angry voice vs. 30% angry face-30% angry voice; 30% angry face-30% angry voice vs. 50% angry face-50% angry voice and so on) to clarify whether each stimulus significantly differed from the other with regards to categorization. After a Bonferroni correction for multiple comparisons (level of significance:  $p < 0.01$ ), we found each of the stimuli significantly differed from the next—10% angry face-10% angry voice vs. 30% angry face-30% [angry voice:  $t_{(17)} = -2.82, p < 0.0125$ ; 30% angry face-30% angry voice vs. 50% angry face-50% angry voice:  $t_{(17)} = -13.7, p < 0.0001$ ; 50% angry face-50% angry voice vs. 70% angry face-70% angry voice:  $t_{(17)} = -10.8, p < 0.0001$ ; 70% angry face-70% angry voice vs. 90% angry face-90% angry voice:  $t_{(17)} = -5.44, p < 0.0001$ ].

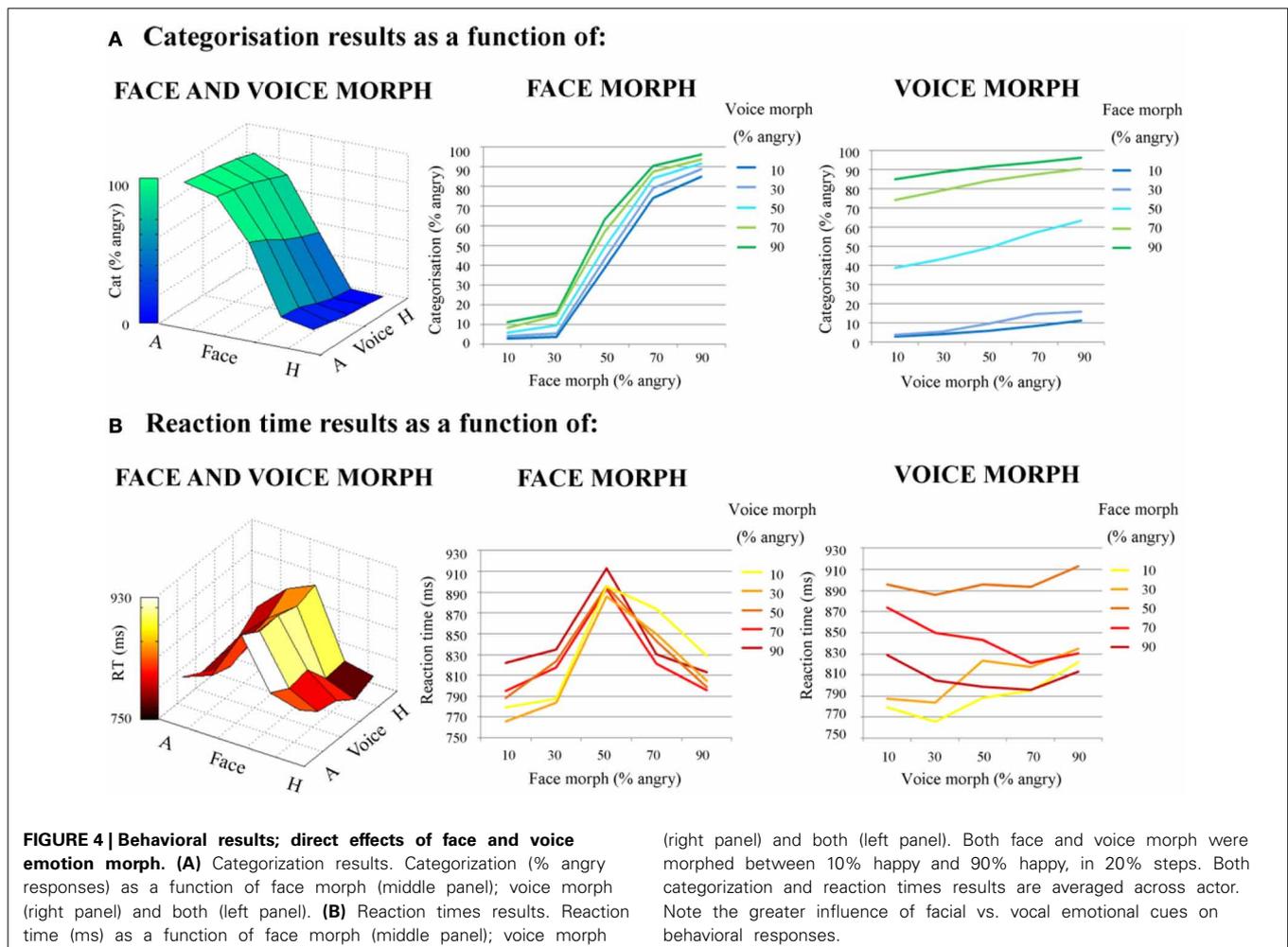
In a series of planned comparisons, we further examined at which points there were significant differences in categorization ratings between stimuli. We proposed that maximum incongruence between Face and Voice (i.e., 80% difference) would cause significant shifts in categorization, as compared to “end point” congruent stimuli (i.e., 10% angry face-10% angry voice; 90% angry face-90% angry voice). In order to test these hypotheses, we performed the following paired sample *t*-tests:

- (i) 10% angry face-10% angry voice vs. 10% angry face-90% angry voice
- (ii) 10% angry face-90% angry voice vs. 90% angry face-90% angry voice
- (iii) 90% angry face-90% angry voice vs. 90% angry face-10% angry voice
- (iv) 90% angry face-10% angry voice vs. 10% angry face-10% angry voice

After a Bonferroni correction for multiple comparisons (level of significance:  $p < 0.0125$ ), all comparisons were significant [ $t_{(17)} = -24.0, p < 0.0001$ ;  $t_{(17)} = -3.42, p < 0.004$ ;  $t_{(17)} = 27.6, p < 0.0001$ ;  $t_{(17)} = 2.87, p < 0.0125$ , respectively].

For an illustration of categorization results, refer to **Figure 4**.

**Reaction time data.** Each participant’s mean reaction time values for each stimulus (collapsed across actor) were firstly submitted to a two factor (face morph and voice morph), fully within subjects repeated measures ANOVA, with 5 levels per factor (percentage of “anger” information in the morph). As with categorical data, this was in order to assess the overall contribution of face and voice emotion morph—or the “direct effects” of face and voice morph—on reaction times.



The mean reaction times for the two end point congruent stimuli were 813 ms ( $\pm 67.4$  ms) and 779 ms ( $\pm 64.4$  ms) (for 10% angry face-10% angry voice and 90% angry face-90% angry voice, respectively). For the 50% angry face-50% angry voice stimulus the mean reaction time was 895 ms ( $\pm 100$  ms). Finally, for the two most incongruent stimuli (10% angry face-90% angry voice; 90% angry face-10% angry voice) these reaction times were 822 ms ( $\pm 101$  ms) and 829 ms ( $\pm 92.5$  ms). The ANOVA of reaction time data highlighted a main effect of voice morph [ $F_{(2.91, 49.6)} = 11.8, p < 0.0001, \eta_p^2 = 0.409$ ] and of face morph [ $F_{(2.34, 39.7)} = 70.6, p < 0.0001, \eta_p^2 = 0.806$ ], and also a significant interaction between the two modalities [ $F_{(2.90, 39.4)} = 7.40, p < 0.0001, \eta_p^2 = 0.303$ ].

*Post-hoc*, we compared reaction time values between each of our completely congruent stimuli (i.e., 10% angry face-10% angry voice; 30% angry face-30% angry voice; 50% angry face-50% angry voice; 70% angry face-70% angry voice; 90% angry face-90% angry voice) in five paired *t*-tests. Each stimulus was compared to the next one (i.e., 10% angry face-10% angry voice vs. 30% angry face-30% angry voice; 30% angry face-30% angry voice vs. 50% angry face-50% angry voice and so on) to see whether each stimulus significantly differed from the other with regards to reaction time. After a Bonferroni correction for multiple comparisons (level of significance:  $p < 0.01$ ), the following comparisons were significant—30% angry face-30% angry voice vs. 50% angry face-50% angry voice:  $t_{(17)} = -7.74, p < 0.0001$ ; 50% angry face-50% angry voice vs. 70% angry face-70% angry voice:  $t_{(17)} = 7.67, p < 0.0001$ . The following comparisons were not significant—10% angry face-10% angry voice vs. 30% angry face-30% angry voice:  $t_{(17)} = -0.638, p = 0.532$ ; 70% angry face-70% angry voice vs. 90% angry face-90% angry voice:  $t_{(17)} = 0.904, p = 0.379$ .

As in our categorization analysis, we proposed that maximum incongruence between Face and Voice (i.e., 80% difference) would take significantly longer to categorize, as compared to “end point” congruent stimuli. However, we also expected that some stimuli that were congruent, but with a lower clarity value (i.e., 50% angry face-50% angry voice), would take longer to categorize than end-point congruent stimuli. In order to test these hypotheses, we performed the following paired sample *t*-tests:

- (i) 10% angry face-10% angry voice vs. 10% angry face-90% angry voice
- (ii) 10% angry face-90% angry voice vs. 90% angry face-90% angry voice
- (iii) 90% angry face-90% angry voice vs. 90% angry face-10% angry voice
- (iv) 90% angry face-10% angry voice vs. 10% angry face-10% angry voice
- (v) 50% angry face-50% angry voice vs. 10% angry face-10% angry voice
- (vi) 50% angry face-50% angry voice vs. 90% angry face-90% angry voice

After a Bonferroni correction for multiple comparisons (level of significance:  $p < 0.0125$ ), comparisons (i), (iv), (v), and (vi) were significant [ $t_{(17)} = -4.72, p < 0.0001$ ;  $t_{(17)} = 3.25, p < 0.006$ ;

$t_{(17)} = 10.67, p < 0.0001$ ;  $t_{(17)} = 6.29, p < 0.0001$ , respectively], but comparisons (ii) and (iii) were not [ $t_{(17)} = 1.30, p = 0.210$ ;  $t_{(17)} = -5.80, p = 0.569$ , respectively].

For an illustration of reaction time results, refer to **Figure 4**.

### Effect of stimulus clarity and incongruence

We also computed a multiple regression analysis to investigate the relative contribution of stimulus and incongruence of our audiovisual stimulus on the reaction times in individual subjects. This analysis confirmed that clarity was significantly related to reaction time ( $\beta = -18.7, t = -4.43, p < 0.0001$ ), with a lower level of clarity resulting in longer reaction times, but that incongruence was not ( $\beta = -7.78, t = -1.83, p = 0.067$ ).

## fMRI RESULTS

### Clarity

After removing the variance associated with incongruence, a positive effect of clarity was found in the right STG/superior temporal sulcus (**Figure 5, Table 1A**). These regions were more active when the audiovisual stimulus was clear. A negative effect was observed in the anterior cingulate gyrus, extending to the supplementary motor area—here, there was greater activation for the more unclear types of stimuli (**Figure 5, Table 1B**).

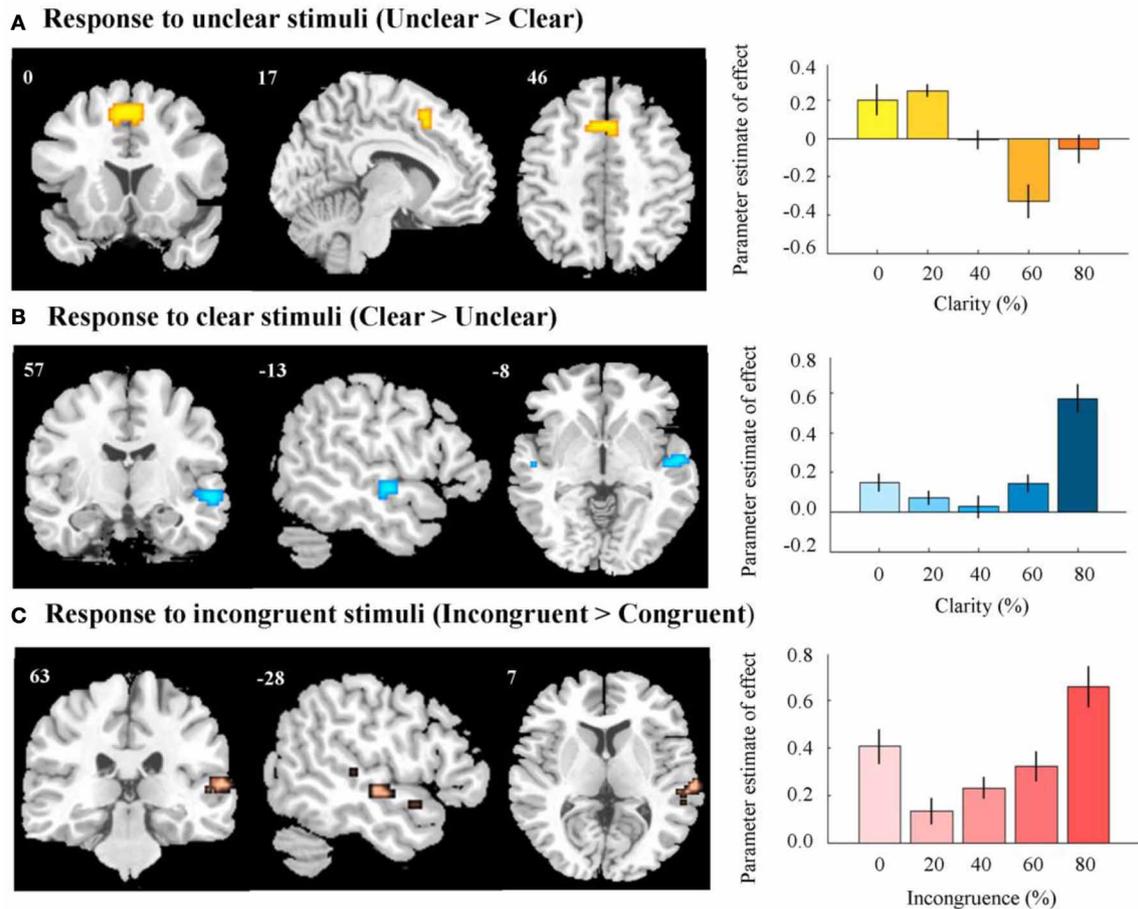
### Incongruence

After the variance associated with clarity values was regressed out, we found a positive effect of incongruence across a wide region of the right STG/STS (**Figure 5, Table 1C**). This region appeared to respond more to incongruent information, as compared to congruent. We observed no negative effect of incongruence (i.e., congruent > incongruent), even at a relatively liberal threshold [ $p < 0.005$  (voxel uncorrected)] (**Table 1D**).

## DISCUSSION

In the present study we used visual and auditory morphing technologies to generate a range of face-voice stimuli parametrically varying in emotion, in conjunction with a continuous carry-over design so to examine the cerebral correlates of face-voice affect perception. Specifically, our main aim was to investigate the multimodal representation of emotion, and potential response to affective conflict, by observing the neural response to emotional incongruency in the face and voice. Furthermore, our intention was to investigate these effects independent of task difficulty, which has not yet been achieved in previous studies.

We firstly observed that emotion categorization, and speed of categorization, were modulated in line with parametric shifts in affective content of the face and voice: that is, the specific degree of morph of the face-voice stimulus had a direct effect on how angry or happy the participant viewed it, and as the information in the combined stimulus became increasingly unclear, the stimulus took longer to classify. Significantly, both modalities affected emotion perception—an integration effect—but face morph exerted a far larger influence on behavioral responses, both categorical and reaction times. This infers that participants found the faces in this study easier to categorize with regards to emotion as compared to voices. This is in line with other studies, where categorization has consistently been found to be



**FIGURE 5 | fMRI results. (A)** Activation in anterior cingulate and supplementary motor area in response to unclear, compared to clear information. Right panel indicates response in peak activated voxel as a result of increasing level of clear information. **(B)** Activation in right STG/STS in response to clear, compared to unclear information. Right panel indicates response in peak activated voxel as a result of increasing level of clear information. **(C)** Activation in right STG/STS in response to incongruent, compared to congruent information. Right panel indicates

response in peak activated voxel as a result of increasing level of incongruence. In all right panels, error bars represent the standard error of the parameter estimate of effect. Contrasts were thresholded to display voxels reaching a significance level of  $p < 0.05$  (FWE voxel-corrected), and an additional minimum cluster size of greater than 5 contiguous voxels. Contrasts were masked by an AV vs. baseline contrast thresholded at  $p < 0.001$  (uncorrected). MNI coordinates and t-scores are from the peak voxel of a cluster.

more accurate and quicker for faces than to voices (e.g., Hess et al., 1988; de Gelder and Vroomen, 2000; Kreifelts et al., 2007; Collignon et al., 2008; Bänziger et al., 2009), although it should be noted that this will naturally vary dependent on the specific stimuli used from study to study.

We then investigated the effect of both stimulus clarity and incongruence on reaction time. Values for each of these dimensions were assigned based on where each stimulus lay in the  $5 \times 5$  audiovisual emotion space: incongruence values related to the degree of discordance between the emotion displayed in the face and voice, whereas clarity values referred to how clear the affective information in the *combined* stimulus was.

We observed a significant effect of stimulus clarity on reaction time, with the more unclear stimuli taking longer to categorize. However, there was no significant effect of stimulus incongruence on reaction time. In similar studies it has been observed that generally, the greater the incongruence between

face and voice, the more time it takes to classify the emotion (e.g., Massaro and Egan, 1996; de Gelder and Vroomen, 2000). However, due to the novel morphing procedure in our study some stimuli that were completely congruent would still have proved difficult for our participants to categorize—for example, those that had a pairing of ambiguous information in both the face and the voice. Thus, in this study it is unsurprising that the level of stimulus clarity was more reflective of task difficulty. This result meant we were able to take stimulus clarity as an indicator of task difficulty, and use these values to disentangle task difficulty from any observed incongruence effects.

At the cerebral level, we observed that there was an effect of both stimulus clarity and incongruence on brain activity. Firstly, we observed a negative effect of clarity in the anterior cingulate gyrus, extending to the supplemental motor area (SMA). In these latter regions, there was heightened activation in response

**Table 1 | Effects of stimulus clarity and incongruence.**

Brain regions	Coordinates (mm)			k	t-statistic
	x	y	z		
<b>(A) CLARITY (POSITIVE EFFECT; CLEAR &gt; UNCLEAR)</b>					
Superior temporal gyrus/ Superior temporal sulcus	57	-13	-8	35	9.14
<b>(B) CLARITY (NEGATIVE EFFECT; UNCLEAR &gt; CLEAR)</b>					
Cingulate gyrus/ Supplementary motor area	0	17	46	70	8.92
<b>(C) INCONGRUENCE (POSITIVE EFFECT; INCONGRUENT &gt; CONGRUENT)</b>					
Superior temporal gyrus/ Superior temporal sulcus	63	-28	7	25	6.96
Superior temporal gyrus/ Superior temporal sulcus	54	-13	-8	24	6.60
<b>(D) INCONGRUENCE (NEGATIVE EFFECT; CONGRUENT &gt; INCONGRUENT)</b>					
<b>NO SIGNIFICANT CLUSTERS</b>					

A,B. Positive and negative effects of clarity value of stimulus; C,D. Positive and negative effects of incongruence value of stimulus. Contrasts were thresholded to display voxels reaching a significance level of  $p < 0.05$  (FWE voxel-corrected), and an additional minimum cluster size of greater than 5 contiguous voxels. Contrasts were masked by an AV vs. baseline contrast thresholded at  $p < 0.001$  (uncorrected). MNI coordinates and t-scores are from the peak voxel of a cluster.

to unclear stimuli (i.e., stimuli that were harder to categorize), as compared to clear stimuli.

In the study of non-emotional conflict, the cingulate gyrus [particularly, the anterior cingulate cortex (ACC)] is amongst the brain regions most frequently reported as being significantly activated when engaging in attentionally or behaviorally demanding cognitive tasks (Paus et al., 1998). A number of studies have also implicated this region in the detection of conflict between different possible responses to a stimulus, event, or situation (e.g., Carter et al., 1999; Kerns et al., 2004; Wendelken et al., 2009). The results of these studies have led to the conflict monitoring hypothesis, which suggests that conflict is detected by the dorsal ACC, which in turn recruits prefrontal regions to increase cognitive control (Botvinick et al., 2004; Kerns et al., 2004; Carter and van Veen, 2007). In our study, stimuli of an unclear or ambiguous nature were more demanding to categorize, requiring more energy for decision making, and thus it is unsurprising that we observed heightened activity in the cingulate in response to this information.

With regards to affective conflict, activation in ACC has been observed for (within-modality) conflicts in the visual and auditory domain (Haas et al., 2006; Ochsner et al., 2009; Wittfoth et al., 2010). Additionally, emotional conflict has been linked to the SMA, a region which plays a major role in voluntary action, cognitive control and initiation/inhibition of motor responses (Sumner et al., 2007; Grefkes et al., 2008; Kasess et al., 2008; Nachev et al., 2008). For example, SMA activation was found for emotional conflict in a study by Ochsner et al. (2009) in an affective flanker task, and in the previously cited study of Müller et al. (2011). These authors suggest that higher SMA activity may reflect increased executive control needed to select

an adequate response in the presence of conflicting (emotional) stimuli.

Interestingly, incongruent stimuli did not elicit activation in these regions, as compared to congruent stimuli. As the clarity value of the stimulus was more linked to task difficulty than the incongruence value, we suggest that the cingulate and SMA respond specifically when there is difficulty in classifying a stimulus, as opposed to incongruence *per se*.

Instead, we observed a positive effect of incongruence across the bilateral STG/STS, in addition to a positive effect of clarity. Such regions have been implicated in auditory-visual processing and multisensory integration for both speech and non-speech stimuli (Calvert et al., 2000; Sekiyama et al., 2003; Beauchamp, 2005; Miller and D'Esposito, 2005). In overlapping regions, there was an increase in activation in response to stimuli that were by nature clear, and interestingly, also an increase in response to stimuli that were classified as incongruent.

With regards to incongruence, we might have expected that this pattern would be the reverse. One of the initial claims for the STS as an audiovisual binding site came from Calvert et al. (2000) who contrasted audiovisual speech to each modality in isolation (i.e., heard words or silent lip-reading). This revealed a super-additive response (i.e., a heightened response relative to the sum of the responses of audio and visual speech information presented alone) in the left pSTS when the audiovisual input was congruent but a sub-additive response when the audiovisual input was incongruent (i.e., showing a reduced response relative to the sum of the responses of audio and visual speech information presented alone). Moving from speech to emotion, Klasen et al. (2011) also found a stronger response to congruent vs. incongruent information in the amygdala and posterior cingulate, leading the authors to propose that these regions may be involved in integrating affective information. In contrast, we did not find any regions that responded more to congruent vs. incongruent information.

However, it should be noted that a number of studies have also produced conflicting results. Indeed, Hocking and Price (2008) stated that at that time they were unable to find any studies that replicated the Calvert et al. (2000) study showing enhanced pSTS activation for congruent relative to incongruent bimodal stimuli. In an fMRI study of the "McGurk effect"—a famous perceptual phenomenon observed in speech perception, where incompatible face-voice information leads to illusory percepts—conducted by Jones and Callan (2003), greater responses in the STS/STG for congruent audiovisual stimuli were not observed over incongruent audiovisual stimuli, as one might predict for a multisensory integration site. With regards to emotional incongruence, Müller et al. (2011) also did not observe a greater effect of congruent affective information over incongruent information in this region, or any others.

Hocking and Price (2008) suggest that potentially, one reason for the inconsistent congruency effects could be due to the fact that attention to one modality only during bimodal presentation elicits sub-additive effects (Talsma and Woldorff, 2005; Talsma et al., 2007). They argue that to minimize interference during incongruent audiovisual speech streams, participants may automatically or attentionally reduce visual processing (Deneve and Pouget, 2004; Ernst and Bulthoff, 2004), particularly in the study

of Calvert et al. (2000) where congruent and incongruent conditions were presented in separate experiments with no instructions to attend to the visual stimuli. This would explain the absence of congruency effects in studies that presented brief stimuli or forced participants to attend to the visual input during incongruent audiovisual conditions.

Hocking and Price (2008) found that when task and stimulus presentation were controlled, a network of regions, including the pSTS, were activated more strongly for incongruent than congruent pairs of stimuli (stimuli were color photographs of objects, their written names, their auditory names and their associated environmental sounds). They suggest that activation reflects processing demand which is greater when two simultaneously presented stimuli refer to different concepts (as in the incongruent condition) than when two stimuli refer to the same object (the congruent condition). They also hypothesize that if participants were able to attend to one input modality whilst suppressing the other, then pSTS activation would be less for incongruent bimodal trials. In contrast, if subjects were forced to attend to both modalities then the pSTS activation would be higher for incongruent bimodal trials that effectively carry twice the information content as congruent trials.

In our study, a key point should be noted: values assigned to stimuli (specifically, those indicating incongruence) on the basis of the face and voice morph information were not necessarily reflective of the perceptual difficulty of classifying emotion. The incongruence value related to the degree of discordance between affect in the face and voice, whereas the clarity value related to how clear the overall, combined face-voice information was. Importantly, only clarity values were correlated with reaction times: the more unclear the combined information in the audiovisual stimulus was, the longer it took to classify. Although some incongruent stimuli resulted in shorted reaction times (e.g., 10% angry face-10% angry voice), some did not (i.e., 50% angry face-50% angry voice). Therefore, we can suggest that the heightened response to incongruent information across the right STS was not due to the perceptual difficulty of classifying the stimulus or processing demand.

In our study participants were instructed to attend both modalities: although we cannot be sure that participants definitely attended to both modalities in the incongruent trials, our behavioral data does suggest they did integrate the two modalities to some degree (indicated by a significant interaction between Face and Voice emotion morph for both categorical and reaction time data, in addition to a main effect of both modality). Therefore, in line with the proposal of Hocking and Price (2008), a tentative explanation is that participants were attending to both modalities and thus the STS activation was higher for incongruent bimodal trials. It is important to note that this is not necessarily reflective of greater perceptual difficulty in categorization (i.e., task difficulty). Rather, we propose this could be caused by the pure recognition that the auditory and visual inputs were different—an error detection.

Finally, Klasen et al. (2011) argue that incongruent emotional information cannot be successfully integrated into a bimodal emotional percept, and propose that regions responding more to congruent information than incongruent are reflective of

an integrative process. However, Campanella and Belin (2007) suggested that conversely, it may be possible for incompatible affective information in the face and voice to be combined in such a way as to create an entirely new emotional percept, one independent of information contained in either modality—an “emotional McGurk effect.” This would imply some form of audiovisual integration, although perhaps one with a nature and mechanisms entirely different from the integration of emotionally congruent information. We are far from being able to conclusively answer this question; nonetheless, our results point to a strong activation in the STG/STS region in response to incongruent information that cannot be explained simply by task difficulty. We suggest that this instead could be due to an audiovisual mismatch detection, underlying the important role of the STG/STS in audiovisual processing.

## REFERENCES

- Aguirre, G. K. (2007). Continuous carry-over designs for fMRI. *Neuroimage* 35, 1480–1494. doi: 10.1016/j.neuroimage.2007.02.005
- Bänziger, T., Grandjean, D., and Scherer, K. R. (2009). Emotion recognition from expressions in face, voice, and body: the multimodal emotion recognition test (MERT). *Emotion* 9, 691–704. doi: 10.1037/a0017088
- Beauchamp, M. S. (2005). Statistical criteria in fMRI studies of multisensory integration. *Neuroinformatics* 3, 93–113. doi: 10.1385/NI:3:2:093
- Besl, P. J., and McKay, N. D. (1992). A method for registration of 3-D shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* 14, 239–255. doi: 10.1109/34.121791
- Bolles, R. C., and Fischler, M. A. (1981). “A RANSAC-based approach to model fitting and its application to finding cylinders in range data,” in *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI)*, (Vancouver, BC), 637–643.
- Botvinick, M., Cohen, J. D., and Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn. Sci.* 8, 539–546. doi: 10.1016/j.tics.2004.10.003
- Calvert, G. A., Campbell, R., and Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr. Biol.* 10, 649–657. doi: 10.1016/S0960-9822(00)00513-3
- Campanella, S., and Belin, P. (2007). Integrating face and voice in person perception. *Trends Cogn. Sci.* 11, 535–543. doi: 10.1016/j.tics.2007.10.001
- Carter, C. S., Botvinick, M. M., and Cohen, J. D. (1999). The contribution of anterior cingulate cortex to executive processes in cognition. *Rev. Neurosci.* 10, 49–57. doi: 10.1515/REVNEURO.1999.10.1.49
- Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., and Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science* 280, 747–749. doi: 10.1126/science.280.5364.747
- Carter, C. S., and van Veen, V. (2007). Anterior cingulate and conflict detection: An update of theory and data. *Cogn. Affect. Behav. Neurosci.* 7, 367–379. doi: 10.3758/CABN.7.4.367
- Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., and Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain Res.* 25, 126–135. doi: 10.1016/j.brainres.2008.04.023
- de Gelder, B., and Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cogn. Emotion* 14, 289–311. doi: 10.1080/026999300378824
- Deneve, S., and Pouget, A. (2004). Bayesian multisensory integration and cross-modal spatial links. *J. Physiol.* 98, 249–258. doi: 10.1016/j.jphysparis.2004.03.011
- Dolan, R. J., Morris, J. S., and de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proc. Natl. Acad. Sci. U.S.A.* 98, 10006–10010. doi: 10.1073/pnas.171288598
- Durston, S., Davidson, M. C., Thomas, K. M., Worden, M. S., Tottenham, N., Martinez, A., et al. (2003). Parametric manipulation of conflict and response competition using rapid mixed-trial event-related fMRI. *Neuroimage* 20, 2135–2141. doi: 10.1016/j.neuroimage.2003.08.004
- Ernst, M. O., and Bulthoff, H. H. (2004). Merging the senses into a robust percept. *Trends Cogn. Sci.* 8, 162–169. doi: 10.1016/j.tics.2004.02.002

- Ethofer, T., Pourtois, G., and Wildgruber, D. (2006a). Investigating audiovisual integration of emotional signals in the human brain. *Prog. Brain Res.* 156, 345–361. doi: 10.1016/S0079-6123(06)56019-4
- Ethofer, T., Anders, S., Erb, M., Droll, C., Royen, L., Saur, R., et al. (2006b). Impact of voice on emotional judgment of faces: an event-related fMRI study. *Hum. Brain Mapp.* 27, 707–714. doi: 10.1002/hbm.20212
- Grefkes, C., Eickhoff, S. B., Nowak, D. A., Dafotakis, M., and Fink, G. R. (2008). Dynamic intraand interhemispheric interactions during unilateral and bilateral hand movements assessed with fMRI and DCM. *Neuroimage* 41, 1382–1394. doi: 10.1016/j.neuroimage.2008.03.048
- Haas, B. W., Omura, K., Constable, R. T., and Canli, T. (2006). Interference produced by emotional conflict associated with anterior cingulate activation. *Cogn. Affect. Behav. Neurosci.* 6, 152–156. doi: 10.3758/CABN.6.2.152
- Hess, U., Kappas, A., and Scherer, K. (1988). “Multichannel communication of emotion: synthetic signal production,” in *Facets of Emotion: Recent Research*, ed K. Scherer (Hillsdale, NJ: Erlbaum), 161–182.
- Hocking, J., and Price, C. J. (2008). The role of the posterior superior temporal sulcus in audiovisual processing. *Cereb. Cortex* 18, 2439–2449. doi: 10.1093/cercor/bhn007
- Jones, J. A., and Callan, D. E. (2003). Brain activity during audiovisual speech perception: an fMRI study of the McGurk effect. *Neuroreport* 14, 1129–1133. doi: 10.1097/00001756-200306110-00006
- Kassess, C. H., Windischberger, C., Cunnington, R., Lanzenberger, R., Pezawas, L., and Moser, E. (2008). The suppressive influence of SMA on M1 in motor imagery revealed by fMRI and dynamic causal modeling. *Neuroimage* 40, 828–837. doi: 10.1016/j.neuroimage.2007.11.040
- Kawahara, H. (2003). “Exemplar-based voice quality analysis and control using a high quality auditory morphing procedure based on straight,” in *VoQual 03: Voice Quality: Functions, Analysis, and Synthesis* (Geneva: ISCA Tutorial and Research Workshop).
- Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A., and Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science* 303, 1023–1026. doi: 10.1126/science.1089910
- Klasen, M., Kenworthy, C. A., Mathiak, K. A., Kircher, T. T., and Mathiak, K. (2011). Supramodal representation of emotions. *J. Neurosci.* 31, 13635–13643. doi: 10.1523/JNEUROSCI.2833-11.2011
- Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., and Wildgruber, D. (2007). Audiovisual integration of emotional signals in voice and face: an event-related fMRI study. *Neuroimage* 37, 1445–1456. doi: 10.1016/j.neuroimage.2007.06.020
- Kreifelts, B., Ethofer, T., Huberle, E., Grodd, W., and Wildgruber, D. (2010). Association of trait emotional intelligence and individual fMRI activation patterns during the perception of social signals from voice and face. *Hum. Brain Mapp.* 31, 979–991. doi: 10.1002/hbm.20913
- Kreifelts, B., Ethofer, T., Shiozawa, T., Grodd, W., and Wildgruber, D. (2009). Cerebral representation of non-verbal emotional perception: fMRI reveals audiovisual integration area between voice- and face-sensitive regions in the superior temporal sulcus. *Neuropsychologia* 47, 3059–3066. doi: 10.1016/j.neuropsychologia.2009.07.001
- MacDonald, A. W. I., Cohen, J. D., Stenger, V. A., and Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science* 288, 1835–1838. doi: 10.1126/science.288.5472.1835
- Massaro, D. W., and Egan, P. B. (1996). Perceiving affect from the voice and the face. *Psychon. Bull. Rev.* 3, 215–221. doi: 10.3758/BF03212421
- Miller, L. M., and D’Esposito, M. (2005). Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. *J. Neurosci.* 25, 5884–5893. doi: 10.1523/JNEUROSCI.0896-05.2005
- Müller, V. I., Habel, U., Derntl, B., Schneider, F., Zilles, K., Turetsky, B. I., et al. (2011). Incongruence effects in crossmodal emotional integration. *Neuroimage* 54, 2257–2266. doi: 10.1016/j.neuroimage.2010.10.047
- Nachev, P., Kennard, C., and Husain, M. (2008). Functional role of the supplementary and presupplementary motor areas. *Nat. Rev. Neurosci.* 9, 856–869. doi: 10.1038/nrn2478
- Nonyane, B. A. S., and Theobald, C. M. (2007). Design sequences for sensory studies: achieving balance for carry-over and position effects. *Br. J. Math. Stat. Psychol.* 60, 339–349. doi: 10.1348/000711006X114568
- Ochsner, K. N., Hughes, B., Robertson, E. R., Cooper, J. C., and Gabrieli, J. D. (2009). Neural systems supporting the control of affective and cognitive conflicts. *J. Cogn. Neurosci.* 21, 1842–1855. doi: 10.1162/jocn.2009.21129
- Paus, T., Koski, L., Caramanos, Z., and Westbury, C. (1998). Regional differences in the effects of task difficulty and motor output on blood flow response in the human anterior cingulate cortex: a review of 107 PET activation studies. *Neuroreport* 9, R37–R47. doi: 10.1097/00001756-199806220-00001
- Pourtois, G., de Gelder, B., Bol, A., and Crommelinck, M. (2005). Perception of facial expressions and voices and of their combination in the human brain. *Cortex* 41, 49–59. doi: 10.1016/S0010-9452(08)70177-1
- Robins, D. L., Hunyadi, E., and Schultz, R. T. (2009). Superior temporal activation in response to dynamic audio-visual emotional cues. *Brain Cogn.* 69, 269–278. doi: 10.1016/j.bandc.2008.08.007
- Sekiya, K., Kanno, I., Miura, S., and Sugita, Y. (2003). Auditory-visual speech perception examined by fMRI and PET. *Neurosci. Res.* 47, 277–287. doi: 10.1016/S0168-0102(03)00214-1
- Sumner, P., Nachev, P., Morris, P., Peters, A. M., Jackson, S. R., Kennard, C., et al. (2007). Human medial frontal cortex mediates unconscious inhibition of voluntary action. *Neuron* 54, 697–711. doi: 10.1016/j.neuron.2007.05.016
- Talsma, D., Doty, T. J., and Woldorff, M. G. (2007). Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration. *Cereb. Cortex* 17, 679–690. doi: 10.1093/cercor/bhk016
- Talsma, D., and Woldorff, M. G. (2005). Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. *J. Cogn. Neurosci.* 17, 1098–1114. doi: 10.1162/0898929054475172
- Weissman, D. H., Warner, L. M., and Woldorff, M. G. (2004). The neural mechanisms for minimizing cross-modal distraction. *J. Neurosci.* 24, 10941–10949. doi: 10.1523/JNEUROSCI.3669-04.2004
- Wendelken, C., Ditterich, J., Bunge, S. A., and Carter, C. S. (2009). Stimulus and response conflict processing during perceptual decision making. *Cogn. Affect. Behav. Neurosci.* 9, 434–447. doi: 10.3758/CABN.9.4.434
- Winder, J., Darvann, T. A., McKnight, W., Magee, J. D. M., and Ramsay-Baggs, P. (2008). Technical validation of the Di3D stereophotogrammetry surface imaging system. *Br. J. Oral Maxillofac. Surg.* 46, 33–37. doi: 10.1016/j.bjoms.2007.09.005
- Wittfoth, M., Schroder, C., Schardt, D. M., Dengler, R., Heinze, H. J., and Kotz, S. A. (2010). On emotional conflict: interference resolution of happy and angry prosody reveals valence-specific effects. *Cereb. Cortex* 20, 383–392. doi: 10.1093/cercor/bhp106

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 01 June 2013; accepted: 18 October 2013; published online: 13 November 2013.

Citation: Watson R, Latinus M, Noguchi T, Garrod O, Crabbe F and Belin P (2013) Dissociating task difficulty from incongruence in face-voice emotion integration. *Front. Hum. Neurosci.* 7:744. doi: 10.3389/fnhum.2013.00744

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2013 Watson, Latinus, Noguchi, Garrod, Crabbe and Belin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

## APPENDIX

### PRE-TEST: STIMULUS VALIDATION

In a pre-test, using the separate group of ten participants, we investigated categorization of our stimuli across the two actors, firstly in order to ensure that expressions were recognized as intended, and secondly to clarify that there were no significant differences in categorization of expressions produced by different actors. Five participants were assigned to the expressions of the male actor, and another five were assigned to the expressions of the female actor. The stimuli were played to participants through a FLASH ([www.adobe.com](http://www.adobe.com)) object interface running on the Mozilla Firefox web browser. For each condition, stimuli were preloaded prior to running the experiment. The conditions were as follows:

#### **Audio only**

In this condition, participants heard a series of voices alone. They were instructed to listen to each voice, and make a forced choice decision on emotion based on the voice they had just heard, where the responses were either “Angry” or “Happy.” Again they indicated their decision via a button press. The five voice morphs were presented 10 times each, in a randomized order in one block consisting of 50 trials.

#### **Video only**

Participants saw all face videos, uncoupled with a voice. They were instructed to watch the screen and indicate their decision regarding emotion in the same way as before. The five faces were presented 10 times each, in randomized order in one block consisting of 50 trials.

Participants could respond either whilst the stimulus was playing, or after it ended. Regardless of when they responded, there was a 100ms wait until the next stimulus began playing. Conditions were counterbalanced between participants, with five participants for each of the two possible orders (collapsing across actor gender).

Categorization data was submitted to two, two factor mixed ANOVAs. In the first, degree of face emotion morph was a within subject factor, whilst the actor (male or female) was a between subject factor. This analysis highlighted a significant effect of face emotion morph on categorization [ $F_{(1.35, 10.8)} = 126, p < 0.0001$ ]. There was no effect of actor on categorization [ $F_{(1, 8)} = 0.949, p = 0.359$ ]. Face categorization results (averaged across actors) yielded the classic sigmoid-like psychometric function from the emotion classification task, with a steeper slope at central portions of the continua. The percentages of anger identification were 96% ( $\pm 2.23\%$ ) for the 90% angry face, and 2% ( $\pm 2.74\%$ ) for the 90% happy face. The 50% angry-happy face was identified as angry 53 times out of 100 ( $\pm 9.75\%$ ). In the second ANOVA, degree of voice emotion morph was a within subject factor, whilst the actor was the between subject factor. This analysis highlighted a significant effect of voice emotion morph on categorization [ $F_{(2.23, 17.7)} = 127, p < 0.0001$ ]. There was no effect of actor on categorization [ $F_{(1, 8)} = 0.949, p = 0.575$ ]. Again, voice categorization results (averaged across actors) yielded the classic sigmoid-like psychometric function from the emotion classification task, with a steeper slope at central portions of the continua. The percentages of anger identification were 96% ( $\pm 6.52\%$ ) for the 90% angry voice, and 0% ( $\pm 0.00\%$ ) for the 90% happy voice. The 50% angry-happy voice was identified as angry 32 times out of 100 ( $\pm 19.4\%$ ).



# Multisensory integration of dynamic emotional faces and voices: method for simultaneous EEG-fMRI measurements

Patrick D. Schelenz<sup>1,2\*</sup>, Martin Klasen<sup>1,2</sup>, Barbara Reese<sup>1,2</sup>, Christina Regenbogen<sup>1,2</sup>, Dhana Wolf<sup>1,2</sup>, Yutaka Kato<sup>1,3</sup> and Klaus Mathiak<sup>1,2</sup>

<sup>1</sup> Department of Psychiatry, Psychotherapy, and Psychosomatics, Medical School, Rheinisch-Westfälische Technische Hochschule Aachen University, Aachen, Germany

<sup>2</sup> Jülich Aachen Research Alliance, Translational Brain Medicine, Aachen, Germany

<sup>3</sup> Department of Neuropsychiatry, Keio University School of Medicine, Tokyo, Japan

## Edited by:

Benjamin Kreifelts, University of Tübingen, Germany

## Reviewed by:

Andy P. Bagshaw, University of Birmingham, UK

Pierre LeVan, University Medical Center Freiburg, Germany

## \*Correspondence:

Patrick D. Schelenz, Department of Psychiatry, Psychotherapy, and Psychosomatics, Medical School, Rheinisch-Westfälische Technische Hochschule Aachen University, Pauwelsstraße 30, Aachen, Germany  
e-mail: patrick.schelenz@rwth-aachen.de

Combined EEG-fMRI analysis correlates time courses from single electrodes or independent EEG components with the hemodynamic response. Implementing information from only one electrode, however, may miss relevant information from complex electrophysiological networks. Component based analysis, in turn, depends on a priori knowledge of the signal topography. Complex designs such as studies on multisensory integration of emotions investigate subtle differences in distributed networks based on only a few trials per condition. Thus, they require a sensitive and comprehensive approach which does not rely on a-priori knowledge about the underlying neural processes. In this pilot study, feasibility and sensitivity of source localization-driven analysis for EEG-fMRI was tested using a multisensory integration paradigm. Dynamic audiovisual stimuli consisting of emotional talking faces and pseudowords with emotional prosody were rated in a delayed response task. The trials comprised affectively congruent and incongruent displays. In addition to event-locked EEG and fMRI analyses, induced oscillatory EEG responses at estimated cortical sources and in specific temporo-spectral windows were correlated with the corresponding BOLD responses. EEG analysis showed high data quality with less than 10% trial rejection. In an early time window, alpha oscillations were suppressed in bilateral occipital cortices and fMRI analysis confirmed high data quality with reliable activation in auditory, visual and frontal areas to the presentation of multisensory stimuli. In line with previous studies, we obtained reliable correlation patterns for event locked occipital alpha suppression and BOLD signal time course. Our results suggest a valid methodological approach to investigate complex stimuli using the present source localization driven method for EEG-fMRI. This novel procedure may help to investigate combined EEG-fMRI data from novel complex paradigms with high spatial and temporal resolution.

**Keywords: emotion, audiovisual integration, emotion integration, methods for EEG-fMRI, affective neuroscience, EEG-fMRI, perceptual processing**

## INTRODUCTION

Combined EEG-fMRI investigates simultaneously neural activity at high spatial and temporal resolution (Mulert and Lemieux, 2010; Ullsperger and Debener, 2010). Early EEG-fMRI studies by Goldman et al. (2002) and Laufs et al. (2003) on the relationship between alpha power and BOLD signal investigated whether alpha power and BOLD signal are related. They correlated the time-series of occipital alpha during resting state with BOLD signal changes and reported an inverse correlation between occipital alpha power and BOLD signal in visual areas. Similar results in the visual cortex have been replicated in recent studies (Becker et al., 2011; Mayhew et al., 2013; Mo et al., 2013). In these studies, involvement of alpha oscillations in working memory (Scheeringa et al., 2009), linear superimposition in visual cortex (Becker et al., 2011) and default mode network (Mayhew et al., 2013; Mo et al., 2013) was investigated.

So far, two methods have been used to integrate alpha power in EEG-fMRI studies: correlation of single electrodes and of EEG components with the BOLD response. Single trial correlations of alpha power of single electrodes have been subject to several studies (Goldman et al., 2002; Laufs et al., 2003; Mo et al., 2013). They investigated neural correlations of occipital alpha oscillations in resting state with prior knowledge about the topography of the EEG signal. Other studies investigate neural correlates of ERPs and correlate single trial variation of one electrode with BOLD response. The EEG signal on the scalp can derive from several sources. Thus, this approach may impact the investigation of networks that are related to that scalp signal as the neural correlates of each electrophysiological source can not be identified.

Scheeringa et al. (2009) used a modified Sternberg paradigm to investigate the neural correlates of posterior alpha

power increase during working memory maintenance. They conducted standard artifact reduction in EEG data and employed an independent component analysis (ICA). The component reflecting posterior alpha increase was chosen individually based on its topographical distribution. This was the first study to investigate single-trial coupling of alpha power and BOLD signal in a paradigm. However, subtle differences between different conditions may not be separable using ICA. Furthermore, if the topographical distribution of alpha power is unknown a priori, an ICA based approach to integrate EEG and fMRI data is not possible: Without a predefined topography, the component of interest cannot be identified. Therefore, we wanted to establish a way to identify neural correlates of electrophysiological oscillations during multisensory integration of emotions.

For complex stimuli – such as dynamic emotional multisensory stimuli, the exact topography of alpha oscillations – and particularly the subtle differences between emotional congruent (CON) and incongruent (INC) audio-visual stimuli is not known a priori. One solution for that is to use a source localization driven approach to identify those small distinctions and then combine EEG and fMRI data. In this pilot study, we will present a data-driven approach using EEG source localisation to analyse such complex multisensory information using EEG-fMRI. In a previous fMRI study on the integration of multisensory emotional stimuli, Klasen et al. (2011) compared CON and INC combinations of emotional dynamic faces and disyllabic pseudowords. They reported reduced workload on a fronto-parietal attention network for emotionally CON multisensory stimuli. In a magnetoencephalography (MEG) study, Chen et al. (2010) combined dynamic emotional faces with affective pseudowords. They reported increased alpha (8–13 Hz) power 200–400 ms after stimulus onset in frontal areas when comparing affectively CON multisensory to unisensory trials and concluded that multisensory integration of emotions occurred in higher order areas rather than in unisensory ones. Both studies report involvement of frontal areas during multisensory integration but lack either temporal (Klasen et al., 2011) or spatial resolution (Chen et al., 2010). Therefore, the exact spatio-temporal investigation of multisensory emotional integration with EEG-fMRI remains a challenging task: complex multisensory integration paradigms provide only a few trials per condition and the MR environment impacts the EEG signal-to-noise ratio (Laufs et al., 2003). Furthermore, the neural difference between emotionally CON and INC multisensory stimuli is expected to be rather small since early sensory processes are ruled out. Therefore, it needs to be established whether a source localization driven analysis of EEG and fMRI data suited to investigate neural processes during multisensory emotion integration on a trial-by-trial basis.

Based on the literature we used the following benchmarks for method validation:

- (1) fMRI analysis will reveal robust activation of visual and auditory cortices after presentation of multisensory stimuli. Likewise, alpha power will be suppressed over occipital areas

to show that alpha power was retained after EEG artifact rejection.

- (2) An inverse relationship of alpha power in occipital areas and BOLD signal has been reliably reported in several EEG-fMRI studies. This relationship can be employed as a benchmark for a valid methodological approach. Thus, we hypothesize that stimulus induced alpha power suppression over occipital areas – as defined by source localization – will inversely correlate with BOLD response. This may confirm the technical feasibility of the here suggested data-driven approach for combining information of EEG and fMRI data.

In a next step, we provide evidence that this new approach may be suited to investigate the neural correlates of multisensory integration of emotions.

## MATERIALS AND METHODS

### SUBJECTS

Data was acquired from three male participants (P1: age: 21, P2: age: 24, P3: age: 22) for method demonstration. They reported normal vision, normal hearing, no contraindications against MR investigations, and no history of neurological or psychiatric illness. The participants were right-handed as assessed with the Edinburgh Handedness Inventory (Oldfield, 1971), german speaking and had normal intelligence according to multiple choice word test (MWT-B; Lehrl, 2005).

The experiment was designed according to the Code of Ethics of the World Medical Association (2008), and the study protocol was approved by the local ethics committee. Written informed consent was obtained and the participants were financially compensated for their participation.

### STIMULI

Audiovisual stimuli were dynamic angry, happy, and neutral virtual characters (avatars) combined with disyllabic pseudowords (angry, happy, or neutral prosody). Visual and auditory channels were combined in emotionally CON or INC fashion; the latter combined different auditory and visual emotions, e.g., a happy face with an angry pseudoword. Animated avatars were created with a 3D animation software package (Poser Pro, Smith Micro Software, CA, USA) and combined with the pseudowords using the incorporated Lip Synchronization Toolbox to assure lip and speech synchronization. The pseudowords followed German phonotactic rules and had no semantic content. Two female and two male avatars were associated with the voices of two male and two female speakers, with each avatar-voice combination being unique. Additionally, the avatars and the pseudowords were presented as visual-only and auditory-only stimuli, respectively. These stimuli have been validated and employed in a previous study (Klasen et al., 2011).

### TASK AND PROCEDURE

All stimuli were displayed using Presentation software (Neurobehavioral Systems, Inc, Albany, California, USA). A hybrid fMRI design of blocks for modality and events for emotions was used. The stimuli were grouped in 32 blocks (8 auditory, 8 visual and 16 audiovisual blocks), separated by a jittered pause of 19–21 s duration. Each block contained 12 trials resulting in a total of

384 stimuli (96 auditory, 96 visual and 192 audiovisual, balanced for emotion and gender). Audiovisual blocks contained randomly distributed CON and INC stimuli. Blocks and trials in each block were presented in pseudo-randomized order.

Each trial started with a stimulus (1–1.2 s) followed by a decision phase (1 s) and a response phase (1 s), during which the participants had to rate the stimulus (delayed response design). The three different response options were displayed during stimulus and decision phase in a white color. The response phase was indicated by a change of colors to green (Figure 1) and the participants were instructed to rate the stimulus as a whole in the response phase as fast as possible. Responses were given by button pressing with index, middle and ring finger of the right hand.

## fMRI

Magnetic resonance imaging was conducted on a 3 Tesla Siemens Trio scanner (Siemens Medical, Erlangen, Germany). One run of echo-planar imaging (EPI) sequence acquired 34 transversal slices (TR = 2000 ms, TE = 28 ms, flip angle = 77°, voxel size = 3 × 3 mm with 64 × 64 matrix, 3 mm slice thickness, 0.75 mm gap). A radio frequency transmit-receive birdcage head coil allowed for simultaneous EEG recording. After the functional measurements, a high resolution, whole brain anatomical image was acquired with a 12-channel head coil (MPRAGE, T1-weighted, TE, 2.52 ms; TR, 1900 ms; flip angle, 9°; FOV, 256 × 256 mm; 1 mm isotropic voxels; 176 sagittal slices).

## fMRI PRE-PROCESSING

Image analysis was performed with *Brain Voyager QX 2.6* (Brain Innovation, Maastricht, the Netherlands). Pre-processing of

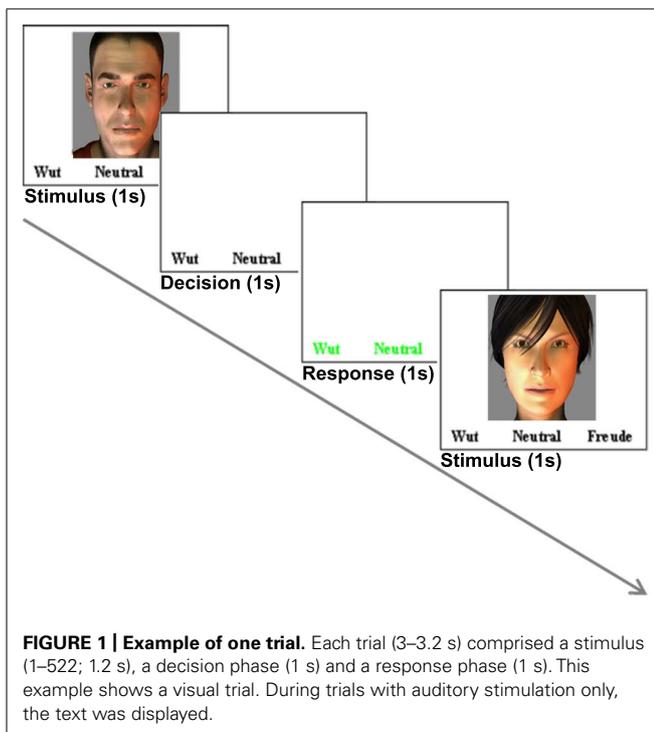
functional MR images included slice scan time correction, 3D motion correction, spatial smoothing (4 mm FWHM), and high-pass filtering including linear trend removal. The first two images were discarded to avoid T1 saturation effects. Functional images were co-registered to 3D anatomical images and transformed into Talairach space. Trials without a response were omitted from further analysis. For auditory, visual, and CON audiovisual trials, only trials with correct responses were included. For INC audiovisual stimuli, all responded trials were included irrespective of correctness. All omitted and incorrect trials were modeled as a confound predictor in the GLM. Cluster threshold was determined with Monte-Carlo simulation (10000 iterations) as implemented in *BrainVoyager QX 2.6*.

## EEG ACQUISITION

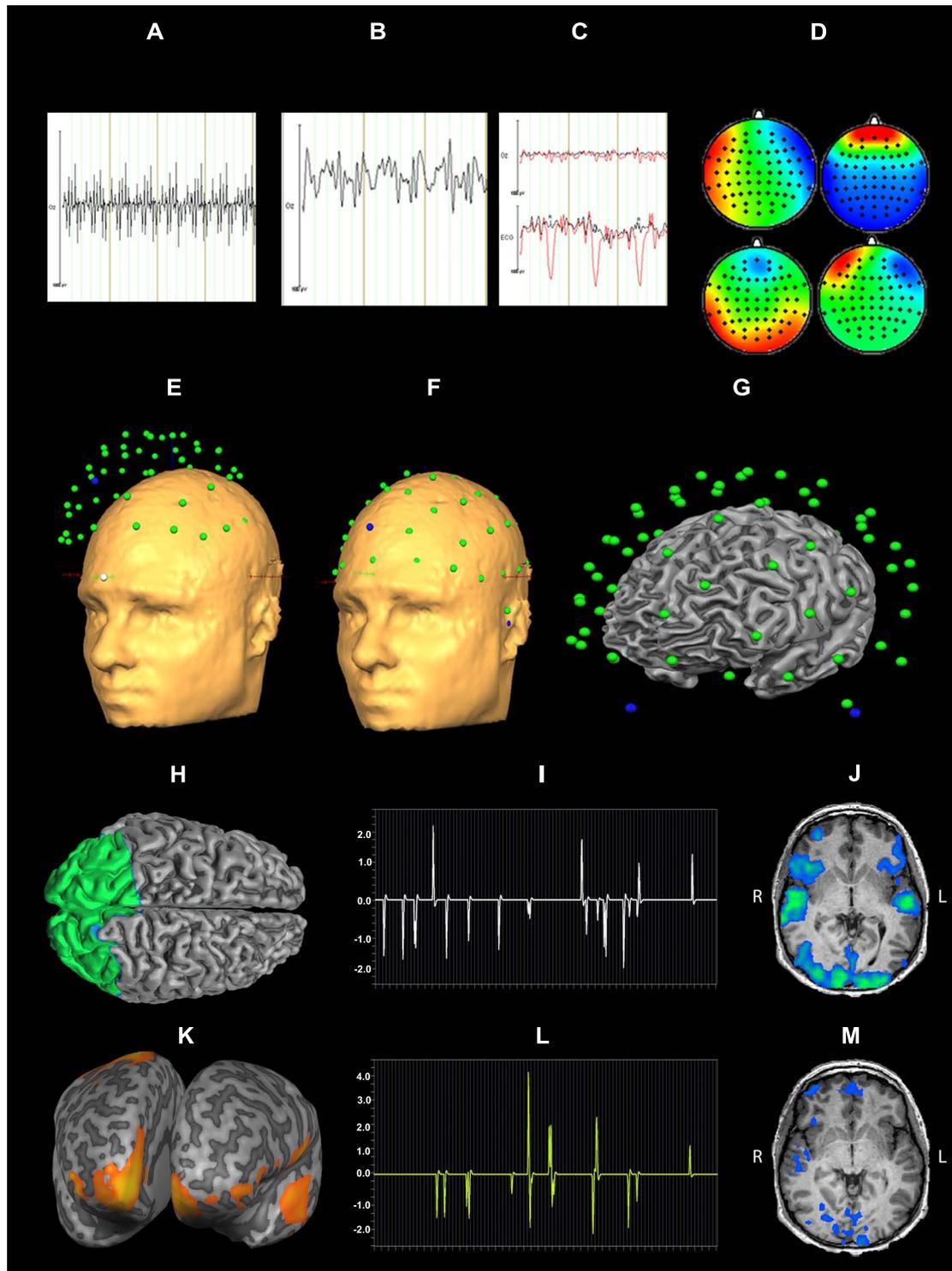
Simultaneously with the fMRI acquisition, EEG was recorded from a 64-channel MR-compatible EEG-cap (Easycap GmbH, Herrsching-Breitbrunn, Germany) connected to a MR-compatible amplifier system (two BrainAmp MR plus 32-channel amplifiers, BrainProducts GmbH, Gilching, Germany). The EEG cap consisted of 64 Ag–AgCl electrodes (5 kΩ resistors), 63 of which covered the 10–20 system and an additional electrocardiogram (ECG) electrode placed below the left collar bone. Midline electrodes anterior and posterior to Fz served as the recording reference and ground channel, respectively. Prior to measurement, all channel positions were digitized using ELGuide V1.8 (Zebris Medical GmbH, Baden-Württemberg, Germany). Channel impedances were kept below 10 kΩ. To improve MR pulse artifact removal, a sync box (BrainProducts GmbH, Gilching, Germany) was used for optimal synchronization of EEG recording with the clock controlling MRI slice acquisition. EEG data were recorded in BrainVision Recorder software (v 1.05, BrainProducts GmbH, Gilching, Germany) at 5000 Hz sampling frequency (0.01–250 Hz analog band-pass filter) and analyzed in BrainVision Analyzer software (Version 2.02, BrainProducts, Gilching, Germany).

## EEG PRE-PROCESSING

Pre-processing of EEG data included gradient artifact removal using a template subtraction algorithm (Allen et al., 2000). After gradient artifact removal (Figures 2A,B), the data were low-pass filtered with a digital infinite impulse response filter (IIR, 70 Hz, 48 dB slope) and down-sampled to 500 Hz. Cardiac pulse correction was carried out based on an automatically detected pulse template in the ECG channel. Cardiac pulse markers were visually confirmed and the BCG artifact was subtracted (Allen et al., 1998; Figure 2C). Data sets were then down-sampled to 250 Hz and artifacts exceeding ± 300 μV were rejected. To remove artifacts due to eye movements, eye blinks and residual BCG artifacts, an ICA was conducted using 63 independent components. Components reflecting artifacts (Figure 2D) were visually identified and rejected based on topography and time course. All EEG channels re-referenced to average reference and pseudo-electrodes AFz and FCz were calculated using spherical interpolation resulting in a total of 65 channels. For EEG analysis, data were processed with BrainVision Analyzer software (Version 2.02, BrainProducts,



**FIGURE 1 | Example of one trial.** Each trial (3–3.2 s) comprised a stimulus (1–1.2 s; 1.2 s), a decision phase (1 s) and a response phase (1 s). This example shows a visual trial. During trials with auditory stimulation only, the text was displayed.



**FIGURE 2 | Data processing for EEG-fMRI integration: EEG artifact removal included subtraction of MR pulse (A: before MR pulse subtraction, B: after MR pulse subtraction), cardioballistic artifact (C, red: before artifact subtraction, black: after artifact subtraction), eye blinks, head movement and residual BCG (D). 3D EEG channel coordinates were coregistered to Talairach space (E,F). Successful co-registration of Talairach transformed EEG channel positions was confirmed by visual inspection (G).**

Source localization of alpha power for audiovisual stimuli (CON + INC) revealed alpha power suppression in occipital areas (H). This cluster was determined as a patch of interest to calculate alpha power time course: this time course was convolved with hemodynamic response function (I) and correlated with BOLD signal (J). Contrast for (CON > INC) was estimated (K) and the alpha power time course in frontal region convolved with HRF (L) was used for correlation with BOLD signal time course (M).

Gilching, Germany). For EEG-fMRI integration, EEG data were exported to BrainVoyager QX 2.6.

### EEG ANALYSIS

Stimulus markers were imported based on Presentation timing log files (Neurobehavioral Systems, Inc, Albany, CA, USA). Trials without a response, errors, presentation uncertainty above 10 ms or amplitudes exceeding  $\pm 125 \mu\text{V}$  were omitted from further analysis. Segmentation was based on stimulus onset ( $-2.000$  to  $+1.000$  ms) for auditory, visual, audiovisual CON and INC stimuli. Frequency decomposition was achieved by continuous wavelet transformation (complex morlet motherwavelet,  $c = 4.2$ ) and baseline corrected ( $-1.500$  to  $-500$  ms). Further, segment's average and standard deviation of alpha power was calculated. For representation of alpha power topography, wavelets with a center frequency of 10.5 Hz (borders:  $\pm 2.5$  Hz, wavelet length = 133 ms) were extracted.

### EEG-fMRI ANALYSIS

For EEG-fMRI coupling, the EMEG toolbox, implemented in *BrainVoyager QX 2.6*, was used. Talairach transformed anatomy was used for individual head surface and cortex mesh reconstruction. Individual 3D EEG channel coordinates were coregistered to head surface mesh for transformation to Talairach space (Figures 2E,F) and successful coregistration of EEG channel positions was visually confirmed (Figure 2G). Cortex meshes for both hemispheres were reconstructed using the outer gray matter boundary and the number of vertices was reduced to 2500 per hemisphere. Lead fields for each EEG channel were estimated assuming a four layer spherical head model (Berg and Scherg, 1994). A combination of three surface maps (in  $x$ ,  $y$  and  $z$  direction) represented the channel specific lead field. To calculate the regularization term of the inverse solution, a  $65 \times 65$  covariance matrix for  $-2000$  to  $0$  ms (stimulus onset) was calculated which contained the spatial distribution of noise and spatial correlation of EEG channels. For estimation of an inverse solution to the EEG inverse problem (Hämäläinen and Ilmoniemi, 1994), we used the weighted-minimum norm solution with noise-based normalization as proposed by Dale et al. (2000). For a given SNR 5, the regularization parameter  $\lambda$  (Tikhonov and Arsenin, 1977) was estimated as 0.34 (accounting for 0.17% of the trace) to minimize noise amplification. Variation of the SNR between 1 and 10 did not yield relevant changes in the regularization.

### EEG-SOURCE ANALYSIS

Time series of alpha power (8–13 Hz) were calculated in *BrainVoyager QX2.6* for emotionally CON and INC multisensory stimuli using short time Fourier transformation (STFFT, Portoff, 1980)

with the following settings: One time window consisted of 500 ms and was shifted for 100 ms resulting in 80% overlap between two neighboring windows. The alpha power values were estimated from  $-2000$  to  $1000$  ms after stimulus onset and baseline corrected from  $-2000$  to  $0$  ms. Statistical maps of distributed EEG sources were estimated for affective multisensory stimuli. Separate Contrasts for audiovisual (CON + INC) trials over baseline and for CON over INC trials (CON > INC) were calculated: The first contrast tested preservation of alpha power after preprocessing and the second one evaluated a putative facilitation effect for CON multisensory information. The resulting clusters for (CON + INC) > baseline and the frontal cluster for (CON > INC; Figures 2H,K) served as a patch of interest (POI) for the correlation of alpha power and BOLD response in affective multisensory trials to provide evidence for a valid and sensitive methodological approach (Figures 2H,I,J).

### SINGLE TRIAL EEG-fMRI COUPLING

Single-trial induced alpha power at each event for 200–400 ms was calculated for the occipital and prefrontal cortex POI. Alpha power values were convolved with the hemodynamic response function and predicted the BOLD signal in a general linear model (GLM; Figures 2I,L). Induced power values of alpha oscillations considered the inverse solution of each POI incorporated the inverse solution and weighted the influence of all electrodes based on the inverse solution. Episodes without an event were set to zero. The definition of entire regions for correlations was based on results of the EEG source localization (see Figures 2H,K). Correlation maps for each occipital alpha power and prefrontal alpha power of CON and INC stimuli were estimated using first level statistics to identify neural sources that were related with the induced alpha in occipital and frontal areas (Figures 2J and M, respectively).

## RESULTS

### BEHAVIOR

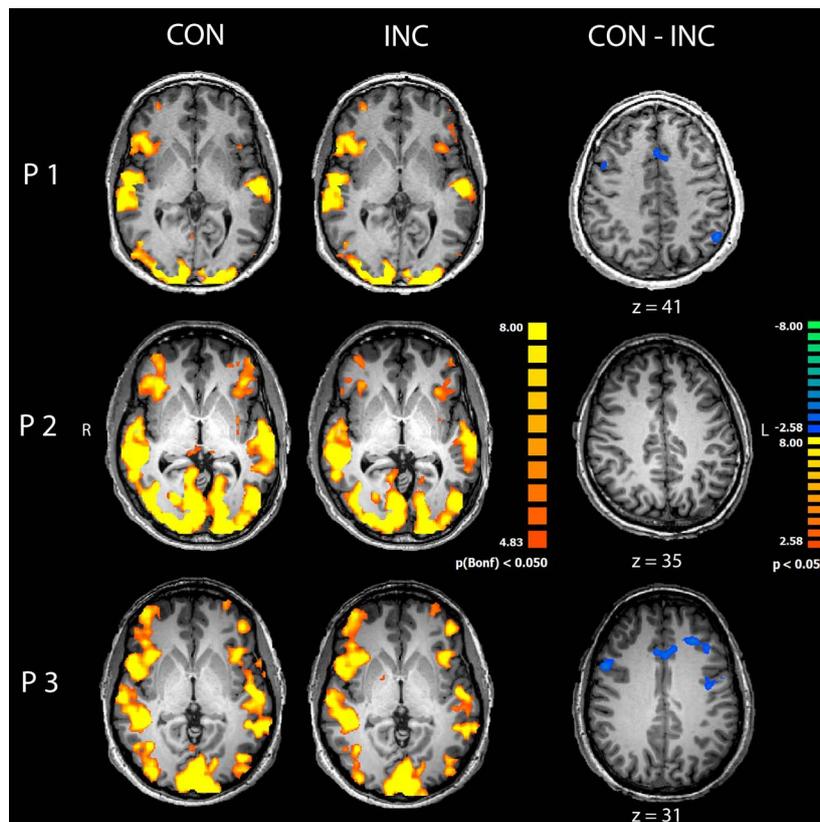
On average, the participants correctly identified emotions in 70.8% of the prosodic trials and in 93.8% of the visual trials. Multisensory CON emotions were correctly classified in 92.9%. In the INC condition, the participants decided in 51.0% of the stimuli according to the facial expression and in 28.4% according to the emotional prosody and in 22.6% for neither facial nor auditory emotion.

### fMRI

Congruent and INC trials compared to baseline yielded significantly increased activity in bilateral visual, auditory, frontal and motor areas ( $p < 0.05$ , Bonferroni corrected; see Figure 3, CON

**Table 1 | Decisions for uni- and multisensory emotional trials.**

	Auditory trials (%)	Visual trials (%)	Congruent trials (%)	Decision for face (%)	Decision for voice (%)
P1	69.3	92.0	96.1	39.6	40.6
P2	75.4	96.2	93.8	57.3	23.6
P3	72.2	95.5	89.8	62.3	21.0



**FIGURE 3 | Comparison of congruent and incongruent trials.**

Congruent (CON) and incongruent (INC) trials elicited highly significant widespread activations in visual and auditory cortices as well as motor and medial frontal areas in participants P1–P3 ( $p < 0.05$ , Bonferroni

corrected). Incongruent compared to congruent trials induced significantly higher activity in mediofrontal and dorsolateral regions in participants P1 and P3 (CON-INC,  $p < 0.05$ , cluster-threshold corrected).

and INC). The comparison of CON and INC trial revealed significantly higher activation for INC trials in ACC and dorsolateral-prefrontal cortex in two out of three participants (Figure 3, CON-INC;  $p < 0.05$ ; cluster-threshold corrected).

### EEG

In each channel, less than 10% of the data points were excluded due to artifact rejection in all participants. EEG analysis suggested alpha power suppression over occipital areas for both CON and INC trials (Figure 4, CON-INC) and elevated induced alpha oscillations for CON stimuli in frontal areas during 200–400 ms after stimulus onset (Figure 4, CON). In contrast to CON stimuli, Incongruency of emotions induced alpha power suppression in a small cluster at the Fz electrode (Figure 4, INC). The difference between induced alpha power of CON and INC stimuli 200–400 ms after stimulus onset yielded a left lateralized frontal cluster (Figure 4, CON-INC).

### EEG source analysis

A cortical constrained minimum-norm-weighted inverse solution estimated the sources (Grech et al., 2008). Alpha power was suppressed over the occipital cortex (OC) about 200–400 ms after onset of audiovisual stimuli (Figure 5 CON and INC).

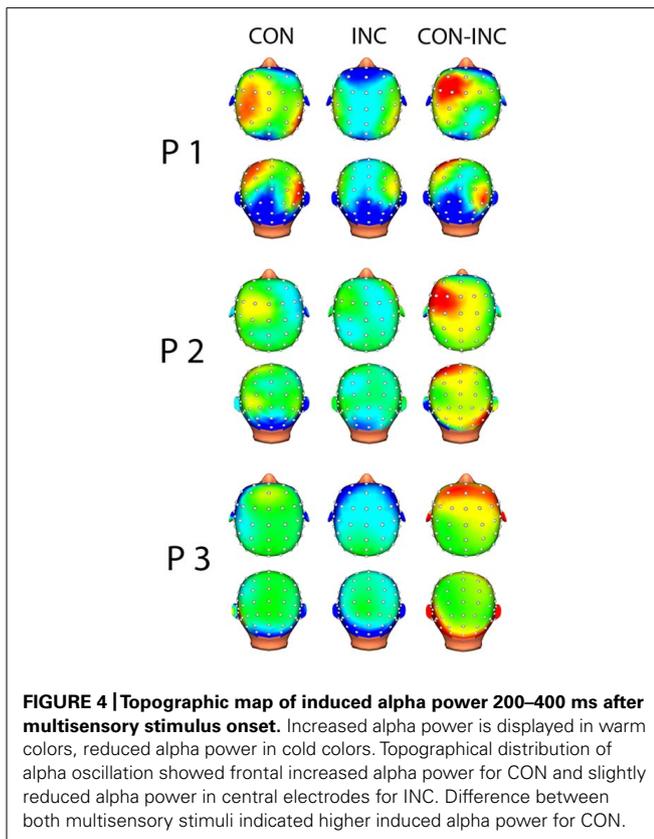
The topography confirmed preservation of the alpha oscillations after EEG preprocessing whilst successful artifact removal, therefore allowing the investigation of differences between induced alpha power of CON and INC stimuli. For CON stimuli, higher alpha power extended bilaterally to fronto-medial areas (Figure 5, CON) whereas INC stimuli induced alpha power suppression at the frontal cortex in participants P2 and P3 (Figure 5, INC).

A contrast for (CON > INC) was estimated to test whether previous results by Chen et al. (2010) could be replicated. It revealed significantly elevated induced alpha power 200–400 ms after stimulus onset for CON stimuli in left prefrontal cortex (Figure 5, CON-INC;  $p < 0.05$ ). No further difference for (CON > INC) was observed during later time windows (600–800 ms and 800–1000 ms). As a next step, the time series of single trial induced alpha power for the OC and frontal areas were individually correlated with whole brain BOLD signal to identify neural networks supporting the event-related changes of alpha oscillations.

### EEG-INFORMED fMRI

#### Correlation analysis of induced alpha power in OC

Single-trial variability of induced alpha power to multisensory stimuli in the OC correlated negatively with BOLD response in



a wide-spread network encompassing visual and auditory cortices, dorsolateral and prefrontal areas as well as bilateral insula (**Figure 6A**;  $p < 0.05$ , cluster-threshold corrected).

Further correlation analysis between time series of induced alpha power during 200–400 ms after presentation onset in the PFC and BOLD response revealed exclusively inverse correlations in mediofrontal, dorsolateral, prefrontal and visual areas (**Figure 6B** and **Table 1**,  $p < 0.05$ , uncorrected) to emotionally INC stimuli.

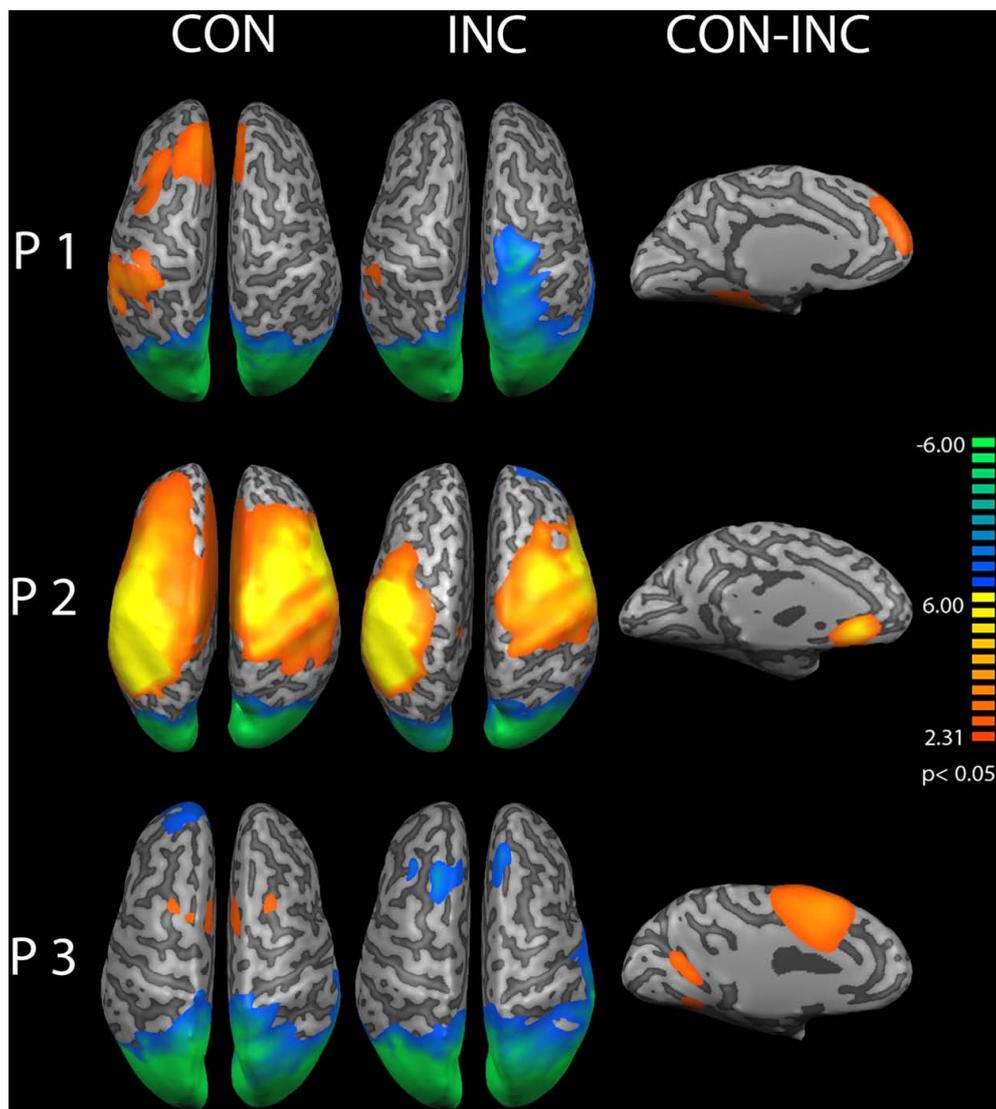
## DISCUSSION

The aim of this pilot study was to investigate the technical feasibility of a source analysis-driven approach to integrate the high spatial accuracy of fMRI and high temporal resolution of EEG using simultaneous EEG-fMRI in a fast, event-related design. We confirmed EEG and fMRI data quality separately by reproducing established response patterns in all participants. During the presentation of multisensory affective trials the fMRI data analysis revealed strong activation in a distributed network encompassing visual, auditory cortex and insula in accordance with previous studies (Ethofer et al., 2006; Klasen et al., 2011) and confirms reliable fMRI data registration despite the simultaneous EEG recording. In a similar vein, it is known that MRI environment is detrimental on EEG data quality (Mullinger and Bowtell, 2011) before various artifact removal steps are being applied (mainly correction of HF scanner artifact, correction of cardiobalistic signal and eye blinks). The low amount of epoch rejection for multisensory stimuli suggested a successful artifact removal. However,

there remains a speculative notion to this conclusion since we did not have any outside scanner data to directly compare data quality. A source analysis was conducted of induced alpha oscillations for multisensory trials to test for the preservation of alpha oscillations after EEG processing. Topography of induced alpha power in occipital areas during 200–400 ms after multisensory stimulus onset confirmed preservation of alpha oscillations and indicated increased cortical activation due to complex sensory input. Shagass (1972) assumed that alpha oscillations are an indirect measurement of cortical activity which has been supported by recent studies (Klimesch, 1999; Nunez et al., 2001; Jokisch and Jensen, 2007; Palva and Palva, 2007). Therefore, reduced alpha power over occipital areas is likely to indicate effective signal processing at sensory cortices (Mathiak et al., 2011). The first EEG-fMRI studies on the relationship of alpha power and BOLD response (Goldman et al., 2002; Laufs et al., 2003) reported an inverse correlation of occipital alpha oscillations and BOLD signal in line with previous hypothesis about the function of the alpha oscillations as an idling or suppression rhythm (Mazaheri and Jensen, 2010). This negative correlation has also been used as a benchmark for a valid methodological approach in recent EEG-fMRI studies (Scheeringa et al., 2009; Mayhew et al., 2013; Mo et al., 2013). In this study, we replicated a reasonable correlation pattern between occipital alpha power suppression and BOLD signal, thus confirming the validity of our source localization driven method.

But for which investigations can the presented procedure be used? Current studies apply two methods: correlation of single electrode time course (e.g., Oz) or individual EEG components with BOLD response. In a sophisticated study, Scheeringa et al. (2009) investigated working memory networks related to alpha power using EEG-fMRI. The authors applied an ICA approach to extract a components reflecting alpha power time course based on its topography. This method enhances SNR due to exclusion of noise which remains in the other components. But this analysis depends critically on a priori knowledge about the topography of the EEG signal which makes this approach unfeasible for investigating new paradigms. The selection of single electrodes is frequently employed in ERP studies to calculate correlations with the BOLD signal (e.g., Eichele et al., 2005; Dubois et al., 2012). In contrast to this more common ERP analysis, the present analysis incorporates multichannel information. Hence, different components of a network as identified by source analysis and their neural correlates may be investigated using EEG-fMRI.

Studies investigating multisensory integration usually identify neural networks and deal with only subtle differences between conditions where the topography of the EEG signal is unknown a priori. Therefore the classical procedures are not suitable and a source localization-driven, sensitive method is necessary to investigate the time course of multisensory integration of emotions with high spatial and temporal resolution using EEG-fMRI. In an explorative analysis, we further tested the feasibility of this method to investigate multisensory integration of emotions. Chen et al. (2010) reported in an MEG a facilitation effect for multisensory emotional stimuli 200–400 ms post stimulus presentation. Therefore we verified our method further in the context of multisensory integration. Our behavioral data are generally

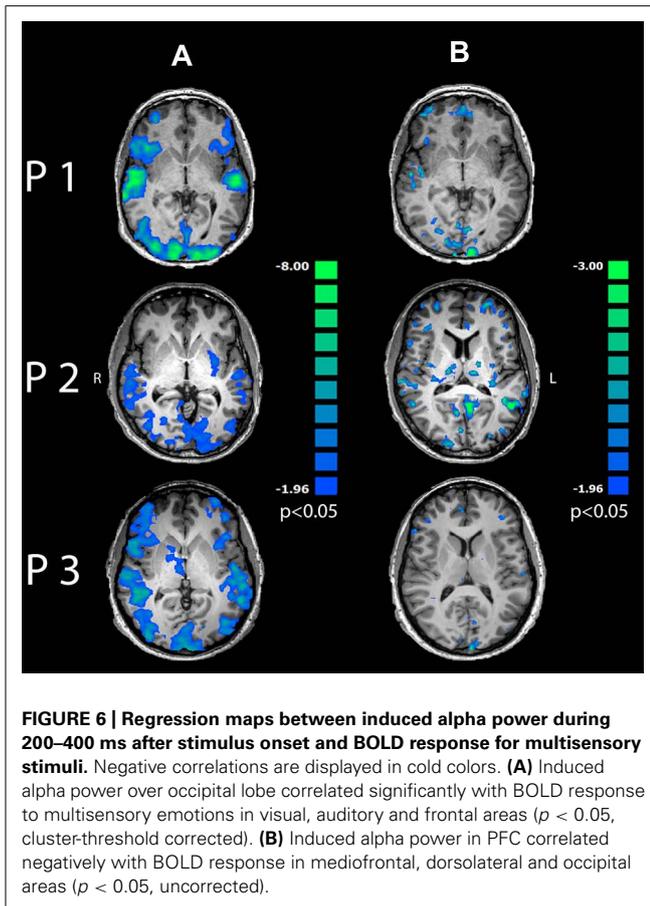


**FIGURE 5 | Source analysis of induced alpha power during 200–400 ms after stimulus onset.** Increased alpha power over baseline is displayed in warm colors, reduced alpha power in cold colors for congruent (CON) and incongruent (INC) stimuli ( $p < 0.05$ ). In participants P1–P3, stimulus presentation suppressed alpha power over the

occipital lobe. Congruent trials elicited higher alpha power in mediofrontal regions. A comparison of congruent over incongruent stimuli showed significant (CON,  $p < 0.05$ ) higher induced alpha power for congruent trials in frontal areas in the left hemisphere (CON-INC,  $p < 0.05$ ).

in line with previous studies reporting high visual and CON multisensory recognition rates and slightly lower ones for the auditory-only stimuli (Vroomen et al., 2001; Campanella and Belin, 2007; Klasen et al., 2011). This confirmed that affective pseudowords, faces and their combination were identified correctly on a behavioral level. The missing facilitation effect between visual-only and CON multisensory stimuli may be attributed to a ceiling effect since recognition rates for both were very high (>90%). Using first level statistics, a significant facilitation effect for CON stimuli was found in PFC for early (200–400 ms) but not for late (600–800 ms) induced alpha oscillations during early perceptual processing. This power difference disappeared after 600 ms, reproducing a previous MEG study (Chen et al., 2010).

They reported a facilitation effect of affective audiovisual processing over both auditory and visual stimuli in a similar time window only (200–450 ms). The here presented reproduction based on only few stimuli in a single subject further corroborates the efficiency of the presented EEG analysis. We suggest that increased induced alpha power for CON stimuli may reflect reduced cognitive demand by stimulus disambiguation (Stein and Meredith, 1993; Ernst and Bühlhoff, 2004) and perceptual processing of affective multisensory stimuli during 200–400 ms after stimulus onset. Essential information on the role of various brain structures in multisensory emotion integration comes from the trial-wise correlations of induced oscillations and BOLD signal, which constitutes the major benefit of simultaneous EEG-fMRI



measurements. Combined EEG-fMRI indicated that the facilitation effect of early induced alpha power for emotional CON multisensory information is not constrained to frontal areas. Further contributions may origin in a fronto-medial and -lateral evaluation network which is known to be involved in processing of multisensory stimuli (for reviews, see Calvert and Thesen, 2004 and Klasen et al., 2012). The reproducible activation in the fronto-medial and -lateral network may reflect cognitive feature evaluation in early processing whereas this may not be necessary for CON stimuli as auditory and visual input confers redundant information.

## CONCLUSION

With our study we provided a novel method to investigate the temporal course of affective multisensory integration. To our knowledge, this is the first study so far that investigated the feasibility of a source localization driven approach of induced alpha power of affective multisensory processing with BOLD response in a fast event related design employing simultaneous EEG-fMRI. We demonstrated that the analysis of combined simultaneous EEG-fMRI recording provided valid and informative EEG, fMRI, as well as EEG-informed fMRI results.

Thus, the results support the technical feasibility of this novel approach and may help to disentangle the neural correlates of perceptual and decisional processing during multisensory integration of affective information.

## LIMITATIONS

This method relies – in contrast to previous ones – on the inverse solution of EEG data. The source localization using this method can deviate substantially from subject to subject. Therefore a high data quality and strict EEG artifact rejection is necessary to employ this method. Furthermore, we did not compare EEG scalp inside and outside the MR environment and interpretation about the topographies remains speculative. But we replicated known response patterns for alpha power after visual stimulus presentation and even confirmed the topography of induced alpha power to multisensory stimuli (Chen et al., 2010). Conceivably, the presented source localization driven method can specify small signal differences in empirical EEG-fMRI studies.

This pilot study included three subjects only. Although the results of EEG-fMRI combination yielded significance, a generalization of these exploratory findings is not possible so far.

## AUTHOR CONTRIBUTIONS

Patrick D. Schelenz and Martin Klasen designed the paradigm. Patrick D. Schelenz, Barbara Reese, Martin Klasen and Dhana Wolf acquired data. Patrick D. Schelenz, Yutaka Kato, Christina Regenbogen and Klaus Mathiak analyzed the data. Patrick D. Schelenz, Barbara Reese, Christina Regenbogen, Dhana Wolf, Yutaka Kato and Klaus Mathiak wrote the paper.

## ACKNOWLEDGMENTS

This study was supported by the German Research Foundation (DFG; IRTG 1328, MA 2631/4-1), IRTG 1328 and the IZKF Aachen (N4-2). Fabrizio Esposito provided valuable support with BrainVoyager data analysis.

## REFERENCES

- Allen, P. J., Josephs, O., and Turner, R. (2000). A method for removing imaging artifact from continuous EEG recorded during functional MRI. *Neuroimage* 12, 230–239. doi: 10.1006/nimg.2000.0599
- Allen, P. J., Polizzi, G., Krakow, K., Fish, D. R., and Lemieux, L. (1998). Identification of EEG events in the MR scanner: the problem of pulse artifact and a method for its subtraction. *Neuroimage* 8, 229–239. doi: 10.1006/nimg.1998.0361
- Becker, R., Reinacher, M., Freyer, F., Villringer, A., and Ritter, P. (2011). How ongoing neuronal oscillations account for evoked fMRI variability. *J. Neurosci.* 31, 11016–11027. doi: 10.1523/JNEUROSCI.0210-11.2011
- Berg, P., and Scherg, M. (1994). A fast method for forward computation of multiple-shell spherical head models. *Electroencephalogr. Clin. Neurophysiol.* 90, 58–64. doi: 10.1016/0013-4694(94)90113-9
- Calvert, G. A., and Thesen, T. (2004). Multisensory integration: methodological approaches and emerging principles in the human brain. *J. Physiol.* 98, 191–205.
- Campanella, S., and Belin, P. (2007). Integrating face and voice in person perception. *Trends Cogn. Sci.* 11, 535–543. doi: 10.1016/j.tics.2007.10.001
- Chen, Y., Edgar, J. C., Holroyd, T., Dammers, J., Thönneßen, H., Roberts, T. P. L., et al. (2010). Neuromagnetic oscillations to emotional faces and prosody. *Eur. J. Neurosci.* 31, 1818–1827. doi: 10.1111/j.1460-9568.2010.07203.x
- Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., et al. (2000). Dynamic statistical parametric mapping: combining fMRI and MEG for high resolution imaging of cortical activity. *Neuron* 26, 55–67. doi: 10.1016/S0896-6273(00)81138-1
- Dubois, C., Otzenberger, H., Gounot, D., Sock, R., and Metz-Lutz, M.-N. (2012). Visemic processing in audiovisual discrimination of natural speech: a simultaneous fMRI-EEG study. *Neuropsychologia* 50, 1316–1326. doi: 10.1016/j.neuropsychologia.2012.02.016
- Eichele, T., Specht, K., Moosmann, M., Jongsma, M. L. A., Quiroga, R. Q., Nordby, H., et al. (2005). Assessing the spatiotemporal evolution of

- neuronal activation with single-trial event-related potentials and functional MRI. *Proc. Natl. Acad. Sci. U.S.A.* 102, 17798–17803. doi: 10.1073/pnas.0505508102
- Ethofer, T., Pourtois, G., and Wildgruber, D. (2006). Investigating audiovisual integration of emotional signals in the human brain. *Prog. Brain Res.* 156, 345–361. doi: 10.1016/S0079-6123(06)56019-4
- Ernst, M. O., and Bühlhoff, H. H. (2004). Merging the senses into a robust percept. *Trends Cogn. Sci.* 162–169. doi: 10.1016/j.tics.2004.02.002
- Goldman, R. I., Stern, J. M., Engel, J. Jr., and Cohen, M. S. (2002). Simultaneous EEG and fMRI of the alpha rhythm. *Neuroreport* 13, 2487–2492. doi: 10.1097/00001756-200212200-00022
- Grech, R., Cassar, T., Muscat, J., Camilleri, K. P., Fabri, S. G., and Zervakis, M. (2008). Review on solving the inverse problem in EEG source analysis. *J. Neuroeng. Rehabil.* 5, 25. doi: 10.1186/1743-0003-5-25
- Hämäläinen, M. S., and Ilmoniemi, R. J. (1994). Interpreting magnetic fields of the brain: minimum norm estimates. *Med. Biol. Eng. Comput.* 32, 35–42. doi: 10.1007/BF02512476
- Jokisch, D., and Jensen, O. (2007). Modulation of gamma and alpha activity during a working memory task engaging the dorsal or ventral stream. *J. Neurosci.* 27, 3244–3251. doi: 10.1523/JNEUROSCI.5399-06.2007
- Klasen, M., Chen, Y., and Mathiak, K. (2012). Multisensory emotions: perception, combination and underlying neural processes. *Rev. Neurosci.* 23, 381–392. doi: 10.1515/revneuro-2012-0040
- Klasen, M., Kenworthy, C. A., Mathiak, K. A., Kircher, T. T., and Mathiak, K. (2011). Supramodal representation of emotions. *J. Neurosci.* 31, 13635–13643. doi: 10.1523/JNEUROSCI.2833-11.2011
- Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Rev.* 29, 169–195. doi: 10.1016/S0165-0173(98)00056-3
- Laufs, H., Kleinschmidt, A., Beyerle, A., Eger, E., Salek-Haddadi, A., Preibisch, C., et al. (2003). EEG-correlated fMRI of human alpha activity. *Neuroimage* 19, 1463–1476. doi: 10.1016/S1053-8119(03)00286-6
- Lehr, S. (2005). *Mehrfachwahl-Wortschatz-Intelligenztest MWT-B*, 5th Edn. Balin-gen: Spitta Verlag.
- Mathiak, K., Ackermann, H., Rapp, A., Mathiak, K. A., Shergill, S., Riecker, A., et al. (2011). Neuromagnetic oscillations and hemodynamic correlates of P50 suppression in schizophrenia. *Psychiatry Res.* 194, 95–104. doi: 10.1016/j.psychres.2011.01.001
- Mayhew, S. D., Ostwald, D., Porcaro, C., and Bagshaw, A. P. (2013). Spontaneous EEG alpha oscillation interacts with positive and negative BOLD responses in the visual-auditory cortices and default-mode network. *Neuroimage* 76, 362–372. doi: 10.1016/j.neuroimage.2013.02.070
- Mazaheri, A., and Jensen, O. (2010). Rhythmic pulsing: linking ongoing brain activity with evoked responses. *Front. Hum. Neurosci.* 4:177. doi: 10.3389/fnhum.2010.0017
- Mo, J., Liu, Y., Huang, H., and Ding, M. (2013). Coupling between visual alpha oscillations and default mode activity. *Neuroimage* 68, 112–118. doi: 10.1016/j.neuroimage.2012.11.058
- Mulert, C., and Lemieux, L. (2010). *EEG-fMRI: Physiological Basis, Technique and Applications*. Berlin: Springer press.
- Mullinger, K. J., and Bowtell, R. (2011). Combining EEG and fMRI. *Methods Mol. Biol.* 711, 303–326. doi: 10.1007/978-1-61737-992-5\_15
- Nunez, P. L., Wingeier, B. M., and Silberstein, R. B. (2001). Spatial-temporal structures of human alpha rhythms: theory, microcurrent sources, multiscale measurements, and global binding of local networks. *Hum. Brain Mapp.* 13, 125–164. doi: 10.1002/hbm.1030
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- Palva, S., and Palva, J. M. (2007). New vistas for alpha-frequency band oscillations. *Trends Neurosci.* 30, 150–158. doi: 10.1016/j.tins.2007.02.001
- Portoff, M. R. (1980). Time-frequency representation of digital signals and systems based on Short-time Fourier analysis. *Acoust. Speech Signal Process.* 28, 55–69. doi: 10.1109/TASSP.1980.1163359
- Scheeringa, R., Petersson, K. M., Oostenveld, R., Norris, D. G., Hagoort, P., and Bastiaansen, M. C. M. (2009). Trial-by-trial coupling between EEG and BOLD identifies networks related to alpha and theta EEG power increases during working memory maintenance. *Neuroimage* 44, 1224–1238. doi: 10.1016/j.neuroimage.2008.08.041
- Shagass, C. (1972). “Electrical activity of the brain,” in *Handbook of Psychophysiology*, eds N. S. Greenfield and R. A. Sternbach (New York: Holt, Rinehart, and Winston), 263–328.
- Stein, B. E., and Meredith, M. A. (1993). *Merging the Senses*. Cambridge: MIT Press.
- Tikhonov, A., and Arsenin, V. (1977). *Solutions to Ill-Posed Problems*. New York: Wiley.
- Ullsperger, M., and Debener, S. (2010). *Simultaneous EEG and fMRI: Recording, Analysis, and Application*. Oxford: Oxford University Press
- Vroomen, J., Driver, J., and de Gelder, B. (2001). Is cross-modal integration of emotional expressions independent of attentional resources? *Cogn. Affect. Behav. Neurosci.* 1, 382–387. doi: 10.3758/CABN.1.4.382
- World Medical Association. (2008). *WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects*. Seoul: World Medical Association. Available at: <http://www.wma.net/en/30publications/10policies/b3/>

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 01 June 2013; accepted: 13 October 2013; published online: 14 November 2013.

Citation: Schelenz PD, Klasen M, Reese B, Regenbogen C, Wolf D, Kato Y and Mathiak K (2013) Multisensory integration of dynamic emotional faces and voices: method for simultaneous EEG-fMRI measurements. *Front. Hum. Neurosci.* 7:729. doi: 10.3389/fnhum.2013.00729

This article was submitted to the journal *Frontiers in Human Neuroscience*. Copyright © 2013 Schelenz, Klasen, Reese, Regenbogen, Wolf, Kato and Mathiak. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Situating emotional experience

Christine D. Wilson-Mendenhall<sup>1\*</sup>, Lisa Feldman Barrett<sup>1†</sup> and Lawrence W. Barsalou<sup>2†</sup>

<sup>1</sup> Department of Psychology, Northeastern University, Boston, MA, USA

<sup>2</sup> Department of Psychology, Emory University, Atlanta, GA, USA

## Edited by:

Martin Klasen, Rheinisch-Westfälische Technische Hochschule Aachen University, Germany

## Reviewed by:

Ruben Gur, University of Pennsylvania School of Medicine, USA  
Cindy Hagan, University of Cambridge, UK

## \*Correspondence:

Christine D. Wilson-Mendenhall,  
Department of Psychology,  
Northeastern University, 125  
Nightingale Hall, Boston, MA 02115,  
USA  
e-mail: cd.wilson@neu.edu

<sup>†</sup> Lisa Feldman Barrett and Lawrence W. Barsalou have joint senior authorship.

Psychological construction approaches to emotion suggest that emotional experience is situated and dynamic. Fear, for example, is typically studied in a physical danger context (e.g., threatening snake), but in the real world, it often occurs in social contexts, especially those involving social evaluation (e.g., public speaking). Understanding *situated* emotional experience is critical because adaptive responding is guided by situational context (e.g., inferring the intention of another in a social evaluation situation vs. monitoring the environment in a physical danger situation). In an fMRI study, we assessed situated emotional experience using a newly developed paradigm in which participants vividly imagine different scenarios from a first-person perspective, in this case scenarios involving either social evaluation or physical danger. We hypothesized that distributed neural patterns would underlie immersion in social evaluation and physical danger situations, with shared activity patterns across both situations in multiple sensory modalities and in circuitry involved in integrating salient sensory information, and with unique activity patterns for each situation type in coordinated large-scale networks that reflect situated responding. More specifically, we predicted that networks underlying the social inference and mentalizing involved in responding to a social threat (in regions that make up the “default mode” network) would be reliably more active during social evaluation situations. In contrast, networks underlying the visuospatial attention and action planning involved in responding to a physical threat would be reliably more active during physical danger situations. The results supported these hypotheses. In line with emerging psychological construction approaches, the findings suggest that coordinated brain networks offer a systematic way to interpret the distributed patterns that underlie the diverse situational contexts characterizing emotional life.

**Keywords:** emotion, situated cognition, affective neuroscience, affect, cognitive neuroscience

## INTRODUCTION

Darwin's *The Expression of the Emotions in Man and Animals* is often used to motivate emotion research that focuses on identifying the biological signatures for five or so emotion categories (Ekman, 2009; Hess and Thibault, 2009). Interestingly, though, the evolution paradigm shift initiated by Darwin and other scientists heavily emphasized *variability*: species are biopopulations in which individuals within a population are unique and in which individual variation within a species is meaningfully tied to variation in the environment (and they are *not* physical types defined by essential features; Barrett, 2013). In other words, an individual organism is best understood by the situational context in which it operates. It is not a great leap, then, to hypothesize that “situatedness” is also a basic principle by which the human mind operates, during emotions and during many other mental phenomena (Barrett, 2013).

Situated approaches to the mind typically view the brain as a coordinated system designed to use information captured during prior situations (and stored in memory) to flexibly interpret and infer what is happening in the current situation – dynamically shaping moment-to-moment responding in the form of perceiving, coordinating action, regulating the body, and organizing thoughts (Glenberg, 1997; Barsalou, 2003, 2009; Aydede

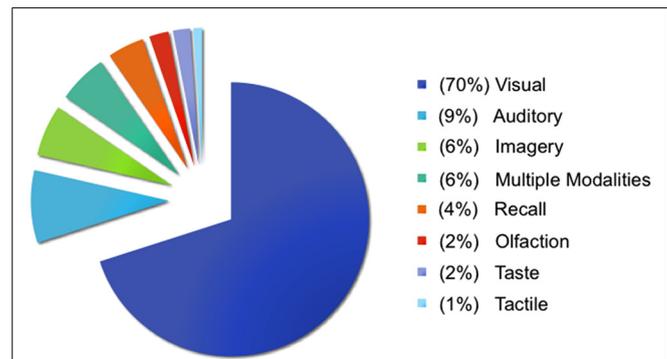
and Robbins, 2009; Mesquita et al., 2010; Barrett, 2013). “Cognitive” research domains (e.g., episodic and semantic memory, visual object recognition, language comprehension) are increasingly adopting a situated view of the mind (for empirical reviews, see Zwaan and Radvansky, 1998; Barsalou, 2003; Bar, 2004; Yeh and Barsalou, 2006; Mesquita et al., 2010). In contrast, emotion research largely remains entrenched in a “stimulus-response” reflexive approach to brain function, which typically views the brain as reacting to the demands of the environment, often in a simple, stereotyped way (cf. Raichle, 2010). Traditional “basic” emotion views often assume that an event (i.e., a stimulus) triggers one of several stereotyped responses in the brain and body that can be classified as either fear, disgust, anger, sadness, happiness, etc. (for a review of basic emotion models, see Tracy and Randles, 2011). Decades of research have revealed substantial variability in the neural, physiological, and behavioral patterns associated with these emotion categories (cf. Barrett, 2006; Lindquist et al., 2012). Whereas basic emotion approaches now focus on trying to identify primitive “core” (and often narrowly defined) instances of these emotions, alternative theoretical approaches to emotion, such as psychological construction, propose taking a situated approach to explaining the variability that exists in the experiences people refer to using words like fear, disgust,

anger, sadness, happiness (and using many other emotion terms; Barrett, 2009b, 2013).

In the psychological construction view that we have developed, emotions are not fundamentally different from other kinds of brain states (Barrett, 2009a, 2012; Wilson-Mendenhall et al., 2011). During emotional experiences and during other kinds of experiences, the brain is using prior experience to dynamically interpret ongoing neural activity, which guides an individual's responding in the situation. We refer to this process, which often occurs without awareness (i.e., it is a fundamental process for making sense of one's relation to the world at any given moment), as *situated conceptualization*. The term *situated* takes on a broad meaning in our view, referring to the distributed neural activity across the modal systems of the brain involved in constructing situations, not just to perception of the external environment or to what might be considered the background. More specifically, situated neural activity reflects the dynamic actions that individuals engage in, and the events, internal bodily sensations, and mentalizing that they experience, as well as the perceptions of the external environmental setting and the physical entities and individuals it contains (Wilson-Mendenhall et al., 2011).

Emotions, like other classes of mental experiences, operate in this situation-specific way because rich, cross-modal knowledge is critical for interpreting, inferring, and responding when similar situations occur in the future. On this view, situational knowledge develops for emotion categories like fear, anger, etc., as it does for other abstract categories of experiences (e.g., situations that involve the abstract categories gossip, modesty, or ambition). Experiences categorized as fear, for example, can occur when delivering a speech to a respected audience or when losing control while driving a car. A situated, psychological construction perspective suggests that it is more adaptive to respond differently in these situations, guided by knowledge of the situation, than to respond in a stereotyped way. Whereas responding in the social speech situation involves inferring what audience members are thinking, responding in the physical car situation involves rapid action and attention to the environment. Stereotyped responding in the form of preparing the body to flee or fight does not address the immediate threat present in either of these situations. A psychological construction approach highlights the importance of studying the situations commonly categorized as emotions like fear or anger, not because these situations merely describe emotions, but because emotions would not exist without them.

A significant challenge in taking a situated approach to studying emotional experience is maintaining a balance between the rich, multimodal nature of situated experiences and experimental control. Immersion in emotional situations through vividly imagined imagery is recognized as a powerful emotion induction method for evoking physiological responses (Lang et al., 1980; Lench et al., 2011). Imagery paradigms were initially developed to study situations thought to be central to various forms of psychopathology (Lang, 1979; Pitman et al., 1987), and remain a focus in clinical psychology (for a review, see Holmes and Mathews, 2010). In contrast, a small proportion of neuroimaging studies investigating emotion in typical populations use these methods. **Figure 1** illustrates the methods used across 397 studies in a



**FIGURE 1 | Methods used to study emotion and affect.** Visual methods typically involved viewing faces, pictures, films, words, sentences, and/or bodies. Auditory methods typically involved listening to voices, sounds, music, words, and/or sentences. Imagery methods typically involved generating imagery using personal memories, sentences, faces, and/or pictures (and are described further in the main text). Recall methods typically involved recalling personal events, words, films, or pictures. Tactile methods involved touch or thermal stimulation, olfaction methods involved smelling odors, and taste methods involved tasting food. Multiple modalities refers to studies that involved two or more of the aforementioned methods in the same study, with visual and auditory methods being the most frequent combination.

database constructed for neuroimaging meta-analyses of affect and emotion (Kober et al., 2008; Lindquist et al., 2012)<sup>1</sup>. Visual methods dominate (70% of studies), with the majority of these studies using faces (42% of visual methods) and pictures (36% of visual methods) like the International Affective Picture System (IAPS; Lang et al., 2008). In contrast, only 6% of studies have used imagery methods<sup>2</sup>. Imagery methods appear to be used more frequently when studying complex socio-emotional experiences that would be difficult to induce with an unfamiliar face or picture and that are often clinically oriented, including angry rumination (Denson et al., 2009), personal anxiety (Bystritsky et al., 2001), competition and aggression (Rauch et al., 1999; Pietrini et al., 2000), social rejection and insult (Kim et al., 2008; Kross et al., 2011), romantic love (Aron et al., 2005), moral disgust (Moll et al., 2005; Schaich Borg et al., 2008), and empathy (Perry et al., 2012).

Imagery-based neuroimaging studies of emotional experience typically take one of two approaches. The most frequent approach is to draw on the personal experiences of the participant, cueing specific, vivid memories in the scanner. Often participants' personal narratives are scripted and vividly imagined (guided by the experimenter) outside the scanner, and then a version of this script

<sup>1</sup>This meta-analytic database has recently been updated to include articles through 2011. The proportions reported here reflect the updated database.

<sup>2</sup>Lindquist et al. (2012) distinguished between "emotion perception" (defined as perception of emotion in others) and "emotion experience" (defined as experience of emotion in oneself) in their meta-analysis. When restricting our analysis of study methods to studies that involved emotion experience (as coded in the database), the use of imagery methods was still minimal (10% of 233 studies). Although emotional imagery is typically thought of as an induction of emotion experience, it seems likely that imagined situations, especially if they are social in nature, involve dynamic emotion perception as well.

is used to induce these memory-based emotional experiences during neuroimaging (e.g., Bystritsky et al., 2001; Marci et al., 2007; Gillihan et al., 2010). Less often, a specific visual stimulus is potent enough to easily evoke personal, emotional imagery in the scanner (e.g., face of a romantic partner; Aron et al., 2005; Kross et al., 2011). The second approach is to present standard prompts (e.g., a sentence) that participants use to generate imagery underlying emotional experiences (e.g., Colibazzi et al., 2010; Costa et al., 2010). A key strength of the first approach is that emotional experiences are tightly tied to situated, real-life memories, whereas a key strength of the second approach is the experimental control afforded by presenting the same prompts to all participants. In both cases, though, the situational context of the emotional experiences is typically lost, either because the situational details are specific to the individual (and thus lost in group-level analyses) or because standard prompts are not designed to cultivate and/or systematically manipulate the situational context of the emotional experience.

Building on the strengths of existing imagery-based approaches, we developed a neuroimaging procedure that would allow us to examine participants' immersion in rich, situated emotional experiences while maximizing experimental control and rigor. In our paradigm, participants first received training

outside the scanner on how to immerse themselves in richly detailed, full paragraph-long versions of emotional scenarios from a first-person perspective. The scenarios reflected two ecologically important situation types in which emotional experiences are often grounded: social evaluation and physical danger. Every scenario was constructed using written templates to induce a social evaluation emotional experience or a physical danger emotional experience (see **Table 1** for examples). Participants listened to audio recordings of the scenarios, which facilitated immersion by allowing participants to close their eyes. In the scanner, participants were prompted with shorter, core (audio) versions of the scenarios in the scanner, so that a statistically powerful neuroimaging design could be implemented.

We hypothesized that immersion across both social evaluation and physical danger situations would be characterized by distributed neural patterns across multiple sensory modalities and across regions involved in detecting and integrating salient sensory information. Much previous research has demonstrated neural overlap between sensorimotor perception/action and sensorimotor imagery (for a review, see Kosslyn et al., 2001). If our scenario immersion method induces richly situated emotional experiences, then the vivid mental imagery generated should be grounded in brain regions underlying sensory perception and action. Perhaps

**Table 1 | Examples of physical danger and social evaluation scenarios used in the experiment.**

Examples of physical danger situations
<p><b>Full version</b></p> <p>(P1) You are driving home after staying out drinking all night. (S1) The long stretch of road in front of you seems to go on forever. (P2A) You close your eyes for a moment. (P2C) The car begins to skid. (S2) You jerk awake. (S3) You feel the steering wheel slip in your hands.</p> <p><b>Core version</b></p> <p>(P1) You are driving home after staying out drinking all night. (P2) You close your eyes for a moment, and the car begins to skid.</p>
<p><b>Full version</b></p> <p>(P1) You are jogging along an isolated lake at dusk. (S1) Thick dark woods surround you as you move along the main well-marked trail. (P2A) On a whim, you veer onto an overgrown unmarked trail. (P2C) You become lost in the dark. (S2) The trees close in around you, and you cannot see the sky. (S3) You feel your pace quicken as you try to run out of the darkness.</p> <p><b>Core version</b></p> <p>(P1) You are jogging along an isolated lake at dusk. (P2) On a whim, you veer onto an overgrown unmarked trail, and become lost in the dark.</p>
Examples of social evaluation situations
<p><b>Full version</b></p> <p>(P1) You are at a dinner party with friends. (S1) A debate about a contentious issue arises that gets everyone at the table talking. (P2A) You alone bravely defend the unpopular view. (P2C) Your comments are met with sudden uncomfortable silence. (S2) Your friends are looking down at their plates, avoiding eye contact with you. (S3) You feel your chest tighten.</p> <p><b>Core version</b></p> <p>(P1) You are at a dinner party with friends. (P2) You alone bravely defend the unpopular view, and your comments are met with sudden uncomfortable silence.</p>
<p><b>Full version</b></p> <p>(P1) You are having drinks at a trendy bar. (S1) The bartender tosses ice cubes into glasses, making a loud clinking sound. (P2A) An attractive stranger strolls by, looks you up and down. (P2C) The stranger walks away smirking. (S2) People around you begin saying that you never meet the right people in bars. (S3) Your cheeks are burning.</p> <p><b>Core version</b></p> <p>(P1) You are having drinks at a trendy bar. (P2) An attractive stranger strolls by, looks you up and down, and walks away smirking.</p>

surprisingly, studies using imagery paradigms to investigate emotional experiences do not typically examine sensorimotor activity, because the goal is often to isolate a category of experience (e.g., anger, disgust) or other “emotion” components. In contrast, our approach is designed to examine the distributed neural patterns that underlie emotional experiences.

Our second, primary hypothesis was motivated by a situated approach to studying the varieties of emotional experience. We hypothesized that unique activity patterns for each situation type would occur in coordinated large-scale networks that reflect situated responding. Whereas networks underlying the social inference and mentalizing involved in responding to a social threat (in regions that make up the “default mode” network) would be reliably more active during social evaluation situations (for reviews of default mode network functions, see Buckner et al., 2008; Barrett and Satpute, 2013)<sup>3</sup>, networks underlying the visuospatial attention and action planning involved in responding to a physical threat would be reliably more active during physical danger situations (for reviews of attention networks, see Chun et al., 2011; Petersen and Posner, 2012; Posner, 2012). These large-scale, distributed networks largely consist of heteromodal regions that engage in the multimodal integration necessary for coordinated interpretation and responding (Sepulcre et al., 2012; Spreng et al., 2013).

As a further test of our second hypothesis, we examined whether participants’ trial-by-trial ratings of immersion during the training session correlated with neural activity, across social evaluation scenarios and across physical danger scenarios. If emotional experience is situated, then feeling immersed in a situation should be realized by neural circuitry that underlies engaging in the specific situation. Whereas immersion in social evaluation situations should occur when affect is grounded in mentalizing about others, immersion in physical danger situations should occur when affect is grounded in taking action in the environment.

## MATERIALS AND METHODS

### PARTICIPANTS

Twenty right-handed, native-English speakers from the Emory community, ranging in age from 20 to 33 (10 female), participated in the experiment. Six additional participants were dropped due to problems with audio equipment (three participants) or excessive head motion in the scanner. Participants had no history of psychiatric illness and were not currently taking any psychotropic medication. They received \$100 in compensation, along with anatomical images of their brain.

### MATERIALS

A full and core form of each scenario was constructed, the latter being a subset of the former (see **Table 1**). The full form served to

provide a rich, detailed, and affectively compelling scenario. The core form served to minimize presentation time in the scanner, so that the number of necessary trials could be completed in the time available. Each full and core scenario described an emotional situation from a first-person perspective, such that the participant could immerse him- or herself in it. As described shortly, participants practiced enriching the core form of the scenario during the training sessions using details from the full form, so that they would be prepared to immerse in the rich situational detail of the full forms during the scanning session when they received the core forms.

Both situation types were designed so the threat described could be experienced as any number of high arousal, negative emotions like fear or anger (and participants’ ratings of the ease of experiencing negative emotions in the two situation types validated this approach; see Wilson-Mendenhall et al., 2011 for details). In social evaluation situations, another person put the immersed participant in a socially threatening situation that involved damage to his or her social reputation/ego. In physical danger situations, the immersed participant put him- or herself in a physically threatening situation that involved impending or actual bodily harm.

Templates were used to systematically construct different scenarios in each situation type (social evaluation and physical danger). **Table 1** provides examples of the social evaluation and physical danger scenarios. Each template for the full scenarios specified a sequence of six sentences: three primary sentences ( $P_1$ ) also used in the related core scenario, and three secondary sentences ( $S_1$ ) not used in the core scenario that provided additional relevant detail. The two sentences in each core scenario were created using  $P_1$  as the first sentence and a conjunction of  $P_{2A}$  and  $P_{2C}$  as the second sentence.

For the social evaluation scenarios, the template specified the following six sentences in order:  $P_1$  described a setting and activity performed by the immersed participant in the setting, along with relevant personal attributes;  $S_1$  provided auditory detail about the setting;  $P_{2A}$  described an action (A) of the immersed participant;  $P_{2C}$  described the consequence (C) of that action;  $S_2$  described another person’s action in response to the consequence;  $S_3$  described the participant’s resulting internal bodily experience. The templates for the physical danger scenarios were similar, except that  $S_1$  provided visual detail about the setting (instead of auditory),  $S_2$  described the participant’s action in response to the consequence (instead of another person’s action), and  $S_3$  described the participant’s resulting external somatosensory experience (on the body surface).

A broad range of real-world situations served as the content of the experimental situations. The physical danger scenarios were drawn from situations that involved vehicles, pedestrians, water, eating, wildlife, fire, power tools, and theft. The social evaluation scenarios were drawn from situations that involved friends, family, neighbors, love, work, classes, public events, and service.

During the training sessions and the critical scan session, 30 social evaluation scenarios and 30 physical danger scenarios were presented. An additional three scenarios of each type were included in the training sessions so participants could practice the scanner task prior to the scan session.

<sup>3</sup>There is substantial evidence that default mode network (DMN) regions are active during tasks that involve social inference and mentalizing (for reviews, see Barrett and Satpute, 2013; Buckner and Carroll, 2007; Van Overwalle and Baetens, 2009) and that the DMN is disrupted in disorders involving social deficits (for reviews, see Menon, 2011; Whitfield-Gabrieli and Ford, 2012). Recent work has directly demonstrated that neural activity during social/mentalizing tasks occurs in the DMN as it is defined using resting state analyses (e.g., Andrews-Hanna et al., 2010) and that resting state connectivity in the DMN predicts individual differences in social processing (e.g., Yang et al., 2012).

## IMAGING DESIGN

The event-related neuroimaging design involved two critical events: (1) immersing in an emotional scenario (either a social evaluation or physical danger scenario) and (2) experiencing the immersed state in one of four ways upon hearing an auditory categorization cue (as emotional: fearful or angry, or as another active state: planning or observing). We will refer to the first event as “immersion” and the second event as “categorization.” Because all neural patterns described here reflect activity during the first immersion event, we focus on this element of the design (for the categorization results and related methodological details, please see Wilson-Mendenhall et al., 2011). This design afforded a unique opportunity to examine the situations in which emotions emerge before the emotional state was explicitly categorized. As will be described later, the participant could not predict which categorization cue would follow the scenario, so the immersion period reflects situated activity that is not tied to a specific emotion category.

In order to separate neural activity during the immersion events from neural activity during the categorization events, we implemented a catch trial design (Ollinger et al., 2001a,b). Participants received 240 complete trials that each contained a social evaluation scenario or a physical danger scenario followed immediately by one of the four categorization cues. Participants also received 120 partial “catch” trials containing only a scenario (with no subsequent categorization cue), which enabled separation of the first scenario immersion event from the second categorization event. The partial trials constituted 33% of the total trials, a proportion in the recommended range for an effective catch trial design. Each of the 30 social evaluation scenarios and the 30 physical danger scenarios was followed once by each categorization cue, for a total of 240 complete trials (60 scenarios followed by 4 categorizations). Each of the 60 scenarios also occurred twice as a partial trial, for a total of 120 catch trials.

During each of 10 fMRI runs, participants received 24 complete trials and 12 partial trials. The complete and partial trials were intermixed with no-sound baseline periods that ranged from 0 to 12 s in increments of 3 s (average 4.5 s) in a pseudo-random order optimized by optseq2 (Greve, 2002). On a given trial, participants could not predict whether a complete or partial trial was coming, a necessary condition for an effective catch trial design (Ollinger et al., 2001a,b). Participants also could not predict the type of situation or the categorization cue they would hear. Across trials in a run, social evaluation and physical danger situations each occurred 18 times, and each of the 4 categorization cues (anger, fear, observe, plan) occurred 6 times, equally often with social evaluation and physical danger scenarios. A given scenario was never repeated within a run.

## PROCEDURE

The experiment contained two training sessions and an fMRI scan session. The first training session occurred 24–48 h before the second training session, followed immediately by the scan. During the training sessions, participants were encouraged to immerse themselves in all scenarios from a first-person perspective, to imagine the scenario in as much vivid detail as possible, and to construct mental imagery as if the scenario events were actually happening

to them. The relation of the full to the core scenarios was also described, and participants were encouraged to reinstate the full scenario whenever they heard a core scenario.

During the first training session, participants listened over computer headphones to the full versions of the 66 scenarios that they would later receive on the practice trials and in the critical scan 24–48 h later, with the social evaluation and physical danger scenarios randomly intermixed. After hearing each full scenario, participants provided three judgments about familiarity and prior experiences, prompted by questions and response scales on the screen. After taking a break, participants listened to the 66 core versions of the scenarios, again over computer headphones and randomly intermixed. While listening to each core scenario, participants were instructed to reinstate the full version that they listened to earlier, immersing themselves fully into the respective scenario as it became enriched and developed from memory. After hearing each core scenario over the headphones, participants rated the vividness of the imagery that they experienced while immersed in the scenario. This task encouraged the participants to develop rich imagery upon hearing the core version. A detailed account of the first training session can be found in Wilson-Mendenhall et al. (2011).

During the second training session directly before the scan, participants first listened to the 66 full scenarios to be used in the practice and critical scans, and rated how much they were able to immerse themselves in each scenario, again hearing the scenarios over computer headphones and in a random order. After listening to each full scenario, the computer script presented the question, “How much did you experience ‘being there’ in the situation?” Participants responded on the computer keyboard, using a 1–7 scale, where one meant not experiencing being there in the situation at all, four meant experiencing being there a moderate amount, and seven meant experiencing being there very much, as if it was actually happening to them. The full scenarios were presented again at this point to ensure that participants were reacquainted with all the details before hearing the core versions later in the scanner. This first phase of the second training session lasted about an hour.

Participants were then instructed on the task that they would perform in the scanner and performed a run of practice trials. During the practice and during the scans, audio events were presented and responses collected using E-prime software (Schneider et al., 2002). On each complete trial, participants were told to immerse in the core version of a scenario as they listened to it, and that they would receive one of four words (anger, fear, observe, plan) afterward. The participant’s task was to judge how easy it was to experience what the word described in the context of the situation. The core scenario was presented auditorily at the onset of a 9 s period, lasting no more than 8 s. The word was then presented auditorily at the onset of a 3 s period, and participants responded as soon as ready. To make their judgments, participants pressed one of three buttons on a button box for not easy, somewhat easy, and very easy. During the practice trials, participants used an E-Prime button box to practice making responses. In the scanner, participants used a Current Designs fiber optic button box designed for high magnetic field environments. Participants were also told that there would be partial trials containing scenarios

and no word cues, and that they were not to respond on these trials.

At the beginning of the practice trials, participants heard the same short instruction that they would hear before every run in the scanner: "Please close your eyes. Listen to each scenario and experience being there vividly. If a word follows, rate how easy it was to have that experience in the situation." Participants performed a practice run equal in length to the runs that they would perform in the scanner. Following the practice run, the experimenter and the participant walked 5 min across campus to the scanner. Once settled safely and comfortably in the scanner, an initial anatomical scan was performed, followed by the 10 critical functional runs, and finally a second anatomical scan. Prior to beginning each functional run, participants heard the same short instruction from the practice run over noise-muffling headphones. Participants took a short break between each of the 8 min 3 s runs. Total time in the scanner was a little over 1.5 h.

### IMAGE ACQUISITION

The neuroimaging data were collected in the Biomedical Imaging Technology Center at Emory University on a research-dedicated 3T Siemens Trio scanner. In each functional run, 163 T2\*-weighted echo planar image volumes depicting BOLD contrast were collected using a Siemens 12-channel head coil and parallel imaging with an iPAT acceleration factor of 2. Each volume was collected using a scan sequence that had the following parameters: 56 contiguous 2 mm slices in the axial plane, interleaved slice acquisition, TR = 3000 ms, TE = 30 ms, flip angle = 90°, bandwidth = 2442 Hz/Px, FOV = 220 mm, matrix = 64, voxel size = 3.44 mm × 3.44 mm × 2 mm. This scanning sequence was selected after testing a variety of sequences for susceptibility artifacts in orbitofrontal cortex, amygdala, and the temporal poles. We selected this sequence not only because it minimized susceptibility artifacts by using thin slices and parallel imaging, but also because using 3.44 mm in the X–Y dimensions yielded a voxel volume large enough to produce a satisfactory temporal signal-to-noise ratio. In each of the two anatomical runs, 176 T1-weighted volumes were collected using a high resolution MPRAGE scan sequence that had the following parameters: 192 contiguous slices in the sagittal plane, single-shot acquisition, TR = 2300 ms, TE = 4 ms, flip angle = 8°, FOV = 256 mm, matrix = 256, bandwidth = 130 Hz/Px, voxel size = 1 mm × 1 mm × 1 mm.

### IMAGE PREPROCESSING AND ANALYSIS

Image preprocessing and statistical analysis were conducted in AFNI (Cox, 1996). The first anatomical scan was registered to the second, and the average of the two scans computed to create a single high-quality anatomical scan. Initial preprocessing of the functional data included slice time correction and motion correction in which all volumes were registered spatially to a volume within the last functional run. A volume in the last run was selected as the registration base because it was collected closest in time to the second anatomical scan, which facilitated later alignment of the functional and anatomical data. The functional data were then smoothed using an isotropic 6 mm full-width half-maximum Gaussian kernel. Voxels outside the brain were removed

from further analysis at this point, as were high-variability low-intensity voxels likely to be shifting in and out of the brain due to minor head motion. Finally, the signal intensities in each volume were divided by the mean signal value for the respective run and multiplied by 100 to produce percent signal change from the run mean. All later analyses were performed on these percent signal change data.

The averaged anatomical scan was corrected for non-uniformity in image intensity, skull-stripped, and then aligned with the functional data. The resulting aligned anatomical dataset was warped to Talairach space using an automated procedure employing the TT\_N27 template (also known as the Colin brain, an averaged dataset from one person scanned 27 times).

Regression analyses were performed on each individual's pre-processed functional data using a canonical, fixed-shape Gamma function to model the hemodynamic response. In the first regression analysis, betas were estimated using the event onsets for 10 conditions: 2 situation immersion conditions (social, physical) and 8 categorization conditions that resulted from crossing the situation with the categorization cue (social-anger, physical-anger, social-fear, physical-fear, social-observe, physical-observe, social-plan, physical-plan). Again, we only present results for the two situation immersion conditions here (see Wilson-Mendenhall et al., 2011 for the categorization results). The two situation immersion conditions were modeled by creating regressors that included scenario immersion events from both the complete trials and the partial trials. Including scenario immersion events from both trial types in one regressor made it possible to mathematically separate the situation immersion conditions from the subsequent categorization conditions (Ollinger et al., 2001a,b). Because scenario immersion events were 9 s in duration, the Gamma function was convolved with a boxcar function for the entire duration to model the situation immersion conditions. Six regressors obtained from volume registration during preprocessing were also included to remove any residual signal changes correlated with movement (translation in the X, Y, and Z planes; rotation around the X, Y, and Z axes). Scanner drift was removed by finding the best-fitting polynomial function correlated with time in the preprocessed time course data.

At the group level, the betas resulting from the each individual's regression analysis were then entered into a second-level, random-effects ANOVA. Two key analyses were computed at this level of analysis using a voxel-wise threshold of  $p < 0.005$  in conjunction with the 41-voxel extent threshold determined by AFNI ClustSim to produce an overall corrected threshold of  $p < 0.05$ . In the first analysis (that assessed our first hypothesis), we extracted clusters that were more active during immersion in social evaluation situations than in the no-sound baseline and clusters that were more active during immersion in physical danger situations than in the no-sound baseline (using the voxel-wise and extent thresholds specified above). We then entered the results of these two contrasts (social evaluation > baseline; physical danger > baseline) into a conjunction analysis to determine clusters shared by the two situation types (i.e., overlapping regions of activity). In the second analysis (that assessed our second hypothesis), we computed a standard contrast to directly compare immersion during social

evaluation situations to immersion during physical danger situations using *t* tests (social evaluation > physical danger; physical danger > social evaluation).

A second individual-level regression was computed to examine the relationship between neural activity and the scenario immersion ratings collected during the training session just prior to the scan session, providing an additional test of our second hypothesis. This regression model paralleled the first regression model with the following exceptions. In this regression analysis, each participant's "being there" ratings were specified trial-by-trial for each scenario in the social evaluation immersion condition and in the physical danger immersion condition. For the two situation immersion conditions (social evaluation and physical danger), both the onset times and ratings were then entered into the regression using the amplitude modulation option in AFNI. This option specified two regressors for each situation immersion condition, which were used to detect: (1) voxels in which activity was correlated with the ratings (also known as a parametric regressor); (2) voxels in which activity was constant for the condition and was not correlated with the ratings.

At the group level, each participant's betas produced from the first parametric regressor for each situation immersion condition (i.e., indicating the strength of the correlation between neural activity and "being there" immersion ratings) were next entered into a second-level analysis. In this analysis, the critical statistic for each condition was a *t* test indicating if the mean across individuals differed significantly from zero (zero indicating no correlation between neural activity and the ratings). In these analyses, a slightly smaller cluster size of 15 contiguous voxels was used in conjunction with the voxel-wise threshold of  $p < 0.005$ .

In summary, this analysis is examining whether scenarios rated as easier to immerse in during the training are associated with greater neural activity in any region of the brain (the individual-level analysis), and whether this relationship between immersion ratings and neural activity is consistent across participants (group-level analysis). We computed this analysis separately for social evaluation and for physical danger situation types to test our hypothesis. This analysis is not examining between-subject individual differences in immersion (i.e., whether participants who generally experience greater immersion across all scenarios also show greater neural activity in specific regions), which is a different question that is not of interest here.

## RESULTS

### COMMON NEURAL ACTIVITY DURING IMMERSION ACROSS SITUATIONS

Our first hypothesis was that neural activity during both situations would be reliably greater than baseline across multiple sensory modalities and across regions involved in detecting and integrating salient sensory information (see **Table 2** for the baseline contrasts). As shown in **Figure 2A**, neural activity was reliably greater than baseline in bilateral primary somatomotor and visual cortex, as well as premotor cortex, SMA, and extrastriate visual cortex, suggesting that participants easily immersed in the situations. The self-reported rating data from the training session confirmed that participants found the social evaluation and physical danger situations relatively easy to immerse in (see **Figure 2B**), with no

significant differences in "being there" ratings between situation types [repeated measures *t* test;  $t(19) = 1.64, p > 0.05$ ]. Because participants listened to the scenarios with their eyes closed and because participants did not make responses while immersing in the scenarios, it is significant that these sensorimotor regions were significantly more active than the no-sound baseline. As would be expected with an auditory, language-based immersion procedure, we observed activity in bilateral auditory cortex and in superior temporal and inferior frontal regions associated with language processing, with more extensive activity in the left frontal regions.

Consistent with the hypothesis that immersion would also generally involve selection, encoding, and integration of salient sensory and other information, we observed activity in bilateral hippocampus and in right amygdala (see **Figure 2C**). Extensive evidence implicates the hippocampus in mnemonic functions (Squire and Zola-Morgan, 1991; Tulving, 2002; Squire, 2004), especially the integration and binding of the multimodal information involved in constructing (and reconstructing) situated memories (Addis and McAndrews, 2006; Kroes and Fernandez, 2012). More recent evidence establishes a central role for this structure in simulating future, imagined situations (Addis et al., 2007; Hassabis et al., 2007; Schacter et al., 2007, 2012), which is similar in nature to our immersion paradigm, and which requires similar integration and binding of concepts established in memory (from prior experience). The amygdala plays a central role in emotional experiences by efficiently integrating multisensory information to direct attention and guide encoding (Costafreda et al., 2008; Bliss-Moreau et al., 2011; Klasen et al., 2012; Lindquist et al., 2012), especially during situations that involve threat (Adolphs, 2008; Miskovic and Schmidt, 2012). As we will see, no differences emerged in the amygdala or in the hippocampus during the social evaluation and physical danger situations, suggesting these structures played a similar role in both types of experiences.

### UNIQUE NEURAL PATTERNS EMERGE FOR SOCIAL EVALUATION AND PHYSICAL DANGER SITUATIONS

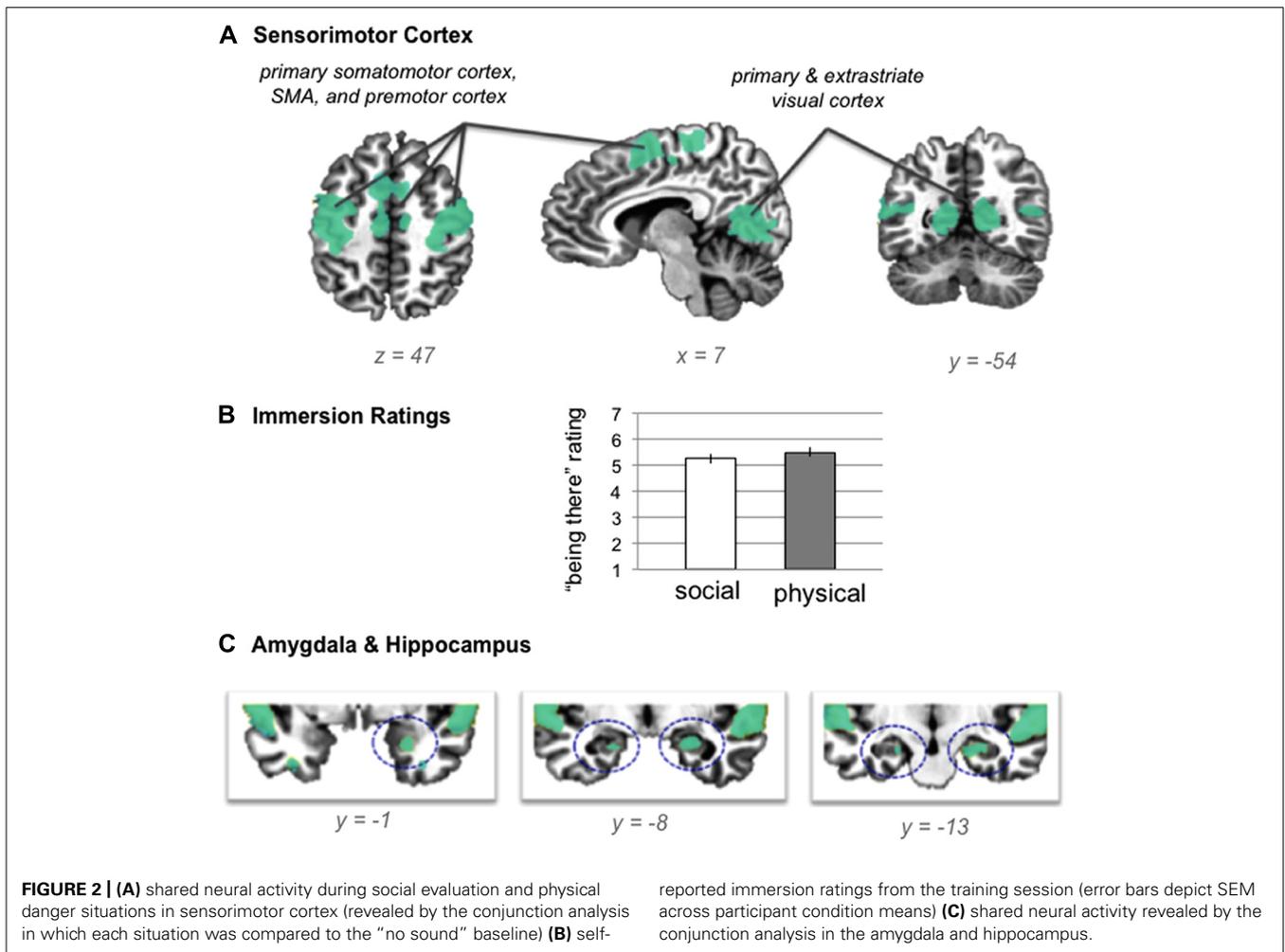
Our second hypothesis was that networks underlying the social inference and mentalizing involved in responding to a social threat would be reliably more active during social evaluation situations, whereas networks underlying visuospatial attention and action planning involved in responding to a physical threat would be reliably more active during physical danger situations. As **Table 3**, together with **Figures 3–5**, illustrate, the neural patterns that emerged when we compared social evaluation situations to physical danger situations are consistent with these predictions. **Figure 3** shows these results on representative 2D slices, with regions showing reliably greater activity during social evaluation in orange, and regions showing reliably greater activity during physical danger in green. **Figures 4** and **5** display these maps projected onto the surface of the brain<sup>4</sup>, and directly compare the maps from this study with

<sup>4</sup>It is important to note that each individual's data were not analyzed on the surface. We are using a standardized (Talairach) surface space for illustration of the group results in comparison to the resting state network maps from a large sample that have been made freely available (Yeo et al., 2011).

**Table 2 | Social evaluation > baseline and physical danger > baseline contrasts.**

Cluster	Brain region	Brodmann area(s)	Mean <i>t</i>	Spatial extent	Peak		
					<i>x</i>	<i>y</i>	<i>z</i>
<b>Social evaluation &gt; baseline</b>							
1	R temp pole/STG/STS	38, 21, 22, 41, 42	4.73	1868	46	-16	-7
	R angular g	39					
	R ITG/fusiform g	37, 19					
	R mid/sup occipital g	19					
2	L temp pole/STG/STS	38, 21, 22, 41, 42	4.83	1780	-45	9	-15
	L angular g	39					
	L mid/sup occipital g	19					
3	L and R calcarine/lingual g	17, 18, 19	4.21	1532	14	-57	14
	L and R posterior cingulate	31					
	L and R parahippocampal g	35, 36					
	L and R hippocampus/amygdala						
4	L premotor/precentral g	6, 4	4.36	921	-37	-6	50
	L postcentral g	2, 3					
	L lateral PFC/Ant insula	44, 45, 46, 9					
5	L and R SMA/precentral g	6, 4	4.60	596	-4	7	48
6	R premotor/precentral g	6, 4	4.52	501	50	-11	51
	R postcentral g	2, 3					
7	mPFC/mOFC	10, 11	4.45	115	-1	34	-8
8	R lateral PFC	45/46	4.09	77	52	19	22
9	L fusiform g	37	4.00	58	-36	-38	-11
<b>Physical danger &gt; baseline</b>							
1	L and R SMA/premotor	6	4.37	6887	-5	7	47
	L and R precentral g	4					
	L and R postcentral g	2, 3					
	L and R mid cingulate	24, 31					
	L lateral PFC/Ant insula	44, 45, 46, 9					
	L and R temp Pole/STG/STS	38, 21, 22, 41, 42					
	L and R MTG	37					
	L ITG/fusiform	37					
	L and R parahippocampal g	35, 36					
	L and R hippocampus/amygdala						
	L and R mid/sup occipital g	19					
	L and R calcarine/lingual g	17, 18, 19					
	L inferior parietal	40, 7					
2	L and R thalamus		4.22	85	-9	-21	2
3	R lateral PFC	45/46	3.93	68	55	19	26

Spatial extent is the number of 23.67 mm<sup>3</sup> functional voxels. L is left and R is right, Ant is anterior, Mid is middle, Sup is superior, m is medial, and g is gyrus. PFC is prefrontal cortex and OFC is orbitofrontal cortex. STG is superior temporal gyrus, STS superior temporal sulcus, MTG is middle temporal gyrus, and ITG is inferior temporal gyrus. SMA is supplementary motor area.



the large-scale networks that have been defined using resting state connectivity techniques across large samples (Yeo et al., 2011).

#### **Heightened activity in the default mode network during social evaluation**

As displayed in **Figure 3** and **Table 3**, robust activity was observed during immersion in social evaluation situations (vs. physical danger situations) in midline medial prefrontal and posterior cingulate regions, as well as lateral temporal regions, in which activity spanned from the temporal pole to the posterior superior temporal sulcus/temporoparietal junction bilaterally, and on the left, extended in to inferior frontal gyrus. This pattern of activity maps onto a network that is often referred to as the “default mode” network (Gusnard and Raichle, 2001; Raichle et al., 2001; Buckner et al., 2008). **Figure 4** illustrates the overlap between the default mode network and the pattern of neural activity that underlies immersing in social evaluation situations here (Yeo et al., 2011). The default mode network has been implicated in mentalizing and social inference (i.e., inferring what others’ are thinking/feeling and how they will act), as well as other socially motivated tasks, including autobiographical memory retrieval, envisioning the

future, and moral reasoning (for reviews, see Buckner et al., 2008; Van Overwalle and Baetens, 2009; Barrett and Satpute, 2013). Consistent with the idea of situated emotional experience, participants engaged in the social inference and mentalizing that would be adaptive in responding to a social threat when immersed in social evaluation situations.

#### **Heightened activity in fronto-parietal attention networks during physical danger**

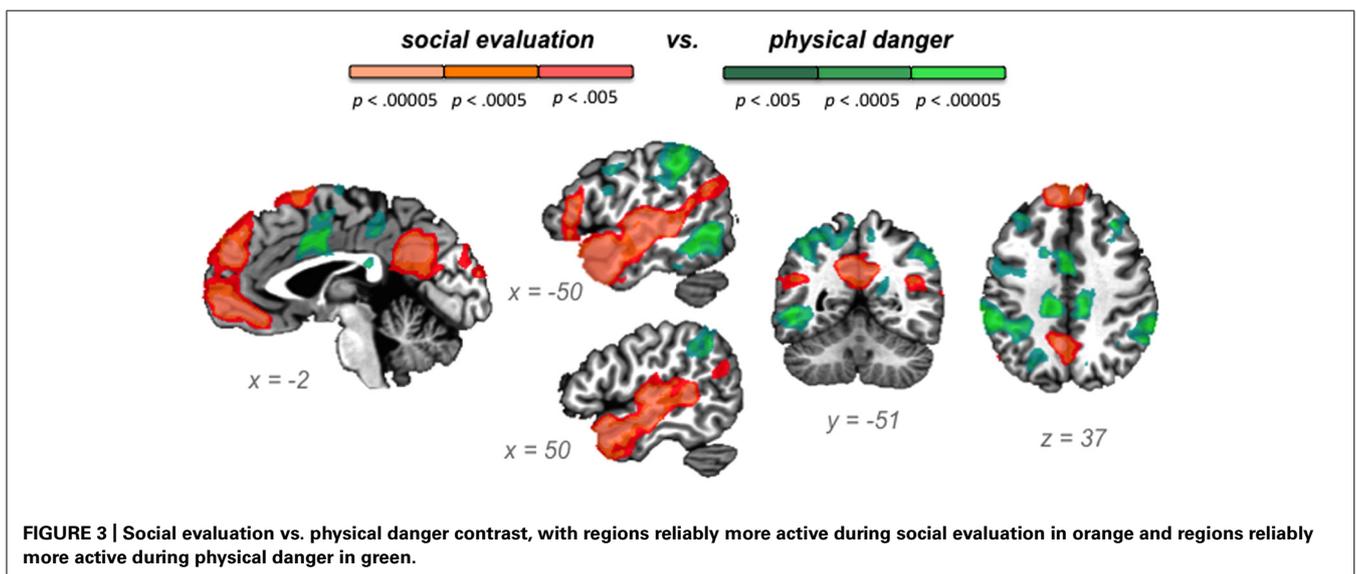
**Figure 3** and **Table 3** show the fronto-parietal patterns of activity observed during immersion in physical danger situations (vs. social evaluation situations). In addition to lateral frontal and parietal regions (including bilateral middle frontal gyrus, bilateral inferior frontal gyrus extending into pars orbitalis, bilateral inferior parietal lobule, and bilateral superior parietal/precuneus), neural activity was also reliably greater in right anterior insula, mid cingulate cortex, and bilateral premotor cortex during immersion in physical danger situations. **Figure 5** illustrates the overlap between this pattern of activity and three networks that have been implicated in attention<sup>5</sup> (Chun et al., 2011; Petersen and

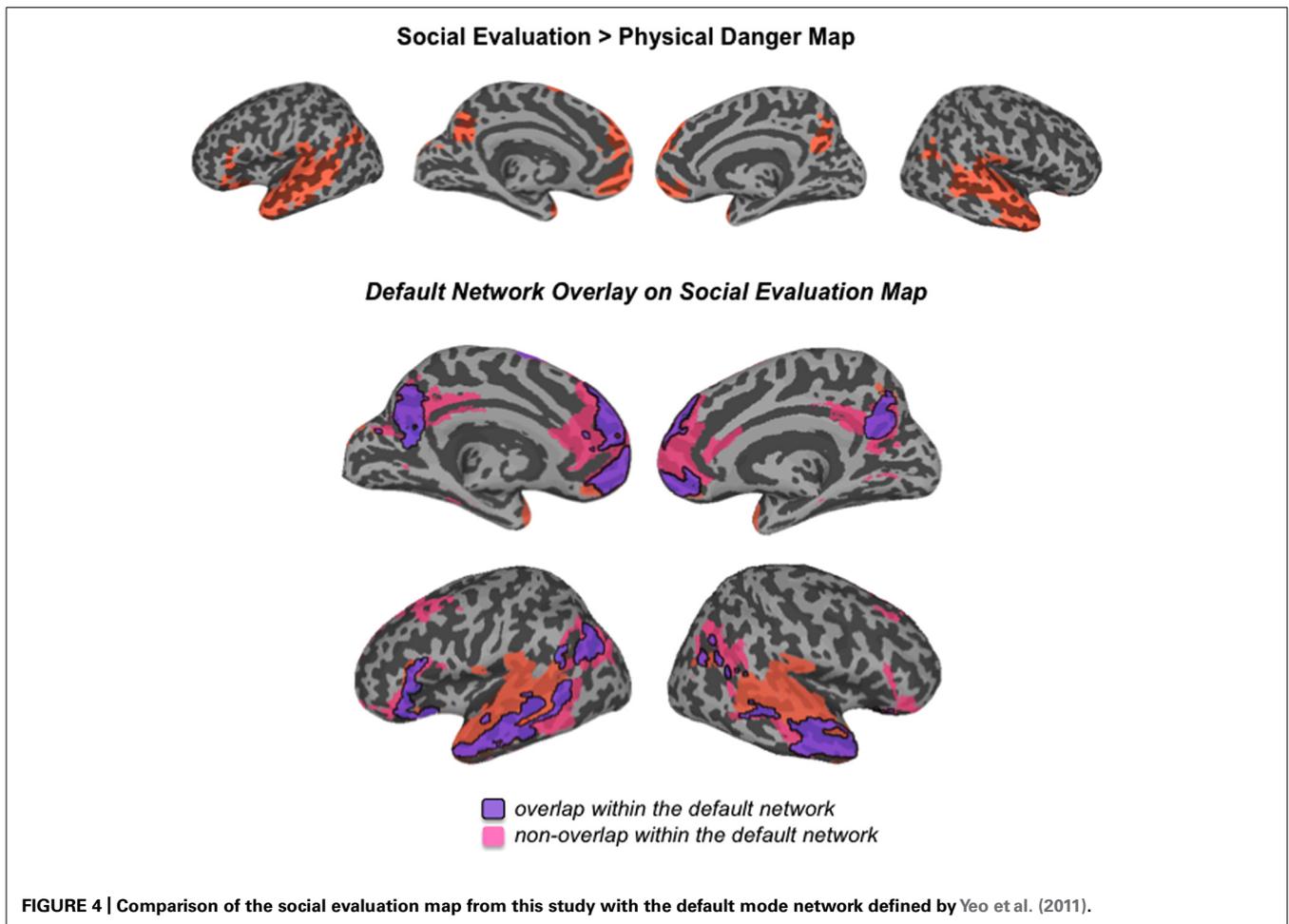
<sup>5</sup> These networks are sometimes referred to by different names, and can take somewhat different forms depending on the methods used to define them (with core

**Table 3 | Brain regions that emerged in the social evaluation vs. physical danger contrast.**

Cluster	Brain region	Brodmann area(s)	Mean <i>t</i>	Spatial extent	Peak		
					<i>x</i>	<i>y</i>	<i>z</i>
<b>Social evaluation &gt; physical danger</b>							
1	L STG/STS/post insula/angular g/temp pole/OFC/IFG	41, 42, 22, 21, 39, 38, 47, 45	5.13	2059	-58	-17	-1
2	R STG/STS/post insula/temp pole	41, 42, 22, 21, 38	4.77	1668	51	9	-20
3	mPFC/mOFC/SMA	10, 11, 9, 8, 6	4.63	1136	4	51	31
4	Post cingulate/precuneus	31, 7	4.73	498	-7	-53	34
5	R STG/STS/angular g	22, 39	3.97	112	40	-49	22
6	L cuneus	18	3.67	57	-7	-95	23
<b>Physical danger &gt; social evaluation</b>							
1	L inf/sup parietal/precuneus	40, 7	4.23	992	-59	-33	38
2	Mid cing/L premotor/L MFG	24, 6	4.20	715	4	6	31
3	L MTG/fusiform g/parahippocampal g	37, 20, 35	4.37	478	-49	-54	0
4	Mid cing	31, 23	4.35	321	-13	-26	37
5	L MFG	46, 9, 10	4.14	266	-37	38	16
6	R MFG/Ant insula/OFC	10	3.99	212	37	44	6
7	R inf parietal	40	4.14	199	59	-37	35
8	R premotor	6	4.16	173	15	2	59
9	R MFG	9	3.95	104	31	30	38
10	R precuneus	7	3.94	74	7	-56	53
11	L OFC	11	3.82	49	-29	44	-5
12	R retrosplenial	29	3.78	42	12	-44	12

Spatial extent is the number of 23.67 mm<sup>3</sup> functional voxels. L is left and R is right. Post is posterior, Ant is anterior, Inf is inferior, Sup is superior, m is medial, and g is gyrus. PFC is prefrontal cortex, OFC is orbitofrontal cortex, Cing is cingulate, and MFG is middle frontal gyrus. STG is superior temporal gyrus, STS is superior temporal sulcus, and MTG is middle temporal gyrus. SMA is supplementary motor area.





Posner, 2012; Posner, 2012). The most significant overlap was observed in the lateral fronto-parietal executive network and the dorsal attention network. These networks are thought to allocate attentional resources to prioritize specific sensory inputs (what is often referred to as “orienting” to the external environment) and to guide flexible shifts in behavior (Dosenbach et al., 2007; Petersen and Posner, 2012). The operations they carry out are critical for maintaining a vigilant state (Tang et al., 2012), which is important during threat. Less overlap was evident in the ventral attention network that is thought to interrupt top-down operations through bottom-up “salience” detection (Corbetta et al., 2008), although robust activity was observed in the mid cingulate regions shown in **Figure 5** that support the action monitoring that occurs, especially, in situations involving physical pain (Morecraft and Van Hoesen, 1992; Vogt, 2005). Taken together, this pattern of results suggests, strikingly, that immersion in the physical danger situations (from a first-person perspective with eyes closed) engaged attention networks that are studied almost exclusively using

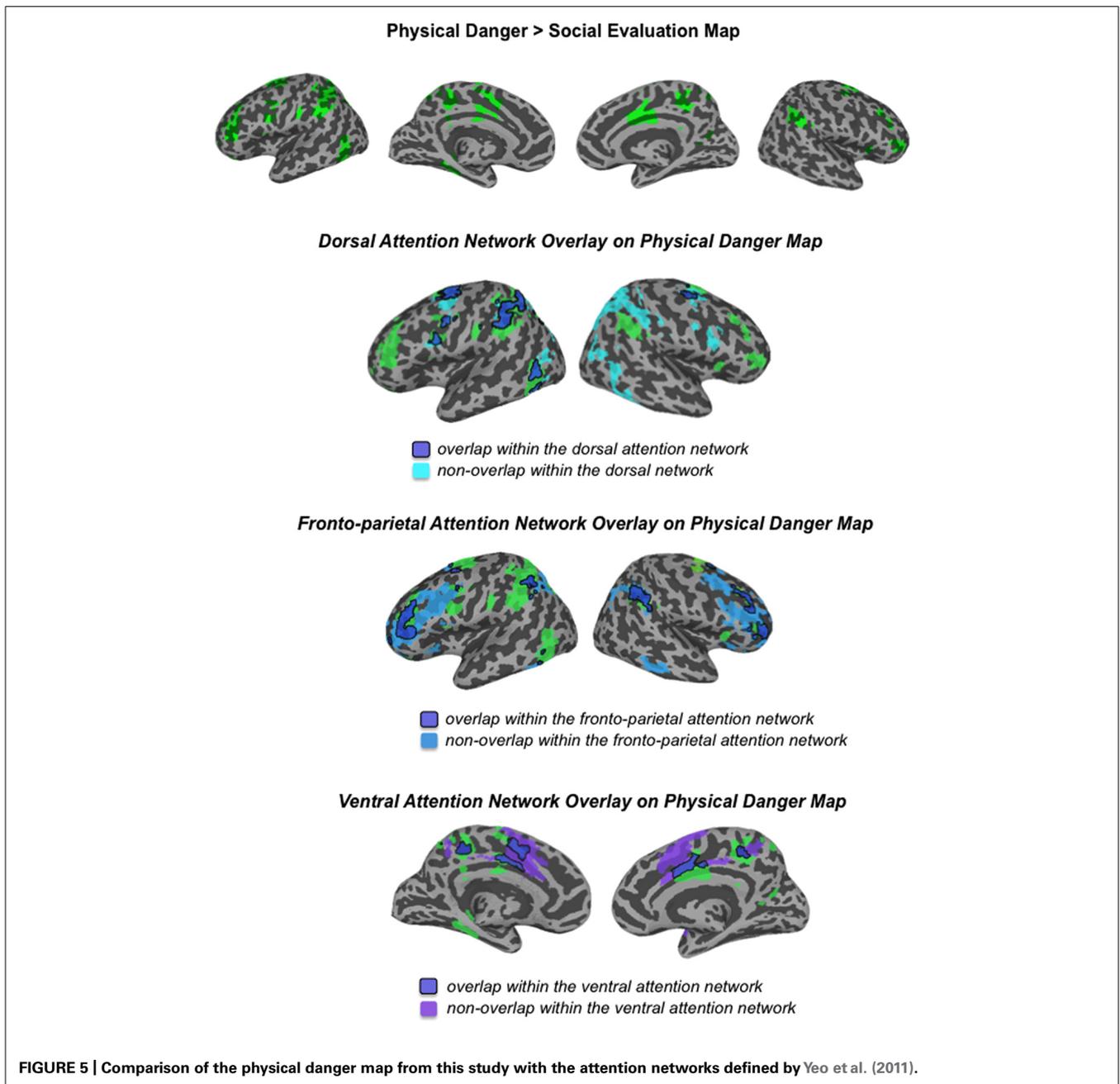
nodes remaining the same). Because the network maps we present here are taken from Yeo et al. (2011), we use their terminology. They note (and thus so do we) that the ventral attention network, especially, is similar to what has been described as the salience network (Seeley et al., 2007) and the cingulo-opercular network (Dosenbach et al., 2007).

external visual cues. Consistent with the idea of situated emotional experience, participants engaged in the monitoring of the environment and preparation for flexible action that would be adaptive in action to a physical threat when immersed in physical danger situations.

#### ***Immersion ratings correlate with activity in different regions during social evaluation vs. physical danger situations***

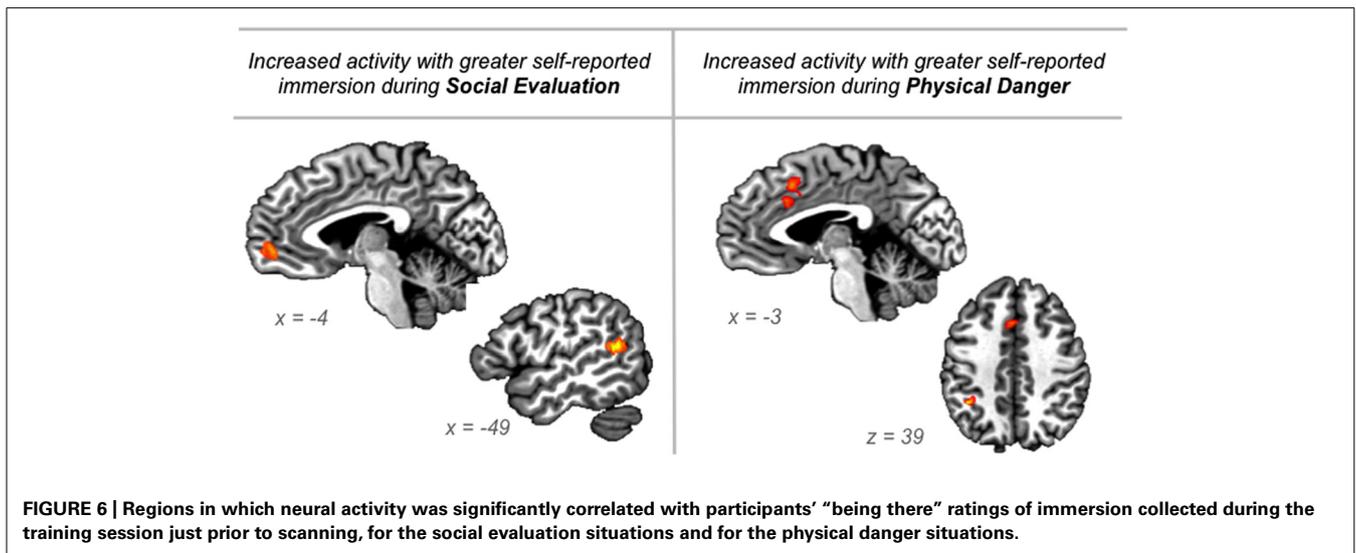
To provide another test of our second hypothesis, we examined whether self-reported immersion ratings of “being there” in the situation (from the training session) were associated with brain activity during the two situation types. If emotional experience is situated, then feeling immersed in a situation should be realized by neural circuitry that underlies engaging in the specific situation. Whereas immersion in social evaluation situations should occur when affect is grounded in mentalizing about others, immersion in physical danger situations should occur when affect is grounded in taking action in the environment. The results displayed in **Figure 6** support this prediction.

During social evaluation situations, participants’ immersion ratings correlated with activity in anterior medial prefrontal cortex (frontal pole area; peak voxel  $-6\ 51\ 0$ ; 23 voxels) and in superior temporal gyrus/sulcus (peak voxel  $-47\ -49\ 14$ ; 24 voxels; see



**Figure 6).** As described above, these regions are part of the default mode network and are central to social perception and mentalizing (Allison et al., 2000; Buckner et al., 2008; Adolphs, 2009; Van Overwalle, 2009). The anterior, frontal pole region of medial prefrontal cortex is considered the anterior hub of the default mode network (Andrews-Hanna et al., 2010) that integrates affective information from the body with social event knowledge (including inferences about others' thoughts) originating in ventral and dorsal aspects of medial prefrontal cortex, respectively (Mitchell et al., 2005; Krueger et al., 2009). This integration may underlie the experience of "personal significance" (Andrews-Hanna et al., 2010) that appears important for immersing in social evaluation situations.

In contrast, during physical danger situations, participants' immersion ratings correlated with activity in dorsal anterior cingulate/mid cingulate (extending into SMA; peak  $-1\ 17\ 40$ ; 40 voxels) and in left inferior parietal cortex (peak  $-36\ -46\ 39$ ; 15 voxels; see **Figure 6**). The robust cluster of activity that emerged in the cingulate is part of the ventral attention "salience" network, and it is anterior to the mid cingulate activity observed in the initial whole-brain contrasts reported above. Because this region has been implicated across studies of emotion, pain, and cognitive control, and because it is anatomically positioned at the intersection of insular-limbic and fronto-parietal sub-networks within the attention system, it may



play an especially important role in specifying goal-directed action based on affective signals originating in the body (Shackman et al., 2011; Touroutoglou et al., 2012). This integration may underlie the experience of action-oriented agency (Craig, 2009) that appears important for immersing in physical danger situations. The significant correlation with activity in left inferior parietal cortex, which supports planning action in egocentric space (e.g., Fogassi and Luppino, 2005), further suggests that immersion in physical danger situations is driven by preparing to act in the environment.

## DISCUSSION

Our novel scenario immersion paradigm revealed robust patterns of neural activity when participants immersed themselves in social evaluation scenarios and in physical danger scenarios. Consistent with participants' high self-reported immersion ratings, neural activity across multiple sensory regions, and across limbic regions involved in the multisensory integration underlying the selection, encoding, and interpretation that influences what is salient and remembered (e.g., amygdala, hippocampus), occurred during both situation types. In addition to this shared activity, distributed patterns unique to each situation type reflected situated responding, with regions involved in mentalizing and social cognition more active during social evaluation and with regions involved in attention and action planning more active during physical danger.

Taken together, these findings suggest that our method produced vivid, engaging experiences during neuroimaging scans and that it could be used to study a variety of emotional experiences. One reason this immersion paradigm may be so powerful is that people often find themselves immersed in imagined situations in day-to-day life. Large-scale experience sampling studies have revealed that people spend much of their time imagining experiences that are unrelated to the external world around them (e.g., Killingsworth and Gilbert, 2010). An important direction for future research will be to understand if, consistent with other imagery-based paradigms, physiological changes occur

during our scenario immersion paradigm and if these physiological changes are associated with subjective experiences of immersion.

The scenarios we developed for this study represent a small subset of the situations that people experience in real life (see also Wilson-Mendenhall et al., 2013). Because emotional experiences vary tremendously, it is adaptive to develop situated knowledge that guides inference and responding when similar situations arise in the future (Barsalou, 2003, 2008, 2009; Barrett, 2013). Here, we focused on immersion in emotion-inducing situations before they were explicitly categorized as an emotion (or another state). From our perspective, the situation plays a critical role in the emergence of an emotion, and it should not be considered a separate phenomenon from it (Barrett, 2009b, 2012; Wilson-Mendenhall et al., 2011). For example, it would be impossible to experience *fear* upon delivering a public speech without inferring others' thoughts. Instead of viewing mentalizing as a "cold" cognitive process that interacts with a primitive "hot" emotion, we view mentalizing as an essential part of the situation in which the emotion emerges. Likewise, it would be impossible to experience *fear* upon getting lost in the woods without focusing attention on the environment (in other words, if one was instead lost in internal thought while traversing the same environment, it is unlikely that this fear would occur). We propose that it will be more productive to study emotional experience as dynamic situated conceptualizations that the brain continually generates to interpret one's current state (based on prior experience), as opposed to temporally constrained cognition-emotion frameworks that often strip away much of the dynamically changing situated context. A situated approach also offers new insights into studying dynamic emotion regulation and dysregulation (Barrett et al., in press).

Network approaches to brain function provide functional frameworks for interpreting the distributed patterns that characterize situated experiences (Cabral et al., 2011; Deco et al., 2011; Lindquist and Barrett, 2012; Barrett and Satpute, 2013). As shown in **Figures 4 and 5**, the patterns unique to each situation type in

this study can be differentiated by the anatomically constrained resting state networks<sup>6</sup> identified in previous work (Raichle et al., 2001; Fox et al., 2005; Vincent et al., 2006; Dosenbach et al., 2007; Fair et al., 2007; Seeley et al., 2007; Yeo et al., 2011; Touroutoglou et al., 2012). Whereas the neural patterns underlying social threat situations primarily map onto the default mode network that supports social inference and mentalizing, the neural patterns underlying physical threat situations primarily map onto attention networks underlying monitoring of the environment and action planning. The neural pattern unique to each situation type reflects adaptive, situated responding. Furthermore, regions traditionally associated with emotion diverged in line with these networks (e.g., ventromedial prefrontal cortex as part of the default mode network; lateral orbitofrontal cortex and cingulate regions as part of the attention networks). Interestingly, these regions appear to be central to immersion in each type of situation, with the anterior medial prefrontal cortex (which is often considered part of ventromedial prefrontal cortex) associated with immersion during social evaluation situations and dorsal anterior cingulate associated with immersion during physical danger situations. These results suggest, strikingly, that the brain realizes immersion differently depending on the situation.

Resting state networks provide a starting point for examining how networks underlie situated experiences, but recent evidence suggests that coordination between regions in these networks dynamically changes during different psychological states (e.g., van Marle et al., 2010; Raz et al., 2012; Wang et al., 2012). In this study, for example, the neural patterns underlying physical danger experiences recruited various aspects of several different attention networks. Attention is primarily studied using simple visual detection tasks that examine external stimuli vs. internal goal dichotomies. Recent reviews emphasize the need for research that examines how attention systems operate during experiences guided by memory (e.g., Hutchinson and Turk-Browne, 2012), which arguably constitute much of our experience. Because inferior parietal cortex and cingulate regions figured prominently in the pattern observed across the attention networks in this study, this particular configuration may reflect the attention operations involved in coordinating bodily actions in space. It is also important to consider that these patterns reflect relative differences between the social and physical threat situations. As we showed initially, the situation types also share patterns of activity that contribute to the overall pattern of situated activity. In our view, it is useful to think about situated neural activity as dynamically changing patterns that are distributed across structurally and functionally distinct networks (see also Barrett and Satpute, 2013). Even within a structurally defined network, different distributed patterns of neural activity may reflect unique functional motifs that underlie different experiences and behaviors (Sporns and Kotter, 2004).

In closing, a psychological construction approach to studying situated emotion motivates different questions than traditional

approaches to studying emotion. It invites shifting research agendas from defining five or so emotion categories to studying the rich situations that characterize emotional experiences.

## ACKNOWLEDGMENTS

Preparation of this manuscript was supported by an NIH Director's Pioneer Award DPI OD003312 to Lisa Feldman Barrett at Northeastern University with a sub-contract to Lawrence Barsalou at Emory University. We thank A. Satpute and K. Lindquist for the meta-analysis codes indicating study methods/tasks.

## REFERENCES

- Addis, D. R., and McAndrews, M. P. (2006). Prefrontal and hippocampal contributions to the generation and binding of semantic associations during successful encoding. *Neuroimage* 33, 1194–1206. doi: 10.1016/j.neuroimage.2006.07.039
- Addis, D. R., Wong, A. T., and Schacter, D. L. (2007). Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. *Neuropsychologia* 45, 1363–1377. doi: 10.1016/j.neuropsychologia.2006.10.016
- Adolphs, R. (2008). Fear, faces, and the human amygdala. *Curr. Opin. Neurobiol.* 18, 166–172. doi: 10.1016/j.conb.2008.06.006
- Adolphs, R. (2009). The social brain: neural basis of social knowledge. *Annu. Rev. Psychol.* 60, 693–716. doi: 10.1146/annurev.psych.60.110707.163514
- Allison, T., Puce, A., and McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends Cogn. Sci.* 4, 267–278. doi: 10.1016/S1364-6613(00)01501-1
- Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., and Buckner, R. L. (2010). Functional-anatomic fractionation of the brain's default network. *Neuron* 65, 550–562. doi: 10.1016/j.neuron.2010.02.005
- Aron, A., Fisher, H., Mashek, D. J., Strong, G., Li, H., and Brown, L. L. (2005). Reward, motivation, and emotion systems associated with early-stage intense romantic love. *J. Neurophysiol.* 94, 327–337. doi: 10.1152/jn.00838.2004
- Aydede, H., and Robbins, P. (eds). (2009). *The Cambridge Handbook of Situated Cognition*. New York, NY: Cambridge University Press.
- Bar, M. (2004). Visual objects in context. *Nat. Rev. Neurosci.* 5, 617–629. doi: 10.1038/nrn1476
- Barrett, L. F. (2006). Emotions as natural kinds? *Perspect. Psychol. Sci.* 1, 28–58. doi: 10.1111/j.1745-6916.2006.00003.x
- Barrett, L. F. (2009a). The future of psychology: connecting mind to brain. *Perspect. Psychol. Sci.* 4, 326–339. doi: 10.1111/j.1745-6924.2009.01134.x
- Barrett, L. F. (2009b). Variety is the spice of life: a psychological construction approach to understanding variability in emotion. *Cogn. Emot.* 23, 1284–1306. doi: 10.1080/02699930902985894
- Barrett, L. F. (2012). Emotions are real. *Emotion* 12, 413–429. doi: 10.1037/a0027555
- Barrett, L. F. (2013). Psychological construction: a Darwinian approach to the science of emotion. *Emot. Rev.* 5, 379–389.
- Barrett, L. F., and Satpute, A. B. (2013). Large-scale brain networks in affective and social neuroscience: towards an integrative functional architecture of the brain. *Curr. Opin. Neurobiol.* 23, 361–372. doi: 10.1016/j.conb.2012.12.012
- Barrett, L. F., Wilson-Mendenhall, C. D., and Barsalou, L. (in press). "A psychological construction account of emotion regulation and dysregulation: the role of situated conceptualizations," *The Handbook of Emotion Regulation*, 2nd Edn, ed. J. J. Gross (New York: Guilford).
- Barsalou, L. W. (2003). Situated simulation in the human conceptual system. *Lang. Cogn. Process* 18, 513–562. doi: 10.1080/01690960344000026
- Barsalou, L. W. (2008). "Situating concepts" in *Cambridge Handbook of Situated Cognition*, eds P. Robbins and M. Aydede (New York: Cambridge University Press), 236–263.
- Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1281–1289. doi: 10.1098/rstb.2008.0319
- Bliss-Moreau, E., Bauman, M. D., and Amaral, D. G. (2011). Neonatal amygdala lesions result in globally blunted affect in adult rhesus macaques. *Behav. Neurosci.* 125, 848–858. doi: 10.1037/a0025757
- Buckner, R. L., Andrews-Hanna, J. R., and Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Ann. N. Y. Acad. Sci.* 1124, 1–38. doi: 10.1196/annals.1440.011

<sup>6</sup>The term "resting state" is often misinterpreted to mean the resting brain. It should not be assumed that the brain is actually "at rest" during these scans, but simply that there is no externally orienting task.

- Buckner, R. L., and Carroll, D. C. (2007). Self-projection and the brain. *Trends Cogn. Sci.* 11, 49–57. doi: 10.1016/j.tics.2006.11.004
- Bystriksky, A., Pontillo, D., Powers, M., Sabb, F. W., Craske, M. G., and Bookheimer, S. Y. (2001). Functional MRI changes during panic anticipation and imagery exposure. *Neuroreport* 12, 3953–3957. doi: 10.1097/00001756-200112210-00020
- Cabral, J., Hugues, E., Sporns, O., and Deco, G. (2011). Role of local network oscillations in resting-state functional connectivity. *Neuroimage* 57, 130–139. doi: 10.1016/j.neuroimage.2011.04.010
- Chun, M. M., Golomb, J. D., and Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annu. Rev. Psychol.* 62, 73–101. doi: 10.1146/annurev.psych.093008.100427
- Colibazzi, T., Posner, J., Wang, Z., Gorman, D., Gerber, A., Yu, S., et al. (2010). Neural systems subserving valence and arousal during the experience of induced emotions. *Emotion* 10, 377–389. doi: 10.1037/a0018484
- Corbetta, M., Patel, G., and Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron* 58, 306–324. doi: 10.1016/j.neuron.2008.04.017
- Costa, V. D., Lang, P. J., Sabatinelli, D., Versace, F., and Bradley, M. M. (2010). Emotional imagery: assessing pleasure and arousal in the brain's reward circuitry. *Hum. Brain Mapp.* 31, 1446–1457. doi: 10.1002/hbm.20948
- Costafreda, S. G., Brammer, M. J., David, A. S., and Fu, C. H. (2008). Predictors of amygdala activation during the processing of emotional stimuli: a meta-analysis of 385 PET and fMRI studies. *Brain Res. Rev.* 58, 57–70. doi: 10.1016/j.brainresrev.2007.10.012
- Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173. doi: 10.1006/cbmr.1996.0014
- Craig, A. D. (2009). How do you feel – now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70. doi: 10.1038/nrn2555
- Deco, G., Jirsa, V. K., and McIntosh, A. R. (2011). Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nat. Rev. Neurosci.* 12, 43–56. doi: 10.1038/nrn2961
- Denson, T. E., Pedersen, W. C., Ronquillo, J., and Nandy, A. S. (2009). The angry brain: neural correlates of anger, angry rumination, and aggressive personality. *J. Cogn. Neurosci.* 21, 734–744. doi: 10.1162/jocn.2009.21051
- Dosenbach, N. U., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A., et al. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proc. Natl. Acad. Sci. U.S.A.* 104, 11073–11078. doi: 10.1073/pnas.0704320104
- Ekman, P. (2009). Darwin's contributions to our understanding of emotional expressions. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 3449–3451. doi: 10.1098/rstb.2009.0189
- Fair, D. A., Dosenbach, N. U., Church, J. A., Cohen, A. L., Brahmbhatt, S., Miezin, F. M., et al. (2007). Development of distinct control networks through segregation and integration. *Proc. Natl. Acad. Sci. U.S.A.* 104, 13507–13512. doi: 10.1073/pnas.0705843104
- Fogassi, L., and Luppino, G. (2005). Motor functions of the parietal lobe. *Curr. Opin. Neurobiol.* 15, 626–631. doi: 10.1016/j.conb.2005.10.015
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., and Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. U.S.A.* 102, 9673–9678. doi: 10.1073/pnas.0504136102
- Gillihan, S. J., Rao, H., Wang, J., Detre, J. A., Breland, J., Sankoorikal, G. M., et al. (2010). Serotonin transporter genotype modulates amygdala activity during mood regulation. *Soc. Cogn. Affect. Neurosci.* 5, 1–10. doi: 10.1093/scan/nsp035
- Glenberg, A. M. (1997). What memory is for. *Behav. Brain Sci.* 20, 1–19; discussion 19–55.
- Greve, D. N. (2002). *Optseq2*. [Computer software]. Available at: <http://surfer.nmr.mgh.harvard.edu/optseq/>
- Gusnard, D. A., and Raichle, M. E. (2001). Searching for a baseline: functional imaging and the resting human brain. *Nat. Rev. Neurosci.* 2, 685–694. doi: 10.1038/35094500
- Hassabis, D., Kumaran, D., Vann, S. D., and Maguire, E. A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. *Proc. Natl. Acad. Sci. U.S.A.* 104, 1726–1731. doi: 10.1073/pnas.0610561104
- Hess, U., and Thibault, P. (2009). Darwin and emotion expression. *Am. Psychol.* 64, 120–128. doi: 10.1037/a0013386
- Holmes, E. A., and Mathews, A. (2010). Mental imagery in emotion and emotional disorders. *Clin. Psychol. Rev.* 30, 349–362. doi: 10.1016/j.cpr.2010.01.001
- Hutchinson, J. B., and Turk-Browne, N. B. (2012). Memory-guided attention: control from multiple memory systems. *Trends Cogn. Sci.* 16, 576–579. doi: 10.1016/j.tics.2012.10.003
- Killingsworth, M. A., and Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science* 330, 932. doi: 10.1126/science.1192439
- Kim, J. W., Choi, E. A., Kim, J. J., Jeong, B. S., Kim, S. E., and Ki, S. W. (2008). The role of amygdala during auditory verbal imagery of derogatory appraisals by others. *Neurosci. Lett.* 446, 1–6. doi: 10.1016/j.neulet.2008.09.019
- Klases, M., Chen, Y. H., and Mathiak, K. (2012). Multisensory emotions: perception, combination and underlying neural processes. *Rev. Neurosci.* 23, 381–392. doi: 10.1515/revneuro-2012-0040
- Kober, H., Barrett, L. F., Joseph, J., Bliss-Moreau, E., Lindquist, K., and Wager, T. D. (2008). Functional grouping and cortical-subcortical interactions in emotion: a meta-analysis of neuroimaging studies. *Neuroimage* 42, 998–1031. doi: 10.1016/j.neuroimage.2008.03.059
- Kosslyn, S. M., Ganis, G., and Thompson, W. L. (2001). Neural foundations of imagery. *Nat. Rev. Neurosci.* 2, 635–642. doi: 10.1038/35090055
- Kroes, M. C., and Fernandez, G. (2012). Dynamic neural systems enable adaptive, flexible memories. *Neurosci. Biobehav. Rev.* 36, 1646–1666. doi: 10.1016/j.neubiorev.2012.02.014
- Kross, E., Berman, M. G., Mischel, W., Smith, E. E., and Wager, T. D. (2011). Social rejection shares somatosensory representations with physical pain. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6270–6275. doi: 10.1073/pnas.1102693108
- Krueger, F., Barbey, A. K., and Grafman, J. (2009). The medial prefrontal cortex mediates social event knowledge. *Trends Cogn. Sci.* 13, 103–109. doi: 10.1016/j.tics.2008.12.005
- Lang, P. J. (1979). Presidential address, 1978. A bio-informational theory of emotional imagery. *Psychophysiology* 16, 495–512. doi: 10.1111/j.1469-8986.1979.tb01511.x
- Lang, P. J., Bradley, M. M., and Curthbert, B. N. (2008). *International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual*. Gainesville, FL: University of Florida.
- Lang, P. J., Kozak, M. J., Miller, G. A., Levin, D. N., and McLean, A. Jr. (1980). Emotional imagery: conceptual structure and pattern of somato-visceral response. *Psychophysiology* 17, 179–192. doi: 10.1111/j.1469-8986.1980.tb00133.x
- Lench, H. C., Flores, S. A., and Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: a meta-analysis of experimental emotion elicitation. *Psychol. Bull.* 137, 834–855. doi: 10.1037/a0024244
- Lindquist, K. A., and Barrett, L. F. (2012). A functional architecture of the human brain: emerging insights from the science of emotion. *Trends Cogn. Sci.* 16, 533–540. doi: 10.1016/j.tics.2012.09.005
- Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., and Barrett, L. F. (2012). The brain basis of emotion: a meta-analytic review. *Behav. Brain Sci.* 35, 121–143. doi: 10.1017/S0140525X11000446
- Marci, C. D., Glick, D. M., Loh, R., and Dougherty, D. D. (2007). Autonomic and prefrontal cortex responses to autobiographical recall of emotions. *Cogn. Affect. Behav. Neurosci.* 7, 243–250. doi: 10.3758/CABN.7.3.243
- Menon, V. (2011). Large-scale brain networks and psychopathology: a unifying triple network model. *Trends Cogn. Sci.* 15, 483–506. doi: 10.1016/j.tics.2011.08.003
- Mesquita, B., Barrett, L. F., and Smith, E. R., (eds). (2010). *The Mind in Context*. New York: Guilford.
- Miskovic, V., and Schmidt, L. A. (2012). Social fearfulness in the human brain. *Neurosci. Biobehav. Rev.* 36, 459–478. doi: 10.1016/j.neubiorev.2011.08.002
- Mitchell, J. P., Banaji, M. R., and MacRae, C. N. (2005). The link between social cognition and self-referential thought in the medial prefrontal cortex. *J. Cogn. Neurosci.* 17, 1306–1315. doi: 10.1162/0898929055002418
- Moll, J., de Oliveira-Souza, R., Moll, F. T., Ignacio, F. A., Bramati, I. E., Caparelli-Daquer, E. M., et al. (2005). The moral affiliations of disgust: a functional MRI study. *Cogn. Behav. Neurosci.* 18, 68–78. doi: 10.1097/01.wnn.0000152236.46475.a7
- Morecraft, R. J., and Van Hoesen, G. W. (1992). Cingulate input to the primary and supplementary motor cortices in the rhesus monkey: evidence for somatotopy in areas 24c and 23c. *J. Comp. Neurol.* 322, 471–489. doi: 10.1002/cne.903220403

- Ollinger, J. M., Corbetta, M., and Shulman, G. L. (2001a). Separating processes within a trial in event-related functional MRI. *Neuroimage* 13, 218–229. doi: 10.1006/nimg.2000.0711
- Ollinger, J. M., Shulman, G. L., and Corbetta, M. (2001b). Separating processes within a trial in event-related functional MRI. *Neuroimage* 13, 210–217. doi: 10.1006/nimg.2000.0710
- Perry, D., Hendler, T., and Shamay-Tsoory, S. G. (2012). Can we share the joy of others? Empathic neural responses to distress vs joy. *Soc. Cogn. Affect. Neurosci.* 7, 909–916. doi: 10.1093/scan/nsr073
- Petersen, S. E., and Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annu. Rev. Neurosci.* 35, 73–89. doi: 10.1146/annurev-neuro-062111-150525
- Pietrini, P., Guazzelli, M., Basso, G., Jaffe, K., and Grafman, J. (2000). Neural correlates of imaginal aggressive behavior assessed by positron emission tomography in healthy subjects. *Am. J. Psychiatry* 157, 1772–1781. doi: 10.1176/appi.ajp.157.11.1772
- Pitman, R. K., Orr, S. P., Foa, D. F., de Jong, J. B., and Claiborn, J. M. (1987). Psychophysiological assessment of posttraumatic stress disorder imagery in Vietnam combat veterans. *Arch. Gen. Psychiatry* 44, 970–975. doi: 10.1001/archpsyc.1987.01800230050009
- Posner, M. I. (2012). Imaging attention networks. *Neuroimage* 61, 450–456. doi: 10.1016/j.neuroimage.2011.12.040
- Raichle, M. E. (2010). Two views of brain function. *Trends Cogn. Sci.* 14, 180–190. doi: 10.1016/j.tics.2010.01.008
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 676–682. doi: 10.1073/pnas.98.2.676
- Rauch, S. L., Shin, L. M., Dougherty, D. D., Alpert, N. M., Orr, S. P., Lasko, M., et al. (1999). Neural activation during sexual and competitive arousal in healthy men. *Psychiatry Res.* 91, 1–10. doi: 10.1016/S0925-4927(99)00020-7
- Raz, G., Winetraub, Y., Jacob, Y., Kinreich, S., Maron-Katz, A., Shaham, G., et al. (2012). Portraying emotions at their unfolding: a multilayered approach for probing dynamics of neural networks. *Neuroimage* 60, 1448–1461. doi: 10.1016/j.neuroimage.2011.12.084
- Schacter, D. L., Addis, D. R., and Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. *Nat. Rev. Neurosci.* 8, 657–661. doi: 10.1038/nrn2213
- Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V. C., Spreng, R. N., and Szpunar, K. K. (2012). The future of memory: remembering, imagining, and the brain. *Neuron* 76, 677–694. doi: 10.1016/j.neuron.2012.11.001
- Schaich Borg, J., Lieberman, D., and Kiehl, K. A. (2008). Infection, incest, and iniquity: investigating the neural correlates of disgust and morality. *J. Cogn. Neurosci.* 20, 1529–1546. doi: 10.1162/jocn.2008.20109
- Schneider, W., Eschman, A., and Zuccolotto, A. (2002). *E-Prime Computer Software and Manual*. Pittsburgh, PA: Psychology Software Tools Inc.
- Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., et al. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* 27, 2349–2356. doi: 10.1523/JNEUROSCI.5587-06.2007
- Sepulcre, J., Sabuncu, M. R., Yeo, T. B., Liu, H., and Johnson, K. A. (2012). Stepwise connectivity of the modal cortex reveals the multimodal organization of the human brain. *J. Neurosci.* 32, 10649–10661. doi: 10.1523/JNEUROSCI.0759-12.2012
- Shackman, A. J., Salomons, T. V., Slagter, H. A., Fox, A. S., Winter, J. J., and Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nat. Rev. Neurosci.* 12, 154–167. doi: 10.1038/nrn2994
- Sorns, O., and Kotter, R. (2004). Motifs in brain networks. *PLoS Biol.* 2:e369. doi: 10.1371/journal.pbio.0020369
- Spreng, R. N., Sepulcre, J., Turner, G. R., Stevens, W. D., and Schacter, D. L. (2013). Intrinsic architecture underlying the relations among the default, dorsal attention, and frontoparietal control networks of the human brain. *J. Cogn. Neurosci.* 25, 74–86. doi: 10.1162/jocn\_a\_00281
- Squire, L. R. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiol. Learn. Mem.* 82, 171–177. doi: 10.1016/j.nlm.2004.06.005
- Squire, L. R., and Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science* 253, 1380–1386. doi: 10.1126/science.1896849
- Tang, Y. Y., Rothbart, M. K., and Posner, M. I. (2012). Neural correlates of establishing, maintaining, and switching brain states. *Trends Cogn. Sci.* 16, 330–337. doi: 10.1016/j.tics.2012.05.001
- Touroutoglou, A., Hollenbeck, M., Dickerson, B. C., and Barrett, L. F. (2012). Dissociable large-scale networks anchored in the right anterior insula subserved affective experience and attention. *Neuroimage* 60, 1947–1958. doi: 10.1016/j.neuroimage.2012.02.012
- Tracy, J. L., and Randles, D. (2011). Four models of basic emotions: a review of Ekman and Cordaro, Izard, Levenson, and Panksepp and Watt. *Emot. Rev.* 3, 397–405. doi: 10.1177/1754073911410747
- Tulving, E. (2002). Episodic memory: from mind to brain. *Annu. Rev. Psychol.* 53, 1–25. doi: 10.1146/annurev.psych.53.100901.135114
- van Marle, H. J., Hermans, E. J., Qin, S., and Fernandez, G. (2010). Enhanced resting-state connectivity of amygdala in the immediate aftermath of acute psychological stress. *Neuroimage* 53, 348–354. doi: 10.1016/j.neuroimage.2010.05.070
- Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Hum. Brain Mapp.* 30, 829–858. doi: 10.1002/hbm.20547
- Van Overwalle, F., and Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage* 48, 564–584. doi: 10.1016/j.neuroimage.2009.06.009
- Vincent, J. L., Snyder, A. Z., Fox, M. D., Shannon, B. J., Andrews, J. R., Raichle, M. E., et al. (2006). Coherent spontaneous activity identifies a hippocampal-parietal memory network. *J. Neurophysiol.* 96, 3517–3531. doi: 10.1152/jn.00048.2006
- Vogt, B. A. (2005). Pain and emotion interactions in subregions of the cingulate gyrus. *Nat. Rev. Neurosci.* 6, 533–544. doi: 10.1038/nrn1704
- Wang, Z., Liu, J., Zhong, N., Qin, Y., Zhou, H., and Li, K. (2012). Changes in the brain intrinsic organization in both on-task state and post-task resting state. *Neuroimage* 62, 394–407. doi: 10.1016/j.neuroimage.2012.04.051
- Whitfield-Gabrieli, S., and Ford, J. M. (2012). Default mode network activity and connectivity in psychopathology. *Annu. Rev. Clin. Psychol.* 8, 49–76. doi: 10.1146/annurev-clinpsy-032511-143049
- Wilson-Mendenhall, C. D., Barrett, L. F., and Barsalou, L. W. (2013). Neural evidence that human emotions share core affective properties. *Psychol. Sci.* 24, 947–956. doi: 10.1177/0956797612464242
- Wilson-Mendenhall, C. D., Barrett, L. F., Simmons, W. K., and Barsalou, L. W. (2011). Grounding emotion in situated conceptualization. *Neuropsychologia* 49, 1105–1127. doi: 10.1016/j.neuropsychologia.2010.12.032
- Yang, X. F., Bossmann, J., Schifffauer, B., Jordan, M., and Immordino-Yang, M. H. (2012). Intrinsic default mode network connectivity predicts spontaneous verbal descriptions of autobiographical memories during social processing. *Front. Psychol.* 3:592. doi: 10.3389/fpsyg.2012.00592
- Yeh, W., and Barsalou, L. W. (2006). The situated nature of concepts. *Am. J. Psychol.* 119, 349–384. doi: 10.2307/20445349
- Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* 106, 1125–1165. doi: 10.1152/jn.00338.2011
- Zwaan, R. A., and Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychol. Bull.* 123, 162–185. doi: 10.1037/0033-2909.123.2.162

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 02 June 2013; accepted: 24 October 2013; published online: 26 November 2013.

Citation: Wilson-Mendenhall CD, Barrett LF and Barsalou LW (2013) Situating emotional experience. *Front. Hum. Neurosci.* 7:764. doi: 10.3389/fnhum.2013.00764

This article was submitted to the journal *Frontiers in Human Neuroscience*. Copyright © 2013 Wilson-Mendenhall, Barrett and Barsalou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Neural networks underlying affective states in a multimodal virtual environment: contributions to boredom

Krystyna A. Mathiak<sup>1,2</sup>, Martin Klasen<sup>2,3</sup>, Mikhail Zvyagintsev<sup>2,3</sup>, René Weber<sup>4</sup> and Klaus Mathiak<sup>2,3</sup>\*

<sup>1</sup> Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, RWTH Aachen University, Aachen, Germany

<sup>2</sup> Department of Psychiatry, Psychotherapy and Psychosomatics, RWTH Aachen University, Aachen, Germany

<sup>3</sup> Jülich-Aachen Research Alliance (JARA)-Translational Brain Medicine, Jülich, Germany

<sup>4</sup> Department of Communication-Media Neuroscience Lab, University of California, Santa Barbara, CA, USA

## Edited by:

Benjamin Kreifelts, University of Tübingen, Germany

## Reviewed by:

Lutz Jäncke, University of Zurich, Switzerland  
Thorsten Fehr, Universität Bremen, Germany

## \*Correspondence:

Klaus Mathiak, Department of Psychiatry, Psychotherapy and Psychosomatics, RWTH Aachen University, Pauwelsstr. 30, 52074 Aachen, Germany  
e-mail: kmathiak@ukaachen.de

The interaction of low perceptual stimulation or goal-directed behavior with a negative subjective evaluation may lead to boredom. This contribution to boredom may shed light on its neural correlates, which are poorly characterized so far. A video game served as simulation of free interactive behavior without interruption of the game's narrative. Thirteen male German volunteers played a first-person shooter game (*Tactical Ops: Assault on Terror*) during functional magnetic resonance imaging (fMRI). Two independent coders performed the time-based analysis of the audio-visual game content. Boredom was operationalized as interaction of prolonged absence of goal-directed behavior with lowered affect in the Positive and Negative Affect Schedule (PANAS). A decrease of positive affect (PA) correlated with response amplitudes in bilateral insular clusters extending into the amygdala to prolonged inactive phases in a game play and an increase in negative affect (NA) was associated with higher responses in bilateral ventromedial prefrontal cortex (vmPFC). Precuneus and hippocampus responses were negatively correlated with changes in NA. We describe for the first time neural contributions to boredom, using a video game as complex virtual environment. Further our study confirmed that PA and NA are separable constructs, reflected by distinct neural patterns. PA may be associated with afferent limbic activity whereas NA with affective control.

**Keywords:** boredom, negative affect, positive affect, video game, PANAS

## INTRODUCTION

Arousal theories define boredom as the state of non-optimal arousal that ensues when there is a mismatch between an individual's needed arousal and the availability of environmental stimulation (e.g., Csikszentmihalyi, 1975, 1990); it is the aversive state that occurs when it is not possible to achieve an optimal level of arousal through engagement with the environment. Boredom is particularly likely to occur when a task provides little external support for keeping attention engaged, such that performance relies instead on self-sustained attention (Eastwood et al., 2012). Considering video games, this may refer to prolonged situations where the player has no apparent task. Other authors however emphasize aversive aspects of boredom, such as feelings of displeasure, sadness, emptiness, anxiety, and even anger (Csikszentmihalyi, 1975; Csikszentmihalyi, 2000; Fahlman et al., 2013).

One aspect of boredom is the interaction of behavior and affect, i.e., reduced affect associated with a lack of goal-oriented behavior. Many researchers suggested that "wishing to, but being unable to, become engrossed in satisfying activity" reflects the state of boredom (for a review, see Eastwood et al., 2012). It is, however, important to remember that task-irrelevant daydreaming or mind wandering is not typically linked with negative mood (Killingsworth and Gilbert, 2010) and rather can be experienced

as pleasant engagement (Eastwood et al., 2012). Therefore, the combination of low goal-directed activity with subsequent deterioration of affect is one contribution to boredom.

Boredom is an important and very common phenomenon that, despite its potential significant psychosocial consequences, is still poorly understood (Eastwood et al., 2012). To the best of our knowledge, no study to date has specifically investigated its neural correlates. Virtual environments, particularly video games, can be used as a model to study neuronal processes involved in semi-naturalistic behavior that in a classical block or event-related functional magnetic resonance imaging (fMRI) paradigm would not be accessible (Mathiak and Weber, 2006). We examined neural contributions to boredom using the interactive virtual reality model of a first-person shooter video game and subjective evaluation of affect change due to game play. Since it is not possible to reliably measure the subjective affective state during game play without interrupting it (Klasen et al., 2008, 2011; Weber et al., 2009a,b), we applied the Positive and Negative Affect Schedule (PANAS; Watson et al., 1988) directly before and after the fMRI measurement to measure the affect change due to game playing. Reduced goal-directed behavior may be accompanied by the subjective feeling of boredom and will be reflected in the increase of the negative affect (NA) or a decrease of the positive

affect (PA). We expect that the changes in NA and PA should evoke separate activation patterns. The increase of brain activity during prolonged inactivity phases in individuals whose NA is increased or PA is decreased after the game will reflect the subjective feeling of boredom. In addition to emotion processing areas, resting state networks were candidate areas.

## MATERIALS AND METHODS

### PARTICIPANTS

We recruited 13 male German volunteers (age 18–26, median 23) by means of ads posted at the local university and in video game stores. All participants were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971) and considered themselves as regular players of video games (> 5 h/wk, 7–28, median 13 h/wk). Individuals who reported in their history contraindication against magnetic resonance (MR) investigations or neurological, psychiatric or ophthalmologic disorders were excluded from the study. All participants gave their written informed consent and the local ethics committee approved the study protocol.

### IMAGING PARADIGM

After getting acquainted with the game and the controllers for at least 30 min, the volunteers played a violent video game “*Tactical Ops: Assault on Terror*” (Infogrames Europe, Villeurbanne, France) during five functional imaging sessions (except for three participants, who played only four sessions). In the game, the players played freely and experienced the action from the perspective of the virtual character that they control (first-person perspective), while other characters were controlled by the computer. An MR-compatible trackball with five buttons was used by the players to control the game. The participants had time to get acquainted with the controller before the fMRI experiment and the game sound level was adjusted individually (for details, see Weber et al., 2006).

During each 12 min session we recorded hemodynamic brain activity with triple-echo single-shot echo-planar imaging (EPI; repetition time TR = 2.25 s; echo times TE = 23, 40, and 62 ms; 64 × 48 matrix with 4 × 4 mm<sup>2</sup> resolution; 24 slices with 4 mm thickness plus 1 mm gap; 220 volumes) using a 3T MR scanner (Magnetom Trio, Siemens, Erlangen, Germany). As compared to the conventional single-echo EPI, this technique may increase sensitivity to the blood oxygenation level dependent (BOLD) effect as well as reduce drop-outs and distortions (Weiskopf et al., 2005). We recorded the video display of the game play with the audio track for content analysis. The synchronization with the fMRI data was provided by recording the scanner pulses as second audio track. The fMRI data have been evaluated previously using a different content analysis (see Klasen et al., 2011). We acquired anatomical data from each participant before the functional sessions, for functional coregistration (*T1*-weighted 3d magnetization-prepared rapid acquisition with gradient echo, MPRAGE, 256 × 224 × 160 matrix with 1 mm isotropic voxels).

### INVENTORIES

Participants completed the PANAS (Watson et al., 1988; German version in Krohne et al., 1996) directly before entering and after

leaving the MR scanner. The questionnaire contains 20 adjectives describing positive or negative emotions. Each item is rated on a 5-point scale ranging from “very slightly or not at all” to “extremely”, with a total score of 10–50 points per scale.

### CONTENT ANALYSIS

Two independent coders and one supervisor performed the time-based analysis of game content at high time resolution (for details, see Weber et al., 2009a,b). Goal oriented behavior can be assumed most of the time course. From a behavioral perspective remarkable phases are prolonged safe situations, with no apparent task. Those phases where the participants have no actual task and do not change it over for more than 10 s, we defined as being absent of goal-directed behavior. The response patterns to the absence of goal-directed behavior were considered for the boredom analysis.

### fMRI DATA ANALYSIS

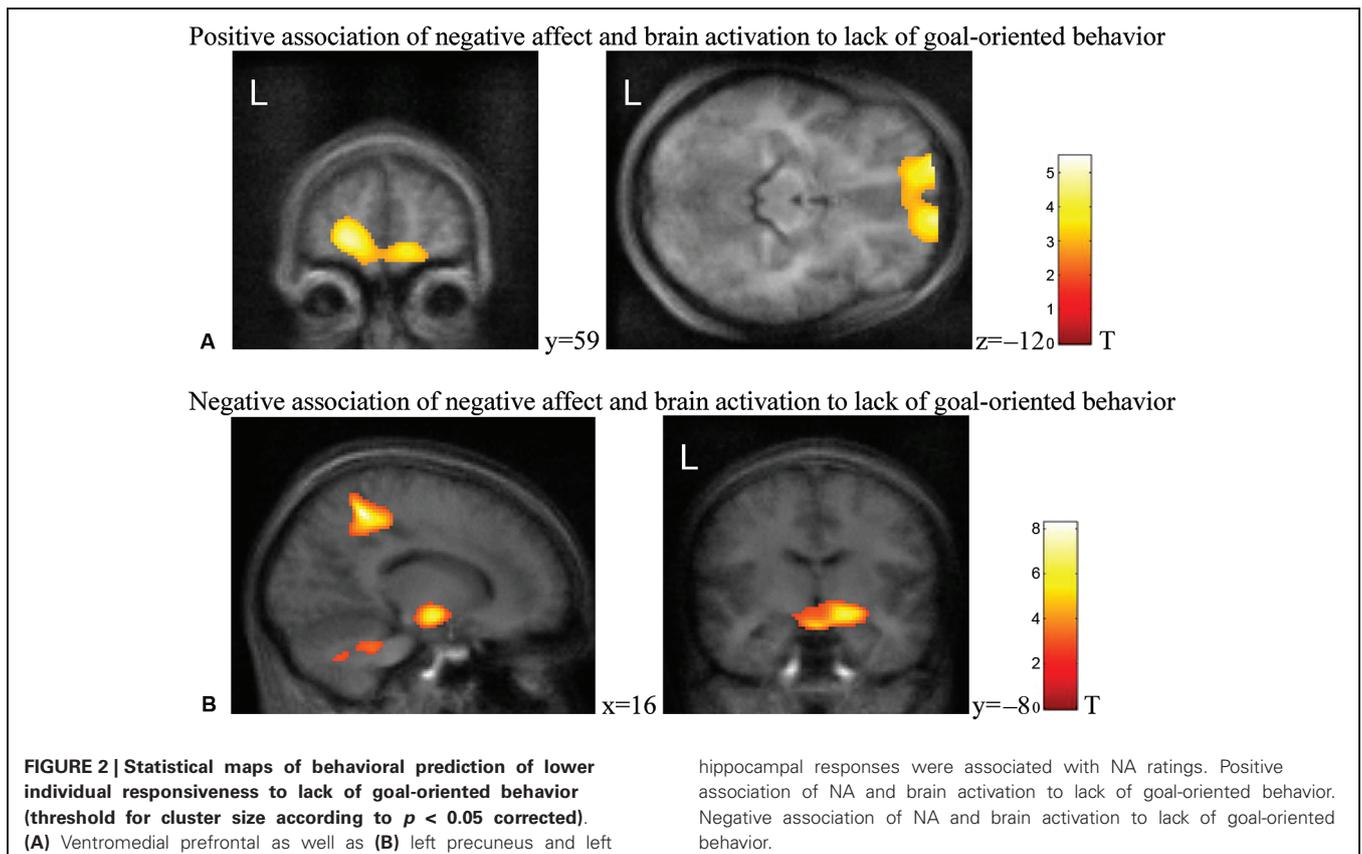
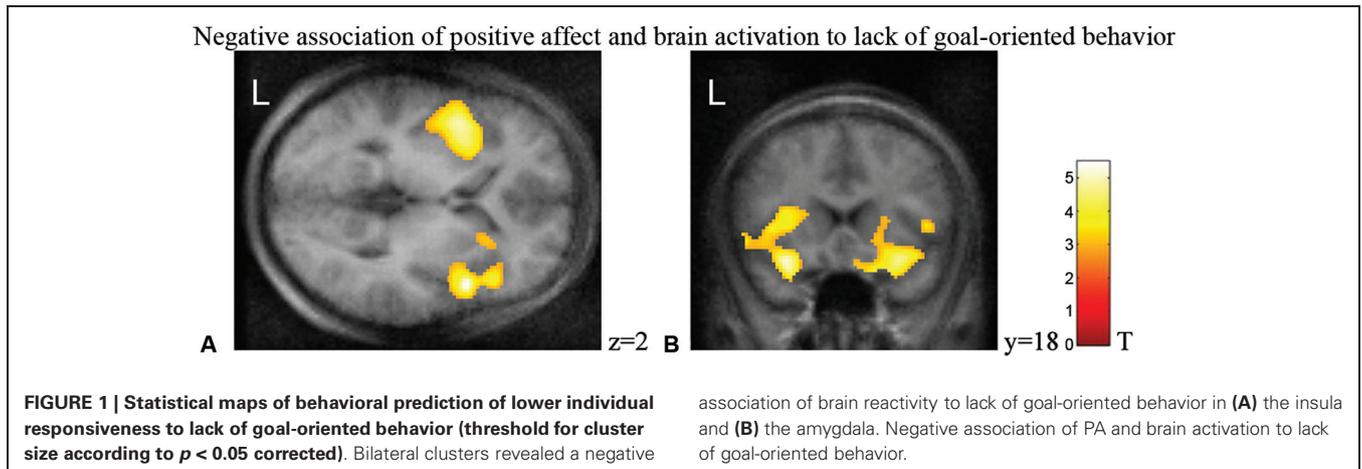
The reconstructed images underwent artifact reduction: construction of dynamic distortion maps from triple-echo EPI with alternating phase-encoding direction as well as subsequent matching of the three echoes (Weiskopf et al., 2005; Mathiak et al., 2012), a combination of the three echoes weighted with TE\**S*<sub>TE</sub> based on expected contrast from the averaged signal decay (Mathiak et al., 2004). We conducted statistical parametric mapping following the standard SPM procedures. Preprocessing comprised motion correction and smoothing after normalization into the Montreal Neurological Institute (MNI; Collins et al., 1994) template space of functional and anatomical data; smoothing with 12 mm full-width at half-maximum Gaussian kernel; general linear model constructed from the coding events convoluted with hemodynamic response function as independent variables; and random effect model for group analysis corrected for multiple testing across the entire brain volume (family wise error (FWE) correction; for further details, see Mathiak et al., 2004).

Neuronal networks were disentangled that activate during phases with lack of goal-directed behavior. The BOLD response was modeled by a generic hemodynamic response to these phases. Therefore for each individual, contrast maps were extracted that represented change of neural activity during phases with low goal-oriented behavior. To investigate their relation with affective evaluation the interaction with affect change was evaluated. Therefore, we calculated the inter-subject regression models with the individual change in PA and in NA measures as predictors for the contrast maps. Considering an high inter-individual variability of networks subserving affective evaluation, we applied a cluster corrected threshold, i.e., we considered only clusters with a size larger than a threshold according to  $p < 0.05$  corrected for multiple comparisons across the brain volume after applying a voxel-wise threshold according to  $p < 0.01$ ; we previously found these parameters most efficient to detect distributed networks rather than circumscribed areas (Mathiak et al., 2011). Calculations were conducted with statistical parametric mapping software (SPM5, Wellcome Department of Imaging Neuroscience, London, UK) and Matlab 7.1 (The Mathworks Inc., Natick, MA, USA).

**RESULTS**

All participants were able to play the game successfully inside the fMRI scanner. The participants scored on average  $30.4 \pm 4.0$  on the positive and  $13.0 \pm 3.2$  on the negative scale of PANAS before the game. After game play, on the scale of PA they reported  $26.5 \pm 5.1$  and on the NA  $11.8 \pm 3.4$ , reflecting in general a slight decrease in the intensity of affect (PA:  $t(12) = 2.90, p = 0.013$ ; NA:  $t(12) = 1.14, p = 0.447$ ). Phases with minimal goal-directed behavior occurred with a frequency of  $10.5 \pm 3.8$  times per 12 min playing session with an average duration of  $14.6 \pm 15.8$  s, resulting in  $17.3 \pm 9.8\%$  of the recorded playing time.

Statistical mapping of the linear prediction of affect change correlating on the hemodynamic responses to low goal-directed behavior were calculated. PA correlated negatively with activation in bilateral insular cluster extending into the amygdala during phases low in goal-directed behavior (Figures 1A, B). Increase in NA was associated with activation in bilateral ventromedial prefrontal cortex (vmPFC) during phases without goal-oriented behavior (Figure 2A) and with right-lateralized deactivation in precuneus and hippocampus (Figure 2B; see Table 1 for the list of clusters associated with the boredom construct). The extent of the activation clusters yielded survival after correction for multiple



**Table 1 | Cluster associated with boredom construct.**

Area	MNI coordinates			t-peak	k <sub>E</sub>	p
	x	y	z			
Reduced PA with low goal-directed behavior						
Insula R	50	12	2	5.34	2489	0.016
Insula L	-36	22	-22	4.91	3455	0.003
Increased NA with low goal-directed behavior						
vmPFC L and R	-26	70	-6	5.48	3370	0.004
Reduced NA with low goal-directed behavior						
Precuneus R	16	-48	50	8.26	2000	0.048
Hippocampus R	16	-8	-12	6.51	3351	0.004

PA: positive affect; NA: negative affect; vmPFC: ventromedial prefrontal cortex; MNI: Montreal Neurological Institute; k<sub>E</sub>: cluster size (voxels); R–right; L–left; p: corrected p-value for cluster size.

comparisons across the volume. Peak *t*-values in contrast would not survive strict thresholds. This is in agreement with previous observation that subjective ratings are associated with rather distributed network activation or that the activation centers vary across individuals (Mathiak et al., 2011).

## DISCUSSION

Virtual reality served as a model for complex social behavior, enabling recording of neural activity accompanying different affective outcomes. Considering boredom as affective outcome of prolonged phases with lowered goal-directed behavior during video game play, neural networks underlying affective control emerged. Increase of PA correlated with deactivation in amygdala and insula. Increase of NA, reflecting the dissatisfaction with the game experience, correlated with activation in vmPFC as well as deactivation of hippocampus and precuneus during prolonged inactive phases during game play.

Emotions elicited by or during the appraisal of external stimuli can be characterized according to different dimensions. In cognitive neuroscience, the most established concept differentiates valence and arousal. Basing on this model, Anders et al. (2004) demonstrated a functional segregation of brain structures underlying peripheral physiologic responses and verbal ratings along the emotional dimensions of valence and arousal. Valence of a stimulus as measured by startle responses correlated with amygdala activity whereas verbal reports of negative emotional valence were associated with insular activity. Further, peripheral physiological and verbal responses along the arousal dimension correlated with activity in thalamus and vmPFC. We adopted an alternative but widely accepted model, assuming that PA and NA dimensions are independent to a large extent (Huebner and Dew, 1996). Indeed, the experimental data support the existence of the separate neural circuits underlying the change of PA and NA and describe the approach system (facilitating appetitive behavior and generating certain types of PA that are approach-related) and withdrawal system (generating certain, withdrawal-related, forms of NA; for a review, see Davidson and Irwin, 1999). The PANAS measures both PA and NA and the correlation between the two scales is low and stable across different time frames (Watson, 1988). In agreement with these findings, changes of the

two constructs were reflected in separate neural networks. NA depended on the activity of the vmPFC, putamen and hippocampus whereas PA correlated with activation of amygdala and insula. Similar to the study by Anders et al. (2004), the activation of the amygdala and insula correlated with one stimulus dimension and ventromedial prefrontal networks with the other one.

The amygdala and the PFC have extensive reciprocal connections and act together to regulate the processing of negative emotions. Diekhof et al. (2011) demonstrated that the vmPFC, accompanied by a concordant reduction of activation in the left amygdala, controlled negative affective responses. The cognitive reappraisal strategies were accompanied by a hyperactivation in the anterior cingulate and the insular cortex. Our study showed the dichotomy among those structures: while the vmPFC was involved in the processing of NA, the amygdala and insula were involved in processing of the PA.

Limbic structures with afferent functions such as amygdala and insula have been implicated in processing of negative emotions such as fear and disgust. Amygdala is a core structure involved in emotional processing, particularly of fear or anger (Dyck et al., 2011; see e.g., Costafreda et al., 2008; for a review). Consequently, increase of amygdala activation in individuals interfered with the experience of PA. In a similar vein, the anterior insula is suggested as being a central structure in mediating interoceptive awareness and the subjective experience of feelings through the representations of bodily reactions, consistent with the James–Lange theory of emotion and the somatic marker hypothesis (Craig, 2002, 2009; Damasio, 2003). It is believed to be responsible chiefly for negative emotions, in particular disgust (for a review, see Bossaerts, 2010). In line with this theory, an inhibition of the insula—similarly to the amygdala—may help to preserve the PA. Indeed, anterior insula may contribute to the mediation of fear-related arousal and negative affective states through its extensive reciprocal connections with the amygdala (Augustine, 1996; Anders et al., 2004). Alternatively, Sterzer and Kleinschmidt (2010) proposed that the anterior insula plays an integrative role in perception-action coupling. Driven by the salience of a sensory event, by task demands, or even by spontaneous activity fluctuations, insular activity mediates states of elevated sensory alertness and readiness for action. Derek (2011) considered a “boredom threshold” yielding reactivation of alternative perceptual concepts. This should render the individual more sensitive and more reactive to any kind of sensory information in situations that pose potential challenges to homeostasis. Accordingly, the game players who failed to decrease the insula activation adequately to lower task demands in the inactive game phases experienced lower PA.

The vmPFC controls emotion experience. This area receives inputs from sensory cortices and has extensive connections with emotional and affective areas including amygdala, striatum, and brainstem (Ridderinkhof et al., 2004) leading to hypotheses on its role in modulation of time course of emotional responding (Davidson, 1998). Further, the vmPFC is proposed to serve as an integrator of external and internal environment, capturing the emotional significance of events and coordinating the appropriate emotional response (for a review, see Barbas, 2000). The vmPFC may be directly involved in the representation of elementary

positive and negative emotional states even in the absence of immediately present incentives (for a review, see Davidson and Irwin, 1999). Diekhof et al. (2011) demonstrated in their meta-analysis that the activation of the vmPFC reduced the degree of subjectively perceived unpleasantness. Contrary, the areas in medial and ventromedial PFC as well as subgenual anterior cingulate cortex activated in healthy participants during sad mood induction (Wang et al., 2006; Paulesu et al., 2010) and were hyperactive in patients with depression (Drevets et al., 2008). Moreover, an excitatory circuit within the vmPFC augmented fear expression, which is located dorsal to fear-inhibiting regions and could be capable of exciting the amygdala (Quirk and Beer, 2006). Diekhof et al. (2011) postulate the vmPFC as a controller of perceived fear and averseness that modulates negative affective responses in phylogenetically older structures of the emotion processing system, such as the amygdala. In this theoretical framework, the increased activation of vmPFC during prolonged inactivity in the game increased NA.

In our study, hippocampal activity seemed to counteract the experience of boredom. The hippocampus is the core structure involved in the formation and temporary storage of episodic and semantic memories, as well as in spatial navigation (for a review, see Stella et al., 2011). Similarly the precuneus was involved in reduction of boredom. A recent study using EEG source localization found a similar areal associated with the feeling of spatial presence during video games (Havranek et al., 2012). Presence in virtual environments (Baumgartner et al., 2008) is related to flow experience (Csikszentmihalyi, 2000; Faiola et al., 2013), which was found to be associated with precuneus activity as well (Klasen et al., 2011). In contrast to our study, the senso-motor network contributed flow (Klasen et al., 2011) and prefrontal networks to activity control in a first person simulation (Havranek et al., 2012), supporting a dissociation of boredom from these constructs. Episodic memory and engagement with the game may counteract subjective experience of boredom.

In a related account, the precuneus, along with adjacent areas within the posteromedial parietal cortex, contribute significantly to the “default mode” of brain function during conscious resting state (Cavanna, 2007). It is considered one of the core structures responsible for consciousness and self-representation (for a review, see Cavanna and Trimble, 2006). The precuneus may be involved in the generation of the spatial information necessary for imagined whole body movements and its activation preceding the beginning of imagined movement (Ogiso et al., 2000). Moreover, activation of precuneus was demonstrated in cognitive tasks requiring mental imagery, including visual rotation, deductive reasoning, music processing and mental navigation (Cavanna and Trimble, 2006). According to Watson et al. (1994), boredom is an externally driven state, the affective result of impoverished external stimuli, conceivably due to a lack of cognitive resources necessary to intrinsically generate interest. The activation of both precuneus and hippocampus may support the planning of coming actions during waiting periods and protect the game players from the feeling of boredom.

Despite the relatively clear findings concerning the neural networks, caution has to be taken with the generalization of the present study. Conceivably only aspects of boredom were

assessed with this methodology. Boredom due to exhaustion such as fatigue cannot be considered in such short time scale nor can it be measured using fMRI. Indeed the considered change of affect as measured by the PANAS may not be sufficiently validated as a measure for mood effects. Therefore we also adhere to the label “change of affect” in reference to the PA and NA labels. A future challenge would be to establish a direct causal link between low activity and mood effects, which is only partially fulfilled in the current experiment.

Methodologically, the low number of participants must be considered an important limitation. In particular in a study with higher power, more networks contributing to boredom can be expected. The survival of clusters at the rather rigorous threshold with FWE-correction, however, indicates rather high effect sizes in the observed clusters. More seriously, the link between the affect measures and the lack of goal-directed behavior is only correlational. PANAS was conducted only directly before and after the fMRI measurements. Therefore additional events may have contributed to the changes in affect. Phases with lack of activity may have only been intercorrelated variables. This, however, is a general disadvantage of naturalistic studies. We studied rather unrestricted gaming behavior. Therefore the observed correlations cannot be directly interpreted as causal. Nevertheless the approach has the advantage that it reflects rather naturalistic behavior which is not hampered by intervening explicit mood assessments or experimental interventions. In future, alternative approaches should assess affect during the game play, e.g., by pop-up questions or peripheral physiological markers such as heart rate.

## CONCLUDING REMARKS

We demonstrated neural contributions to boredom in video games. Conceivably, deactivation in putamen and hippocampus reflected decreased task-related mind-wandering and action planning while the increased activity in vmPFC were associated with the accompanying increase of NA and the decreased activity in amygdala and insula-improved PA. Moreover, our study confirmed that PA and NA are separate constructs, represented by distinct neural patterns, with vmPFC involvement in the processing of NA and the amygdala and insula in processing of PA. The results of our study shed new light on the mechanisms of emotional processing. Understanding better the concept of independent PA and NA as well as their neural correlates will improve our understanding of the emotional system in the brain.

## ACKNOWLEDGMENTS

Krystyna A. Mathiak was supported by the Habilitationsstipendium of the Faculty of Medicine, RWTH Aachen. This research project was supported by the German Research Foundation (DFG, IRTG 1329 and MA 2631/4-1, the START-Program of the Faculty of Medicine, RWTH Aachen, and the IZKF Aachen N4-2).

## REFERENCES

- Anders, S., Lotze, M., Erb, M., Grodd, W., and Birbaumer, N. (2004). Brain activity underlying emotional valence and arousal: a response-related fMRI study. *Hum. Brain Mapp.* 23, 200–209. doi: 10.1002/hbm.20048

- Augustine, J. R. (1996). Circuitry and functional aspects of the insular lobe in primates including humans. *Brain Res. Brain Res. Rev.* 22, 229–244. doi: 10.1016/s0165-0173(96)00011-2
- Barbas, H. (2000). Connections underlying the synthesis of cognition, memory and emotion in primate prefrontal cortices. *Brain Res. Bull.* 52, 319–330. doi: 10.1016/s0361-9230(99)00245-2
- Baumgartner, T., Speck, D., Wettstein, D., Masnari, O., Beeli, G., and Jäncke, O. (2008). Feeling present in arousing virtual reality worlds: prefrontal brain regions differentially orchestrate presence experience in adults and children. *Front. Hum. Neurosci.* 2:8. doi: 10.3389/fnhum.2008.008.2008
- Bossaerts, P. (2010). Risk and risk prediction error signals in anterior insula. *Brain Struct. Funct.* 214, 645–653. doi: 10.1007/s00429-010-0253-1
- Cavanna, A. E. (2007). The precuneus and consciousness. *CNS Spectr.* 12, 545–552.
- Cavanna, A. E., and Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain* 129, 564–583. doi: 10.1093/brain/awl004
- Collins, D. L., Neelin, P., Peters, T. M., and Evans, A. C. (1994). Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J. Comput. Assist. Tomogr.* 18, 192–205. doi: 10.1097/00004728-199403000-00005
- Costafreda, S. G., Brammer, M. J., David, A. S., and Fu, C. H. (2008). Predictors of amygdala activation during the processing of emotional stimuli: a meta-analysis of 385 PET and fMRI studies. *Brain Res. Rev.* 58, 57–70. doi: 10.1016/j.brainresrev.2007.10.012
- Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* 3, 655–666. doi: 10.1038/nrn894
- Craig, A. D. (2009). How do you feel—now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70. doi: 10.1038/nrn2555
- Csikszentmihalyi, M. (1975). *Beyond Boredom and Anxiety*. San Francisco, CA: Jossey-Bass.
- Csikszentmihalyi, M. (1990). *Finding Flow: The Psychology of Engagement with Everyday Life*. New York, NY: Basic Books.
- Csikszentmihalyi, M. (2000). *Beyond Boredom and Anxiety. Experiencing Flow in Work and Play*. 25th Anniversary Edn. San Francisco: Jossey-Bass.
- Damasio, A. R. (2003). *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*. Orlando: Hartcourt.
- Davidson, R. J. (1998). Affective style and affective disorders: perspectives from affective neuroscience. *Cogn. Emot.* 12, 307–320. doi: 10.1080/026999398379628
- Davidson, R. J., and Irwin, W. (1999). The functional neuroanatomy of emotion and affective style. *Trends Cogn. Sci.* 3, 11–21. doi: 10.1016/s1364-6613(98)01265-0
- Derek, H. A. (2011). I agree: binocular rivalry stimuli are common but rivalry is not. *Front. Hum. Neurosci.* 5:157. doi: 10.3389/fnhum.2011.00157
- Diekhof, E. K., Geier, K., Falkai, P., and Gruber, O. (2011). Fear is only as deep as the mind allows: a coordinate-based meta-analysis of neuroimaging studies on the regulation of negative affect. *Neuroimage* 58, 275–285. doi: 10.1016/j.neuroimage.2011.05.073
- Drevets, W. C., Price, J. L., and Furey, M. L. (2008). Brain structural and functional abnormalities in mood disorders: implications for neurocircuitry models of depression. *Brain Struct. Funct.* 213, 93–118. doi: 10.1007/s00429-008-0189-x
- Dyck, M., Loughhead, J., Kellermann, T., Boers, F., Gur, R. C., and Mathiak, K. (2011). Cognitive versus automatic mechanisms of mood induction differentially activate left and right amygdala. *Neuroimage* 54, 2503–2513. doi: 10.1016/j.neuroimage.2010.10.013
- Eastwood, J. D., Frischen, A., Fenske, M. J., and Smilek, D. (2012). The unengaged mind: defining boredom in terms of attention. *Perspect. Psychol. Sci.* 7, 482–495. doi: 10.1177/1745691612456044
- Fahlman, S. A., Mercer-Lynn, K. B., Flora, D. B., and Eastwood, J. D. (2013). Development and validation of the multidimensional state boredom scale. *Assessment* 20, 68–85. doi: 10.1177/1073191111421303
- Faiola, A., Newlon, C., Pfaff, M., and Smyslova, O. (2013). Correlating the effects of flow and telepresence in virtual worlds: enhancing our understanding of user behavior in game-based learning. *Comp. Hum. Behav.* 29, 1113–1121. doi: 10.1016/j.chb.2012.10.003
- Havranek, M., Langer, N., Cheetham, M., and Jäncke, L. (2012). Perspective and agency during video gaming influences spatial presence experience and brain activation patterns. *Behav. Brain Funct.* 8:34. doi: 10.1186/1744-9081-8-34
- Huebner, E. S., and Dew, T. (1996). The interrelationships of positive affect, negative affect and life satisfaction in an adolescent sample. *Soc. Indic. Res.* 38, 129–137. doi: 10.1007/bf00300455
- Killingsworth, M. A., and Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science* 330:932. doi: 10.1126/science.1192439
- Klasen, M., Weber, R., Mathiak, K. A., and Mathiak, K. (2011). Neural contributions to flow experience during video game playing. *Soc. Cogn. Affect. Neurosci.* 7, 485–495. doi: 10.1093/scan/nsr021
- Klasen, M., Zvyagintsev, M., Weber, R., Mathiak, K. A., and Mathiak, K. (2008). Think aloud during fMRI: neuronal correlates of subjective experience in video games. *Lect. Notes Comput. Sci.* 5294, 132–138. doi: 10.1007/978-3-540-88322-7\_13
- Krohne, H. W., Egloff, B., Kohlmann, C. W., and Tausch, A. (1996). Untersuchung mit einer deutschen Form der Positive and Negative Affect Schedule (PANAS). *Diagnostica* 42, 139–156.
- Mathiak, K., Hertrich, L., Grodd, W., and Ackermann, H. (2004). Discrimination of temporal information at the cerebellum: functional magnetic resonance imaging of non-verbal auditory memory. *Neuroimage* 21, 154–162. doi: 10.1016/j.neuroimage.2003.09.036
- Mathiak, K. A., Klasen, M., Weber, R., Ackermann, H., Shergill, S. S., and Mathiak, K. (2011). Reward system and temporal pole contributions to affective evaluation during a first person shooter video game. *BMC Neurosci.* 12:66. doi: 10.1186/1471-2202-12-66
- Mathiak, K., and Weber, R. (2006). Toward brain correlates of natural behavior: fMRI during violent video games. *Hum. Brain Mapp.* 27, 948–956. doi: 10.1002/hbm.20234
- Mathiak, K. A., Zvyagintsev, M., Ackermann, H., and Mathiak, K. (2012). Lateralization of amygdala activation in fMRI may depend on phase-encoding polarity. *MAGMA* 25, 177–182. doi: 10.1007/s10334-011-0285-4
- Ogiso, T., Kobayashi, K., and Sugishita, M. (2000). The precuneus in motor imagery: a magnetoencephalographic study. *Neuroreport* 11, 1345–1349. doi: 10.1097/00001756-200004270-00039
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- Paulesu, E., Samburgaro, E., Torti, T., Danelli, L., Ferri, F., Scialfa, G., et al. (2010). Neural correlates of worry in generalized anxiety disorder and in normal controls: a functional MRI study. *Psychol. Med.* 40, 117–124. doi: 10.1017/s0033291709005649
- Quirk, G. J., and Beer, J. S. (2006). Prefrontal involvement in the regulation of emotion: convergence of rat and human studies. *Curr. Opin. Neurobiol.* 16, 723–727. doi: 10.1016/j.conb.2006.07.004
- Ridderinkhof, K. R., van den Wildenberg, W. P. M., Segalowitz, S. J., and Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring and reward-based learning. *Brain Cogn.* 56, 129–140. doi: 10.1016/j.bandc.2004.09.016
- Stella, F., Cerasti, E., Si, B., Jezek, K., and Treves, A. (2011). Self-organization of multiple spatial and context memories in the hippocampus. *Neurosci. Biobehav. Rev.* 36, 1609–1625. doi: 10.1016/j.neubiorev.2011.12.002
- Sterzer, P., and Kleinschmidt, A. (2010). Anterior insula activations in perceptual paradigms: often observed but barely understood. *Brain Struct. Funct.* 214, 611–622. doi: 10.1007/s00429-010-0252-2
- Wang, L., LaBar, K. S., and McCarthy, G. (2006). Mood alters amygdala activation to sad distractors during an attentional task. *Biol. Psychiatry* 60, 1139–1146. doi: 10.1016/j.biopsych.2006.01.021
- Watson, D. (1988). The vicissitudes of mood measurement: effects of varying descriptors, time frames and response formats on measures of positive and negative affect. *J. Pers. Soc. Psychol.* 55, 128–141. doi: 10.1037//0022-3514.55.1.128
- Watson, D., Clark, L. A., and Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *J. Pers. Soc. Psychol.* 54, 1063–1070. doi: 10.1037/0022-3514.54.6.1063
- Watson, P. J., Hickman, S. E., Morris, R. J., Stutz, N. L., and Whitting, L. (1994). Complexity of self-consciousness subscales: correlations of factors with self-esteem and dietary restraint. *J. Soc. Behav. Pers.* 9, 761–774.
- Weber, R., Behr, K., Tamborini, R., Ritterfeld, U., and Mathiak, K. (2009b). What do we really know about first person shooter games: an event-related, high-resolution content analysis. *J. Comput. Mediat. Commun.* 14, 1016–1037. doi: 10.1111/j.1083-6101.2009.01479.x

- Weber, R., Ritterfeld, U., and Mathiak, K. (2006). Does playing violent video games induce aggression? Empirical evidence of a functional magnetic resonance imaging study. *Media Psychol.* 8, 39–60. doi: 10.1207/S1532785XMEP0801\_4
- Weber, R., Tamborini, R., Westcott-Baker, A., and Kantor, B. (2009a). Theorizing flow and media enjoyment as cognitive synchronization of attention and reward networks. *Commun. Theory* 19, 397–422. doi: 10.1111/j.1468-2885.2009.01352.x
- Weiskopf, N., Klose, U., Birbaumer, N., and Mathiak, K. (2005). Single-shot compensation of image distortions and BOLD contrast optimization using multi-echo EPI for real-time fMRI. *Neuroimage* 24, 1068–1079. doi: 10.1016/j.neuroimage.2004.10.012

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 29 May 2013; accepted: 12 November 2013; published online: 28 November 2013.*

*Citation: Mathiak KA, Klasen M, Zvyagintsev M, Weber R and Mathiak K (2013) Neural networks underlying affective states in a multimodal virtual environment: contributions to boredom. Front. Hum. Neurosci. 7:820. doi: 10.3389/fnhum.2013.00820.*

*This article was submitted to the journal Frontiers in Human Neuroscience.*

*Copyright © 2013 Mathiak, Klasen, Zvyagintsev, Weber and Mathiak. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*



# Multisensory integration mechanisms during aging

Jessica Freiherr<sup>1\*</sup>, Johan N. Lundström<sup>2,3,4</sup>, Ute Habel<sup>5,6</sup> and Kathrin Reetz<sup>6,7,8</sup>

<sup>1</sup> Diagnostic and Interventional Neuroradiology, RWTH Aachen University, Aachen, Germany

<sup>2</sup> Department of Clinical Neuroscience, Karolinska Institute, Stockholm, Sweden

<sup>3</sup> Monell Chemical Senses Center, Philadelphia, PA, USA

<sup>4</sup> Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA

<sup>5</sup> Department of Psychiatry, Psychotherapy, and Psychosomatics, RWTH Aachen University, Aachen, Germany

<sup>6</sup> JARA BRAIN – Translational Brain Medicine, RWTH Aachen University, Aachen, Germany

<sup>7</sup> Department of Neurology, RWTH Aachen University, Aachen, Germany

<sup>8</sup> Institute of Neuroscience and Medicine (INM-4), Research Center Jülich, Jülich, Germany

## Edited by:

Yu-Han Chen, University of New Mexico, USA

## Reviewed by:

Wen Li, University of Wisconsin-Madison, USA  
Jennifer L. Mozolic, Warren Wilson College, USA  
Jeannette R. Mahoney, Albert Einstein College of Medicine, USA

## \*Correspondence:

Jessica Freiherr, Diagnostic and Interventional Neuroradiology, RWTH Aachen University, Pauwelsstr. 30, 52074 Aachen, Germany  
e-mail: jfreiherr@ukaachen.de

The rapid demographical shift occurring in our society implies that understanding of healthy aging and age-related diseases is one of our major future challenges. Sensory impairments have an enormous impact on our lives and are closely linked to cognitive functioning. Due to the inherent complexity of sensory perceptions, we are commonly presented with a complex multisensory stimulation and the brain integrates the information from the individual sensory channels into a unique and holistic percept. The cerebral processes involved are essential for our perception of sensory stimuli and becomes especially important during the perception of emotional content. Despite ongoing deterioration of the individual sensory systems during aging, there is evidence for an increase in, or maintenance of, multisensory integration processing in aging individuals. Within this comprehensive literature review on multisensory integration we aim to highlight basic mechanisms and potential compensatory strategies the human brain utilizes to help maintain multisensory integration capabilities during healthy aging to facilitate a broader understanding of age-related pathological conditions. Further our goal was to identify where further research is needed.

**Keywords:** crossmodal sensory integration, cognition, multimodal, aging neuroscience, elderly population, multisensory integration

## MULTISENSORY INTEGRATION

Each of our sensory systems provides us with a qualitatively distinct subjective complementary impression of our surrounding, which are of critical importance for perception, cognitive processing and control of action and can occur in a highly automatized manner (Meredith and Stein, 1983, 1985; Stein and Meredith, 1990). However, most of our everyday percepts are conveyed by multiple sensory systems like the olfactory, auditory, visual, gustatory, and tactile system. Our brain has the remarkable ability to integrate even disparate and complex multisensory information into a unique and coherent percept. The computational mechanisms responsible for this integration assure that the signal is amplified and any kind of accompanied noise is filtered out, thereby promoting the saliency of ecologically meaningful events. Input from two sensory channels, in comparison to a single one, increases the likelihood and speed of detecting and correctly identifying events (Gottfried and Dolan, 2003; Dematte et al., 2006, 2009) and also enhances sensory sensitivity (Dalton et al., 2000; Chen and Spence, 2011). It has also been indicated that higher cognitive sensory processing, like pleasantness evaluation of an odor, is enhanced when it is combined with a congruent auditory stimulus (Seo and Hummel, 2011). Overall, numerous results indicate that multisensory integration plays an important

role in our daily life by facilitating and improving our perceptual capacities.

In more detail, multisensory integration is governed by four different principles. First, unimodal sensory stimuli need to be applied within a certain temporal sequence and second, spatial concordance is necessary in order to achieve optimal integration results (King and Palmer, 1985; Meredith and Stein, 1986). Thus, sensory stimuli of different modalities have to coincide with regards to time and space in order to be integrated. Third, contextual, semantic congruency, or correspondence, is fundamental for efficient multisensory integration (Spence, 2011). When those preconditions are fulfilled, the sensory stimuli appear as if emanating from the same object. Importantly, multisensory integration processes follow the principle of inverse effectiveness—cross- or multisensory integration is most effective and therefore elicit maximal behavioral enhancements when less intense or weak and ambiguous individual stimuli are applied (Stein and Stanford, 2008).

Initial research on multisensory integration observed the superior colliculus of anesthetized cats as an important part of the neural network involved (Meredith and Stein, 1983, 1985; Stein and Meredith, 1990). Nowadays, it has been demonstrated that the traditional cortical network supporting multisensory integration

in humans consists primarily of the superior temporal sulcus (STS) and the intraparietal sulcus (IPS); it has been suggested that the STS is associated with the integration and labeling of object identity, whereas the IPS is involved in a low-level spatial information processing (Calvert, 2001; Stein and Stanford, 2008). Further, the insula is considered a key area for detection of crossmodal coincidence and matching (Calvert, 2001). However, orbital and ventral areas of the frontal cortex and hippocampal areas have also been identified as neuronal correlates of bimodal multisensory integration involving the olfactory and gustatory domains (Gottfried and Dolan, 2003; Small and Prescott, 2005). Those areas might be responsible for attention and memory processing and are involved in novelty detection, congruency assessment, or task difficulty evaluation (Calvert, 2001). Within frontal areas, a dissociation with regard to object identity exists: the orbital network seems to be especially important for the integration of multisensory input related to food items while the ventrolateral prefrontal region mediates the assessment of non-food objects (Price, 2008).

The existing functional imaging research regarding multisensory integration processes can be divided into two topics: multisensory integration of (i) object observation and (ii) emotional perception. For the investigation of multisensory processes during object recognition, combinations of two different modalities—combined visual-auditory stimulations—is the most utilized. A concurrent combination of three different and congruent sensory stimuli has yet to be applied within a neuroimaging setting. However, when analyzing areas of overlap between the three senses of touch, sound, and vision, both the STS and IPS demonstrate a considerable overlap in processing (Bremmer et al., 2001; Beauchamp et al., 2008; Langner et al., 2012). In addition, the left fusiform gyrus (FG) seems to play a putative role within trimodal sensory object manipulation (Kassuba et al., 2011). Most of these studies are, however, based upon a combination of simple stimuli or a conjunction analysis of brain activation related to unimodal sensory stimulation. Multisensory integration mechanisms are of special importance during the perception and processing of emotions. Research from our own and other groups provides evidence for a neural network involving the amygdala, insula, frontal areas, FG, and STS which are responsible for integration of cross- or multisensory information related to emotional perception during stimulation with dynamic stimulus material of different modalities (Ethofer et al., 2006; Kreifelts et al., 2009; Seubert et al., 2010a,b; Klasen et al., 2011, 2012; Muller et al., 2011, 2012; Regenbogen et al., 2012a,b). Although we gained novel insights into multisensory integration processes with regards to object and emotion perception during the last years, a systematic investigation of multisensory integration in relation to differences between age groups and age-related pathologies using functional imaging means is still missing. Since emotional perception accounts for a major part of our everyday wellbeing and multisensory processes are heavily involved in emotional processes we want to draw attention towards a better understanding of multisensory processes and the underlying neural connections during emotional perception. Therefore, the aim of this review is to outline the knowledge about multisensory integration across the lifespan with a special focus on behavioral and neural correlates of multisensory integration

during aging and age-related diseases. We further aim to identify areas where further research is needed in order to shed light onto the mechanisms of multisensory integration.

## MULTISENSORY INTEGRATION DURING AGING

In light of the evolving demographic changes of our society, one important future task for the research communities is to further our understanding of lifelong healthy aging. Aging is a multifactorial and multidimensional process involving physiological, psychological, and social alterations. As part of ongoing degeneration throughout life, there is a progressive deterioration of physical function leading to loss of viability and an increase in vulnerability. Particularly, sensory impairments have an enormous impact on our lives and are closely related to intellectual functioning. Because we experience our environment through multiple sensory systems, which ultimately ensure our everyday safety, quality of life, and social adjustment, it is of interest to understand how multisensory integration processing changes as a function of healthy aging. In facing these challenges, the following question is noteworthy: what age-dependent changes in the neuronal integration of multisensory stimuli occurs in individuals experiencing healthy aging?

Effectiveness of multisensory integration depends upon functioning of the peripheral sensory organs as well as higher cognitive processes. As we age, we experience a decline of function in all our five senses—e.g., visual acuity decreases (Spear, 1993) and auditory thresholds increase (Liu and Yan, 2007). Olfactory capabilities are known to decline during aging as well (Rawson, 2006), however this deterioration can be attributed to a poor medical status and cognitive decline in the elderly (Nordin et al., 2012). Typically, motor speed and executive functions (Falkenstein et al., 2006), as well as working memory and attention control deteriorate as well (Fabiani, 2012). Previous structural neuroimaging studies, including voxel-based morphometry (VBM), deformation field morphometry (DBM), cortical thickness analyses, manual tracing techniques, and diffusion-weighted magnetic resonance imaging (MRI) have given some insights into the complexity of age-related structural brain changes (Sowell et al., 2004; Raz and Rodrigue, 2006; Sala et al., 2012). In a recent review, Hedman et al. (2012) concluded that apart from brain volume increases during childhood and adolescence, a continuous volume decrease of 0.2% per year can be observed, which accelerates to an annual brain volume loss of 0.5% at age 60 and more than 0.5% above the age of 60 years. Thereby, evidence is provided for a volume loss of gray matter and cortical thinning during aging. However, not only structural changes occur during lifespan. Evidence exists that older participants exhibit altered patterns of functional activation during cognitive tasks. The elderly engage brain areas (especially frontal areas) to a greater extent than young adults; this is most likely to compensate for impaired function in other brain areas (Posterior-Anterior Shift with Aging, PASA; Grady et al., 1994; Davis et al., 2008; Grady, 2012). In contrast to task-based methods, the task-independent approach of a resting-state analysis is appealing due to its ability to assess altered brain function independent of the participant's active involvement and task understanding as well as independent of their sensory performance. The resting-state

approach assesses altered brain functions caused by the summation of subtle physiological and pathological changes across the lifespan, which, in turn, can be linked to respective structural changes as noted above (Reetz et al., 2012). Given the extensive changes of perceptual and cognitive processes and the underlying structural and functional brain changes during healthy aging, it seems reasonable that multisensory integration performance as well is altered throughout the lifespan.

Research on multisensory integration in aging in relation to young adults mainly focused on visual-auditory paradigms and employed mostly static stimuli. These early studies indicated that older adults, when compared to younger adults, do not benefit from multisensory conditions (Stine et al., 1990; Walden et al., 1993; Sommers et al., 2005) or even report a suppressed cortical multisensory integration response in the elderly (Stephen et al., 2010). More recent studies, however, point towards an enhancement of multisensory integration effects also in older adults. Among others, shorter reaction times in response to multisensory events have been reported (Helfer, 1998; Laurienti et al., 2006; Diederich et al., 2008; Mahoney et al., 2011, 2012; Diaconescu et al., 2013). A potential reason for the earlier negative results and more recent positive results can be that the early studies focused on very complex auditory-visual speech perception and thus involve sensory as well as higher-order cognitive processes whereas later studies mostly used simpler combinations of stimuli originating from objects. Another explanation for the conflicting results is the use of different multisensory testing and data analysis techniques. It is also possible that the basic principles for a successful multisensory integration (temporal and spatial accordance, inverse effectiveness, semantic congruency) have been violated within the earlier studies.

Different aspects were discussed as basis for this improvement in multisensory integrative function in elder adults (Mozolic et al., 2012). General sensorimotor and cognitive slowing during aging is obviously not able to explain response times acceleration (Laurienti et al., 2006; Peiffer et al., 2007). However, it was recently demonstrated that older adults have a broader time window of integration as a consequence of increased response times and a wider distribution or range of response times; the combined effect aids older adults to separate stimuli in time (Diederich et al., 2008). In this study, older adults demonstrated a lower probability of integration due to the broader time window; however, in case of a successful integration the gain of older adults was larger compared to younger adults. Age-dependent deficits in top-down selective attention to incoming sensory information do not provide an explanation for the enhancement of multisensory integration in older subjects (Mozolic et al., 2008; Hugenschmidt et al., 2009). One plausible explanation is the principle of inverse effectiveness, i.e., that reduced sensitivity in the individual sensory systems (e.g., rigidity of the lenses, loss of hair cells in the ear and olfactory receptors) combined with age-related alterations in cognitive processing increases the relative magnitude of multisensory enhancements (Hairston et al., 2003). Thereby, multisensory integration becomes more important during aging as it helps to counteract the often-destructive consequences of unisensory deterioration. Mozolic et al. (2012) recently proposed a second possible explanation. They suggested that elderly do

not adequately filter sensory noise and hence are more prone to distraction than their younger counterparts. As soon as the extraneous sensory information becomes relevant, however, older adults benefit from this enhanced processing of sensory background information. Evidence for a higher level of background sensory processing in the elderly was also provided by several resting-state studies pointing towards an increased default mode network (DMN) activity (Grady et al., 2006; Li et al., 2007). Furthermore, in elderly, the observed decrease in visual memory and visuo-constructive functions seems to be strongly associated with an age-dependent increase of functional connectivity specifically in the temporal lobe (Schlee et al., 2012).

Age-related changes during information processing (compensatory reallocation, neural compensation, dedifferentiation, inhibition) are based upon basic circuitry of the sensory systems involving several interactive neuronal loops such as prefronto-thalamo-cortical gating between the thalamus and the neocortex in order to effectively process sensory and higher-order cognitive information (Mahoney et al., 2011). Using magnetoencephalography, Diaconescu et al. (2013) indicated that sensory-specific regions showed an increased activity after visual-auditory stimulation in young and old participants, whereas inferior parietal and medial prefrontal areas responded preferentially in older subjects. Further, activation of the latter areas was related to faster detection of multisensory stimuli. This relation was mediated by age-related reductions in gray matter volume in those regions (Diaconescu et al., 2013). The authors propose that posterior parietal and medial prefrontal activity is the basis for the integrated response in older adults. This hypothesis is supported by the theory of PASA described above as well as the theory of cortical dedifferentiation stating that healthy aging is accompanied by decreased specificity of neurons in the prefrontal cortex (Park and Reuter-Lorenz, 2009).

Thus, the neural network governing multisensory integration displays clear age-dependent alterations and age-related changes in cognitive function have clear implications for multisensory processing. That said, although age is the number one major risk factor for a large range of degenerative diseases, it remains to be determined how multisensory integration is affected in different states of age-related diseases, and in particular, diseases of neurodegenerative nature. Unfortunately, to date, there are few published studies on this topic. Any potential changes due to neurodegenerative states might, however, be highly relevant as most of the age-related neurodegenerative disorders are preceded by long presymptomatic periods. Neuropathobiological changes occur many years, even decades, before the clinical manifestation of the disease and numerous compensatory mechanisms and neuroplastic capacities of the human brain remain to be clarified.

The most frequent age-related neurodegenerative disorder is Alzheimer's disease (AD). Due to aging of populations in both developed and developing societies, AD affects 24.3 million people worldwide and has become one of the most severe socio-economical and medical burdens (Ferri et al., 2005). AD, the most common form of dementia, is a complex disease characterized by an accumulation of  $\beta$ -amyloid plaques and neurofibrillary tangles composed of tau amyloid fibrils. These changes are associated with synapse loss and neurodegeneration leading to a general and

progressive loss of cognitive functions, initially predominantly manifested as memory impairment. As studies have revealed that increased DMN activity in healthy aging is associated with a higher level of background sensory processing (Grady et al., 2006; Li et al., 2007), functional changes in AD are of particular interest. Recent studies have mainly demonstrated task-induced deactivations of the DMN as well as a decreased DMN functional connectivity along a continuum from normal aging to pathological conditions (Andrews-Hanna et al., 2007; Hafkemeijer et al., 2012). Resting state functional MRI demonstrated that the regional coherence of brain activity is significantly altered in patients with AD [e.g., Filippi and Agosta, 2011]. Furthermore, abnormalities in the precuneus of patients with amnesic mild cognitive impairment compared to controls were found while AD patients showed alterations of large-scale functional brain networks extending well beyond the DMN. Moreover, episodic-memory related task-based functional neuroimaging studies in AD have revealed increased activity in the precuneus (Brown dyke et al., 2013); an area which plays a crucial role in the DMN and has the highest metabolic and blood perfusion rates during resting conditions (Gusnard and Raichle, 2001). The clinical relevance of resting state networks beyond the DMN is notable because the degree of connectivity in the resting state networks may predict individual cognitive and emotional functions (Wang et al., 2010a). Specifically, interhemispheric functional connectivity of the hippocampi (Wang et al., 2010b), as well as connectivity between the hippocampus and the posteromedial cortices, predicts memory performance in healthy individuals (Wang et al., 2010a). The available research demonstrates that suppression of the DMN during task performance is important. However, conceptualizing the network in terms of suppression implies that the DMN is a sort of nuisance network where its importance to voluntary cognitive functions lies primarily in minimizing its activity during tasks. Consequently, patients with AD that demonstrate hypometabolism in the precuneus/posterior cingulate component of the DMN (Bradley et al., 2002; Chetelat et al., 2008; Langbaum et al., 2009; Petrie et al., 2009; Schroeter et al., 2009) should have very good cognitive functioning during purposeful tasks because the DMN is not active. However, this is clearly not the case. Therefore, further studies with a focus on functional connectivity analyses might foster our understanding of changes in the context of multisensory integration in healthy populations and pathological conditions.

Overall, very little is known about behavioral benefits during multisensory integration in patients with age-related neurodegenerative disorders. Using bimodal stimulation (audio-visual speech presentation), patients demonstrate a limited ability to benefit from concurrent perceptual and linguistic cues compared to healthy aged subjects (Phillips et al., 2009). Given the fact that we have a poor understanding of multisensory integration in neurodegenerative disorders, further studies regarding behavioral and neuronal substrates of multisensory integration are warranted. The knowledge of mechanisms involved in multisensory integration in neurodegenerative disorders, with AD representing the most common one, is of high value given the fact that multisensory stimulation has an inherently high potential for early intervention and thus, also therapeutic application (Staal et al., 2003).

## CONCLUSION

Advances are being made in disentangling the multisensory integration mechanisms in elderly, thus encouraging future larger longitudinal studies needed to understand the specific neurobiological and neuropathological basis of multisensory integration in health and disease. The current state of research on multisensory integration in healthy aging and age-related neurodegenerative disorders would greatly benefit from further studies as we aim to understand basic mechanisms and potential compensatory strategies of the human brain that help maintain multisensory integration capabilities during both healthy and pathological aging. Early identification of changes in multisensory integration will help to better inform choice of therapy and aid a personalized approach to clinical treatment. Future studies are warranted to determine the clinical translational value of multisensory integration processes in the elderly.

## ACKNOWLEDGMENTS

Jessica Freiherr is supported by the Medical Faculty of RWTH Aachen University (IZKF—Interdisciplinary Centre for Clinical Research, START program). Johan N. Lundström is funded by the Knut and Alice Wallenberg Foundation (KAW 2012.0141) and the Swedish Research Council (VR 2009:3356). Kathrin Reetz is funded by the excellence Initiative of the German Research Foundation (DFG ZUK32/1).

## REFERENCES

- Andrews-Hanna, J. R., Snyder, A. Z., Vincent, J. L., Lustig, C., Head, D., Raichle, M. E., et al. (2007). Disruption of large-scale brain systems in advanced aging. *Neuron* 56, 924–935. doi: 10.1016/j.neuron.2007.10.038
- Beauchamp, M. S., Yasar, N. E., Frye, R. E., and Ro, T. (2008). Touch, sound and vision in human superior temporal sulcus. *Neuroimage* 41, 1011–1020. doi: 10.1016/j.neuroimage.2008.03.015
- Bradley, K. M., O'sullivan, V. T., Soper, N. D., Nagy, Z., King, E. M., Smith, A. D., et al. (2002). Cerebral perfusion SPET correlated with Braak pathological stage in Alzheimer's disease. *Brain* 125, 1772–1781. doi: 10.1093/brain/awf185
- Bremmer, F., Schlack, A., Shah, N. J., Zafiris, O., Kubischik, M., Hoffmann, K., et al. (2001). Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron* 29, 287–296. doi: 10.1016/s0896-6273(01)00198-2
- Brown dyke, J. N., Giovanello, K., Petrella, J., Hayden, K., Chiba-Falek, O., Tucker, K. A., et al. (2013). Phenotypic regional functional imaging patterns during memory encoding in mild cognitive impairment and Alzheimer's disease. *Alzheimers Dement.* 9, 284–294. doi: 10.1016/j.jalz.2011.12.006
- Calvert, G. A. (2001). Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cereb. Cortex* 11, 1110–1123. doi: 10.1093/cercor/11.12.1110
- Chen, Y. C., and Spence, C. (2011). Crossmodal semantic priming by naturalistic sounds and spoken words enhances visual sensitivity. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 1554–1568. doi: 10.1037/a0024329
- Chetelat, G., Desgranges, B., Landeau, B., Mezenge, F., Poline, J. B., De La Sayette, V., et al. (2008). Direct voxel-based comparison between grey matter hypometabolism and atrophy in Alzheimer's disease. *Brain* 131, 60–71. doi: 10.1093/brain/awm288
- Dalton, P., Doolittle, N., Nagata, H., and Breslin, P. A. (2000). The merging of the senses: integration of subthreshold taste and smell. *Nat. Neurosci.* 3, 431–432. doi: 10.1038/74797
- Davis, S. W., Dennis, N. A., Daselaar, S. M., Fleck, M. S., and Cabeza, R. (2008). Que PASA? The posterior-anterior shift in aging. *Cereb. Cortex* 18, 1201–1209. doi: 10.1093/cercor/bhm155
- Dematte, L. M., Sanabria, D., and Spence, C. (2006). Cross-modal associations between odors and colors. *Chem. Senses* 31, 531–538. doi: 10.1093/chemse/bjj057

- Dematte, M. L., Sanabria, D., and Spence, C. (2009). Olfactory discrimination: when vision matters? *Chem. Senses* 34, 103–109. doi: 10.1093/chemse/bjn055
- Diaconescu, A. O., Hasher, L., and McIntosh, R. A. (2013). Visual dominance and multisensory integration changes with age. *Neuroimage* 65, 152–166. doi: 10.1016/j.neuroimage.2012.09.057
- Diederich, A., Colonius, H., and Schomburg, A. (2008). Assessing age-related multisensory enhancement with the time-window-of-integration model. *Neuropsychologia* 46, 2556–2562. doi: 10.1016/j.neuropsychologia.2008.03.026
- Ethofer, T., Pourtois, G., and Wildgruber, D. (2006). Investigating audiovisual integration of emotional signals in the human brain. *Prog. Brain Res.* 156, 345–361. doi: 10.1016/s0079-6123(06)56019-4
- Fabiani, M. (2012). It was the best of times, it was the worst of times: a psychophysiological view of cognitive aging. *Psychophysiology* 49, 283–304. doi: 10.1111/j.1469-8986.2011.01331.x
- Falkenstein, M., Yordanova, J., and Kolev, V. (2006). Effects of aging on slowing of motor-response generation. *Int. J. Psychophysiol.* 59, 22–29. doi: 10.1016/j.ijpsycho.2005.08.004
- Ferri, C. P., Prince, M., Brayne, C., Brodaty, H., Fratiglioni, L., Ganguli, M., et al. (2005). Global prevalence of dementia: a Delphi consensus study. *Lancet* 366, 2112–2117. doi: 10.1016/s0140-6736(05)67889-0
- Filippi, M., and Agosta, F. (2011). Structural and functional network connectivity breakdown in Alzheimer's disease studied with magnetic resonance imaging techniques. *J. Alzheimers Dis.* 24, 455–474. doi: 10.3233/JAD-2011-101854
- Gottfried, J. A., and Dolan, R. J. (2003). The nose smells what the eye sees: crossmodal visual facilitation of human olfactory perception. *Neuron* 39, 375–386. doi: 10.1016/s0896-6273(03)00392-1
- Grady, C. (2012). The cognitive neuroscience of ageing. *Nat. Rev. Neurosci.* 13, 491–505. doi: 10.1038/nrn3256
- Grady, C. L., Maisog, J. M., Horwitz, B., Ungerleider, L. G., Mentis, M. J., Salerno, J. A., et al. (1994). Age-related changes in cortical blood flow activation during visual processing of faces and location. *J. Neurosci.* 14, 1450–1462.
- Grady, C. L., Springer, M. V., Hongwanishkul, D., McIntosh, A. R., and Winocur, G. (2006). Age-related changes in brain activity across the adult lifespan. *J. Cogn. Neurosci.* 18, 227–241. doi: 10.1162/jocn.2006.18.2.227
- Gusnard, D. A., and Raichle, M. E. (2001). Searching for a baseline: functional imaging and the resting human brain. *Nat. Rev. Neurosci.* 2, 685–694. doi: 10.1038/35094500
- Hafkemeijer, A., Van Der Grond, J., and Rombouts, S. A. (2012). Imaging the default mode network in aging and dementia. *Biochim. Biophys. Acta* 1822, 431–441. doi: 10.1016/j.bbadis.2011.07.008
- Hairston, W. D., Laurienti, P. J., Mishra, G., Burdette, J. H., and Wallace, M. T. (2003). Multisensory enhancement of localization under conditions of induced myopia. *Exp. Brain Res.* 152, 404–408. doi: 10.1007/s00221-003-1646-7
- Hedman, A. M., Van Haren, N. E., Schnack, H. G., Kahn, R. S., and Hulshoff Pol, H. E. (2012). Human brain changes across the life span: a review of 56 longitudinal magnetic resonance imaging studies. *Hum. Brain Mapp.* 33, 1987–2002. doi: 10.1002/hbm.21334
- Helfer, K. S. (1998). Auditory and auditory-visual recognition of clear and conversational speech by older adults. *J. Am. Acad. Audiol.* 9, 234–242.
- Hugenschmidt, C. E., Mozolic, J. L., and Laurienti, P. J. (2009). Suppression of multisensory integration by modality-specific attention in aging. *Neuroreport* 20, 349–353. doi: 10.1097/wnr.0b013e328323ab07
- Kassuba, T., Klinge, C., Holig, C., Menz, M. M., Ptito, M., Roder, B., et al. (2011). The left fusiform gyrus hosts trisensory representations of manipulable objects. *Neuroimage* 56, 1566–1577. doi: 10.1016/j.neuroimage.2011.02.032
- King, A. J., and Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Exp. Brain Res.* 60, 492–500. doi: 10.1007/bf00236934
- Klasen, M., Chen, Y. H., and Mathiak, K. (2012). Multisensory emotions: perception, combination and underlying neural processes. *Rev. Neurosci.* 23, 381–392. doi: 10.1515/revneuro-2012-0040
- Klasen, M., Kenworthy, C. A., Mathiak, K. A., Kircher, T. T., and Mathiak, K. (2011). Supramodal representation of emotions. *J. Neurosci.* 31, 13635–13643. doi: 10.1523/jneurosci.2833-11.2011
- Kreifelts, B., Ethofer, T., Shiozawa, T., Grodd, W., and Wildgruber, D. (2009). Cerebral representation of non-verbal emotional perception: fMRI reveals audiovisual integration area between voice- and face-sensitive regions in the superior temporal sulcus. *Neuropsychologia* 47, 3059–3066. doi: 10.1016/j.neuropsychologia.2009.07.001
- Langbaum, J. B., Chen, K., Lee, W., Reschke, C., Bandy, D., Fleisher, A. S., et al. (2009). Categorical and correlational analyses of baseline fluorodeoxyglucose positron emission tomography images from the Alzheimer's Disease Neuroimaging Initiative (ADNI). *Neuroimage* 45, 1107–1116. doi: 10.1016/j.neuroimage.2008.12.072
- Langner, R., Kellermann, T., Eickhoff, S. B., Boers, F., Chatterjee, A., Willmes, K., et al. (2012). Staying responsive to the world: modality-specific and -nonspecific contributions to speeded auditory, tactile and visual stimulus detection. *Hum. Brain Mapp.* 33, 398–418. doi: 10.1002/hbm.21220
- Laurienti, P. J., Burdette, J. H., Maldjian, J. A., and Wallace, M. T. (2006). Enhanced multisensory integration in older adults. *Neurobiol. Aging* 27, 1155–1163. doi: 10.1016/j.neurobiolaging.2005.05.024
- Li, C. S., Yan, P., Bergquist, K. L., and Sinha, R. (2007). Greater activation of the “default” brain regions predicts stop signal errors. *Neuroimage* 38, 640–648. doi: 10.1016/j.neuroimage.2007.07.021
- Liu, X. Z., and Yan, D. (2007). Ageing and hearing loss. *J. Pathol.* 211, 188–197. doi: 10.1002/path.2102
- Mahoney, J. R., Li, P. C. C., Oh-Park, M., Verghese, J., and Holtzer, R. (2011). Multisensory integration across the senses in young and old adults. *Brain Res.* 1426, 43–53. doi: 10.1016/j.brainres.2011.09.017
- Mahoney, J. R., Verghese, J., Dumas, K., Wang, C., and Holtzer, R. (2012). The effect of multisensory cues on attention in aging. *Brain Res.* 1472, 63–73. doi: 10.1016/j.brainres.2012.07.014
- Meredith, M. A., and Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science* 221, 389–391. doi: 10.1126/science.6867718
- Meredith, M. A., and Stein, B. E. (1985). Descending efferents from the superior colliculus relay integrated multisensory information. *Science* 227, 657–659. doi: 10.1126/science.3969558
- Meredith, M. A., and Stein, B. E. (1986). Visual, auditory and somatosensory convergence on cells in superior colliculus results in multisensory integration. *J. Neurophysiol.* 56, 640–662.
- Mozolic, J. L., Hugenschmidt, C. E., Peiffer, A. M., and Laurienti, P. J. (2008). Modality-specific selective attention attenuates multisensory integration. *Exp. Brain Res.* 184, 39–52. doi: 10.1007/s00221-007-1080-3
- Mozolic, J. L., Hugenschmidt, C. E., Pfeiffer, A. M., and Laurienti, P. J. (2012). “Multisensory integration and aging,” in *The Neural Bases of Multisensory Processes*, eds M. M. Murray and M. T. Wallace (Boca Raton, FL: CRC Press). <http://www.ncbi.nlm.nih.gov/books/NBK92841/>
- Muller, V. I., Cieslik, E. C., Turetsky, B. I., and Eickhoff, S. B. (2012). Cross-modal interactions in audiovisual emotion processing. *Neuroimage* 60, 553–561. doi: 10.1016/j.neuroimage.2011.12.007
- Muller, V. I., Habel, U., Derntl, B., Schneider, F., Zilles, K., Turetsky, B. I., et al. (2011). Incongruence effects in crossmodal emotional integration. *Neuroimage* 54, 2257–2266. doi: 10.1016/j.neuroimage.2010.10.047
- Nordin, S., Almkvist, O., and Berglund, B. (2012). Is loss in odor sensitivity inevitable in the aging individual? A study of “successfully aged” elderly. *Chem. Percept.* 5, 188–196. doi: 10.1007/s12078-011-9102-8
- Park, D. C., and Reuter-Lorenz, P. (2009). The adaptive brain: aging and neurocognitive scaffolding. *Annu. Rev. Psychol.* 60, 173–196. doi: 10.1146/annurev.psych.59.103006.093656
- Peiffer, A. M., Mozolic, J. L., Hugenschmidt, C. E., and Laurienti, P. J. (2007). Age-related multisensory enhancement in a simple audiovisual detection task. *Neuroreport* 18, 1077–1081. doi: 10.1097/wnr.0b013e3281e72ae7
- Petrie, E. C., Cross, D. J., Galasko, D., Schellenberg, G. D., Raskind, M. A., Peskind, E. R., et al. (2009). Preclinical evidence of Alzheimer changes: convergent cerebrospinal fluid biomarker and fluorodeoxyglucose positron emission tomography findings. *Arch. Neurol.* 66, 632–637. doi: 10.1001/archneurol.2009.59
- Phillips, N. A., Baum, S., and Taler, V. (2009). “Audio-visual speech perception in mild cognitive impairment and healthy elderly controls,” in *AVSP 2009 - International Conference on Audio-Visual Speech Processing*. Norwich, UK.
- Price, J. L. (2008). Multisensory convergence in the orbital and ventrolateral prefrontal cortex. *Chemosens. Percept.* 1, 103–109. doi: 10.1007/s12078-008-9013-5
- Rawson, N. E. (2006). Olfactory loss in aging. *Sci. Aging Knowledge Environ.* 2006:pe6. doi: 10.1126/sageke.2006.5.pe6

- Raz, N., and Rodrigue, K. M. (2006). Differential aging of the brain: patterns, cognitive correlates and modifiers. *Neurosci. Biobehav. Rev.* 30, 730–748. doi: 10.1016/j.neubiorev.2006.07.001
- Reetz, K., Nickl-Jockschat, T., Kleiman, A., Schulz, J. B., Scheider, F., Laird, A. R., et al. (2012). Neuroanatomic changes and their association with cognitive decline in mild cognitive impairment: a meta-analysis. *Brain Struct. Funct.* 217, 115–125. doi: 10.1007/s00429-011-0333-x
- Regenbogen, C., Schneider, D. A., Finkelmeyer, A., Kohn, N., Derntl, B., Kellermann, T., et al. (2012a). The differential contribution of facial expressions, prosody and speech content to empathy. *Cogn. Emot.* 26, 995–1014. doi: 10.1080/02699931.2011.631296
- Regenbogen, C., Schneider, D. A., Gur, R. E., Schneider, F., Habel, U., and Kellermann, T. (2012b). Multimodal human communication - Targeting facial expressions, speech content and prosody. *Neuroimage* 60, 2346–2356. doi: 10.1016/j.neuroimage.2012.02.043
- Sala, S., Agosta, F., Pagani, E., Copetti, M., Comi, G., and Filippi, M. (2012). Microstructural changes and atrophy in brain white matter tracts with aging. *Neurobiol. Aging* 33, 488–498, e482. doi: 10.1016/j.neurobiolaging.2010.04.027
- Schlee, W., Leirer, V., Kolassa, S., Thurm, F., Elbert, T., and Kolassa, I. T. (2012). Development of large-scale functional networks over the lifespan. *Neurobiol. Aging* 33, 2411–2421. doi: 10.1016/j.neurobiolaging.2011.11.031
- Schroeter, M. L., Stein, T., Maslowski, N., and Neumann, J. (2009). Neural correlates of Alzheimer's disease and mild cognitive impairment: a systematic and quantitative meta-analysis involving 1351 patients. *Neuroimage* 47, 1196–1206. doi: 10.1016/j.neuroimage.2009.05.037
- Seo, H. S., and Hummel, T. (2011). Auditory-olfactory integration: congruent or pleasant sounds amplify odor pleasantness. *Chem. Senses* 36, 301–309. doi: 10.1093/chemse/bjq129
- Seubert, J., Kellermann, T., Loughhead, J., Boers, F., Brensinger, C., Schneider, F., et al. (2010a). Processing of disgusted faces is facilitated by odor primes: a functional MRI study. *Neuroimage* 53, 746–756. doi: 10.1016/j.neuroimage.2010.07.012
- Seubert, J., Loughhead, J., Kellermann, T., Boers, F., Brensinger, C. M., and Habel, U. (2010b). Multisensory integration of emotionally valenced olfactory-visual information in patients with schizophrenia and healthy controls. *J. Psychiatry Neurosci.* 35, 185–194. doi: 10.1503/jpn.090094
- Small, D. M., and Prescott, J. (2005). Odor/taste integration and the perception of flavor. *Exp. Brain Res.* 166, 345–357. doi: 10.1007/s00221-005-2376-9
- Sommers, M. S., Tye-Murray, N., and Spehar, B. (2005). Auditory-visual speech perception and auditory-visual enhancement in normal-hearing younger and older adults. *Ear Hear.* 26, 263–275. doi: 10.1097/00003446-200506000-00003
- Sowell, E. R., Thompson, P. M., and Toga, A. W. (2004). Mapping changes in the human cortex throughout the span of life. *Neuroscientist* 10, 372–392. doi: 10.1177/1073858404263960
- Spear, P. D. (1993). Neural bases of visual deficits during aging. *Vision Res.* 33, 2589–2609. doi: 10.1016/0042-6989(93)90218-1
- Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Atten. Percept. Psychophys.* 73, 971–995. doi: 10.3758/s13414-010-0073-7
- Staal, J., Pinkney, L., and Roane, D. (2003). Assessment of stimulus preference in Snoezelen therapy for elderly with dementia. *Br. J. Occup. Ther.* 66, 542–551.
- Stein, B. E., and Meredith, M. A. (1990). Multisensory integration. Neural and behavioral solutions for dealing with stimuli from different sensory modalities. *Ann. N Y Acad. Sci.* 608, 51–65. doi: 10.1111/j.1749-6632.1990.tb48891.x
- Stein, B. E., and Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nat. Rev. Neurosci.* 9, 255–266. doi: 10.1038/nrn2331
- Stephen, J. M., Knoefel, J. E., Adair, J., Hart, B., and Aine, C. J. (2010). Aging-related changes in auditory and visual integration measured with MEG. *Neurosci. Lett.* 484, 76–80. doi: 10.1016/j.neulet.2010.08.023
- Stine, E. A., Wingfield, A., and Myers, S. D. (1990). Age differences in processing information from television news: the effects of bisensory augmentation. *J. Gerontol.* 45, P1–P8. doi: 10.1093/geronj/45.1.p1
- Walden, B. E., Busacco, D. A., and Montgomery, A. A. (1993). Benefit from visual cues in auditory-visual speech recognition by middle-aged and elderly persons. *J. Speech Hear. Res.* 36, 431–436.
- Wang, L., Laviolette, P., O'keefe, K., Putcha, D., Bakkour, A., Van Dijk, K. R., et al. (2010a). Intrinsic connectivity between the hippocampus and posteromedial cortex predicts memory performance in cognitively intact older individuals. *Neuroimage* 51, 910–917. doi: 10.1016/j.neuroimage.2010.02.046
- Wang, L., Negreira, A., Laviolette, P., Bakkour, A., Sperling, R. A., and Dickerson, B. C. (2010b). Intrinsic interhemispheric hippocampal functional connectivity predicts individual differences in memory performance ability. *Hippocampus* 20, 345–351. doi: 10.1002/hipo.20771

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 12 June 2013; accepted: 26 November 2013; published online: 13 December 2013.

Citation: Freiherr J, Lundström JN, Habel U and Reetz K (2013) Multisensory integration mechanisms during aging. *Front. Hum. Neurosci.* 7:863. doi: 10.3389/fnhum.2013.00863

This article was submitted to the journal *Frontiers in Human Neuroscience*. Copyright © 2013 Freiherr, Lundström, Habel and Reetz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Cross-modal integration of emotions in the chemical senses

Moustafa Bensafi<sup>1\*</sup>, Emilia Iannilli<sup>2</sup>, Valentin A. Schriever<sup>2</sup>, Johan Poncelet<sup>1</sup>, Han-Seok Seo<sup>3</sup>, Johannes Gerber<sup>4</sup>, Catherine Rouby<sup>1</sup> and Thomas Hummel<sup>2</sup>

<sup>1</sup> CNRS UMR5292, INSERM U1028, Lyon Neuroscience Research Center, University Lyon, Lyon, France

<sup>2</sup> Smell and Taste Clinic, Department of Otorhinolaryngology, University of Dresden Medical School, Dresden, Germany

<sup>3</sup> Department of Food Science, University of Arkansas, Fayetteville, AR, USA

<sup>4</sup> Department of Neuroradiology, University of Dresden Medical School, Dresden, Germany

## Edited by:

Martin Klasen, RWTH Aachen University, Germany

## Reviewed by:

Maria G. Veldhuizen, The John B Pierce Laboratory, USA  
Jessica Freiherr, RWTH Aachen University, Germany

## \*Correspondence:

Moustafa Bensafi, CNRS UMR5292, INSERM U1028, Lyon Neuroscience Research Center, University Lyon, 50 Avenue Tony Garnier, Lyon F-69366, France  
e-mail: bensafi@olfac.univ-lyon1.fr

Although the brain structures involved in integrating odorant and trigeminal stimuli are well-documented, there is still a need to clarify (1) how emotional response is represented in the human brain during cross-modal interaction between odors and trigeminal stimuli, and (2) whether the degree of congruency between the two types of stimuli influences these emotional responses and their neural processing. These questions were explored combining psychophysics, event-related potentials (ERP) and fMRI in the same group of 17 subjects under a “congruent condition” (intranasal carbon dioxide mixed with the smell of orange, a combination found in soda drinks, for example), and an “incongruent condition” (intranasal carbon dioxide mixed with the smell of rose, a combination not encountered in everyday life). Responses to the 3 constituent stimuli (carbon dioxide, orange, and rose) were also measured. Hedonic and intensity ratings were collected for all stimulations. The congruent bimodal stimulus was rated as more pleasant than the incongruent. This behavioral effect was associated with enhanced neural activity in the hippocampus and anterior cingulate gyrus, indicating that these brain areas mediate reactivation of pleasant and congruent olfactory-trigeminal associations.

**Keywords:** olfaction, trigeminal, emotion, fMRI, congruency

## INTRODUCTION

Chemosensation comprises three main sensory modalities: olfaction and gustation, involved in discrimination and identification of, respectively, odorant and tastant stimuli, and the trigeminal system involved in detecting the irritating, fresh or painful component of chemosensory stimuli [see Lundström et al. (2011) for a review]. Past and current studies have detailed the functioning of each of these systems (Anderson et al., 2003; Small et al., 2003; Boyle et al., 2007b; Hummel et al., 2009a,b), but their interactions (although numerous and very close) have been much less studied (Small and Prescott, 2005; Boyle et al., 2007a; Bensafi et al., 2012). Moreover, one important transversal aspect is the strong emotional component of chemosensory perception. Firstly, a particular odor, taste or trigeminal stimulus can provide an early warning of toxic substances (spoiled or toxic food, industrial pollutants). Secondly, olfaction, taste and the trigeminal system combine to play a major role in hedonic experience: orangeade, with its orange odor, sweet taste and fresh gaseous components, can be best appreciated on a hot summer's day or after a sports effort.

The chemical senses thus provide a special window onto the cross-modal integration of emotion: chemosensory stimuli are mixtures of various compounds stimulating the olfactory, gustatory and trigeminal systems; each system may evoke particular affective states. The mechanisms and brain structures involved in the neural integration of odors and tastes have been well-documented in the last decade (Dalton et al., 2000; De Araujo et al., 2003; Small et al., 2004; Small and Prescott, 2005),

but there is still a need to understand how emotional responses are represented in the human brain during cross-modal interaction of odors and trigeminal stimuli. Psychophysical and neuroimaging studies have highlighted the role of congruency in this cross-modal integration. Regarding food in particular, congruency has been defined as the extent to which sensory stimuli can appropriately combine in eating or drinking a given foodstuff (Schifferstein and Verlegh, 1996). Past and recent studies suggest that congruency is a key factor in modulating the cross-modal integration of chemosensory stimuli, especially when the sensory cues belong to the same object.

Schifferstein and Verlegh (1996), studying odor-taste interactions, showed that the pleasantness of odor-taste mixtures correlates positively with the degree of congruency between the two types of stimulus. For these authors, two components need to form a harmonious (or congruent) combination in order to be pleasant.

Pleasantness, congruency and harmony are nevertheless linked to familiarity, which may explain why adding an unpleasant stimulus (salt to chocolate, or CO<sub>2</sub> to a beverage) can increase the overall pleasantness of the combination. Thus, a less rewarding stimulus (such as salt or pepper in chocolate) or an intrinsically painful stimulus (such as CO<sub>2</sub> in a beverage) becomes part of the integrated percept of a familiar food. As suggested by Rozin and colleagues, the memory of a food may inhibit the painful or warning value of the trigeminal input, and even make it desirable (Rozin et al., 1982).

On the neural level, De Araujo and colleagues showed that a congruent odor-taste combination (strawberry/sucrose) was perceived as more pleasant than an incongruent one (strawberry/monosodium glutamate) and that increasing congruency correlated positively with antero-medial orbitofrontal activity (De Araujo et al., 2003). Likewise, Small et al. observed that a congruent odor-taste mixture (vanilla/sweet) was perceived as more pleasant than an incongruent mixture (vanilla/salt) (Small et al., 2004). Moreover, the congruent odor-taste mixture induced greater activation than its components in the anterior cingulate cortex, insula, posterior orbitofrontal cortex, prefrontal cortex and parietal cortex, whereas the same brain areas were not activated in a similar comparative analysis of perception of the incongruent odor-taste mixture.

Such a congruency effect was also reported in odor-vision interaction. Gottfried and Dolan showed that congruent pairs of visual and olfactory stimuli were detected faster than incongruent pairs, and activated the rostro-medial orbitofrontal cortex and the hippocampus (Gottfried and Dolan, 2003). Even color has an effect on the perception of smells. In an fMRI study, Osterbauer and colleagues scanned human subjects exposed to smells and colors, in isolation and in congruent or incongruent combinations: activity in the posterior orbitofrontal cortex and insula increased as a function of the congruency of the smell-color pairs (Osterbauer et al., 2005). Using olfactory event-related potentials, Seo et al. showed that a congruent abstract visual symbol enhanced the intensity of the smell of rose compared to presentation of no symbol; it increased the pleasantness of rose odor and the unpleasantness of an unpleasant odor; and congruent symbols induced significantly higher amplitudes and shorter latencies in the N1 component of olfactory event-related potentials than did incongruent symbols (Seo et al., 2010).

Finally, Seo and Hummel extended this effect of congruency to odor-sound integration, demonstrating that even auditory cues can modulate odor pleasantness. Subjects were presented with congruent, incongruent or neutral sounds before and during the presentation of a smell: the olfactory stimuli were rated more pleasant in the presence of a congruent than an incongruent sound (Seo and Hummel, 2011).

Thus, congruency affects perception at different levels of processing, from detection (Gottfried and Dolan, 2003) to intensity (Seo et al., 2010) or pleasantness (Schiffstein and Verlegh, 1996; De Araujo et al., 2003; Small et al., 2004; Seo and Hummel, 2011). In addition, this perceptual modulation is associated with neural changes in a set of sensory and heteromodal areas, including the orbito-frontal cortex, cingulate cortex, insula, hippocampus, prefrontal cortex, and parietal cortex.

The first aim of the present study was to examine the influence of congruency on the (1) pleasantness and intensity of olfacto-trigeminal mixtures and (2) brain activity in the above-cited central structures in response to bimodal odor-trigeminal stimulation.

Moreover, congruency seems to affect the temporal processing of chemosensory cross-modal integration, as suggested by the chemosensory event-related potential (CSERP) study by Seo et al. (2010). In human adults, CSERPs usually include two

main components: (1) a negative component (N1) at around 400 ms; and (2) a late positive component (P2) at around 600 ms. Congruency has been shown to affect N1 latency and amplitude; its effect on the P2 component, however, is not clear. P2 amplitude increases as a function of emotional intensity and P2 latency has been shown to decrease with odor pleasantness (Pause and Krauel, 2000; Lundstrom et al., 2006; Poncelet et al., 2010).

The second aim of the present study was to test the influence of congruency on both the N1 and P2 CSERP components in response to odor-trigeminal stimuli. Psychophysics, fMRI and electroencephalography were combined in the same subjects under congruent and incongruent conditions particularly relevant to food: in the “congruent” condition, intranasal carbon dioxide was mixed with the smell of orange (a combination found in soda drinks), and in the “incongruent condition” intranasal carbon dioxide was mixed with the smell of rose (a combination not encountered in everyday life). Responses to the 3 unimodal stimuli (carbon dioxide, orange and rose) were also measured. Pleasantness and intensity ratings and hemodynamic responses (fMRI) from all five conditions were measured. After functional imaging, EEG responses (CSERP) to these conditions were collected from all participants.

## MATERIALS AND METHODS

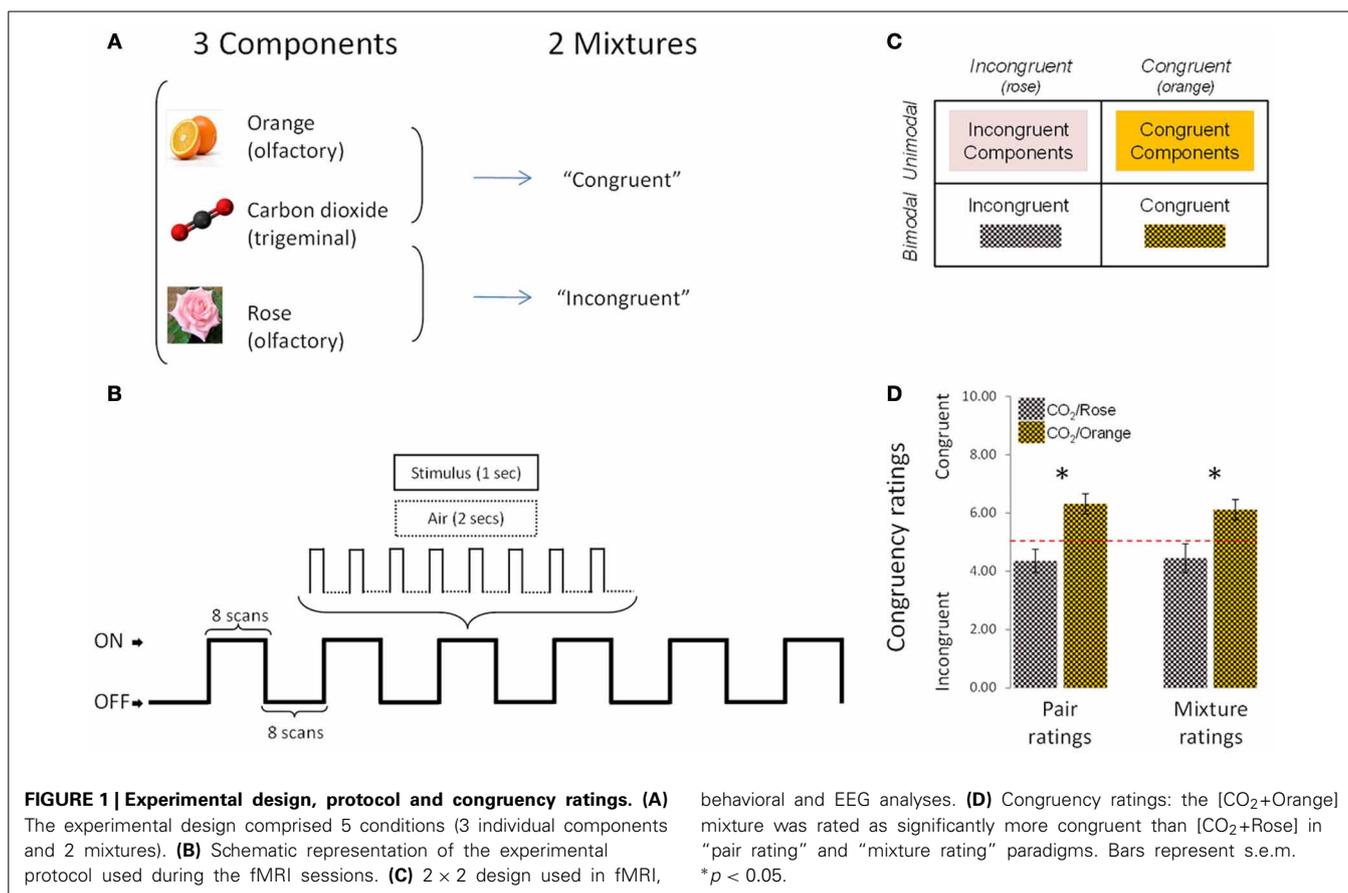
### SUBJECTS

Participants were 17 right-handed volunteers; mean age:  $23.58 \pm 1.97$  years; 4 male, 13 female. They received 20 Euros for participation. The recording procedure was explained in great detail to the subjects, who provided written consent prior to participation. The study was conducted according to the Declaration of Helsinki and was approved by the ethics committee of the University of Dresden. Instructions consisted of an explanation of the experimental design, which included functional, anatomical and EEG sessions. Subjects were instructed to not move during the fMRI experiment. Detailed medical history combined with nasal endoscopy of the nasal cavity and odor perception assessment by the “Sniffin’ Sticks” test (Hummel et al., 1997) ascertained that subjects were in good health and had normal sense of smell.

### STIMULI AND OLFACTOMETER

The stimuli were rose odor (“Ros,” phenyl ethyl alcohol, 20%, Aldrich Chemie GmbH, Riedstraße 2, Stauheim, Germany; CAS # 60-12-8), orange odor (“Ora,” 20%, Orange aroma oil; Frey and Lau, Henstedt-Ulzburg, Germany), carbon dioxide (“CO<sub>2</sub>,” 40%, Praxair, Dresden, Germany), an incongruent mixture of rose odor + CO<sub>2</sub> (“Inc,” 20 + 40%) and a congruent mixture of orange odor + CO<sub>2</sub> (“Cong,” 20 + 40%) (Figure 1A). Stimuli were mixed before dilution, so that the number of molecules per odorant could be presumed to be identical in the mixtures and in the individual stimuli.

A Burghart OM6b pulsed olfactometer was used to deliver rectangular-shaped chemical stimuli with controlled stimulus onset. Mechanical stimulation was avoided by embedding the stimuli in a constant flow of odorless humidified air at controlled temperature (80% relative humidity; total flow 6 L/min; 36°C) (Kobal, 1981). Prior to the experiment, subjects were trained in



**FIGURE 1 | Experimental design, protocol and congruency ratings. (A)** The experimental design comprised 5 conditions (3 individual components and 2 mixtures). **(B)** Schematic representation of the experimental protocol used during the fMRI sessions. **(C)** 2 × 2 design used in fMRI,

behavioral and EEG analyses. **(D)** Congruency ratings: the [CO<sub>2</sub>+Orange] mixture was rated as significantly more congruent than [CO<sub>2</sub>+Rose] in "pair rating" and "mixture rating" paradigms. Bars represent s.e.m. \**p* < 0.05.

the lab to breathe through the mouth without concomitant nasal airflow (velopharyngeal closure; Kobal, 1981) in order to avoid respiratory airflow in the nasal cavity during the chemosensory stimulation. A thermally insulated Teflon™ cannula directed the gaseous stimulus from the olfactometer to the subject's nose in the MRI and EEG rooms.

#### fMRI EXPERIMENT

The study started with the fMRI experiment, which was performed on a 1.5 Tesla MR-scanner (Siemens Sonata, Erlangen, Germany) and lasted approximately 60 min (from the arrival to the departure of the subject). Unlike the EEG experiment, which could be performed in a single session, the fMRI study was divided into 5 functional sessions to allow participants to take a break from the noisy fMRI environment every 5 min. Sessions were randomized, one per stimulus condition: "CO<sub>2</sub>," "Ros," "Ora," "Inc," and "Con." Each functional session in turn comprised 6 on/off-blocks, with 24-s blocks presented alternately in the On (stimulus-on) and Off (stimulus-off or "Air") conditions (Figure 1B). During the "On" conditions (lasting 24 s), odorants were presented 8 times, for 1 s followed by no-odor diffusion for 2 s. The fMRI data were collected in 96 volumes/session with a 36 axial-slice matrix 2D SE/EP sequence (Matrix: 64 × 64; TR: 3 s; TE: 35 ms; FA: 90°; voxel size: 3 × 3 × 3.75 mm). Total duration of the functional sessions was 24 min. In the 6 min immediately

following, a high-resolution T1-weighted image of the brain (3D IR/GR sequence: TR = 2180ms/TE = 3.93 ms) was acquired.

During the functional sessions, subjects were instructed to breathe through the mouth without concomitant nasal airflow (velopharyngeal closure, as described above), were not cued for any stimulus presentation and were not aware of the identity of the stimuli. They were not asked to perform any detection or cognitive tasks during stimulus presentation. For each session, following the 6 on/off-blocks, participants were asked to evaluate the stimulus in terms of intensity (on a scale from "0" = "not perceived" to "10" = "extremely intense") and of pleasantness (on a scale from "-5" = "extremely unpleasant" to "+5" = "extremely pleasant"). One intensity rating and one pleasantness rating were collected for each session.

Statistical analysis of fMRI data used SPM8 software (Statistical Parametric Mapping; Wellcome Department of Cognitive Neurology, London, UK) implemented in Matlab 7.1 (MathWorks Inc., Natick, MA, USA). Spatial pre-processing comprised registering, realignment with motion parameters included later in the model, co-registration between functional and structural images, normalization in stereotaxic space, and smoothing by means of a 7\*7\*7 mm<sup>3</sup> FWHM Gaussian kernel (Ashburner and Friston, 2003); first-level statistical analysis was then implemented with canonical hemodynamic response functions. For each subject, the following individual contrasts were performed: ["Odors" vs. "Air"], ["CO<sub>2</sub>" vs. "Air"], ["Cong" vs. "Air"], ["Inc"

vs. “Air”), [“Cong” vs. its components], and [“Inc” vs. its components], where “Air” corresponds to the stimulus-off period of each condition.

Group analyses used a random-effects model (Penny et al., 2003). In total 4 types of second-level analysis were performed: (1) [“Odors” vs. “Air”] to examine brain areas responding to odors; (2) [“CO<sub>2</sub>” vs. “Air”] to examine brain areas responding to the pure trigeminal stimulus; (3) [“Cong” vs. “Air”] vs. [“Inc” vs. “Air”] (and vice versa) to examine the differential activation of the congruent and the incongruent conditions; and (4) [“Cong” vs. its components] vs. [“Inc” vs. its components] (and vice versa) to examine the differential activation of the congruent condition (minus its components) and the incongruent condition (minus its component).

Activation coordinates were presented in MNI space. We report here results for brain areas in which a congruency effect had previously been demonstrated, as described in the Introduction: the hippocampus, insula, OFC, prefrontal cortex and cingulate gyrus. The primary olfactory cortex (including olfactory areas, amygdala and entorhinal cortex) and somatosensory areas (post-central gyrus) were also included. Activation loci were thus identified within this brain network of interest, delineated by an inclusive mask created with the WFU PickAtlas toolbox (Maldjian et al., 2003).

Areas of significant activation were identified at cluster level for values exceeding a *p*-value of 0.001 (5 voxels, uncorrected). Additionally, small volume corrections (SVC) were implemented, using coordinates from previously published studies, to determine the significance of predicted peaks in anterior cingulate gyrus (15, 33, 27; -10, 48, 4), posterior OFC (36, 18, -12; -16, 32, -4), prefrontal cortex (57, 36, 9), antero-medial OFC (-3, 39, -18; -12, 42, -18), hippocampus (-27, -12, -24) and insula (-32, 22, -8) (Gottfried and Dolan, 2003; Small et al., 2004; Osterbauer et al., 2005).

For behavioral data, because of the nature of the subjective scale (10-point ordinal), the non-parametric Wilcoxon test was applied to intensity and pleasantness ratings. Data were entered into a 2 × 2 design (Figure 1C) and analyzed in two ways. Firstly, congruent and incongruent stimuli were compared directly, for the bimodal condition (mixtures) on the one hand and the unimodal condition (components) on the other. Thus, for intensity and pleasantness ratings, the following comparisons were performed: [“Cong” vs. “Inc”] (bimodal comparison) and [Sum of “Cong” components vs. Sum of “Inc” components] (unimodal comparison).

Secondly, to enable comparison with the fMRI analyses, the following comparison was made for intensity and pleasantness ratings: [“Cong” vs. sum of its components] vs. [“Inc” vs. sum of its components], thereby analyzing differential intensity and pleasantness ratings between the congruent condition (minus its components) and the incongruent condition (minus its components).

## EEG EXPERIMENT

One to 3 days after the fMRI experiment, participants were asked to take part in an EEG experiment lasting approximately 2 h. The 5 stimulus conditions were presented in random order. Each stimulus was presented 15 times with a stimulus duration of 200 ms

and an inter-stimulation interval of 40 s. During the experiment, subjects received white noise through headphones to mask the switching clicks of the stimulation device. To stabilize vigilance, subjects performed a tracking task on a computer screen: using a joystick, they had to keep a small square inside a larger one which moved unpredictably (Bensafi et al., 2007a).

EEG was recorded at positions F3, F4, Fz, C3, C4, Cz, P3, P4, and Pz, of the 10/20 system [referenced against linked earlobes (A1 + A2)] using a 16 channel amplifier (Brain Star AC-2000; Schabert instruments, Röttenbach, Germany). Sintered silver-chlorided silver disc electrodes (electrode diameter, 5 mm) were attached to the cleansed skin (“Skin Pure” prepping cream; Nihon Kohden, Tokyo, Japan) using self-adhesive cream (“EC2 Grass Electrode Cream”; Grass, Warwick, RI, USA).

Eye blinks were monitored via the Fp2 lead. Sampling frequency was 250 Hz. Recording time was 2048 ms per recording (bandpass 0.02–30 Hz, with a pre-trigger baseline period of 530ms). Recordings were additionally filtered off-line (low-pass 15 Hz).

CSERPs were averaged after discarding recordings contaminated by motor artifacts or blinks (>50 μV at Fp2), detected by an experienced investigator. The minimum number of trials for each condition remaining after artifact rejection was *n* = 8 (Kobal, 1981; Hummel and Kobal, 2001). Peak amplitudes and latencies (N1 and P2) were measured heuristically by an experienced observer using EPEvaluate 4.2.2 software (Kobal, Erlangen, Germany). To enable comparison with the behavioral and fMRI data, four experimental conditions were included in the 2 × 2 design (Figure 1C): “Cong,” “Inc,” “Components of Cong” (i.e., “CO<sub>2</sub>” and “Ora”) and “Components of Inc” (i.e., “CO<sub>2</sub>” and “Ros”). Because of the continuous nature of the measurements, these conditions were then entered into ANOVAs (rather than non-parametrical tests) including “modality” (2: unimodal, bimodal) and “congruency” (2: incongruent, congruent) as within-subject factors. Additionally a within-subject “electrode” factor (9: Cz, C3, C4, Fz, F3, F4, Pz, P3, P4) was included in the analysis. This 2 × 2 × 9 ANOVA was performed for the amplitudes and latencies of both the N1 and P2 components. The sum of the components was calculated on N1 and P2 amplitudes (to examine a potential hyperadditivity effect), and the mean of the components was calculated for N1 and P2 latencies (as there was no reason to assume any additive effect for EEG latencies).

## CONTROL FOR PERCEPTUAL CONGRUENCY

To ensure that the two mixtures were indeed rated as congruent and incongruent by subjects, a psychophysical study was performed in a separate set of 13 healthy subjects (27.77 ± 3.39 years; 5 men) with normal sense of smell [ascertained by the “Sniffin’ Sticks” test (Hummel et al., 1997)]. Here, congruency was assessed on two protocols. Firstly, participants were asked to rate the congruency of two stimuli presented separately (“pair ratings”): CO<sub>2</sub> was presented first (for 200 ms), followed by a rest period of 10 s, followed by the smell of orange or rose (for 200 ms). Participants had to estimate the congruency between the two stimuli after delivery of the second one (the odor). Instructions were: “You will be presented two stimuli one after the other. Your task will be to estimate how these 2 stimuli match

or are congruent: i.e., how far they can be associated in real life (in nature, food, drink, perfume, an object, etc.). To this end, please use the following scale: 0 (no association, match or congruency) to 10 (very associated, matched, congruent).” Each pair of stimuli (CO<sub>2</sub> followed by orange, or CO<sub>2</sub> followed by rose) was presented 5 times. The 10 trials were presented in random order with a 1-min interval between pairs of stimuli.

Secondly, subjects were asked to evaluate the congruency between stimuli presented together in mixtures (“mixture ratings”): CO<sub>2</sub> presented simultaneously with the smell of orange or rose (for 200 ms). The instructions were: “You will be presented a mixture composed of carbon dioxide plus a smell. Your task will be to estimate how far these 2 stimulations match or are congruent and how far they can be associated in real life (in nature, food, drink, perfume, an object, etc.). To this end, please use the following scale: 0 (no association, match or congruency) to 10 (very associated, matched, congruent).” Each mixture (CO<sub>2</sub>+Orange or CO<sub>2</sub>+Rose) was presented 5 times. The 10 trials were presented in random order with a 1-min inter-stimulus interval. All stimuli and mixtures were presented at the same concentrations as in the main study.

As expected, results revealed that the mixture composed of CO<sub>2</sub>+Orange was rated as significantly more congruent than the mixture composed of CO<sub>2</sub>+Rose in both the first paradigm (pair ratings;  $p < 0.01$ , Wilcoxon test) and the second paradigm (mixture ratings;  $p < 0.03$ , Wilcoxon test) (**Figure 1D**).

## RESULTS

### fMRI EXPERIMENT

Confirmatory analyses examined the main effect of odors and trigeminal stimuli. For odors, the odorant conditions were summed and contrasted with their odorless baseline conditions. Activity was observed in the piriform cortex and inferior frontal gyrus (Supplementary table 1; Supplementary figure 1). The same analysis was performed for the pure trigeminal stimulus (carbon dioxide), and results revealed neural activity in the inferior, middle and superior frontal gyrus, pre- and post-central gyri, cingulate gyrus, frontomarginal gyrus and insula (Supplementary table 1; Supplementary figure 1). These findings replicate previous studies [see Albrecht et al. (2010); Seubert et al. (2012) for reviews] and indicate that our methodology did induce neural activation in the olfactory and trigeminal systems.

To examine the effect of congruency in the perception of bimodal olfacto-trigeminal mixtures, two types of analysis were performed. First, the activation induced by the congruent mixture (“Cong” vs. “Air”) was compared to that resulting from the incongruent mixture (“Inc” vs. “Air”). Results revealed significant activation in the anterior cingulate gyrus ( $p < 0.05$  after SVC) (**Figure 2A**; **Table 1**). The opposite contrast (incongruent vs. congruent), on the other hand, did not show any significant activation. In the second analysis, the activation induced by the congruent mixture minus its components was compared to that resulting from the incongruent mixture minus its components. In this case, significant activation was observed in the hippocampus ( $p < 0.05$  after SVC), accompanied by additional activation in the anterior cingulate gyrus bordering the upper part of the

posterior orbitofrontal gyrus ( $p < 0.05$  after SVC) (**Figure 2B**; **Table 1**). The opposite contrast, on the other hand, did not show any significant activation.

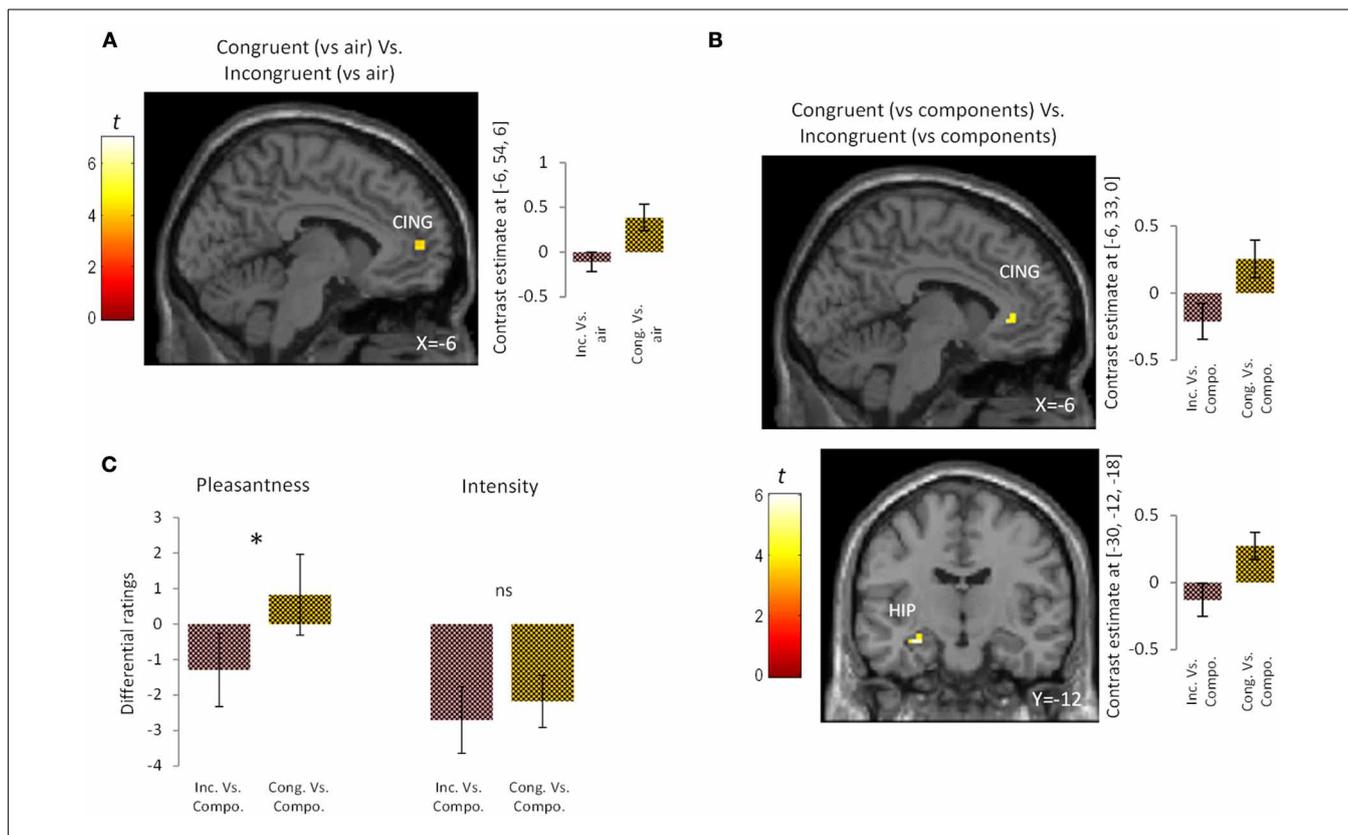
At a perceptual level, the congruent bimodal mixture was perceived as more intense than the incongruent bimodal mixture (mean  $\pm$  s.e.m.: congruent mixture =  $7.47 \pm 0.59$ ; incongruent mixture =  $6.29 \pm 0.59$ ;  $p = 0.05$ ) and as more pleasant (mean  $\pm$  s.e.m.: congruent mixture =  $1.64 \pm 0.58$ ; incongruent mixture =  $-0.23 \pm 0.65$ ;  $p < 0.03$ ). Comparison of the unimodal components of the congruent and incongruent stimuli revealed a trend toward greater intensity for the congruent vs. the incongruent components (mean  $\pm$  s.e.m.: congruent components =  $9.64 \pm 0.73$ ; incongruent components =  $9.00 \pm 0.80$ ;  $p = 0.06$ ) but no significant difference in pleasantness (mean  $\pm$  s.e.m.: congruent components =  $0.82 \pm 1.10$ ; incongruent components =  $1.06 \pm 0.84$ ;  $p = 0.83$ ). Finally, direct comparison between the congruent condition (minus its components) and the incongruent condition (minus its components) revealed that the congruent mixture was perceived as more pleasant ( $p < 0.05$ ) but not more intense ( $p = 0.53$ ) than the incongruent mixture (**Figure 2C**).

### EEG EXPERIMENT

Significant effects of “electrode” were observed for N1 latency [ $F_{(8, 128)} = 2.59$ ,  $p < 0.01$ ] and P2 amplitude [ $F_{(8, 128)} = 2.63$ ,  $p < 0.02$ ]. In all the analyses, no significant effect of congruency was observed ( $p > 0.05$  in all cases). However, significant effects of “modality” were observed for N1 amplitude [mean  $\pm$  s.e.m.: unimodal =  $-10.38 \pm 1.23$ ; bimodal =  $-5.63 \pm 0.89$ ;  $F_{(1, 16)} = 28.11$ ,  $p < 0.0001$ ], N1 latency [mean  $\pm$  s.e.m.: unimodal =  $439.69 \pm 13.77$ ; bimodal =  $382.49 \pm 11.00$ ;  $F_{(1, 16)} = 17.48$ ,  $p < 0.0008$ ] and P2 latency [mean  $\pm$  s.e.m.: unimodal =  $564.56 \pm 18.05$ ; bimodal =  $495.46 \pm 15.01$ ;  $F_{(1, 16)} = 23.38$ ,  $p < 0.0003$ ], reflecting the fact that the bimodal mixtures evoked shorter N1 and P2 latencies and smaller N1 amplitudes than their individual components (**Figure 3**). Interactions between factors did not reach significance ( $p > 0.05$ ) except for P2 latency, where an “electrode”-by-“congruency” interaction was observed [ $F_{(8, 128)} = 2.12$ ,  $p < 0.04$ ]. Nevertheless, paired comparison within each electrode site did not reveal any effect of congruency ( $p > 0.05$  in all 9 cases).

### DISCUSSION

The aim of the present study was to explore the effect of congruency on the perception and neural responses induced by combined odor-trigeminal stimuli. The first result of interest was that congruency affected the perceptual ratings of bimodal chemosensory stimuli: the congruent mixture was perceived as more pleasant and more intense than the incongruent mixture. Analysis of the cumulative effect of the congruent and incongruent mixtures compared to their individual components showed that the two mixtures differed in terms of pleasantness but not intensity. These findings agree with a large set of psychophysical experiments showing that congruency enhances the intensity and/or pleasantness of bimodal stimuli (Schiffman and Verlegh, 1996; De Araujo et al., 2003; Small et al., 2004; Seo et al., 2010; Seo and Hummel, 2011). Analysis of chemosensory



**FIGURE 2 | Brain activation and perceptual ratings. (A)** Brain activation and contrast estimates in response to the congruent mixture (vs. air) vs. the incongruent mixture (vs. air): responses were observed in anterior cingulate cortex (CING,  $p < 0.05$  SVC). **(B)** Brain activation and contrast estimates in response to the congruent mixture (vs. its components) vs. the incongruent mixture (vs. its components):

responses were observed in the hippocampus (HIP,  $p < 0.05$  SVC) and anterior cingulate cortex (CING,  $p < 0.05$  SVC). **(C)** Differential ratings, showing pleasantness and intensity ratings for the congruent and incongruent mixtures vs. their individual components (Compo). \* $p < 0.05$ ; ns = non-significant difference at the 0.05 threshold. Bars represent s.e.m.

**Table 1 | Activation in response to (1) the congruent mixture (vs. air) vs. the incongruent mixture (vs. air) and (2) to the congruent mixture (vs. its components) vs. the incongruent mixture (vs. its components).**

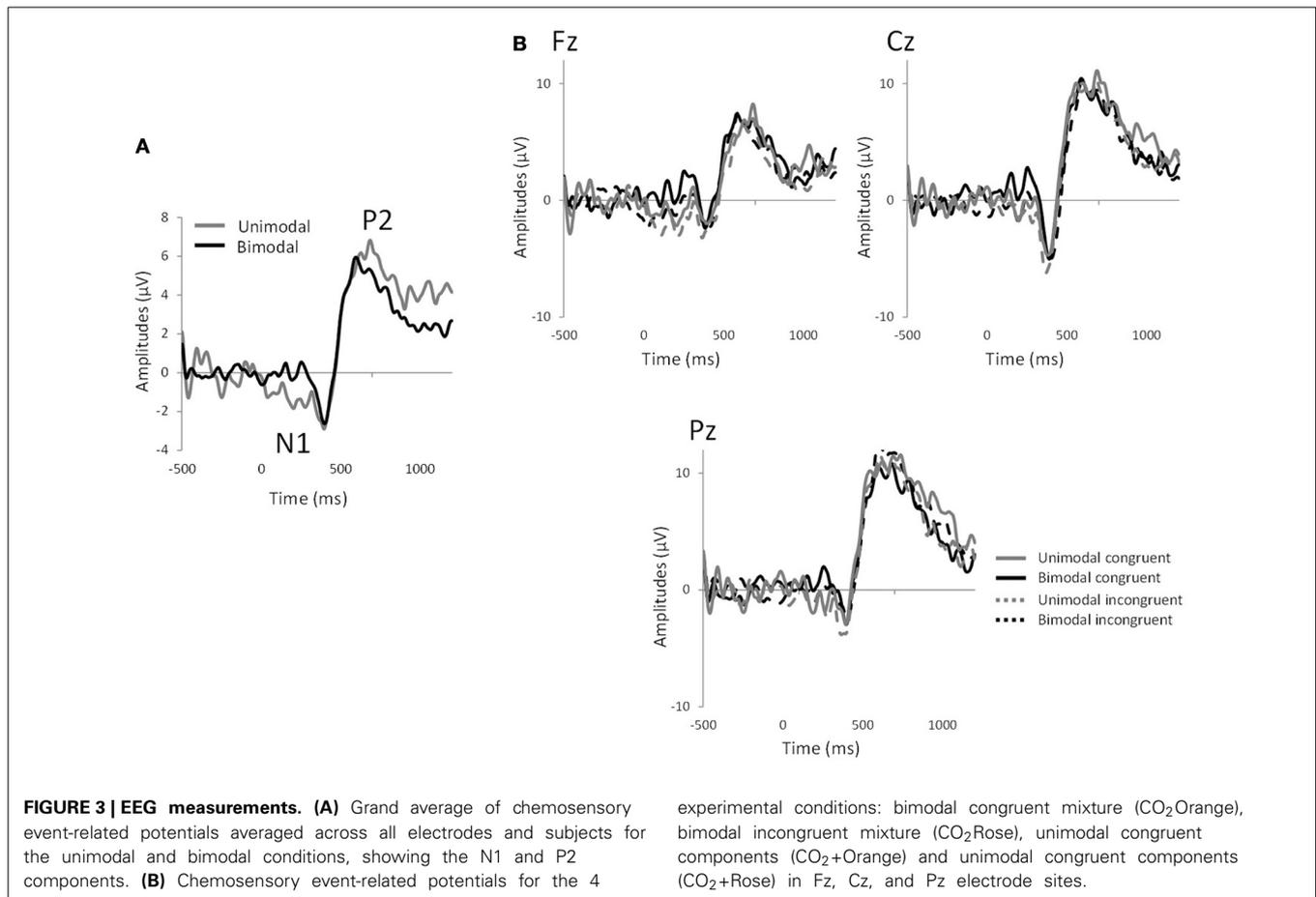
	K	T-value	x	y	z	Brain areas
Congruent mixture (vs. air) vs. the Incongruent mixture (vs. air)	7	4.31	-6	54	6	Cingulate gyrus
Congruent mixture (vs. components) vs. the Incongruent mixture (vs. components)	7	6.02	-30	-12	-18	Hippocampus
	5	4.55	-6	33	0	Cingulate gyrus

K is the cluster size. Statistical t-values are presented. MNI coordinates of activated brain areas are presented in x, y, and z.

event-related potentials revealed that both binary mixtures (congruent and incongruent) induced shorter N1 and P2 latencies and smaller N1 amplitudes compared to their individual components. Although studies in the non-chemosensory domain showed an

effect of congruency on the N400 component of event-related potential (Kutas and Federmeier, 2011), the present EEG study did not show any direct temporal difference between congruent and incongruent mixtures. Because CSERPs were recorded from a small number of electrodes, it is not unlikely that spatial differences exist but are not reflected by our EEG measures. Indeed, fMRI data revealed that congruency affected the spatial processing of chemosensory cross-modal integration.

A major result of the present study was the differential activation pattern seen during perception of the congruent compared to the incongruent mixture, notably in the cingulate gyrus and hippocampus. Past and more recent studies have revealed neural activity in cingulate gyrus in response to olfactory and trigeminal stimuli (Iannilli et al., 2007; Bensafi et al., 2008; Seubert et al., 2012), and a previous study suggested that the cingulate cortex is a multi-integrative structure in processing chemosensory stimuli (Small et al., 2004). Anatomically, cyto-architectural studies of the cingulate gyrus support a multiple-region model, with anterior, middle and posterior parts (Vogt, 2005). The functioning of these sub-regions is not homogeneous and their involvement in cross-modal integration may differ according to modality. Klasen and colleagues showed that the ventral posterior cingulate cortex was activated during integration of congruent audiovisual



stimuli (Klasen et al., 2011), while others showed that pleasant olfacto-verbal associations activated the anterior part of the cingulate cortex (De Araujo et al., 2005). The present findings are in line with the above results, highlighting a role of this brain area in binding olfactory and trigeminal representations of environmental objects.

With regard to the hippocampus, many investigations revealed a functional role of this brain area in chemosensory processing, especially odor processing. For example, positive correlation was observed between the volume of the hippocampus and odor threshold performances of healthy subjects (Smitka et al., 2012). Moreover, compared to sighted volunteers, congenitally blind subjects showed stronger hippocampal activation during a detection task (Kupers et al., 2011). The hippocampus is also involved in more cognitive olfactory tasks. For instance, hippocampal activity increases significantly as a function of odor identifiability (Kjelvik et al., 2012) and amnesic patients with atrophy of the hippocampus are impaired for odor-place associative memory (Goodrich-Hunsaker et al., 2009). The role of the hippocampus in binding between different stimuli has been established by previous studies (see for reviews Squire et al., 2004; Eichenbaum et al., 2007). For example, in the non-chemosensory domains, it has been observed that memory for congruous events, defined as events whose constituent elements match along particular attributes, recruit a network involving the inferior frontal gyrus

and the hippocampus (Staresina et al., 2009). Moreover, the hippocampus was shown by Gottfried and Dolan to be activated by congruent pairs of visual and olfactory stimuli, suggesting that it is a key component of the network underlying the binding of semantic information from different modalities (Gottfried and Dolan, 2003). Combined with the above, our results therefore suggest that emotional information during cross-modal integration of congruent pairs of odors and trigeminal stimuli also merges in the hippocampus, this area being potentially involved in binding the associations of both unimodal stimuli to form a harmonious mnemonic representation.

As mentioned in the introduction, such harmony or congruency is linked to familiarity, explaining why a bimodal stimulus comprising a painful trigeminal stimulus and a pleasant odor is appetitive for the tested individuals: even when one unimodal stimulus (here, intranasal carbon dioxide) arouses a sensation of pain, this intrinsically painful feature becomes part of the integrated percept of a familiar object. To sum up, congruency is likely the fruit of experience and culture, since it is based on the formation of previous associations in particular contexts; congruency effects influence the perceptual, emotional and cognitive processing of sensory stimuli, probably via expectancy (see Schifferstein and Verlegh, 1996). It is known that, when expectancies are evoked by colors (Osterbauer et al., 2005), tastants (Yeomans, 2006; Barkat et al., 2008), verbal labels (Herz, 2003; Bensafi et al.,

2007b, 2013; Rinck et al., 2011) or sounds (Seo and Hummel, 2011), they can alter the intensity and/or pleasantness ratings of chemosensory stimuli. The present study is the first to highlight such effects of congruency on odor-trigeminal integration.

Although the present study provides evidence for neural modulation by congruency, some of the findings deserve discussion. Certain aspects might be explained not only by the integration of the two harmonious congruent stimuli but also by probability summation mechanisms. Evidence that perceptual response is higher in a bimodal condition compared to unimodal conditions does not necessarily mean that the two stimuli interact perceptually: instead, it may be that two sensory stimuli are more likely than single signals to induce perceptual response. This concept was originally introduced for sensory thresholds, and recent studies suggest that it should be considered for detection of bimodal olfacto-gustatory stimuli (Veldhuizen et al., 2011). In the present case, nevertheless, the stimulations were relatively intense and no additive effect on perceived intensity was observed. On the contrary, the two mixtures were significantly less intense than the sum of their components, reflecting a hypo-additivity effect. Moreover, at a neural level, activation always resulted from a comparison between two bimodal stimuli.

In conclusion, the present study showed that a congruent association between an odor and a trigeminal stimulus was perceived as more pleasant than an incongruent association. This behavioral effect was associated with enhanced neural activity in the hippocampus and anterior cingulate gyrus, indicating that these brain areas mediate reactivation of pleasant and congruent olfactory-trigeminal associations. Taken together, these results are in line with the general view that, when a stimulus is encoded, the percept that emerges does not simply come from hierarchical processing in a single modality, from sensory transduction to the creation of a single mental representation: rather, chemosensory integration depends on other available stimuli (olfactory and trigeminal in the present case), and congruency between these cues is a prominent factor in the emotional perception of objects, as observed for the integration of smells with other sensory stimuli.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fnhum.2013.00883/abstract>

## REFERENCES

- Albrecht, J., Kopietz, R., Frasnelli, J., Wiesmann, M., Hummel, T., and Lundström, J. N. (2010). The neuronal correlates of intranasal trigeminal function—an ALE meta-analysis of human functional brain imaging data. *Brain Res. Rev.* 62, 183–196. doi: 10.1016/j.brainresrev.2009.11.001
- Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D. G., Glover, G., et al. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nat. Neurosci.* 6, 196–202. doi: 10.1038/nn1001
- Ashburner, J., and Friston, K. (2003). “Spatial normalization using basis function,” in *Human Brain Function 2nd Edn.*, ed R. S. J. Frackowiak (Amsterdam: Academic Press), 655–672.
- Barkat, S., Poncelet, J., Landis, B. N., Rouby, C., and Bensafi, M. (2008). Improved smell pleasantness after odor-taste associative learning in humans. *Neurosci. Lett.* 434, 108–112. doi: 10.1016/j.neulet.2008.01.037
- Bensafi, M., Croy, I., Phillips, N., Rouby, C., Sezille, C., Gerber, J., et al. (2013). The effect of verbal context on olfactory neural responses. *Hum. Brain Mapp.* doi: 10.1002/hbm.22215. [Epub ahead of print].
- Bensafi, M., Frasnelli, J., Reden, J., and Hummel, T. (2007a). The neural representation of odor is modulated by the presence of a trigeminal stimulus during odor encoding. *Clin. Neurophysiol.* 118, 696–701. doi: 10.1016/j.clinph.2006.10.022
- Bensafi, M., Rinck, F., Schaal, B., and Rouby, C. (2007b). Verbal cues modulate hedonic perception of odors in 5-year-old children as well as in adults. *Chem. Senses* 32, 855–862. doi: 10.1093/chemse/bjm055
- Bensafi, M., Iannilli, E., Gerber, J., and Hummel, T. (2008). Neural coding of stimulus concentration in the human olfactory and intranasal trigeminal systems. *Neuroscience* 154, 832–838. doi: 10.1016/j.neuroscience.2008.03.079
- Bensafi, M., Iannilli, E., Poncelet, J., Seo, H. S., Gerber, J., Rouby, C., et al. (2012). Dissociated representations of pleasant and unpleasant olfacto-trigeminal mixtures: an fMRI study. *PLoS ONE* 7:e38358. doi: 10.1371/journal.pone.0038358
- Boyle, J. A., Frasnelli, J., Gerber, J., Heinke, M., and Hummel, T. (2007a). Cross-modal integration of intranasal stimuli: a functional magnetic resonance imaging study. *Neuroscience* 149, 223–231. doi: 10.1016/j.neuroscience.2007.06.045
- Boyle, J. A., Heinke, M., Gerber, J., Frasnelli, J., and Hummel, T. (2007b). Cerebral activation to intranasal chemosensory trigeminal stimulation. *Chem. Senses* 32, 343–353. doi: 10.1093/chemse/bjm004
- Dalton, P., Doolittle, N., Nagata, H., and Breslin, P. A. (2000). The merging of the senses: integration of subthreshold taste and smell. *Nat. Neurosci.* 3, 431–432. doi: 10.1038/74797
- De Araujo, I. E., Rolls, E. T., Kringelbach, M. L., McGlone, F., and Phillips, N. (2003). Taste-olfactory convergence, and the representation of the pleasantness of flavour, in the human brain. *Eur. J. Neurosci.* 18, 2059–2068. doi: 10.1046/j.1460-9568.2003.02915.x
- De Araujo, I. E., Rolls, E. T., Velazco, M. I., Margot, C., and Cayeux, I. (2005). Cognitive modulation of olfactory processing. *Neuron* 46, 671–679. doi: 10.1016/j.neuron.2005.04.021
- Eichenbaum, H., Yonelinas, A. P., and Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annu. Rev. Neurosci.* 30, 123–152. doi: 10.1146/annurev.neuro.30.051606.094328
- Goodrich-Hunsaker, N. J., Gilbert, P. E., and Hopkins, R. O. (2009). The role of the human hippocampus in odor-place associative memory. *Chem. Senses* 34, 513–521. doi: 10.1093/chemse/bjp026
- Gottfried, J. A., and Dolan, R. J. (2003). The nose smells what the eye sees: cross-modal visual facilitation of human olfactory perception. *Neuron* 39, 375–386. doi: 10.1016/S0896-6273(03)00392-1
- Herz, R. S. (2003). The effect of verbal context on olfactory perception. *J. Exp. Psychol. Gen.* 132, 595–606. doi: 10.1037/0096-3445.132.4.595
- Hummel, T., Iannilli, E., Frasnelli, J., Boyle, J., and Gerber, J. (2009a). Central processing of trigeminal activation in humans. *Ann. N. Y. Acad. Sci.* 1170, 190–195. doi: 10.1111/j.1749-6632.2009.03910.x
- Hummel, T., Oehme, L., Van Den Hoff, J., Gerber, J., Heinke, M., Boyle, J. A., et al. (2009b). PET-based investigation of cerebral activation following intranasal trigeminal stimulation. *Hum. Brain Mapp.* 30, 1100–1104. doi: 10.1002/hbm.20573
- Hummel, T., and Kobal, G. (2001). “Olfactory event-related potentials,” in *Methods and Frontiers in Chemosensory Research*, eds S. A. Simon and M. A. Nicolelis (Boca Raton, FL: CRC press), 429–464.
- Hummel, T., Sekinger, B., Wolf, S. R., Pauli, E., and Kobal, G. (1997). ‘Sniffin’ sticks’: olfactory performance assessed by the combined testing of odor identification, odor discrimination and olfactory threshold. *Chem. Senses* 22, 39–52. doi: 10.1093/chemse/22.1.39
- Iannilli, E., Gerber, J., Frasnelli, J., and Hummel, T. (2007). Intranasal trigeminal function in subjects with and without an intact sense of smell. *Brain Res.* 1139, 235–244. doi: 10.1016/j.brainres.2006.12.082
- Kjelvik, G., Evensmoen, H. R., Brezova, V., and Haberg, A. K. (2012). The human brain representation of odor identification. *J. Neurophysiol.* 108, 645–657. doi: 10.1152/jn.01036.2010
- Klasen, M., Kenworthy, C. A., Mathiak, K. A., Kircher, T. T., and Mathiak, K. (2011). Supramodal representation of emotions. *J. Neurosci.* 31, 13635–13643. doi: 10.1523/JNEUROSCI.2833-11.2011
- Kobal, G. (1981). *Elektrophysiologische Untersuchungen des Menschlichen Geruchssinns*. Stuttgart: Thieme Verlag.

- Kupers, R., Beaulieu-Lefebvre, M., Schneider, F. C., Kassuba, T., Paulson, O. B., Siebner, H. R., et al. (2011). Neural correlates of olfactory processing in congenital blindness. *Neuropsychologia* 49, 2037–2044. doi: 10.1016/j.neuropsychologia.2011.03.033
- Kutas, M., and Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621–647. doi: 10.1146/annurev.psych.093008.131123
- Lundstrom, J. N., Boesveldt, S., and Albrecht, J. (2011). Central processing of the chemical senses: an overview. *ACS Chem. Neurosci.* 2, 5–16. doi: 10.1021/cn1000843
- Lundstrom, J. N., Seven, S., Olsson, M. J., Schaal, B., and Hummel, T. (2006). Olfactory event-related potentials reflect individual differences in odor valence perception. *Chem. Senses* 31, 705–711. doi: 10.1093/chemse/bjl012
- Maldjian, J. A., Laurienti, P. J., Kraft, R. A., and Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* 19, 1233–1239. doi: 10.1016/S1053-8119(03)00169-1
- Osterbauer, R. A., Matthews, P. M., Jenkinson, M., Beckmann, C. F., Hansen, P. C., and Calvert, G. A. (2005). Color of scents: chromatic stimuli modulate odor responses in the human brain. *J. Neurophysiol.* 93, 3434–3441. doi: 10.1152/jn.00555.2004
- Pause, B. M., and Krauel, K. (2000). Chemosensory event-related potentials (CSERP) as a key to the psychology of odors. *Int. J. Psychophysiol.* 36, 105–122. doi: 10.1016/S0167-8760(99)00105-1
- Penny, W. D., Holmes, A. P., and Friston, K. J. (2003). “Random effects analysis,” in *Human Brain Function*, eds R. S. J. Frackowiak, K. J. Friston, C. Frith, R. Dolan, C. J. Price, S. Zeki, et al. (New York, NY: Academic Press).
- Poncelet, J., Rinck, F., Bourgeat, F., Schaal, B., Rouby, C., Bensafi, M., et al. (2010). The effect of early experience on odor perception in humans: psychological and physiological correlates. *Behav. Brain Res.* 208, 458–465. doi: 10.1016/j.bbr.2009.12.011
- Rinck, F., Barkat-Defradas, M., Chakirian, A., Joussain, P., Bourgeat, F., Thevenet, M., et al. (2011). Ontogeny of odor liking during childhood and its relation to language development. *Chem. Senses* 36, 83–91. doi: 10.1093/chemse/bjq101
- Rozin, P., Ebert, L., and Schull, J. (1982). Some like it hot: a temporal analysis of hedonic responses to chili pepper. *Appetite* 3, 13–22. doi: 10.1016/S0195-6663(82)80033-0
- Schifferstein, H. N., and Verlegh, P. W. (1996). The role of congruency and pleasantness in odor-induced taste enhancement. *Acta Psychol. (Amst.)* 94, 87–105. doi: 10.1016/0001-6918(95)00040-2
- Seo, H. S., Arshamian, A., Schemmer, K., Scheer, I., Sander, T., Ritter, G., et al. (2010). Cross-modal integration between odors and abstract symbols. *Neurosci. Lett.* 478, 175–178. doi: 10.1016/j.neulet.2010.05.011
- Seo, H. S., and Hummel, T. (2011). Auditory-olfactory integration: congruent or pleasant sounds amplify odor pleasantness. *Chem. Senses* 36, 301–309. doi: 10.1093/chemse/bjq129
- Seubert, J., Freiherr, J., Djordjevic, J., and Lundstrom, J. N. (2012). Statistical localization of human olfactory cortex. *Neuroimage* 66C, 333–342. doi: 10.1016/j.neuroimage.2012.10.030
- Small, D. M., Gregory, M. D., Mak, Y. E., Gitelman, D., Mesulam, M. M., and Parrish, T. (2003). Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron* 39, 701–711. doi: 10.1016/S0896-6273(03)00467-7
- Small, D. M., and Prescott, J. (2005). Odor/taste integration and the perception of flavor. *Exp. Brain Res.* 166, 345–357. doi: 10.1007/s00221-005-2376-9
- Small, D. M., Voss, J., Mak, Y. E., Simmons, K. B., Parrish, T., and Gitelman, D. (2004). Experience-dependent neural integration of taste and smell in the human brain. *J. Neurophysiol.* 92, 1892–1903. doi: 10.1152/jn.00050.2004
- Smitka, M., Puschmann, S., Buschhueter, D., Gerber, J. C., Witt, M., Honeycutt, N., et al. (2012). Is there a correlation between hippocampus and amygdala volume and olfactory function in healthy subjects? *Neuroimage* 59, 1052–1057. doi: 10.1016/j.neuroimage.2011.09.024
- Squire, L. R., Stark, C. E., and Clark, R. E. (2004). The medial temporal lobe. *Annu. Rev. Neurosci.* 27, 279–306. doi: 10.1146/annurev.neuro.27.070203.144130
- Staresina, B. P., Gray, J. C., and Davachi, L. (2009). Event congruency enhances episodic memory encoding through semantic elaboration and relational binding. *Cereb. Cortex* 19, 1198–1207. doi: 10.1093/cercor/bhn165
- Veldhuizen, M. G., Shepard, T. G., Wang, M. F., and Marks, L. E. (2011). Coactivation of gustatory and olfactory signals in flavor perception. *Chem. Senses* 35, 121–133. doi: 10.1093/chemse/bjp089
- Vogt, B. A. (2005). Pain and emotion interactions in subregions of the cingulate gyrus. *Nat. Rev. Neurosci.* 6, 533–544. doi: 10.1038/nrn1704
- Yeomans, M. R. (2006). Olfactory influences on appetite and satiety in humans. *Physiol. Behav.* 87, 800–804. doi: 10.1016/j.physbeh.2006.01.029

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 31 May 2013; accepted: 04 December 2013; published online: 20 December 2013.

Citation: Bensafi M, Iannilli E, Schriever VA, Poncelet J, Seo HS, Gerber J, Rouby C and Hummel T (2013) Cross-modal integration of emotions in the chemical senses. *Front. Hum. Neurosci.* 7:883. doi: 10.3389/fnhum.2013.00883

This article was submitted to the journal *Frontiers in Human Neuroscience*. Copyright © 2013 Bensafi, Iannilli, Schriever, Poncelet, Seo, Gerber, Rouby and Hummel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Subclinical alexithymia modulates early audio-visual perceptive and attentional event-related potentials

Dyna Delle-Vigne, Charles Kornreich, Paul Verbanck and Salvatore Campanella\*

Laboratoire de Psychologie Médicale et d'Addictologie, ULB Neuroscience Institute, CHU Brugmann-Université Libre de Bruxelles, Brussels, Belgium

## Edited by:

Martin Klasen, RWTH Aachen University, Germany

## Reviewed by:

Kai Wang, Anhui Medical University, China

Michiko Kano, Tohoku University Graduate School of Medicine, Japan

## \*Correspondence:

Salvatore Campanella, The Belgian Fund for Scientific Research (FN.R.S.), CHU Brugmann, Psychiatry Secretary, 4 Place Vangehuchten - B-1020 Brussels, Belgium

e-mail: salvatore.campanella@

chu-brugmann.be;

salvatore.campanella@ulb.ac.be

**Introduction:** Previous studies have highlighted the advantage of using audio-visual oddball tasks (instead of unimodal ones) in order to electrophysiologically index subclinical behavioral differences. Since alexithymia is highly prevalent in the general population, we investigated whether the use of various bimodal tasks could elicit emotional effects in low- vs. high-alexithymic scorers.

**Methods:** Fifty students (33 females and 17 males) were split into groups based on low and high scores on the Toronto Alexithymia Scale (TAS-20). During event-related potential (ERP) recordings, they were exposed to three kinds of audio-visual oddball tasks: neutral-AVN—(geometrical forms and bips), animal-AVA—(dog and cock with their respective shouts), or emotional-AVE—(faces and voices) stimuli. In each condition, participants were asked to quickly detect deviant events occurring amongst a train of repeated and frequent matching stimuli (e.g., push a button when a sad face-voice pair appeared amongst a train of neutral face-voice pairs). P100, N100, and P300 components were analyzed: P100 refers to visual perceptive and attentional processing, N100 to auditory ones, and the P300 relates to response-related stages, involving memory processes.

**Results:** High-alexithymic scorers presented a particular pattern of results when processing the emotional stimulations, reflected in early ERP components by increased P100 and N100 amplitudes in the emotional oddball tasks [P100:  $F_{(2, 48)} = 20, 319, p < 0.001$ ; N100:  $F_{(2, 96)} = 8, 807, p = 0.001$ ] as compared to the animal or neutral ones. Indeed, regarding the P100, subjects exhibited a higher amplitude in the AVE condition ( $8.717 \mu\text{V}$ ), which was significantly different from that observed during the AVN condition ( $4.382 \mu\text{V}$ ,  $p < 0.001$ ). For the N100, the highest amplitude was found in the AVE condition ( $-4.035 \mu\text{V}$ ) and the lowest was observed in the AVN condition ( $-2.687 \mu\text{V}$ ,  $p = 0.003$ ). However, no effect was found on the later P300 component.

**Conclusions:** Our findings suggest that high-alexithymic scorers require heightened early attentional resources in comparison to low scorers, particularly when confronted with emotional bimodal stimuli.

**Keywords:** event-related potentials (ERPs), bimodal, alexithymia, emotion, subclinical

## INTRODUCTION

Event-related potentials (ERPs) offer a sensitive method of monitoring brain electrical activity, through a high temporal resolution in the order of milliseconds (Rugg and Coles, 1995). By this way, it is possible to observe the different electrophysiological components needed for healthy subjects to reach a “normal” performance during distinct cognitive tasks. ERPs also allow identifying the electrophysiological component(s) associated with the onset of a dysfunction in pathological populations. In this regard, ERPs constitute an unique technique for investigating even minor cognitive limitations, which might be undetectable at the behavioral level (e.g., Maurage et al., 2009), as indexing the neuro-cognitive origin of a deficit may give to clinicians important indications about the relevant impaired cognitive stages that should be rehabilitated (Campanella, 2013).

Among others, a classical paradigm used to evoke ERP components consists in an “oddball target detection task,” where the subject has to detect as quickly as possible (typically by pressing a button) a deviant-rare stimulus (e.g., a 1000 Hz sound) among a train of frequent stimuli (e.g., a 500 Hz sound). The oddball task evokes robust and reliable phenomena that have been used as markers of cognitive function (Rodríguez Holguín et al., 1999). There is increasing evidence to support that a number of early and late neuroelectric features in the information-processing stream can be anomalous in various psychiatric populations, the common finding being P300 abnormalities (Hansenne, 2000). The P3b component arises approximately 300–350 ms after an auditory stimulus and from 400 to 450 ms following visual input in the centro-parietal areas of subjects detecting deviant stimuli in oddball tasks (i.e., when the subject has to actively evaluate,

categorize, and make decisions regarding relevant stimuli). The discrimination of the target from the standard stimulus (deviant minus frequent) generates a robust P300, due to the novelty of the deviant stimulus, compared to the repeated frequent ones. (Polich, 2007). This positive deflection appears to be related to memory processing (Polich, 2007), since its amplitude does not seem to be modulated either by the motor requirements of the task or the physical features of the stimulus (Duncan et al., 2009). However, although a large number of studies have provided evidence on the relevance of the P300 component as a biological marker of mental illness, its clinical sensitivity has been hampered by the fact that its parameters (amplitude and latency) are diagnostically unspecific and unreliable (Pogarell et al., 2007). In other words, even though differences in P300 amplitude and latency can indicate the severity or evolution of a clinical state, the clinical value of this component as a diagnostic index is low. Therefore, a current and important challenge for neuro-physiologists is to discover novel and appropriate procedures for enhancing the clinical applicability and sensitivity of the P300 component (e.g., Bruder et al., 2009).

For this purpose, we recently proposed the use of more ecological stimulations (namely synchronized congruent audio-visual stimuli) in oddball tasks, instead of unimodal stimulations. In these studies (2010; 2012), two groups of participants were compared: one group was composed entirely of healthy individuals, whereas the other consisted of healthy people displaying anxious and depressive tendencies in the absence of full-blown clinical symptoms. Both groups were submitted to unimodal (visual and auditory separately) and bimodal oddball tasks. The principal findings suggested that although the two study groups differed in their subclinical level of comorbid anxiety and depression, this difference could not be detected by their P300 amplitudes during unimodal visual and auditory oddball tasks; however, *bimodal tasks did allow for the detection of subclinical symptoms*. The investigators hypothesized that the bimodal processing of multimodal information involves complex associative processes, including the integration of unimodal visual and auditory products into a single, coherent representation (Campanella and Belin, 2007 for a review). Therefore, a deficit in these associative processes might explain why the bimodal procedure enabled detection of a subclinical difference, which was impossible during unimodal conditions.

Nevertheless, despite the potential relevance of these studies to be applied in real clinical settings, the findings raised several questions, two of which formed the background for the present study. First, in the study by Campanella et al. (2012), individuals with subclinical anxious-depressive tendencies were compared to healthy subjects using two kinds of bimodal oddball tasks, consisting of emotional and neutral conditions (i.e., geometric forms and simple sounds) and two unimodal tasks (visual and auditory). In contrast to the unimodal tasks, both bimodal conditions elicited P300 amplitude differences (i.e., lower P300 amplitudes in the subclinical group compared to the control group), which were independent of the emotional content of the stimulations. In other words, the bimodal P300 helps to disclose differences between subclinical and control participants that did not appear through unimodal tasks; however, no major difference emerged

from the emotional-bimodal P300, as compared with the non-emotional one. This was quite surprising, as many psychiatric diseases show a deficit in emotional processing, which is indexed by decreased performance rates, longer response latencies compared to controls (Power and Dalgleish, 1997), as well as by decreased and delayed P300 component (Campanella et al., 2002, 2005, 2006; Mejias et al., 2005; Rossignol et al., 2005, 2008, 2012; Maurage et al., 2007a, 2008). Therefore, we would have expected to have greater bimodal P300 differences between groups when emotional stimuli were used. A possible explanation for this absence of effect could be that a major affective dimension, well-known to affect emotion processing, was not taken into account in these previous studies, i.e., alexithymia. In fact, alexithymia is often considered to be a stable personality trait that reflects a deficit in the conscious experience of emotion (Lane et al., 2000), involving an inability to recognize, regulate, and describe emotions (Luminet et al., 2001; Swart et al., 2009). Alexithymia can also describe difficulties in differentiating feelings from bodily sensations and decreased capacity for symbolization of emotion, such as in fantasy or dreams (Sifneos, 1973; Bagby and Taylor, 1997). Currently, this disorder is considered to be a risk and/or maintenance factor for various medical and psychiatric conditions, including anxiety and depression, even at a subclinical level (e.g., Berthoz et al., 1999; Luminet et al., 2001). Therefore, some differences between the non-emotional and the emotional bimodal oddball tasks might have been obscured in our previous study (Campanella et al., 2012) since we did not measure this factor. In the present study, we added the measure of alexithymia to verify this hypothesis.

Second, it is currently well accepted that ERP data should not “exclusively” focus on the P300 component, as numerous studies have shown that P300 deficits may be correlated with previous “early” ERP alterations (Maurage and Campanella, 2013). For instance, Maurage et al. (2007a) demonstrated in a visual-emotional oddball task that P300 modulations were correlated with deficits in earlier P100 and N170 components. Similarly, a previous study by our group (Campanella et al., 2006) confirmed earlier results obtained by Foxe et al. (2001), which indicated that schizophrenic patients displayed reduced amplitude and longer latencies in early visual components (e.g., P100, N100, and N170). The necessity to investigate other ERP components than the P300 was also supported by the fact that, up to now, reports describing the influence of alexithymia on the P300 component have been inconsistent. While some authors have described no effect of alexithymia on P300 amplitudes (Vermeulen et al., 2008; Walker et al., 2011), others have reported distinct P300 alterations in alexithymic individuals. For example, Bermond et al. (2008) found larger P300 amplitudes related to emotional stimuli in comparison to neutral stimuli, as well as for women compared to men. Also, they elicited a significant gender x alexithymia effect: high-scoring females on TAS-20 exhibited reduced P300 amplitudes compared to low-alexithymic women for negative pictures, whereas the opposite was true for males. In addition, Pollatos and Gramann (2011) reported lower P300 amplitudes for high-scorers on the TAS-20 in response to unpleasant pictures, while Franz et al. (2004) obtained a contradictory result.

In the present study, we intend therefore to analyze bimodal P300 components as well as earlier bimodal P100 and N100 components. Both N100 and P100 components are related to visual and auditory stimuli, but the N100 is more related to the auditory processing (e.g., Jessen et al., 2012; Liu et al., 2012), while the P100 is more visual (e.g., Allison et al., 1999; Singhal et al., 2012; Peschard et al., 2013). Specifically, the P100 wave is a positive component recorded around 100 ms at occipital sites in response to any visual stimulation and is associated with early spatial attention processes. It reflects activity in the extra-striate visual cortex and is sensitive to top-down control mechanisms (Martínez et al., 1999). As a result, the P100 wave is related to the physical-visual properties of stimuli, constituting a more automatic perceptual level of the information-processing stream (Heinze and Mangun, 1995). Usually, when emotional images are displayed, the early/sustained attention increases, easing the effect of the stimuli. This outcome is echoed by amplitude modulations in ERPs, especially the P100 component (Singhal et al., 2012). Regarding the N100 component, this negative deflection peaks around 100 ms following the onset of a prosodic stimulus. It is generated in the bilateral secondary auditory cortex (Engelien et al., 2000) and reflects the extraction of acoustic cues (frequency and intensity) during early auditory treatment. Typically, its amplitude increases based on the amount of attention allocated to an acoustic stimulus (Alho et al., 1994; Rinne et al., 2005).

Some electrophysiological data linking alexithymia to early ERPs abnormalities already existed, but involved, to our knowledge, only visual or auditory stimulations separately. Pollatos and Gramann (2011) reported that processing emotional pictures led to reductions in P100 amplitudes for high-scorers on the TAS-20. This finding was especially associated with neutral and positive pictures, with early P100 deficits linked to later variations in P300 amplitudes. Also, when subjects with subclinical tendencies of alexithymia were asked to detect and identify deviant stimuli in emotional prosody, Goerlich et al. (2012) found N100 alterations in high TAS-20 scorers. In particular, larger N100 amplitudes were observed in relation to disgusted prosody, whereas no behavioral differences were found between the two groups (also see Schäfer et al., 2007). Overall, it appears that alexithymic subjects may present a global deficit in perceiving all kinds of basic emotions, and these impairments can be neurophysiologically indexed by disrupted visual (P100 amplitude) and/or auditory (N100 amplitude) processing. These deficiencies likely extend not only to static and dynamic emotional facial expressions (EFEs) (Kätšyri et al., 2008), but also to neutral facial expressions (Montebarocci et al., 2011). Likewise, alexithymic subjects display a diminished capacity to recognize emotions related to non-verbal stimuli and responses (Lane et al., 1996). Since spatially-degraded or briefly-presented ( $\leq 1$  s) stimuli are more difficult to interpret for alexithymic subjects, it has been suggested that these individuals may need more time and/or more information to correctly identify EFEs (Franz et al., 2004, 2008; Parker et al., 2005; Kätšyri et al., 2008; Prkachin et al., 2009).

Overall, the main objective of the present study was to determine: (1) whether subclinical alexithymia results in particular emotional effects (i.e., amplitude and/or latency modulations) during the performance of emotional bimodal oddball tasks

compared to non-emotional ones; and (2) whether modulations on the P300 amplitude and/or on earlier components can be observed under these conditions. In order to reach this double objective, participants will be confronted to bimodal emotional and non-emotional (geometric forms and simple sounds) oddball tasks, similar than those used in Campanella et al.'s study (2012). However, we also added an "animal" bimodal condition, as this feature allowed us to have an authentic bimodal semantic association, which involved "non-emotional" stimuli existing in the subject's environment (i.e., a dog barking and a cock crowing). Indeed, this "meaningful" but "non-emotional" condition, is considered as a semantic condition, which is known to produce higher neural responses (Bookheimer, 2002). We postulated that emotional stimulation would require more processing than other types of stimuli, irrespective of the alexithymia, as emotional conditions command attentional resources and are prioritized due to significance (Vuilleumier, 2005). Therefore, when analyzing the effect of alexithymia, we believe that an "emotional effect" might exist, meaning that high scorers on the TAS-20 would experience more difficulties in processing emotional conditions compared to low scorers. However, no differences would be expected between the emotional and animal stimuli for the low-scoring group. Also, we hypothesize that deficiencies in the alexithymia-high group should be associated with higher early ERP components (P100 and/or N100 amplitudes), as more attentional resources will be required to accomplish tasks (e.g., Franz et al., 2004). We have to outline that we only used negative (sad) and neutral stimuli in the emotional condition because Mann et al. (1994) reported that high-scorers performed especially worse than low-scorers in labeling tasks when sad pictures were presented. Also, alexithymia is negatively correlated with the propensity to experience positive emotions (Bagby et al., 1994b), and some neuroimaging studies tend to confirm this fact. For example, Zald (2003) described improved cerebral features when using negative stimulations. Mantani et al. (2005) reported reduced activation of the posterior cingulate cortex in response to past and future happy situations in high alexithymic subjects, and Eichmann et al. (2008) found that masked sad faces were associated with greater bilateral activation of the fusiform gyrus in this kind of subjects. Furthermore, Zhang et al. (2012) highlighted that alexithymic subjects were particularly affected at the neural level when identifying negative emotions (anger, sadness, and fear). We also propose that the neutral conditions, having no "semantic content," will be the easiest to process for both groups. Finally, we found it interesting that some studies failed to exhibit performance differences between healthy and clinical/subclinical alexithymic populations. For example, in a study by Mann et al. (1995), high-alexithymic substance abusers performed similarly to a control group in labeling EFEs. Pandey and Mandal (1997) failed to elicit behavioral differences between high and low TAS-20 scorers using a labeling task with sadness, happiness, fear, anger, disgust, and surprise EFEs. Also, Galderisi et al. (2008) did not find differences between patients with panic disorder and high TAS-20 scores when compared to healthy controls. Therefore, as no behavioral difference was expected between groups (mainly due to the facility of the oddball tasks), we hypothesized that any alexithymia-related effect will be disclosed

by modulations on the response-related stages, indexed by P300 component, as in Vermeulen et al. (2008).

## METHODS

### PARTICIPANTS

Fifty students (33 females and 17 males) enrolled at the Free University of Brussels (18–27 years old) participated in this study. These individuals displayed normal/corrected vision and normal hearing. In addition, they were not on any medications and had no history of neurological/psychiatric disease. Screened through questionnaires, students presenting a heavy social drinking behavior, using drugs (mainly cannabis), or smoking >10 cigarettes per day were excluded from the study, as these variables are known to affect the P300 component (Solowij et al., 1995; Polich and Criado, 2006; Mobascher et al., 2010). The local ethics committee at the Brugmann Hospital approved the study, and informed written consent was obtained from each participant.

### TASK AND PROCEDURE

Before the ERP task, the subjects were asked to complete various self-reported questionnaires: the Beck Depression Inventory (13 items, BDI, Beck and Steer, 1987; French version: Collet and Cottraux, 1986) to assess depression tendencies; scores between 0 and 4 signify absence of depression, while scores between 8 and 15 displayed a subclinical level of moderate depression (Beck and Beck, 1972); the State and Trait Anxiety Inventory (STAI, Spielberger et al., 1983; French version: Bruchon-Schweitzer and Paulhan, 1993) for anxiety tendencies; scores below 36 reflect a very low anxiety, 36–45 low anxiety, 46–55 a normal anxiety, 56–65 high, and more than 65 very high anxiety; and the TAS-20 (Bagby et al., 1994a,b; French version: Loas et al., 1996) to measure alexithymic propensities; a score of less than 51 is not considered as alexithymia, and a score equal or higher than 61 indicates alexithymia (Taylor et al., 1997). The TAS-20 contains 20 items, each rated on a 5-point Likert scale. The test is composed of 3 subscales, focusing on difficulties in identifying feelings, difficulties in describing feelings, and an assessment of externally-oriented thinking style.

All participants carried out 6 bimodal (synchronized presentations of visual and auditory stimulations) oddball tasks [two “emotional” (AVE), two “animal” (AVA), and two “neutral” (AVN) tasks]. In each exercise, the participants were required to detect as quickly as possible deviant events occurring amongst a train of repeated and frequent matching stimuli by clicking a button with their right forefinger. This experimental set-up is similar to that used in Campanella et al. (2010, 2012). In the “emotional” bimodal auditory–visual oddball task (AVE), pairs of synchronized and congruent faces and voices were displayed to participants [frequent stimulus: a neutral face and a neutral voice pronouncing the word “*papier*” (French for “paper”); deviant stimulus: a sad face with a sad voice; the frequent and deviant stimuli were inverted in the second block]. Faces were selected from Ekman and Friesen’s set of standardized pictures (1976), and voices were chosen from the validated battery of vocal emotional expressions (Maurage et al., 2007b). In the “animal” bimodal auditory–visual oddball task (AVA), pairs of synchronized and congruent pictures of a dog and a cock with their respective

shouts were displayed to participants (again, in the two blocks of this condition, the frequent stimulus and the deviant one were inverted). Finally, in the “neutral” bimodal auditory–visual oddball task (AVN), pairs of synchronized geometrical figures and tones were shown to participants (frequent stimulus: a square and a 750 Hz sound; deviant stimulus: a triangle with a 1000 Hz sound; stimuli were once again inverted in the second block). The various stimuli are illustrated in **Figure 1**.

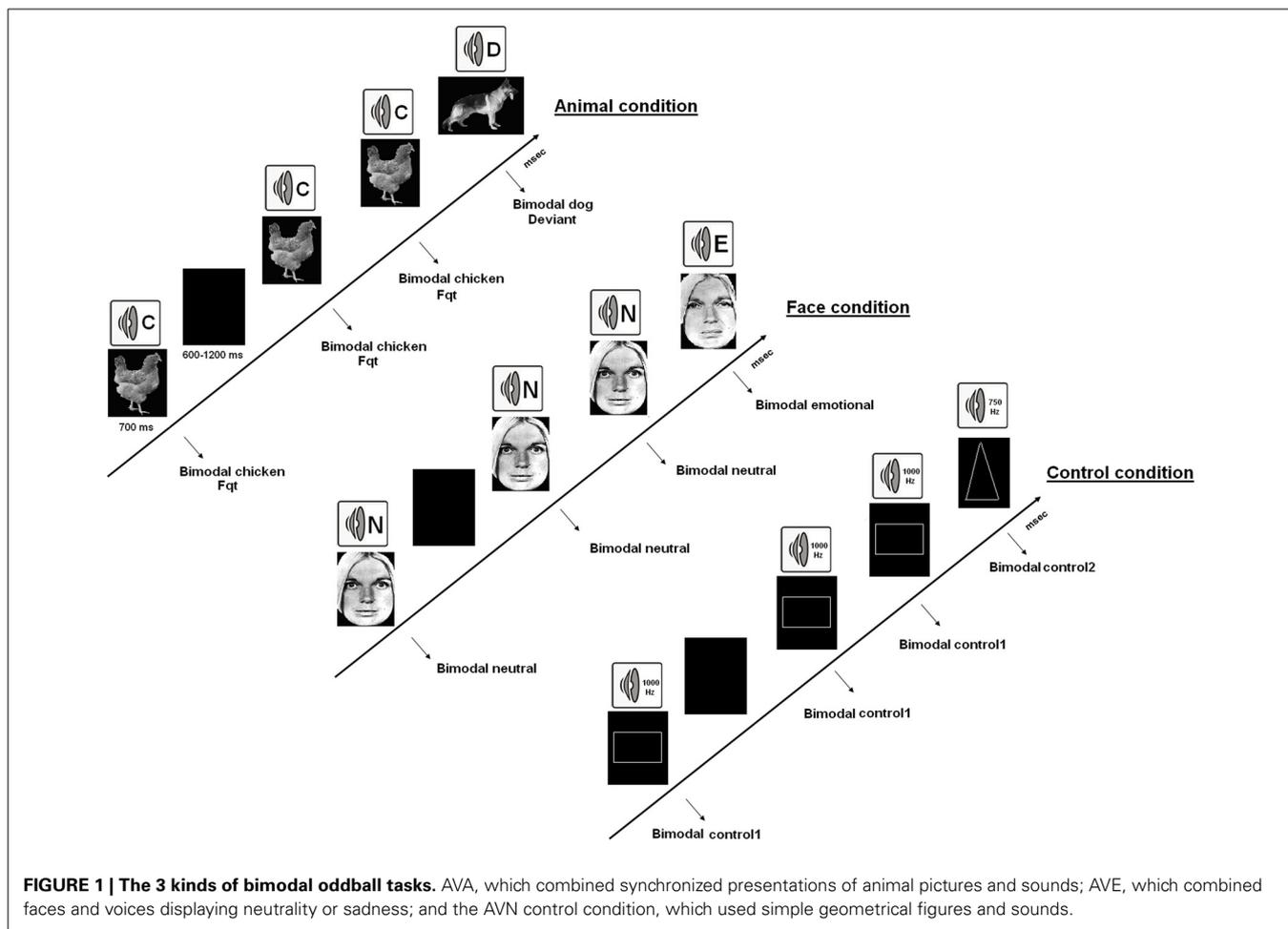
Each block included a total of 130 stimuli (100 frequent and 30 deviant), and every participant completed 6 blocks (2 emotional, 2 animal, and 2 neutral; approximately 5 min for each block). Participants were informed about upcoming exercises during the intervals between the blocks. The order of the 6 blocks was counterbalanced among the subjects.

During the ERP recordings, each participant sat alone in a dark room on a chair placed one meter from the screen, with his or her head restrained on a chin rest. The visual stimuli subtended a visual angle of 3° × 4°. Each stimulus was presented for 700 ms, and a black screen was displayed between stimuli for a random duration (600–1200 ms). From the onset of each stimulus, the participants had at least 1300 ms to respond. Reaction times and error rates were recorded. There were two categories of errors: omission (i.e., not pressing the answer key when a deviant stimulus appeared) and false recognition (i.e., pressing the answer key when a standard stimulus appeared). Participants were informed that speed was important, but not at the cost of accuracy. Only correct answers (i.e., deviant stimuli for which the subject pressed the answer key) were used in the analysis of reaction times and ERPs.

### EEG RECORDING AND ANALYSIS

The electroencephalography (EEG) was recorded by 32 electrodes mounted in an electrode Quick-Cap. Electrode positions included the standard 10–20 system locations and intermediate positions. Recordings were made with a linked mastoid physical reference but were re-referenced using a common average (Bertrand et al., 1985). The EEG was amplified by battery-operated A.N.T.® amplifiers with a gain of 30,000 and a band-pass of 0.01–100 Hz. The impedance of the electrodes was kept below 20 k $\Omega$ . The EEG was continuously recorded (sampling rate 500 Hz; A.N.T. Eeprobe software) and trials that were contaminated by electrooculogram (EOG) artifacts (mean of 15%) were eliminated offline, using a procedure developed by Semlitsch et al. (1986). In brief, an average artifact response was computed for each individual based on a percentage of the maximum eye movement potential (generally recorded on prefrontal electrodes). The EOG response was thereby subtracted from the EEG channels on a sweep-by-sweep, point-by-point basis in order to obtain ocular artifact-free data. Epochs beginning 200 ms prior to the stimulus onset and continuing for 800 ms were created. Three parameters were coded for every stimulus: (1) the modality of the task (AVE, AVA, and AVN), (2) the type of stimulus (deviant vs. frequent), and (3) the response type (keypress for deviant stimuli, no keypress for frequent ones). Data were filtered with a 30 Hz, low-pass filter.

For each modality and each subject, the component of interest (P100, N100, and P300) was investigated by gathering individual



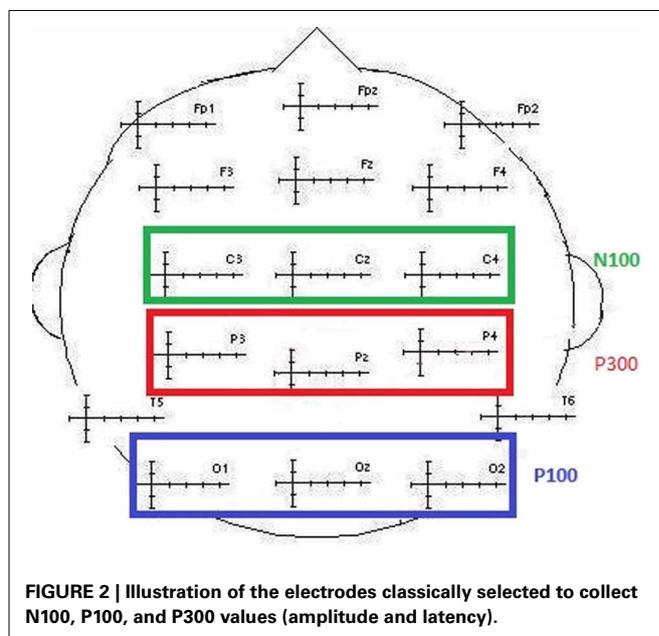
values of the maximum peak amplitudes and latencies separately for frequent and for deviant stimuli. These amplitudes were further averaged when no effect of the electrodes was found. These data were obtained from the classic electrodes used to define the P100 component, recording the maximum amplitudes (O1, Oz, and O2; maximal peak values between 90 and 160 ms). The same was true for the N100 (C3, Cz, and C4; maximal peak values between 90 and 160 ms) and the P300 (P3, Pz, and P4; maximal peak values between 250 and 600 ms) components (see **Figure 2**). The data were explored using repeated measures of analysis of variance (ANOVA) with the Greenhouse–Geisser correction applied when appropriate, using S.P.S.S. 21.0®.

## RESULTS

### BEHAVIORAL DATA

The participants' responses were 99.8% correct. Therefore, only the correct response latencies were statistically analyzed. The characteristics of the sample are shown in **Table 1**.

To examine whether the TAS-20 scores had an influence on the subjects' performances, we computed an ANCOVA on response times (RTs) for correct responses with modality (neutral, animal, and emotional) set as within-variables and inventory scores (BDI, STAI, and TAS-20) as covariates. We also performed Pearson's correlations between the different tests. It



appeared that all of the tests were intercorrelated, except for the TAS-20 (see **Table 2**). We observed a significant modality effect [ $F(2, 90) = 6.182, p = 0.003$ ] but no effect of the covariates

( $p > 0.05$ ). *Post-hoc* Bonferroni tests revealed longer reaction times for the emotional stimuli (449 ms) compared to the animal condition (396 ms), whereas the shortest RTs were observed for neutral stimuli (384 ms). Results are shown in **Tables 2** and **3**.

Although, as expected, no differences related to subclinical alexithymia were observed with the RTs, it might be interesting to compute statistical analyses based on the amplitude and latency values of the ERP components. Indeed, ERPs have been shown to detect even minor neurocognitive restrictions, which are undetectable at the behavioral level (Maurage et al., 2009; Campanella et al., 2012). These further analyses are presented below, and ERP components of interest (P100, N100, and P300) are illustrated in **Figure 3**.

**Table 1 | Means and standard deviations (in parentheses) for the whole sample ( $n = 50$ ) for age, BDI and STAI and TAS-20 psychological tests.**

$n = 50$	Mean	S.D
Age	21.02	2.095
BDI	5.68	5.464
STAI A	50.22	13.826
STAI B	48.68	12.953
TAS-20	48.74	10.583
F1 tas	18.16	5.64
F2 tas	14.98	4.48
F3 tas	17.18	10.58

**Table 2 | Correlations between the different covariates ( $n = 50$ ).**

	BDI	STAI-A	STAI-B	TAS-20
BDI				0.234
$p$				0.103
STAI-A	0.695			0.078
$p$	<0.001			0.592
STAI-B	0.788	0.766		0.176
$p$	<0.001	<0.001		0.223
TAS-20	0.234	0.078	0.176	
$p$	0.103	0.592	0.223	

**Table 3 | Means and standard deviations (in parentheses) of reaction times (ms) for “emotional” bimodal-AVE, “animal” bimodal-AVA and “neutral” bimodal task-AVN, independently of the group.**

Condition	Reaction times (ms)
AVN	384 (58.74)
AVA	396 (53.94)
AVE	449 (56.99)

## ERP DATA

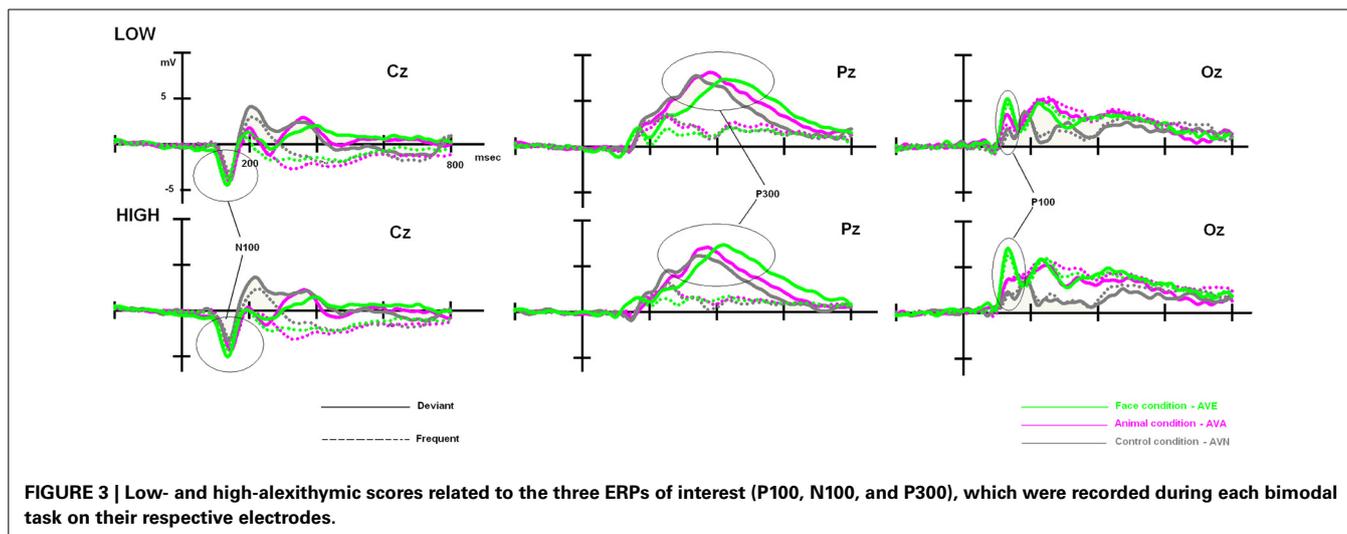
### P100

We computed an ANCOVA on P100 amplitudes, with stimulus (frequent and deviant), electrodes (O1, Oz, and O2), and modality (neutral, animal, and emotional) set as within-variables. Inventory scores (BDI, STAI, and TAS-20) served as covariates. Using this analysis, we found some significant interactions with alexithymia. In particular, we observed a tendency for modality  $\times$  TAS [ $F_{(2, 90)} = 3.094, p = 0.061$ ] and a significant interaction for stimulus  $\times$  TAS [ $F_{(1, 45)} = 4.538, p = 0.039$ ], while the triple interaction for modality  $\times$  stimulus  $\times$  TAS was not significant [ $F_{(2, 90)} = 2.37, p = \text{NS}$ ]. The other covariates (BDI and STAI) did not show significant effects. We also performed Pearson's Correlations between the covariates, according to the groups, and all tests were intercorrelated, except for the TAS-20 (see **Table 4**).

To clarify the significant stimulus  $\times$  TAS interaction, the participants ( $n = 50$ ) were split into LOW and HIGH groups based on their TAS-20 scores, and divided according to the stimulus type (frequent vs. deviant). We performed an ANCOVA on P100 amplitudes, with electrodes (O1, Oz, and O2), and modality (neutral, animal, and emotional) set as within-variables. Inventory scores (BDI, STAI, and TAS-20) served as covariates. For frequent, we obtained a non-significant result [modality  $\times$  TAS =  $F_{(2, 90)} = 1.446, p = \text{NS}$ ]. No other covariates (BDI and STAI) effects were found. For deviants, we obtained a significant modality  $\times$  TAS interaction [ $F_{(2, 90)} = 4.392, p = 0.019$ ]. Further analyses were thus computed for deviants only. We divided the sample into two groups, based on the median TAS-20 value: low alexithymic scores (TAS-20  $\leq 50$ ;  $n = 25$ ) and high-scorers (TAS-20  $> 50$ ;  $n = 25$ ). These two groups displayed alexithymia differences [ $t_{(48)} = -10.980, p < 0.001$ ] but were matched in gender [ $\chi^2_{(1)} = 0.089, p = \text{NS}$ ], which is well-known to modulate early visual ERP components during emotion processes (e.g., Proverbio et al., 2006), and age [ $t_{(48)} = 1.152, p = \text{NS}$ ]. The data are summarized in **Table 5**. Therefore, we computed a  $2 \times 3 \times 3$  ANOVA on P100 amplitude values, with group (LOW and HIGH) set as the between-subject factor and modality (AVE, AVA, and AVN) and electrode (O1, Oz, and O2).

Once again, we observed a significant effect of modality [ $F_{(2, 96)} = 26.630, p < 0.001$ ]. *Post-hoc* Bonferroni tests showed that the distinct conditions were associated with differing P100 amplitudes (independent of the group). The highest amplitude was associated with the AVE condition (7.521  $\mu\text{V}$ ,  $p < 0.001$ ), whereas the lowest amplitude was observed during the AVN condition (4.279  $\mu\text{V}$ ,  $p < 0.001$ ). AVA (5.393  $\mu\text{V}$ ) differed from both AVN ( $p = 0.037$ ) and AVE ( $p < 0.001$ ) conditions. We also observed a significant modality  $\times$  group interaction [ $F_{(2, 96)} = 3.649, p = 0.030$ ]. In order to better characterize this interaction, further analyses were conducted on the averaged amplitudes (mean amplitude of O1, Oz, and O2), as no electrode effects were found [modality  $\times$  electrode:  $F_{(4, 192)} = 1.593, p = \text{NS}$ ; modality  $\times$  electrode  $\times$  group:  $F_{(4, 192)} = 0.745, p = \text{NS}$ ].

For the LOW group, an ANOVA 3 with modality set as the within-variable revealed a significant effect of modality [ $F_{(2, 48)} = 7.181, p = 0.002$ ]. *Post-hoc* Bonferroni tests



showed that the highest amplitude was observed in the AVE condition ( $6.324 \mu\text{V}$ ), which differed significantly from that in the AVN condition ( $4.177 \mu\text{V}$ ,  $p = 0.004$ ). No differences were found between the AVE and AVA ( $5.210 \mu\text{V}$ ,  $p > 0.05$ ) conditions or the AVN and AVA conditions ( $p > 0.05$ ). For the HIGH group, the significant effect of modality was also demonstrated [ $F_{(2, 48)} = 20.319$ ,  $p < 0.001$ ]. *Post-hoc* Bonferroni tests showed that the highest amplitude was found in the AVE condition ( $8.717 \mu\text{V}$ ), which was significantly different from that observed during the AVN condition ( $4.382 \mu\text{V}$ ,  $p < 0.001$ ). In this group, the AVA amplitude ( $5.576 \mu\text{V}$ ) was also significantly lower than the AVE amplitude ( $p < 0.001$ ), suggesting a specific emotional effect. However, the AVA condition was not significantly different from the AVN ( $p > 0.05$ ). These P100-related differences between LOW- and HIGH-alexithymic scorers are illustrated in **Figure 4** and means and standard deviations are found in **Table 6**.

Taken together, no behavioral differences emerged between the two groups. Moreover, in the LOW group, we observed that the AVA and AVE deviant stimuli were handled identically (no emotional specificity), both producing higher P100 amplitudes than the neutral stimuli (AVN). However, in the HIGH group, amplitudes during the AVE condition were higher than those in the AVA and AVN conditions, suggesting that more attentional resources were required to handle the emotional stimuli. Similar analyses were conducted on P100 latencies, but no significant results were obtained ( $p > 0.05$ ).

### N100 AMPLITUDES

We analyzed the N100 amplitudes in the same manner as described above for the P100. First, we computed an ANCOVA on N100 amplitudes, with stimulus (frequent and deviant), electrodes (C3, Cz, and C4), and modality (neutral, animal, and emotional) set as within-variables. Once again, inventory scores (BDI, STAI, and TAS-20) were covariates. Through this analysis,

**Table 4 |** Correlations between the different covariates according to the groups.

	BDI	STAI-A	STAI-B	TAS-20
<b>LOW</b>				
BDI		0.624	0.816	0.227
<i>p</i>		0.001	<0.001	0.276
STAI-A			0.636	0.310
<i>p</i>			0.001	0.131
STAI-B				0.290
<i>p</i>				0.160
TAS-20	0.2274	0.310	0.290	
<i>p</i>	0.276	0.131	0.160	
<b>HIGH</b>				
BDI		0.761	0.768	0.074
<i>p</i>		<0.001	<0.001	0.726
STAI-A			0.891	-0.073
<i>p</i>			<0.001	0.727
STAI-B				-0.041
<i>p</i>				0.845
TAS-20	0.074	-0.73	-0.41	
<i>p</i>	0.726	0.727	0.845	

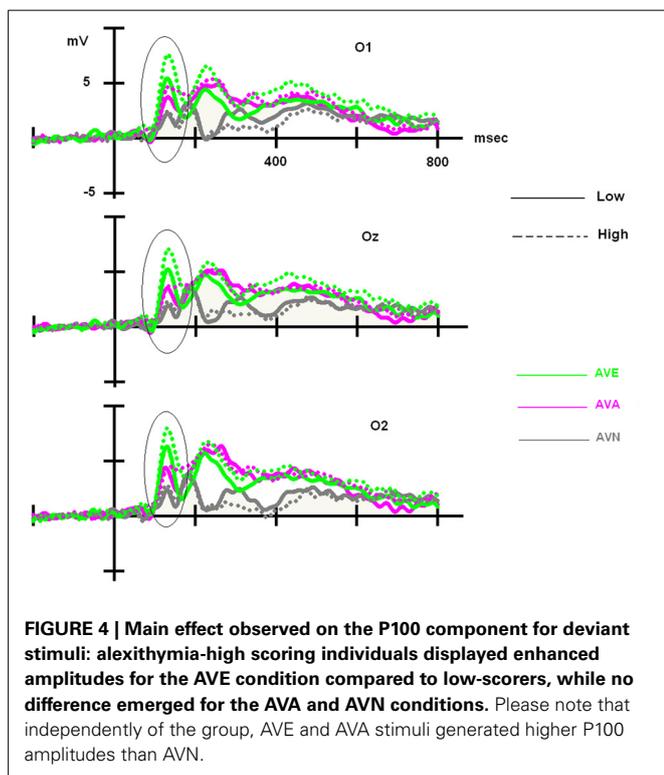
we only obtained a significant modality  $\times$  stimulus  $\times$  TAS [ $F_{(2, 90)} = 3.839$ ,  $p = 0.027$ ] interaction.

To clarify the significant modality  $\times$  stimulus  $\times$  TAS interaction, the participants were split into LOW and HIGH groups based on their TAS-20 scores. We then performed a  $2 \times 3 \times 3 \times 2$  ANOVA on N100 amplitude values, with group (LOW and HIGH) as the between-subject factor, and modality (neutral, animal, and emotional), electrode (C3, Cz, and C4), and stimulus (frequent and deviant) set as within-subject variables. We observed significant effects for modality [ $F_{(2, 96)} = 8.807$ ,  $p < 0.001$ ] and stimulus [ $F_{(1, 192)} = 44.723$ ,  $p < 0.001$ ], as well as a significant modality  $\times$  group interaction [ $F_{(2, 96)} = 5.082$ ,  $p = 0.014$ ].

**Table 5 | Means and standard deviations (in parentheses) of the low group (LOW) and high (HIGH) groups' scores for age, BDI and STAI and TAS-20 psychological tests, and reaction times (ms) for the 3 conditions.**

	Gender (F/M)	Age	BDI	STAI-A	STAI-B	TAS-20	AVN RT	AVA RT	AVE RT
LOW	(17/8)	21.36 (2.307)	4.68 (4.862)	50 (12.806)	47.16 (13.133)	<b>39.88</b> (6.24)	382 (58.373)	394 (54.05)	447 (55.421)
HIGH	(16/9)	20.68 (1.842)	6.68 (5.935)	50.44 (15.039)	50.2 (12.855)	<b>57.6</b> (5.115)	386 (60.24)	398 (54.862)	451 (59.599)
<i>p</i>	0.765	0.255	0.199	0.912	0.412	<b>&lt;0.001</b>	0.999	0.999	0.999

Significant results are indicated in bold.



As no significant result was obtained for the stimulus  $\times$  modality  $\times$  group [ $F_{(2, 96)} = 0.994, p = \text{NS}$ ] and no electrode effects were found [modality  $\times$  electrode  $\times$  stim:  $F_{(4, 192)} = 0.605, p = \text{NS}$ ; modality  $\times$  electrode  $\times$  stim  $\times$  group:  $F_{(4, 192)} = 1.198, p = \text{NS}$ ], we performed further analyses on the averaged amplitudes (mean of C3, Cz, and C4), as well as the averaged amplitudes with regard to the stimulus (mean of the frequent and deviant amplitude for AVE, AVA, and AVN). In order to better define these results, we subsequently performed a repeated measures ANOVA for each group (between factor), according to the modality (within factor). We obtained a modality effect [ $F_{(2, 96)} = 8.807, p = 0.001$ ], a modality  $\times$  group interaction [ $F_{(2, 96)} = 5.082, p = 0.014$ ], and no group effect [ $F_{(1, 48)} = 0.027, p = \text{NS}$ ]. *Post-hoc* Bonferroni tests showed that the highest amplitude was found in the AVE condition ( $-3.799 \mu\text{V}$ ) and the lowest was observed in the AVN condition ( $-3.005 \mu\text{V}, p = 0.004$ ). The AVA condition ( $-3.281 \mu\text{V}$ ) differed from the AVN ( $p = 0.002$ ) but not the AVE ( $p = 0.454$ ) condition. Regarding

the modality  $\times$  group interaction, no differences were found for the LOW group [ $F_{(1, 48)} = 1.609, p = \text{NS}$ ]. For the HIGH group, the highest amplitude was found in the AVE condition ( $-4.035 \mu\text{V}$ ) and the lowest was observed in the AVN condition ( $-2.687 \mu\text{V}, p = 0.003$ ). The AVA condition ( $-3.464 \mu\text{V}$ ) differed both from the AVN ( $p = 0.0286$ ) and AVE ( $p = 0.016$ ) conditions (see **Table 7**).

These findings suggest that more attentional resources were needed in the HIGH group to interpret “semantic” stimuli vs. neutral stimuli, which was more obvious when dealing with emotional stimuli. No differences were found in the LOW group regarding the modality of the condition, while in contrast, it appeared that more attentional resources were needed to decipher “semantic” stimuli in the HIGH group, which was even more obvious when dealing with emotional stimuli. Thus, it appeared that the N100 amplitude modulation was present, but only in the HIGH group. Similar analyses were conducted on N100 latencies but no significant results were obtained ( $p > 0.05$ ).

### P300

We performed ANCOVA on the P300 amplitudes, using stimulus (frequent and deviant), electrodes (P3, Pz, and P4), and modality (neutral, animal, and emotional) as within-variables. Inventory scores (BDI, STAI, and TAS-20) served as covariates. No significant results were obtained for either the amplitudes or latencies ( $p > 0.05$ ; see **Table 6** for values), which was consistent with the absence of behavioral effects. Indeed, P300 values are known to correlate with RTs, functionally referring to response-related stages (Polich, 2007).

### Complementary analysis

**Rejected Trials.** We performed a  $3 \times 2 \times 2$  ANOVA on the rejected trials. Modality and stimulus were set as within-variables and the group served as the between factor. We observed no group interactions through this analysis: modality  $\times$  group [ $F_{(2, 96)} = 0.993, p > 0.05$ ] and modality  $\times$  group  $\times$  stimulus [ $F_{(2, 96)} = 2.537, p > 0.05$ ]. These data indicate that the same number of “artifact” trials was rejected in each group.

**Correlations.** The TAS-20 is composed of 3 subscales: (1) difficulty identifying feelings, (2) difficulty describing feelings, and (3) externally-oriented thinking (Bagby et al., 1994a,b). We wanted to test whether some of these subscales might be related to specifically the P100 and/or N100 modulations that we observed in the AVE condition. To examine this possibility, we performed

**Table 6 | Means and standard deviations (in parentheses) of P100, N100 and P300 amplitudes for frequent and deviant stimuli, in the “emotional” bimodal-AVE, “animal” bimodal-AVA and “neutral” bimodal task-AVN, for each group.**

Condition	P100 Fqt	P100 Dev	N100 Fqt	N100 Dev	P300 Fqt	P300 Dev
<b>LOW</b>						
AVN	2.815 (2.863)	<b>4.177 (2.503)</b>	-2.977 (1.843)	-3.671 (2.352)	2.342 (1.434)	6.552 (2.576)
AVA	4.146 (2.917)	5.21 (3.287)	-2.896 (1.635)	-3.300 (1.781)	3.158 (2.008)	7.087 (3.840)
AVE	5.648 (3.414)	<b>6.324 (2.809)</b>	-3.269 (1.393)	-3.856 (1.464)	2.583 (1.422)	6.401 (2.758)
<b>HIGH</b>						
AVN	3.798 (2.552)	<b>4.382 (3.103)</b>	<b>-2.482 (1.260)</b>	<b>-2.893 (2.137)</b>	1.549 (1.431)	5.886 (2.717)
AVA	5.098 (3.339)	<b>5.576 (3.974)</b>	<b>-3.104 (1.267)</b>	<b>-3.824 (1.569)</b>	2.782 (2.657)	7.157 (2.730)
AVE	7.706 (4.971)	<b>8.718 (5.389)</b>	<b>-3.720 (1.811)</b>	<b>-4.350 (1.855)</b>	2.783 (1.844)	6.935 (2.968)

Significant amplitudes are indicated in bold.

**Table 7 | Means and standard deviations (in parentheses) of N100 amplitudes for frequent and deviant stimuli together, in the “emotional” bimodal-AVE, “animal” bimodal-AVA and “neutral” bimodal task-AVN, for each group.**

Modality	Mean
<b>LOW</b>	
AVN	-3.324 (2.07)
AVA	-3.098 (1.58)
AVE	-3.563 (1.35)
<b>HIGH</b>	
AVN	<b>-2.687 (1.64)</b>
AVA	<b>-3.464 (1.33)</b>
AVE	<b>-4.035 (1.79)</b>

Significant amplitudes are indicated in bold.

Pearson's correlations based on the 3 TAS-20 subscores and the P100/N100 amplitude means obtained from the AVE stimuli. Results are shown in **Table 8**.

For the LOW group, a tendency for a negative correlation ( $r = -0.362$ ,  $p = 0.075$ ) was seen between the externally-oriented thinking subscale (F3) and P100 amplitudes for the AVE deviants (i.e., the higher the F3 score, the lower the P100 amplitude). For the HIGH scorers, F3 was significantly correlated ( $r = 0.458$ ,  $p = 0.021$ ) with the P100 amplitude for the AVE deviants (i.e., the higher the F3 score, the higher the P100 amplitude). The F3 was also negatively correlated ( $r = -0.484$ ,  $p = 0.014$ ) with the N100 amplitude for the AVE frequent stimuli (i.e., the higher the F3 score, the lower the N100 amplitude). Thus, it seems that the influence of the TAS-20 can be mostly explained by the F3 factor (externally-oriented thinking), at least with regard to the HIGH group for the P100 amplitude in the AVE deviants ( $R^2 = 20.9\%$ ) and the N100 amplitudes in the AVE frequent stimuli ( $R^2 = 23.4\%$ ).

## DISCUSSION

In this study we investigated whether subclinical alexithymia could disclose particular emotional effects through the use of bimodal oddball tasks. In order to achieve this, we used a more

**Table 8 | Pearson correlations and  $R^2$  between the 3 sub-scores on the TAS-20.**

		P100AVE dev	N100AVE dev	N100AVE Freq
<b>LOW</b>				
F1	Pearson correlation	0.115	0.226	0.209
	<i>p</i>	0.585	0.278	0.317
F2	Pearson correlation	-0.003	0.342	0.276
	<i>p</i>	0.989	0.094	0.182
F3	Pearson correlation	<b>-0.362</b>	0.269	0.259
	<i>p</i>	<b>0.075</b>	0.193	0.211
<b>HIGH</b>				
F1	Pearson correlation	-0.260	0.323	0.343
	<i>p</i>	0.209	0.115	0.093
F2	Pearson correlation	-0.065	-0.138	-0.039
	<i>p</i>	0.757	0.509	0.853
F3	Pearson correlation	<b>0.458</b>	-0.232	<b>-0.484</b>
		<b>(<math>R^2 = 20.9\%</math>)</b>		<b>(<math>R^2 = 23.4\%</math>)</b>
	<i>p</i>	<b>0.021</b>	0.264	<b>0.014</b>

F1, difficulty identifying feelings; F2, difficulty describing feelings; F3, externally-oriented thinking and the P100/N100 amplitude means in the AVE situation for frequent and deviant stimulations. Significant correlations are indicated in bold, tendencies in bold and italic.

sensitive paradigm than the classic unimodal oddball task, namely an audio-visual oddball task. The use of this tool already allowed us to index the subtlest subclinical differences (anxio-depressive tendencies) at the electrophysiological level, not revealed when unimodal stimuli were employed (Campanella et al., 2010, 2012). Through three different kinds of bimodal oddball tasks (emotional, animal, and neutral), we were able to explore the effect of high and low subclinical alexithymia on P100, N100, and P300 components.

Previous studies (Mann et al., 1995; Galderisi et al., 2008; Vermeulen et al., 2008) demonstrated that no behavioral differences could be detected when low- and high-alexithymic scorers were confronted with cognitive tasks, including emotional conditions. These findings suggest that at a subclinical level, an emotional processing deficit could remain behaviorally invisible, stressing the importance of using sensitive imaging tools, such as

ERPs. Indeed, even in the absence of behavioral modifications, subtle alterations in ERPs have been shown to index minor cognitive restrictions (Rugg and Coles, 1995). In this study, examining the interaction between the TAS-20 values and the distinct bimodal tasks revealed specific ERP modulations. We found that high-alexithymic scorers presented greater amplitudes for early perceptive ERP components as compared with low-scorers when confronted with emotional oddball tasks (AVE). Moreover, our findings confirmed that no specific emotional effect was driven by depressive (BDI) and anxious (STAI) tendencies, as shown in Campanella et al. (2012).

We compared 3 kinds of bimodal stimuli [emotional (AVE), animal (AVA), and neutral (AVN)]. We observed that high-scorers displayed an enhanced visual P100 component in response to deviant stimuli in the AVE condition, which did not occur in the AVN and AVA conditions. However, no difference between the AVE and AVA conditions emerged from the low-scorers. High-scorers also showed enhanced N100 amplitudes for both frequent and deviant stimuli presented in the AVE condition, whereas again no difference was observed when analyzing low-scorers. Furthermore, independently of the group, AVN was the condition that generated the lowest amplitudes for P100 and N100, suggesting that semantic conditions (AVA and AVE) require more attentional resources to be processed. Additionally, in both groups emotional stimuli systematically exhibited higher amplitudes. This result can be interpreted by the emotional salience of the stimuli in terms of survival, reproduction, and procreation, compared to other kinds of stimulations (Schupp et al., 2007).

Overall, the alexithymia-high group demonstrated modified processing of the emotional condition, which translated into higher amplitudes in the early visual treatment and extended to the early auditory processing of the stimuli. These results are in line with findings by Wehmer et al. (1995) and Franz et al. (2004) about unimodal emotional situations and supports the idea that, contrary to behavioral data suggesting an incapacity to detect and process emotions, ERPs data suggest that these subjects are not blind, at a perceptive level, to emotional content, but may request some kind of “hyperactivation” (physiological or electrophysiological) in order to process this information and correctly perform. Thus, they are able to experience emotions, but at the cost of a modified central processing of it, as exhibited by the ERPs modulations. To compensate for their emotional disturbances, subclinical alexithymic subjects actually allocate abnormal attentional resources to successfully perform tasks. For this reason, no behavioral differences (reaction times and correct response rate) could be identified. Indeed, several studies have indicated that emotional processing is not entirely automatic but competes with attentional demands. Thus, emotional stimuli reciprocally influence attentional processes (Pessoa et al., 2002, 2005; Bradley et al., 2003; Sabatinelli et al., 2005). Therefore, alexithymic individuals can perceive emotional stimuli and are not “blind” to emotional information; however, they require a deeper, more intense cognitive process for processing the stimuli. This could be due to doubts in their minds with regard to the meaning of the emotion presented, or perhaps they are less

interested in emotions (less familiar with them) (Grynberg et al., 2012).

Interestingly, P100 modulations were only present for deviant stimulations, with some differences occurring between the high- and low-scoring groups. For the alexithymia-low group, the only distinction that could be made was between the neutral condition and the emotional one, suggesting that the “semantic” content of the stimuli required deeper processing (regardless of whether the stimuli was emotional or animal). Indeed, high-complexity pictures (compared to low) are known to require additional attentional resources for processing details (Bradley et al., 2007). However, for the high group, an “emotional effect” occurred (i.e., the emotional condition was harder to deal with because more attentional resources were necessary to interpret or decipher the emotional content). This effect could be indexed by larger bimodal P100 amplitudes in response to emotional conditions, compared to the animal and neutral stimuli, as has been previously described in the literature for unimodal visual situations (Pollatos and Gramann, 2011; Singhal et al., 2012). Therefore, our data suggest that a very early attention modulation occurs in response to emotional stimuli, as high-alexithymic-scorers displayed larger amplitudes to pictures that were more “complex to process.” Likewise, processing was easier for stimuli that were lower in affective relevance (AVA and AVN) and did not require selective processing (Pastor et al., 2008).

This deficit extended to the auditory domain, but only for the alexithymia-high group (for both frequent and deviant stimulations). Indeed, the N100 amplitudes were gradually larger from the AVN to AVE conditions, suggesting a specific “emotional effect.” These enhanced N100 amplitudes appear to confirm the idea that deficits in emotional prosody exist in alexithymic subjects, at a behavioral level (Schäfer et al., 2007). For instance, when Goerlich et al. (2011) used emotional prosody (music and words with emotional connotations), they did not observe any behavioral differences; however, using affective categorization of happy and sad prosody and music targets they could elicit a negative correlation between alexithymia and the N400 amplitude. Another recent study of Goerlich et al. (2012) also demonstrated larger N100 amplitudes in high TAS-20 scorers in response to deviant emotional prosodies, whereas no behavioral differences were found between the low- and high-alexithymic groups. They also found that alexithymia is related to generally blunted neural responses to speech prosody (Goerlich-Dobre et al., 2013). Our results are consistent with the theory that subclinical alexithymic subjects exhibit attenuated basic emotional processing, involving a reduced early detection of emotional salience that requires more attention for detection of changes in emotional acoustic cues.

Another interesting finding of this study was that only deviant stimuli led to changes in the visual P100 component, whereas both frequent and deviant cues led to N100 modulations. A possible interpretation of this result is that, as mentioned by Joassin et al. (2004), the quantity of available information at stimulus onset is not equal for faces and voices. It has been proposed that auditory information unfolds over time for voices and that this time-based asynchrony can lead to an interference effect

(Calvert et al., 2001). This might be why high-scoring require more resources to process general auditory stimulations, particularly for the more complex “emotional” condition (AVE). Further studies should be designed to investigate this topic in-depth.

At later stages of the information processing stream, no differences in P300 modulations were observable between low- and high-alexithymic individuals. Since P300 functionally corresponds to response-related stages (Polich, 2007), these data are in perfect agreement with the absence of behavioral modifications. This suggests that subclinical alexithymic subjects are able to compensate for their emotional deficits in later decisional levels, as they performed the task equally to low-scoring. However, this “normal” performance by alexithymic individuals requires more visual and auditory attentional resources devoted at early stages of the information processing stream. Further studies analyzing alexithymic subjects with clinical TAS-20 values (>61) should be performed in order to verify whether early deficits would extend to later ERP components in these more extreme cases. Indeed, it would be interesting to apply our bimodal emotional task in clinical populations, as the early emotional alterations (P100 and N100) could constitute a marker of predisposition for the development of later clinical alexithymia. Thus, these data may have an important clinical relevance. If the behavioral deficits presented by clinical alexithymic individuals have an attentional origin, then cognitive remediation procedures targeting the deeper attentional processing of emotions could be envisioned.

Despite the potential clinical relevance of the data, the limitations of our study must be outlined. First, our sample size was modest (50 participants). Second, the TAS-20 scale is an auto-evaluative questionnaire, which could be paradoxical for individuals suffering from difficulties in identifying and describing their own emotions. To address this important issue, future studies could employ observer evaluations, such as those conducted in the Toronto Structured Interview for Alexithymia (TSIA, Bagby et al., 2006) or the Observer Alexithymia Scale (OAS, Haviland et al., 2000). Third, we made the subjects fill out the questionnaires (BDI, STAI, and TAS-20 assessing emotions) before the ERP session, so it is possible that their attention toward emotional processing could have influenced the way that they processed the emotional condition. Fourth, we only used negative (sad) and neutral stimuli in this study; however, alexithymia has been described to relate to a large variety of basic emotions (e.g., sadness, happiness, fear, anger, disgust, and surprise) (Jessimer and Markham, 1997; Prkachin et al., 2009), as well as neutral stimuli (Lane et al., 2000). Thus, it would be interesting to link these particular difficulties with specific ERP modulations. Finally, it is possible that alexithymia could negatively impact performance in the fast processing of emotional information, as described in our paradigm. With regard to this, a study by Parker et al. (2005) used a signal-detection design that allowed participants to judge neutral or negative facial expressions under slow and quick presentation conditions. Indeed, the alexithymia component of difficulty in describing feelings was inversely correlated with the capacity to detect negative emotional expressions in the rapid condition.

Our findings support the utility of a more sophisticated odd-ball (i.e., an audio-visual design) for revealing subtle subclinical differences in healthy populations. Our results are in line with previous reports that have suggested atypical attentional resource allocation in alexithymic individuals during negative and neutral bimodal stimuli. Notably, this allocation would allow these subjects to compensate a perceptive deficit (auditory and visual) and successfully perform tasks. Therefore, we highly recommend the use of bimodal stimulations for future studies, as they allow the extraction of more accurate knowledge about the cognitive processing of emotions (Maurage and Campanella, 2013). This is necessary to create and develop an adequate rehabilitation plan, as alexithymic patients unsuccessfully respond to psychological treatments that center on introspection, emotional consciousness, and/or close alliances with therapists (Lumley et al., 2007). Moreover, they rarely engage in treatment recommendations (Ogrodniczuk et al., 2011). As a matter of fact, our correlational results revealed that the P100 and N100 modulations were mostly associated with the operational thinking subscale of the TAS-20 (e.g., prefer talking about daily activities rather than feelings). As mentioned before, recent ERPs data, including the present study, reveal that alexithymics, in any case subclinical subjects, are finally not blind to emotion, and are able to detect and process it, so the whole concept of alexithymia is challenged. As suggested by Franz et al. (2004), perhaps alexithymics avoid the processing or the expression of affective states. Nevertheless, the externally-oriented thinking style seems to remain a stable feature of the disorder, and therapeutic interventions could successfully target this characteristic. Knowing that alexithymics are more likely to seek healthcare (Joukamaa et al., 1996), this information could be useful for engaging alexithymic individuals in externally-focused interventions that might guarantee greater adherence to structured exercises and behavioral counseling. From this perspective, alexithymia could be redefined as a deficit in emotional aptitudes that could be learned or trained, among other things, through treatment (Lumley et al., 2007). For example, Levant (2001) developed a cognitive-behavioral method in which alexithymics were taught to learn emotional vocabulary, label affective situations, observe their own symptoms, and connect emotional labels to their symptoms. It is important that future longitudinal studies investigate the long-term effects of such cognitive-behavioral techniques in alexithymia treatment. In this regard, it would be interesting to evaluate whether early ERP amplitudes to emotional stimuli decreased over time during these follow-up studies.

## ACKNOWLEDGMENTS

Salvatore Campanella is Research Associate at the Belgian Fund for Scientific Research (F.N.R.S., Belgium).

## REFERENCES

- Alho, K., Teder, W., Lavikainen, J., and Näätänen, R. (1994). Strongly focused attention and auditory event-related potentials. *Biol. Psychol.* 38, 73–90. doi: 10.1016/0301-0511(94)90050-7
- Allison, T., Puce, A., Spencer, D. D., and McCarthy, G. (1999). Electrophysiological studies of human face perception. I: potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cereb. Cortex* 9, 415–430. doi: 10.1093/cercor/9.5.415

- Bagby, R. M., Parker, J. D., and Taylor, G. J. (1994a). The twenty-item Toronto Alexithymia Scale—I. Item selection and cross-validation of the factor structure. *J. Psychosom. Res.* 38, 23–32. doi: 10.1016/0022-3999(94)90005-1
- Bagby, R. M., Parker, J. D., and Taylor, G. J. (1994b). The twenty-item Toronto Alexithymia Scale—II. Convergent, discriminant, and concurrent validity. *J. Psychosom. Res.* 38, 23–32. doi: 10.1016/0022-3999(94)90005-1
- Bagby, R. M., and Taylor, G. J. (1997). “Affect dysregulation and alexithymia,” in *Disorders of Affect Regulation: Alexithymia in Medical and Psychiatric Illness*, eds G. J. Taylor, R. M. Bagby, and J. D. Parker (Cambridge: Cambridge University Press). 26–45. doi: 10.1017/CBO9780511526831.005
- Bagby, R. M., Taylor, G. J., Parker, J. D. A., and Dickens, S. E. (2006). The development of the Toronto Structured Interview for Alexithymia: item selection, factor structure, reliability and concurrent validity. *Psychother. Psychosom.* 75, 25–39. doi: 10.1159/000089224
- Beck, A. T., and Beck, R. W. (1972). Screening depressed patients in family practice. A rapid technic. *Postgrad. Med.* 52, 81–85.
- Beck, A. T., and Steer, R. A. (1987). *BDI, Beck Depression Inventory: Manual*. San Antonio, TX: Psychological Corporation.
- Bermond, B., Righart, R., Ridderinkhof, K. R., and Moormann, P. P. (2008). Alexithymia and the brain potential P300. *Neth. J. Psychol.* 64, 65–77. doi: 10.1007/BF03076408
- Berthoz, S., Consoli, S., Perez-Diaz, F., and Jouvent, R. (1999). Alexithymia and anxiety: compounded relationships? A psychometric study. *Eur. Psychiatry* 14, 372–378. doi: 10.1016/S0924-9338(99)00233-3
- Bertrand, O., Perrin, F., and Pernier, J. (1985). A theoretical justification of the average reference in topographic evoked potential studies. *Electroencephalogr. Clin. Neurophysiol.* 62, 462–464. doi: 10.1016/0168-5597(85)90058-9
- Bookheimer, S. (2002). Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annu. Rev. Neurosci.* 25, 151–188. doi: 10.1146/annurev.neuro.25.112701.142946
- Bradley, M. M., Hamby, S., Löw, A., and Lang, P. J. (2007). Brain potentials in perception: picture complexity and emotional arousal. *Psychophysiology* 44, 364–373. doi: 10.1111/j.1469-8986.2007.00520.x
- Bradley, M. M., Sabatinelli, D., Lang, P. J., Fitzsimmons, J. R., King, W., and Desai, P. (2003). Activation of the visual cortex in motivated attention. *Behav. Neurosci.* 117, 369–380. doi: 10.1037/0735-7044.117.2.369
- Bruchon-Schweitzer, M., and Paulhan, I. (1993). *Manuel de l'inventaire d'anxiété état-trait forme Y (STAI-Y)*. Paris: Editions du Centre de Psychologie Appliquée.
- Bruder, G. E., Kroppmann, C. J., Kayser, J., Stewart, J. W., McGrath, P. J., and Tenke, C. E. (2009). Reduced brain responses to novel sounds in depression: P3 findings in a novelty oddball task. *Psychiatry Res.* 170, 218–223. doi: 10.1016/j.psychres.2008.10.023
- Calvert, G. A., Hansen, P. C., Iversen, S. D., and Brammer, M. J. (2001). Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *Neuroimage* 14, 427–438. doi: 10.1006/nimg.2001.0812
- Campanella, S. (2013). Why it is time to develop the use of cognitive event-related potentials in the treatment of psychiatric diseases. *Neuropsychiatr. Dis. Treat.* 9, 1835–1845. doi: 10.2147/NDT.S53687
- Campanella, S., and Belin, P. (2007). Integrating face and voice in person perception. *Trends Cogn. Sci. (Regul. Ed.)* 11, 535–543. doi: 10.1016/j.tics.2007.10.001
- Campanella, S., Bruyer, R., Froidbise, S., Rossignol, M., Joassin, F., Kornreich, C., et al. (2010). Is two better than one? A cross-modal oddball paradigm reveals greater sensitivity of the P300 to emotional face-voice associations. *Clin. Neurophysiol.* 121, 1855–1862. doi: 10.1016/j.clinph.2010.04.004
- Campanella, S., Delle-Vigne, D., Kornreich, C., and Verbanck, P. (2012). Greater sensitivity of the P300 component to bimodal stimulation in an event-related potentials oddball task. *Clin. Neurophysiol.* 123, 937–946. doi: 10.1016/j.clinph.2011.10.041
- Campanella, S., Gaspard, C., Debatisse, D., Bruyer, R., Crommelinck, M., and Guerit J-M. (2002). Discrimination of emotional facial expressions in a visual oddball task: an ERP study. *Biol. Psychol.* 59, 171–186. doi: 10.1016/S0301-0511(02)00005-4
- Campanella, S., Montedoro, C., Strel, E., Verbanck, P., and Rosier, V. (2006). Early visual components (P100, N170) are disrupted in chronic schizophrenic patients: an event-related potentials study. *Neurophysiologie clinique/Clinical Neurophysiol.* 36, 71–78. doi: 10.1016/j.neucli.2006.04.005
- Campanella, S., Vanhoolandt, M. E., and Philippot, P. (2005). Emotional deficit in subjects with psychopathic tendencies as assessed by the Minnesota Multiphasic Personality Inventory-2: an event-related potentials study. *Neurosci. Lett.* 373, 26–31. doi: 10.1016/j.neulet.2004.09.061
- Collet, L., and Cottraux, J. (1986). Inventaire abrégé de la dépression de Beck (13 items). Etude de la validité concurrente avec les échelles de Hamilton et de ralentissement de Widlöcher. *L'encéphale* 12, 77–79.
- Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., et al. (2009). Event-related potentials in clinical research: guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clin. Neurophysiol.* 120, 1883–1908. doi: 10.1016/j.clinph.2009.07.045
- Eichmann, M., Kugel, H., and Suslow, T. (2008). Difficulty identifying feelings and automatic activation in the fusiform gyrus in response to facial emotion. *Percept. Mot. Skills.* 107, 915–922. doi: 10.2466/PMS.107.7.915-922
- Ekman, P., and Friesen, W. V. (1976). *Pictures of Facial Affect Consulting*. Palo Alto, CA: Psychologists Press.
- Engelien, A., Schulz, M., Ross, B., Arolt, V., and Pantev, C. (2000). A combined functional *in vivo* measure for primary and secondary auditory cortices. *Hear. Res.* 148, 153–160. doi: 10.1016/S0378-5955(00)00148-9
- Foxe, J. J., Doniger, G. M., and Javitt, D. C. (2001). Early visual processing deficits in schizophrenia: impaired P1 generation revealed by high-density electrical mapping. *Neuroreport* 12, 3815–3820. doi: 10.1097/00001756-200112040-00043
- Franz, M., Popp, K., Schaefer, R., Sitte, W., Schneider, C., Hardt, J., et al. (2008). Alexithymia in the German general population. *Soc. Psychiatry Psychiatr. Epidemiol.* 43, 54–62. doi: 10.1007/s00127-007-0265-1
- Franz, M., Schaefer, R., Schneider, C., Sitte, W., and Bachor, J. (2004). Visual event-related potentials in subjects with alexithymia: modified processing of emotional aversive information? *Am. J. Psychiatry.* 161, 728–735. doi: 10.1176/appi.ajp.161.4.728
- Galderisi, S., Mancuso, F., Mucci, A., Garramone, S., Zamboli, R., and Maj, M. (2008). Alexithymia and cognitive dysfunctions in patients with panic disorder. *Psychother. Psychosom.* 77, 182–188. doi: 10.1159/000119738
- Goerlich-Dobre, K. S., Witteman, J., Schiller, N. O., van Heuven, V. J., Aleman, A., and Martens, S. (2013). Blunted feelings: alexithymia is associated with a diminished neural response to speech prosody. *Soc. Cogn. Affect. Neurosci.* doi: 10.1093/scan/nst075. [Epub ahead of print].
- Goerlich, K. S., Aleman, A., and Martens, S. (2012). The sound of feelings: electrophysiological responses to emotional speech in alexithymia. *PLoS ONE* 7:e36951. doi: 10.1371/journal.pone.0036951
- Goerlich, K. S., Witteman, J., Aleman, A., and Martens, S. (2011). Hearing feelings: affective categorization of music and speech in alexithymia, an ERP study. *PLoS ONE* 6:e19501. doi: 10.1371/journal.pone.0019501
- Grynberg, D., Chang, B., Corneille, O., Maurage, P., Vermeulen, N., Berthoz, S., et al. (2012). Alexithymia and the processing of emotional facial expressions (EFEs): systematic review, unanswered questions and further perspectives. *PLoS ONE* 7:e42429. doi: 10.1371/journal.pone.0042429
- Hansenne, M. (2000). The p300 cognitive event-related potential. I. Theoretical and psychobiologic perspectives. *Neurophysiol. Clin.* 30, 191–210. doi: 10.1016/S0987-7053(00)00223-9
- Haviland, M. G., Warren, W. L., and Riggs, M. L. (2000). An observer scale to measure alexithymia. *Psychosomatics* 41, 385–392. doi: 10.1176/appi.psy.41.5.385
- Heinze, H. J., and Mangun, G. R. (1995). Electrophysiological signs of sustained and transient attention to spatial locations. *Neuropsychologia* 33, 889–908. doi: 10.1016/0028-3932(95)00023-V
- Jessen, S., Obleser, J., and Kotz, S. A. (2012). How bodies and voices interact in early emotion perception. *PLoS ONE* 7:e36070. doi: 10.1371/journal.pone.0036070
- Jessimer, M., and Markham, R. (1997). Alexithymia: a right hemisphere dysfunction specific to recognition of certain facial expressions? *Brain Cogn.* 34, 246–258. doi: 10.1006/brcg.1997.0900
- Joassin, F., Maurage, P., Bruyer, R., Crommelinck, M., and Campanella, S. (2004). When audition alters vision: an event-related potential study of the cross-modal interactions between faces and voices. *Neurosci. Lett.* 369, 132–137. doi: 10.1016/j.neulet.2004.07.067
- Joukamaa, M., Karlsson, H., Sholman, B., and Lehtinen, V. (1996). Alexithymia and psychological distress among frequent attendance patients in health care. *Psychother. Psychosom.* 65, 199–202. doi: 10.1159/000289075

- Kätsyri, J., Saalasti, S., Tiippana, K., von Wendt, L., and Sams, M. (2008). Impaired recognition of facial emotions from low-spatial frequencies in Asperger syndrome. *Neuropsychologia* 46, 1888–1897. doi: 10.1016/j.neuropsychologia.2008.01.005
- Lane, R. D., Sechrest, L., Reidel, R., Weldon, V., Kaszniak, A., and Schwartz, G. E. (1996). Impaired verbal and nonverbal emotion recognition in alexithymia. *Psychosom. Med.* 58, 203–210.
- Lane, R. D., Sechrest, L., Riedel, R., Shapiro, D. E., and Kaszniak, A. W. (2000). Pervasive emotion recognition deficit common to alexithymia and the repressive coping style. *Psychosom. Med.* 62, 492–501.
- Levant, R. F. (2001). “Desperately seeking language: understanding, assessing, and treating normative male alexithymia,” in *The New Handbook of Psychotherapy and Counseling with Men: A Comprehensive Guide to Settings, Problems and Treatment Approaches*, eds G. R. Brooks, and G. E. Good (San Francisco, CA: Jossey-Bass/Pfeiffer), 424–443.
- Liu, T., Pinheiro, A., Zhao, Z., Nestor, P. G., McCarley, R. W., and Niznikiewicz, M. A. (2012). Emotional cues during simultaneous face and voice processing: electrophysiological insights. *PLoS ONE* 7:e31001. doi: 10.1371/journal.pone.0031001
- Loas, G., Otmani, O., Verrier, A., Fremaux, D., and Marchand, M. P. (1996). Factor analysis of the French version of the 20-Item Toronto Alexithymia Scale (TAS-20). *Psychopathology* 29, 139–144.
- Luminet, O., Bagby, R. M., and Taylor, G. J. (2001). An evaluation of the absolute and relative stability of alexithymia in patients with major depression. *Psychother. Psychosom.* 70, 254–260. doi: 10.1159/000056263
- Lumley, M. A., Neely, L. C., and Burger, A. J. (2007). The assessment of alexithymia in medical settings: implications for understanding and treating health problems. *J. Pers. Assess.* 89, 230–246. doi: 10.1080/00223890701629698
- Mann, L. S., Wise, T. N., Trinidad, A., and Kohanski, R. (1994). Alexithymia, affect recognition, and the five-factor model of personality in normal subjects. *Psychol. Rep.* 74, 563–567. doi: 10.2466/pr0.1994.74.2.563
- Mann, L. S., Wise, T. N., Trinidad, A., and Kohanski, R. (1995). Alexithymia, affect recognition, and five factors of personality in substance abusers. *Percept. Mot. Skills.* 81, 35–40. doi: 10.2466/pms.1995.81.1.35
- Mantani, T., Okamoto, Y., Shirao, N., Okada, G., and Yamawaki, S. (2005). Reduced activation of posterior cingulate cortex during imagery in subjects with high degrees of alexithymia: a functional magnetic resonance imaging study. *Biol. Psychiatry* 57, 982–990. doi: 10.1016/j.biopsych.2005.01.047
- Martínez, A., Anllo-Vento, L., Sereno, M. I., Frank, L. R., Buxton, R. B., Dubowitz, D. J., et al. (1999). Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nat. Neurosci.* 2, 364–369. doi: 10.1038/7274
- Maurage, P., and Campanella, S. (2013). Experimental and clinical usefulness of crossmodal paradigms in psychiatry: an illustration from emotional processing in alcohol-dependence. *Front. Hum. Neurosci.* 7:394. doi: 10.3389/fnhum.2013.00394
- Maurage, P., Campanella, S., Philippot, P., de Timary, P., Constant, E., Gauthier, S., et al. (2008). Alcoholism leads to early perceptive alterations, independently of comorbid depressed state: an ERP study. *Neurophysiol. Clin.* 38, 83–97. doi: 10.1016/j.neucli.2008.02.001
- Maurage, P., Joassin, F., Philippot, P., and Campanella, S. (2007b). A validated battery of vocal emotional expressions. *Neuropsychological Trends.* 2:63e74. Available online at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.112.8260&rep=rep1&type=pdf#page=59>
- Maurage, P., Pesenti, M., Philippot, P., Joassin, F., and Campanella, S. (2009). Latent deleterious effects of binge drinking over a short period of time revealed only by electrophysiological measures. *J. Psychiatry Neurosci.* 34, 111–118. Available online at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2647570/>
- Maurage, P., Philippot, P., Verbanck, P., Noel, X., Kornreich, C., Hanak, C., et al. (2007a). Is the P300 deficit in alcoholism associated with early visual impairments (P100, N170)? *An oddball paradigm. Clin Neurophysiol.* 118, 633–644. doi: 10.1016/j.clinph.2006.11.007
- Mejias, S., Rossignol, M., Debatisse, D., Streel, E., Servais, L., Guérit, J. M., et al. (2005). Event-related potentials (ERPs) in ecstasy (MDMA) users during a visual oddball task. *Biol. Psychol.* 69, 333–352. doi: 10.1016/j.biopsycho.2004.11.010
- Mobascher, A., Brinkmeyer, J., Warbrick, T., Wels, C., Wagner, M., Gründer, G., et al. (2010). The P300 event-related potential and smoking—a population-based case-control study. *Int. J. Psychophysiol.* 77, 166–175. doi: 10.1016/j.ijpsycho.2010.06.002
- Montebarocci, O., Surcinelli, P., Rossi, N., and Baldaro, B. (2011). Alexithymia, verbal ability and emotion recognition. *Psychiatr. Q.* 82, 245–252. doi: 10.1007/s11126-010-9166-7
- Ogrodniczuk, J. S., Piper, W. E., and Joyce, A. S. (2011). Effect of alexithymia on the process and outcome of psychotherapy: a programmatic review. *Psychiatry Res.* 190, 43–48. doi: 10.1016/j.psychres.2010.04.026
- Pandey, R., and Mandal, M. K. (1997). Processing of facial expressions of emotion and alexithymia. *Br. J. Clin. Psychol.* 36(Pt 4), 631–633. doi: 10.1111/j.2044-8260.1997.tb01269.x
- Parker, P. D., Prkachin, K. M., and Prkachin, G. C. (2005). Processing of facial expressions of negative emotion in alexithymia: the influence of temporal constraint. *J. Pers.* 73, 1087–1107. doi: 10.1111/j.1467-6494.2005.00339.x
- Pastor, M. C., Bradley, M. M., Löw, A., Versace, F., Moltó, J., and Lang, P. J. (2008). Affective picture perception: emotion, context, and the late positive potential. *Brain Res.* 1189, 145–151. doi: 10.1016/j.brainres.2007.10.072
- Peschard, V., Philippot, P., Joassin, F., and Rossignol, M. (2013). The impact of the stimulus features and task instructions on facial processing in social anxiety: an ERP investigation. *Biol. Psychol.* 93, 88–96. doi: 10.1016/j.biopsycho.2013.01.009
- Pessoa, L., McKenna, M., Gutierrez, E., and Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *Proc. Natl. Acad. Sci. U.S.A.* 99, 11458–11463. doi: 10.1073/pnas.172403899
- Pessoa, L., Padmala, S., and Morland, T. (2005). Fate of unattended fearful faces in the amygdala is determined by both attentional resources and cognitive modulation. *Neuroimage.* 28, 249–255. doi: 10.1016/j.neuroimage.2005.05.048
- Pogarell, O., Mulert, C., and Hegerl, U. (2007). Event-related potentials in psychiatry. *Clin. EEG Neurosci.* 38, 25–34. doi: 10.1177/155005940703800108
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019
- Polich, J., and Criado, J. R. (2006). Neuropsychology and neuropharmacology of P3a and P3b. *Int. J. Psychophysiol.* 60, 172–185. doi: 10.1016/j.ijpsycho.2005.12.012
- Pollatos, O., and Gramann, K. (2011). Electrophysiological evidence of early processing deficits in alexithymia. *Biol. Psychol.* 87, 113–121. doi: 10.1016/j.biopsycho.2011.02.016
- Power, M., and Dalgleish, T. (1997). *Cognition and Emotion: From Order to Disorder*. Hove: Psychology Press.
- Prkachin, G. C., Casey, C., and Prkachin, K. M. (2009). Alexithymia and perception of facial expressions of emotion. *Pers. Individ. Dif.* 46, 412–417. doi: 10.1016/j.paid.2008.11.010
- Proverbio, A. M., Brignone, V., Matarazzo, S., Del Zotto, M., and Zani, A. (2006). Gender differences in hemispheric asymmetry for face processing. *BMC Neurosci.* 7:44. doi: 10.1186/1471-2202-7-44
- Rinne, T., Pekkola, J., Degerman, A., Autti, T., Jääskeläinen, I. P., Sams, M., et al. (2005). Modulation of auditory cortex activation by sound presentation rate and attention. *Hum. Brain Mapp.* 26, 94–99. doi: 10.1002/hbm.20123
- Rodríguez Holguín, S., Porjesz, B., Chorlian, D. B., Polich, J., and Begleiter, H. (1999). Visual P3a in male subjects at high risk for alcoholism. *Biol. Psychiatry* 46, 281–291. doi: 10.1016/S0006-3223(98)00247-9
- Rossignol, M., Philippot, P., Bissot, C., Rigoulot, S., and Campanella, S. (2012). Electrophysiological correlates of enhanced perceptual processes and attentional capture by emotional faces in social anxiety. *Brain Res.* 1460, 50–62. doi: 10.1016/j.brainres.2012.04.034
- Rossignol, M., Philippot, P., Crommelinck, M., and Campanella, S. (2008). Visual processing of emotional expressions in mixed anxious-depressed subclinical state: an event-related potential study on a female sample. *Neurophysiol. Clin.* 38, 267–275. doi: 10.1016/j.neucli.2008.07.007
- Rossignol, M., Philippot, P., Douilliez, C., Crommelinck, M., and Campanella, S. (2005). The perception of fearful and happy facial expression is modulated by anxiety: an event-related potential study. *Neurosci. Lett.* 377, 115–120. doi: 10.1016/j.neulet.2004.11.091
- Rugg, M. D., and Coles, M. G. (1995). *Electrophysiology of Mind: Event-related Brain Potentials and Cognition*. Oxford University Press. Available online at: <http://www.getcited.org/?PUB=103210581&showStat= Ratings>
- Sabatinelli, D., Bradley, M. M., Fitzsimmons, J. R., and Lang, P. J. (2005). Parallel amygdala and inferotemporal activation reflect emotional intensity and fear relevance. *Neuroimage.* 24, 1265–1270. doi: 10.1016/j.neuroimage.2004.12.015

- Schäfer, R., Schneider, C., Tress, W., and Franz, M. (2007). Cortical augmenting in alexithymic subjects after unpleasant acoustic stimulation. *J. Psychosom. Res.* 63, 357–364. doi: 10.1016/j.jpsychores.2007.03.015
- Schupp, H. T., Stockburger, J., Codispoti, M., Junghöfer, M., Weike, A. I., and Hamm, A. O. (2007). Selective visual attention to emotion. *J. Neurosci.* 27, 1082–1089. doi: 10.1523/JNEUROSCI.3223-06.2007
- Semlitsch, H. V., Anderer, P., Schuster, P., and Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. *Psychophysiology* 23, 695–703. doi: 10.1111/j.1469-8986.1986.tb00696.x
- Sifneos, P. E. (1973). The prevalence of «alexithymic» characteristics in psychosomatic patients. *Psychother. Psychosom.* 22, 255–262. doi: 10.1159/000286529
- Singhal, A., Shafer, A. T., Russell, M., Gibson, B., Wang, L., Vohra, S., et al. (2012). Electrophysiological correlates of fearful and sad distraction on target processing in adolescents with attention deficit-hyperactivity symptoms and affective disorders. *Front. Integr. Neurosci.* 6:119. doi: 10.3389/fnint.2012.00119
- Solowij, N., Michie, P. T., and Fox, A. M. (1995). Differential impairments of selective attention due to frequency and duration of cannabis use. *Biol. Psychiatry.* 37, 731–739. doi: 10.1016/0006-3223(94)00178-6
- Spielberger, C. D., Gorsuch, R. L., and Luthene, R. E. (1983). *Manual for the State and Trait Anxiety Inventory STAI*. Palo Alto, CA: Consulting Psychologist Press.
- Swart, M., Kortekaas, R., and Aleman, A. (2009). Dealing with feelings: characterization of trait alexithymia on emotion regulation strategies and cognitive-emotional processing. *PLoS ONE* 4:e5751. doi: 10.1371/journal.pone.0005751
- Taylor, G. J., Bagby, R. M., and Parker, J. D. A. (1997). *Disorders of Affect Regulation: Alexithymia in Medical and Psychiatric Illness*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511526831
- Vermeulen, N., Luminet, O., de Sousa, M. C., and Campanella, S. (2008). Categorical perception of anger is disrupted in alexithymia: Evidence from a visual ERP study. *Cogn. Emot.* 22, 1052–1067. doi: 10.1080/02699930701597635
- Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends Cogn. Sci. (Regul. Ed.)* 9, 585–594. doi: 10.1016/j.tics.2005.10.011
- Walker, S., O'Connor, D. B., and Schaefer, A. (2011). Brain potentials to emotional pictures are modulated by alexithymia during emotion regulation. *Cogn. Affect. Behav. Neurosci.* 11, 463–475. doi: 10.3758/s13415-011-0042-1
- Wehmer, F., Brejnak, C., Lumley, M., and Stettner, L. (1995). Alexithymia and physiological reactivity to emotion-provoking visual scenes. *J. Nerv. Ment. Dis.* 183, 351–357. doi: 10.1097/00005053-199506000-00001
- Zald, D. H. (2003). The human amygdala and the emotional evaluation of sensory stimuli. *Brain Res. Brain Res. Rev.* 41, 88–123. doi: 10.1016/S0165-0173(02)00248-5
- Zhang, L., Ye, R., Yu, F., Cao, Z., Zhu, C., Cai, Z., et al. (2012). How does emotional context modulate response inhibition in alexithymia: electrophysiological evidence from an ERP study. *PLoS ONE* 7:e51110. doi: 10.1371/journal.pone.0051110

**Conflict of Interest Statement:** The last author is funded by the Belgian Fund for Scientific Research (F.N.R.S., Belgium), but this fund did not exert any editorial direction or censorship on any part of this article. The other authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 09 August 2013; accepted: 11 February 2014; published online: 03 March 2014.

Citation: Delle-Vigne D, Kornreich C, Verbanck P and Campanella S (2014) Subclinical alexithymia modulates early audio-visual perceptive and attentional event-related potentials. *Front. Hum. Neurosci.* 8:106. doi: 10.3389/fnhum.2014.00106

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2014 Delle-Vigne, Kornreich, Verbanck and Campanella. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Multimodal emotion perception after anterior temporal lobectomy (ATL)

Valérie Milesi<sup>1,2\*</sup>, Sezen Cekic<sup>1,2</sup>, Julie Péron<sup>1,2</sup>, Sascha Frühholz<sup>1,2</sup>, Chiara Cristinzio<sup>1,2,3</sup>, Margitta Seeck<sup>4</sup> and Didier Grandjean<sup>1,2</sup>

<sup>1</sup> Swiss Center for Affective Sciences, University of Geneva, Geneva, Switzerland

<sup>2</sup> Neuroscience of Emotion and Affective Dynamics Laboratory, Department of Psychology, Faculty of Psychology and Educational Sciences, University of Geneva, Geneva, Switzerland

<sup>3</sup> Laboratory for Neurology and Imaging of Cognition, Department of Neurology and Department of Neuroscience, Medical School, University of Geneva, Geneva, Switzerland

<sup>4</sup> Epilepsy Unit, Department of Neurology, Geneva University Hospital, Geneva, Switzerland

## Edited by:

Benjamin Kreifelts, University of Tübingen, Germany

## Reviewed by:

Thomas Ethofer, University Tübingen, Germany

Martin Klasen, RWTH Aachen University, Germany

## \*Correspondence:

Valérie Milesi, Swiss Center for Affective Sciences, University of Geneva, Rue des Batoirs 7, 1205 Geneva, Switzerland  
e-mail: valerie.milesi@unige.ch

In the context of emotion information processing, several studies have demonstrated the involvement of the amygdala in emotion perception, for unimodal and multimodal stimuli. However, it seems that not only the amygdala, but several regions around it, may also play a major role in multimodal emotional integration. In order to investigate the contribution of these regions to multimodal emotion perception, five patients who had undergone unilateral anterior temporal lobe resection were exposed to both unimodal (vocal or visual) and audiovisual emotional and neutral stimuli. In a classic paradigm, participants were asked to rate the emotional intensity of angry, fearful, joyful, and neutral stimuli on visual analog scales. Compared with matched controls, patients exhibited impaired categorization of joyful expressions, whether the stimuli were auditory, visual, or audiovisual. Patients confused joyful faces with neutral faces, and joyful prosody with surprise. In the case of fear, unlike matched controls, patients provided lower intensity ratings for visual stimuli than for vocal and audiovisual ones. Fearful faces were frequently confused with surprised ones. When we controlled for lesion size, we no longer observed any overall difference between patients and controls in their ratings of emotional intensity on the target scales. Lesion size had the greatest effect on intensity perceptions and accuracy in the visual modality, irrespective of the type of emotion. These new findings suggest that a damaged amygdala, or a disrupted bundle between the amygdala and the ventral part of the occipital lobe, has a greater impact on emotion perception in the visual modality than it does in either the vocal or audiovisual one. We can surmise that patients are able to use the auditory information contained in multimodal stimuli to compensate for difficulty processing visually conveyed emotion.

**Keywords:** amygdala, emotion perception, multimodal, prosody, facial expression

## INTRODUCTION

The ability to decode emotional information is crucial in everyday life, allowing us to adapt our behaviors when confronted with salient information, both for survival and for social adaptiveness purposes. The emotional features of objects in the environment have been shown to bring about an increase in the neuronal response, compared with the processing of non-emotional information (for a review, see Phan et al., 2002). The role that different brain regions play in decoding emotional information appears to depend on the modality. Furthermore, research has shown that both the primary and secondary sensory regions are modulated by emotion. For example, visual extrastriate regions are modulated by emotions conveyed by facial expressions (e.g., Morris et al., 1998; Pourtois et al., 2005a; Vuilleumier and Pourtois, 2007), while temporal voice-sensitive areas have been shown to be modulated by emotional prosody (e.g., Mitchell et al., 2003; Grandjean et al., 2005; Schirmer and Kotz, 2006; Wildgruber et al., 2006; Frühholz et al., 2012).

According to Haxby's face perception model (Haxby et al., 2000), visual information is processed along a ventral pathway leading from the primary visual cortex (V1) to the fusiform face area (FFA) and inferior temporal cortex (ITC). Face perception is sufficient to activate the FFA (see, for example, Pourtois et al., 2005a; Kanwisher and Yovel, 2006; Pourtois et al., 2010), but the activity of this structure is enhanced when the facial information is emotional (see, for example, Breiter et al., 1996; Dolan et al., 2001; Vuilleumier et al., 2001; Williams et al., 2004; Vuilleumier and Pourtois, 2007). Another structure whose activity increases when decoding emotional facial information is the amygdala (see, for example, Haxby et al., 2000; Calder and Young, 2005; Phelps and LeDoux, 2005; Adolphs, 2008). In monkeys, this structure has been shown to project to almost every step along the visual ventral pathway (Amaral et al., 2003). Human studies, meanwhile, have suggested that connectivity between the amygdala and the FFA is modulated by emotion perception (Morris et al., 1998; Dolan et al., 2001; Vuilleumier et al.,

2004; Sabatinelli et al., 2005; Vuilleumier, 2005; Vuilleumier and Pourtois, 2007).

Regarding the amygdala's role in emotion perception, the current hypothesis is that this structure detects salience, a general feature of emotion (for a discussion, see Sander et al., 2003; Armony, 2013; Pourtois et al., 2013), through reciprocal connections with the cortex (Amaral et al., 2003). Its main function is to facilitate attention and perception processing (e.g., Armony and Ledoux, 1997, 1999; Whalen, 1998; Vuilleumier et al., 2001) without explicit voluntary attention (for a review, see Vuilleumier and Pourtois, 2007). According to Ledoux's (2007) model, the amygdala's output is directed both to regions that modulate bodily responses (via the endocrine system related to the autonomic system), and to the primary and associative cortices. These encompass regions modulated by emotion such as in the extrastriate visual system, the FFA for face perception, and the voice area in the superior temporal gyrus (STG; including the primary auditory region).

Further insight into emotional face perception and its subcortical bases has been provided by studies of patients with lesions of the amygdala. More specifically, studies have assessed patients with temporal lobe epilepsy whose lesions are linked either to the epileptogenic disease itself or else to its surgical treatment (see, for example, Cristinzio et al., 2007). These studies included patients with congenital or acquired diseases resulting in bilateral lesions, and patients with unilateral epilepsy arising from mesial temporal sclerosis who had undergone lobectomy with amygdalectomy. Patients with bilateral damage have been found to display impaired fearful face perception (Adolphs et al., 1994, 1995; Young et al., 1995; Calder et al., 1996; Broks et al., 1998) and deficits in the perception of surprise and anger (Adolphs et al., 1994). Unilateral lesions have yielded either no differences (Adolphs et al., 1995; prior surgery, Batut et al., 2006) or else a deficit for patients with right-sided lesions covering either a range of emotions (Anderson et al., 2000; Adolphs and Tranel, 2004) or solely fearful faces (prior surgery, Meletti et al., 2003). Palermo et al. (2010) found that both left- and right-lesion groups exhibited a deficit in fear intensity perception, but the left-lesion group was more impaired for fear detection. Anterior temporal lobectomy with amygdalectomy is generally expected to affect the perceived intensity of facial emotional expressions. The functional explanation for this is a lack of modulation by the amygdala of the ventral visual processing network and, more specifically in the case of emotional faces, of the FFA.

In addition to visual emotional information, the amygdala has been shown to be associated with different responses to emotional vocalizations. According to Schirmer et al. (2012), the processing of auditory information takes place along three streams in the temporal lobe: a posterior stream passing through the posterior part of the superior temporal sulcus (pSTS) for sound embodiment; a ventral stream directed toward the middle temporal gyrus (MTG) for concept processing; and an anterior stream extending as far as the temporal pole (TmP) for the perceptual domain (i.e., semantic processing). Another specificity of emotional vocalization perception is the hemispheric specificity modeled by Schirmer and Kotz (2006). In their model, the left

temporal lobe has a higher temporal resolution for processing information than the right hemisphere, and is more involved in linguistic signal processing (segmental information), with suprasegmental analysis taking place in the right hemisphere. The amygdala has been shown to be modulated by emotional vocalizations, including onomatopoeia (e.g., Morris et al., 1999; Fecteau et al., 2007; Plichta et al., 2011), and emotional prosody consisting either of pseudowords (e.g., Grandjean et al., 2005; Sander et al., 2005; Frühholz and Grandjean, 2012, 2013), or of words and sentences (e.g., Ethofer et al., 2006, 2009; Wiethoff et al., 2009).

In contrast to research on emotional face perception, studies of auditory emotion processing in patients with bilateral amygdala lesions have produced divergent results. Some have failed to find any effect at all on emotion recognition (semantically neutral sentences: Adolphs and Tranel, 1999; names and onomatopoeia: Anderson and Phelps, 1998). Others have reported either a general impairment (counting sequences: Brierley et al., 2004) or specific impairments for fear (semantically neutral sentences: Scott et al., 1997; non-verbal vocalizations: Dellacherie et al., 2011), surprise (Dellacherie et al., 2011), anger (Scott et al., 1997), or sadness perception (musical excerpt: Gosselin et al., 2007). There is a similar divergence for unilateral lesions, with either no effects (Adolphs and Tranel, 1999; Adolphs et al., 2001) or a specific impairment for fear (counting sequences: Brierley et al., 2004; meaningless words: Sprengelmeyer et al., 2010; non-verbal vocalizations: Dellacherie et al., 2011). To sum up current knowledge about auditory emotion processing, there is a strong hypothesis about right hemispheric involvement for emotional prosody. The amygdala appears to be involved in prosody perception, but may also be sensitive to the proximal context of the stimulus presentation (for a discussion, see Frühholz and Grandjean, 2013).

In the case of face-voice emotion integration, studies featuring audiovisual emotional stimuli have replicated the response facilitation effect at the behavioral level, namely an increase in perceptual sensitivity and reduced reaction times (e.g., Massaro and Egan, 1996; De Gelder and Vroomen, 2000; Dolan et al., 2001; Kreifelts et al., 2007), that has already been demonstrated in non-emotional studies (e.g., Miller, 1982; Schröger and Widmann, 1998). Responsibility for the behavioral improvement has been mainly attributed to various cortical substrates, including the left MTG (e.g., Pourtois et al., 2005b), the posterior STG (pSTG; e.g., Ethofer et al., 2006; Kreifelts et al., 2007), and, interestingly, the amygdala, either bilaterally (e.g., Klasen et al., 2011) or the left side (e.g., Dolan et al., 2001; Ethofer et al., 2006; Müller et al., 2012). Animal studies have yielded a more detailed multimodal model, with different levels of integration. For instance, a rhinal cortex lesion, as opposed to a direct lesion of the amygdala, is sufficient to disrupt associative mechanisms (Goulet and Murray, 2001). Meanwhile, a comparison of the roles of the perirhinal cortex (PRC) and the pSTS led Taylor et al. (2006) to suggest that the pSTS plays a presemantic integration role, while the PRC integrates higher level conceptual representations.

In summary, studies of the amygdala's modal specificity have reported impairments in patients with temporal lobectomy or specific amygdalectomy for faces and either voices (Scott et al., 1997; Sprengelmeyer et al., 1999; Brierley et al., 2004) or emotion in

music (Gosselin et al., 2007, 2011). However, some patients seem to have a specific deficit for visual emotional stimuli (Adolphs et al., 1994, 2001; Anderson and Phelps, 1998; Adolphs and Tranel, 1999). Discrepancies between studies have been explained by a number of different factors, including the date of epilepsy onset (e.g., McClelland et al., 2006), the nature and context of the stimuli (e.g., face presentation duration; Graham et al., 2007; Palermo et al., 2010). The fear specificity of amygdala processing has also been strongly called into question (for a discussion, see Cahill et al., 1999; Murray, 2007; Morrison and Salzman, 2010). To the best of our knowledge, however, the role of lesion size has not been taken in account thus far.

Our aim in the present study was to test whether the categorization and intensity perception of unimodal (i.e., either visual or non-verbal auditory emotional stimuli), as opposed to bimodal (i.e., audiovisual) emotional stimuli is modified in patients who have undergone unilateral temporal anterior lobectomy with amygdalectomy. The impact of anterior temporal lobe ablation is assumed to differ with modality. Regarding the auditory network, above and beyond the absence of voice area modulations owing to amygdala resection, Schirmer et al. (2012) suggests that the anterior temporal lobe is more involved in semantic processing, representing the final temporal step before the processing shifts to the frontal regions associated with emotion evaluation. We would therefore expect disruption of this input to have an impact on categorization, with patients making more mistakes or confusing more items than matched controls. For the visual modality, we would expect to find the same kind of deficit, stemming from the lack of emotion-related modulation of visual cortical input. Finally, for audiovisual material, we would expect to observe either a better preserved ability for correct detection and perceived intensity, if an intact pSTS and a more dorsal pathway toward the frontal lobe are sufficient to integrate audiovisual information, or no improvement because of the PRC lesion.

Participants rated the intensity of brief onomatopoeic vocalizations produced by actors (Bänziger et al., 2012) and animated synthetic faces (Roesch et al., 2011) on visual analog scales. At the group level, we expected the patients to have a higher error rate than controls when it came to identifying unimodal emotional stimuli. This has been shown to be the case in the visual modality for fearful faces (bilateral lesion: Adolphs et al., 1994, 1995; Young et al., 1995; Calder et al., 1996; Broks et al., 1998; unilateral lesion: Anderson et al., 2000; McClelland et al., 2006), and in the auditory modality for both fearful voices (bilateral lesion: Scott et al., 1997; Adolphs and Tranel, 2004; unilateral lesion: Scott et al., 1997; Brierley et al., 2004; Sprengelmeyer et al., 2010; Dellacherie et al., 2011) and angry voices (bilateral lesion: Scott et al., 1997). For the audiovisual stimuli, we expected to observe a higher error rate for fear identification, arising from the combined effects of the unimodal deficits in each modality. Regarding intensity perception, we expected to observe similar patterns, even after controlling for the extent of the lesion along the ventral pathway. Finally, we investigated the effects of lesion size on emotion recognition. We predicted that perception of emotion intensity would be modulated by the size of the lesion, with more extensive lesions resulting in impairment at different levels of information processing. We developed

an additional hypothesis to explain the discrepant findings of previous studies.

## MATERIALS AND METHODS

### PARTICIPANTS

We recruited five patients who had undergone unilateral antero-medial temporal lobectomy together with the unilateral removal of the amygdala. One patient (JP) had a lesion that extended to the occipital and posterior parietal lobes. The surgery had been performed to control the patients' medically intractable seizures (see **Figure 1** for the location and extent of their lesions): four on the left side (FB, 23 years old; CG, 37 years old; JP, 45 years old; and RS, 62 years old) and one on the right (CM, 31 years old). CG was the only woman in the patient group, and FB the only left-handed patient. Controls were recruited via local advertisements: 12 were matched with FB, CM, and CG for sex, handedness, and age; six with JP; and three with RS (see **Table 1** for a summary and **Table 2** for a detailed description of each patient). Patients did not exhibit any gnosis deficit in their respective neuropsychological tests. The study was approved by the local ethics committee, and all the participants gave their written informed consent. The controls received financial compensation (CHF 15) for taking part in the experiment.

### LESION DELIMITATION AND DESCRIPTION

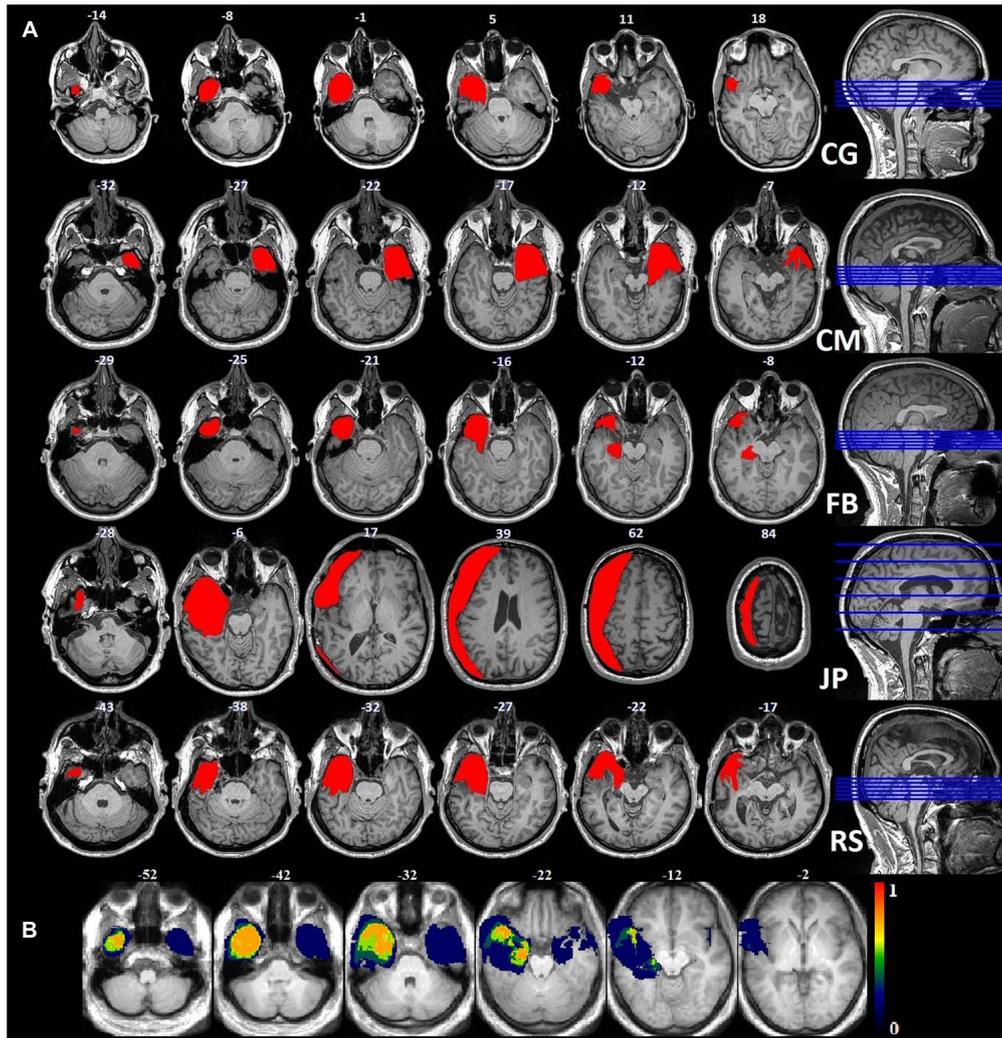
In order to compute the lesion size of each patient, anatomical images were segmented and normalized using a unified segmentation approach (Ashburner and Friston, 2005) together with the Clinical toolbox<sup>1</sup>. Because of the cost function masking purpose (Andersen et al., 2010), lesion masks drawn on the patients' anatomical scans were included in the brain segmentation. Structural images and lesion masks were normalized to MNI space with the DARTEL toolbox, using individual flow fields, which were estimated on the basis of the segmented gray (GM) and white matter (WM) tissue classes. The normalized lesion masks were used to calculate the lesion size for each patient in standard space.

CG had a left anterior temporal lesion with an intact inferior temporal gyrus (ITG) and lateral occipitotemporal gyrus (LOTG). The lesion area included the periamygdaloid cortex (PAM), entorhinal cortex (Ent), medial occipitotemporal gyrus (MOTG), inferior part of the hippocampus (Hi), parahippocampal gyrus (PHG), and amygdala, and ended in the lateral anterior portion of the temporal lobe, in the MTG and TmP.

CM had a right anterior temporal lesion extending to the middle and ventromedial part of the temporal lobe, including the inferior temporal pole (ITmP), ITG, Ent, PAM, PRC, amygdala, inferior Hi, STG, anterior fusiform gyrus (FuG), and rhinal sulcus. In the posterior part of the lesion, the PPo (planum polare), STG and STS were intact.

FB had a left anterior temporal lesion that included the TmP, MTG, MOTG, Ent, Hi, PAM, amygdala, anterior STG, and posterior temporal cortex (PTe). The lesion ended in two separate tails: one in the lateral anterior part of the temporal lobe, the other in the medial part.

<sup>1</sup><http://www.mccauslandcenter.sc.edu/CRNL/clinical-toolbox>



**FIGURE 1 | (A)** Anatomical images of the lesions for each patient: each lesion was delineated manually on the axial plane and corrected using the coronal plane. **(B)** Probability map for the normalized lesion size.

**Table 1 | Participants.**

Patient	Lateralization	Age (years)	Sex	Lesion (vx) (volume, normalized volume)	Controls <i>n</i>	Mean age (SD) of controls
CG	Right-handed	37	Female	Left (31'564, 40'683)	12	33.42 (3.63)
CM	Right-handed	31	Male	Right (36'836, 48'833)	12	32.83 (2.48)
FB	Left-handed	23	Male	Left (21'596, 30'208)	12	22.08 (1.88)
JP	Right-handed	45	Male	Left (434'284, 113'254)	6	45.17 (2.32)
RS	Right-handed	62	Male	Left (44'011, 53'920)	3	60 (4)

JP had an extended left lateral resection including the temporal, frontal, parietal, and occipital lobes. The temporal part included the TmP, MTG, Ent, MOTG, ITG, PTe, anterior STG, and PHG. The frontal part included the lateral inferior and superior frontal gyri, precentral gyrus and postcentral gyrus. Finally, part of the

lateral superior posterior occipital gyrus had been removed, but the FFA was intact.

RS had a left anterior temporal lesion encompassing the TmP, STG, MTG, MOTG, ITG, FuG, amygdala, anterior Hi, Ent, PAM, FuG and anterior PHG. It ended in the lateral anterior

**Table 2 | Patient description.**

Patient	Years since lobectomy	Age at epilepsy onset	Comorbidity	Diagnosis
CG	5	12 years	–	Left temporal partial complex epilepsy with hippocampal sclerosis
CM	6	11 years	Anxiety	Left hemiplegia with right hemisphere hypoplasia
FB	10	–	–	Major left hippocampal sclerosis
JP	11	5 years	Depressed feeling (at time of surgery)	Left temporal epilepsy
RS	7	6 months	–	Left hippocampal sclerosis

part. See **Figure 1** for visual descriptions of the patients' brain damage.

### STIMULI AND PROCEDURE

Non-verbal auditory expressions were drawn from the validated Geneva Multimodal Emotion Portrayal (GEMEP) corpus (Bänziger et al., 2012). We selected angry, joyful, and fearful non-verbal sounds ("ah") produced by two male and two female actors, on the basis of the recognition rate established in a previous pilot study. For the neutral stimuli, we chose the most neutrally rated vocal expressions produced by the same actors (neutrality rating:  $M = 26.5$ ,  $SD = 15.67$ ), and the fundamental frequency was flattened using Praat (Boersma and Weenink, 2011). Sounds were cut and/or stretched to achieve a duration of 1 s (mean duration before time stretch = 0.92 s,  $SD = 0.30$  s) with SoundForge<sup>2</sup>, and 0.025 s fade-ins and fade-outs were included using Audacity<sup>3</sup>. The dynamic faces were created with FACSGen (Roesch et al., 2011), which allows for the parametric manipulation of 3D emotional facial expressions according to the Facial Action Coding System (Ekman and Friesen, 1978). They were selected on the basis of results of a previous study in which participants assessed the gender and believability of each avatar (Roesch et al., 2011). The lips were animated to match the intensity contour of each different sound for both unimodal visual and audiovisual items. The action units (AUs) for each emotion began at 0.25 s and ended at 0.75 s after onset, with their apex at 0.5 s (100% intensity). VirtualDub<sup>4</sup> was used to generate the image sequences and to combine the voiced sounds with them at a rate of 26 frames per second (the final image was a dark screen).

After signing the consent form, participants completed the behavioral inhibition system (BIS)/behavioral approach system (BAS) scales and the state trait anxiety inventory (STAI) on a web interface. They then rated the intensity of 216 items in unimodal [auditory (A), or visual (V)] and audiovisual (AV; congruent: same information in both modalities; incongruent: one modality emotional, the other neutral) conditions. The unimodal and congruent audiovisual stimuli could either express

the emotions of anger, fear, or joy, or be neutral (control condition). Each condition (modality, emotion, or congruency) was repeated 12 times. Items were presented using E-Prime (standard v2.08.90<sup>5</sup>) in a pseudorandomized order to avoid repetition of the same stimulus (i.e., synthetic face or actor's voice) or condition. The participants gave their answers by clicking on a continuous line between *Not intense* and *Very intense* for six different emotions (disgust, joy, anger, surprise, fear, sadness), plus neutral. In each trial, they could provide ratings on one or more scales. At the end of the experiment, they completed a debriefing questionnaire.

### STATISTICAL ANALYSIS

Since multiple intensity scales were used to collect the answers, our data mostly contained zero ratings. To assess the interactions, we therefore ran a zero-inflated mixed model on congruent trials only, using the glmmADMB package for R<sup>6</sup>. This allowed the excess zeroes and remaining values to be modeled as binomial responses, and modeled the distribution as a generalized linear model (GLM) following a negative binomial distribution. Main effects were tested for group (control vs. patient), modality (audio, visual, audiovisual), and emotion (anger, fear, or joy, plus neutral). Contrasts were performed to test specific hypotheses.

The first hypothesis we tested was a group effect for a specific emotion on the target scale (e.g., fearful item ratings on the fear scale) for each modality (A, V, AV). Sex, age, and normalized lesion size were added as control variables. Participant and stimulus ID were added as random effects. A different model was run for each of the three emotions, plus neutral. Second, four different models, one for each emotion, plus neutral, were tested in order to compare the impact of the three different modalities in each group. For instance, for angry item ratings on the anger scale, the modalities were tested in pairs (AV-A, AV-V, A-V) for the patient group, and individually for the control group. For this second set of models, we added the same control and random variables as for the first model. The third model was run to investigate the lateralization effect of the lesion for a specific modality and a specific emotion, controlling for handedness, age, and sex, and with random effect variables for participant ID and

<sup>2</sup><http://www.sonycreativesoftware.com/soundforge>

<sup>3</sup><http://audacity.sourceforge.net/>

<sup>4</sup><http://www.virtualdub.org/>

<sup>5</sup><http://www.pstnet.com/eprime.cfm>

<sup>6</sup><http://www.r-project.org/>

stimulus ID. Owing to the limited size of our patient sample, this comparison was of a purely descriptive and exploratory nature. In order to test whether the effects we found in the different modalities were perceptual or emotional, we ran a complementary analysis to compare emotional versus neutral items in each modality and each group, adding age, sex, and normalized lesion size as control variables, and participant ID and stimulus ID as random effects. Intergroup effects were also tested for emotional versus neutral items in each modality (A, AV, V), with the same control variables. Finally, we tested the impact of lesion size by including the number of voxels in a separate linear model for each emotion and each modality. In this final set of models, random effect variables (participant ID and stimulus ID) were added.

## RESULTS

### CATEGORICAL RESPONSES

Participants could rate the intensity of each item on six different scales (anger, disgust, fear, surprise, joy, sadness, and neutral). For each item, we identified the scale with the highest rating, and calculated a proportional corrected score for each participant (Heberlein et al., 2004; Dellacherie et al., 2011), by looking at how many other members of the participant's group (patient or control) had given the same response. This score could range from 0, meaning that nobody else in the group had chosen the same scale, to 1, meaning that everyone in the group had chosen the same scale. This type of correction is used to weight labeling errors, bearing in mind that some errors are more correct than others. For instance, it is easier to confuse visual fear and surprise (see, for example, Etcoff and Magee, 1992) than it is to confuse fear and anger, as the first two expressions share a number of AUs. For vocal expressions, confusion is also possible, but between different pairs of emotions (see, for example, Banse and Scherer, 1996; Belin et al., 2008; Bänziger et al., 2009).

Using these corrected scores, we looked for possible differences between the two groups. As our data violated the assumptions of homoscedasticity and normal distribution, we ran non-parametric tests for multiple groups. In order to pinpoint differences between the groups within a specific emotion in a specific modality, we used the Kruskal–Wallis test, calculating  $z$  scores and  $p$  values corrected for multiple comparisons of mean ranks ( $z'$ ). These multiple comparisons are summarized in **Figure 2**. The control group was more accurate than the patient group in recognizing joy, whether it was expressed vocally ( $z' = 3.02, p < 0.005$ ), visually ( $z' = 3.17, p < 0.005$ ), or bimodally ( $z' = 3.19, p < 0.005$ ). Greater accuracy within the control group was also observed for visual anger ( $z' = 2.99, p < 0.005$ ), vocal fear ( $z' = 2.78, p < 0.01$ ) and - marginally - visual ( $z' = 1.69, p = 0.08$ ) and bimodal fear ( $z' = 1.89, p = 0.058$ ). Finally, a reverse group effect was observed for the neutral vocal ( $z' = 3.64, p < 0.001$ ) and audiovisual ( $z' = 3.64, p < 0.001$ ) stimuli.

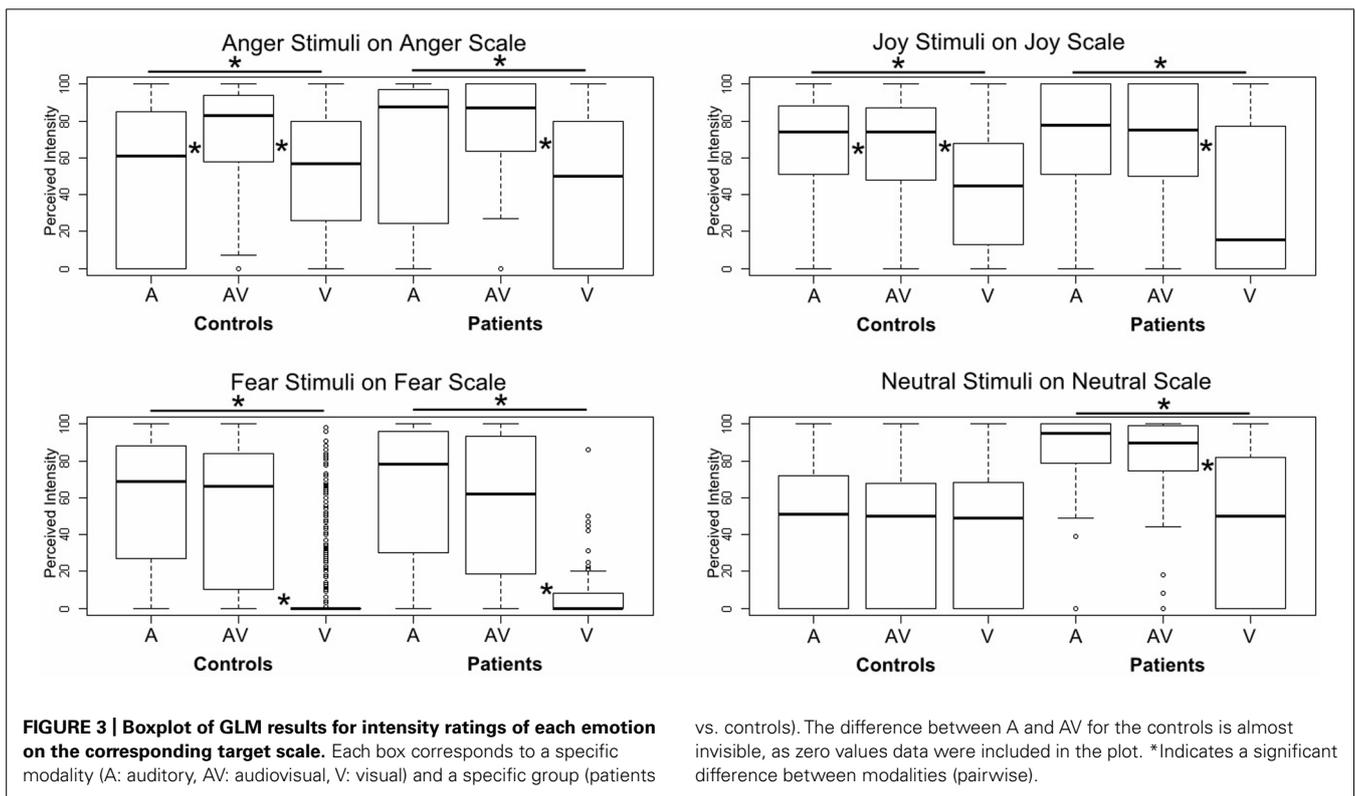
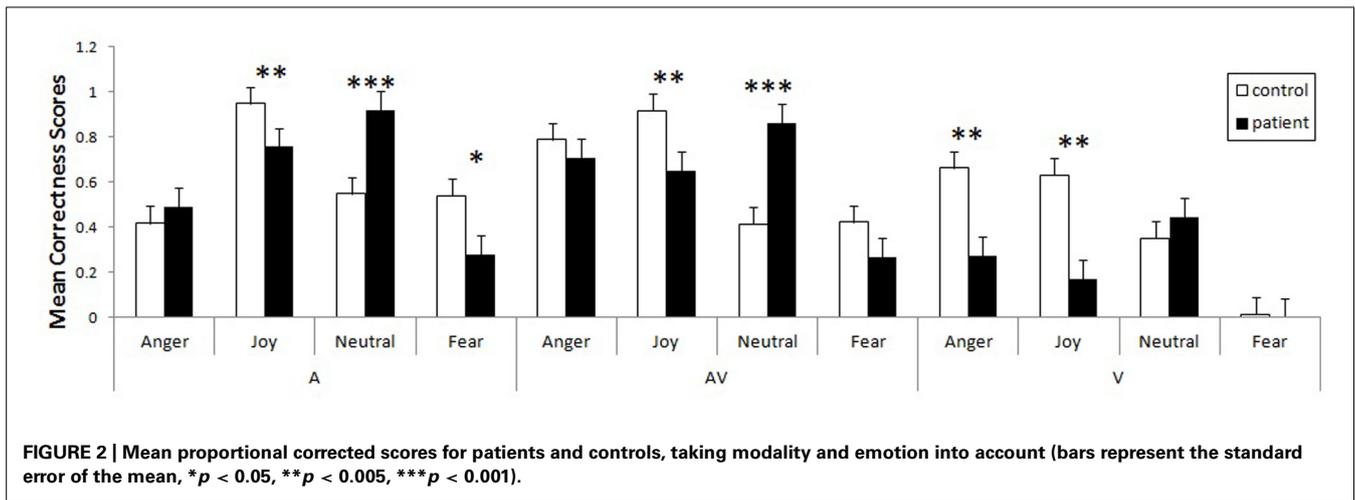
Finally, we tested the impact of lesion size on the corrected hit rate for emotion recognition. We ran supplementary analyses using a GLM to test this effect with the modality (A, AV, V) and emotion (anger, joy, fear, neutral) factors, and added the normalized lesion size as a covariate. The control variables were age, sex,

and lateralization. We observed a significant linear relationship between normalized lesion size and corrected hit score for visual anger ( $z = -2.91, p < 0.005$ ), visual joy ( $z = -2.37, p < 0.05$ ) and visual neutral stimuli ( $z = -3.52, p < 0.001$ ). All the linear regressions were negative, meaning that the more extensive the lesion, the lower the corrected score. We observed no such effect for fear in the visual modality, as patients did not recognize this emotion (their corrected score was equal to 0), confusing it with surprise.

### INTENSITY PERCEPTION

Using a GLM, we first compared the two groups on each specific emotion in each specific modality, controlling for sex, age, and normalized lesion size, and adding participant and stimulus ID as random effects. No significant results were observed, even for the fear items. However, when we ran pairwise comparisons of the modalities for a specific emotion on its target scale and for a specific group, we did observe significant effects, especially for the three emotions (see **Figure 3**). Patients provided higher intensity ratings of audiovisual versus unimodal visual information for angry ( $z = -4.14, p < 0.001$ ), joyful ( $z = -6.14, p < 0.001$ ), fearful ( $z = -8.45, p < 0.001$ ), and neutral ( $z = -5.61, p < 0.001$ ) items. They also provided higher intensity ratings of auditory versus visual information for the same emotions (anger:  $z = -4.14, p < 0.001$ ; joy:  $z = -6.14, p < 0.001$ ; fear:  $z = -8.45, p < 0.001$ ; neutral:  $z = -5.61, p < 0.001$ ). The differences between audiovisual and unimodal auditory information were not significant for any of the emotions ( $p > 0.15$ ). In the control group, a slightly different pattern emerged for anger and joy. Anger was given a higher intensity rating in the audiovisual condition than in either the auditory ( $z = -3.27, p < 0.001$ ) or visual ( $z = -10.93, p < 0.001$ ) condition, and a higher rating in the auditory condition than in the visual one ( $z = -6.94, p < 0.001$ ). For joy, audiovisual information was perceived of as more intense than visual information ( $z = -12.69, p < 0.001$ ), but auditory information was given a higher intensity rating than both audiovisual information ( $z = 3.07, p < 0.005$ ) and visual information ( $z = -15.58, p < 0.001$ ). Finally, fear stimuli were rated as more intense in the audiovisual modality than in the visual one ( $z = -11.74, p < 0.001$ ), and also more intense in the auditory modality than in the visual one ( $z = -12.33, p < 0.001$ ). No significant differences were observed between the modalities for neutral stimuli ( $p > 0.4$ ).

In order to ascertain whether the results were perceptual or emotional, we tested another model contrasting emotional versus neutral stimuli for each group and each modality (A, V, AV). Controls rated emotional auditory items as more intense than neutral auditory items ( $z = 5.61, p < 0.001$ ), and this was also the case for audiovisual information ( $z = 5.15, p < 0.001$ ). By contrast, the patients provided higher intensity ratings for neutral items than they did for emotional items in the auditory ( $z = -2.64, p < 0.01$ ) and audiovisual ( $z = -2.42, p < 0.05$ ) modalities. In the visual modality, patients ( $z = -3.56, p < 0.001$ ) and controls ( $z = -8.40, p < 0.001$ ) alike gave higher intensity ratings for neutral items than for emotional ones. When we compared the two groups on emotional and neutral items in each modality, we found that the patients rated



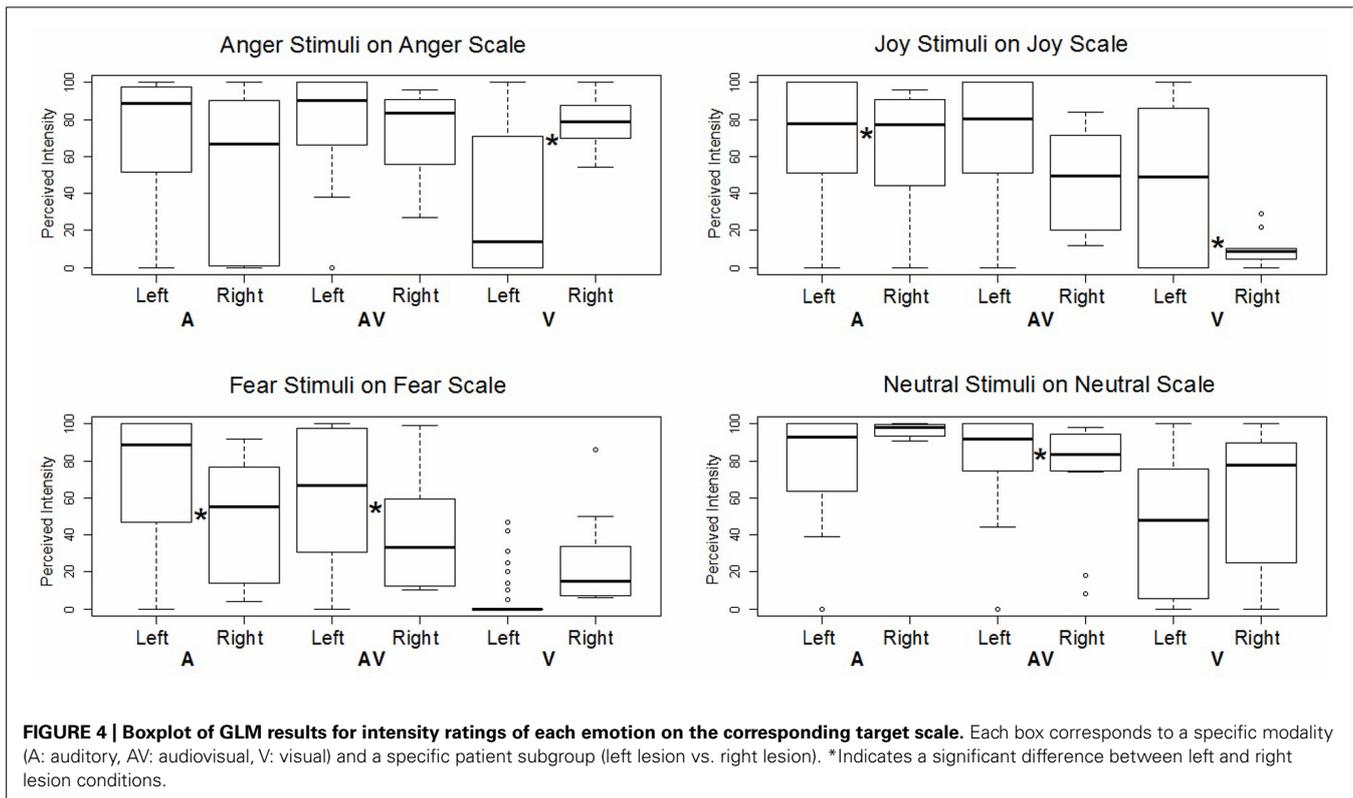
the intensity of the neutral items more highly than the controls did in the auditory modality ( $z = 2.24, p < 0.025$ ). For the audiovisual modality, the effect was only marginal ( $z = 1.86, p = 0.062$ ).

**INTENSITY PERCEPTION AND LESION EFFECT**

We then assessed the impact of lesion lateralization for each specific emotion in each specific modality. In this GLM analysis, we compared the patients’ ratings on the target scale according to the side of their lesion, controlling for handedness, age, and sex, and adding participant and stimulus ID as random effects (see Figure 4). The patient with a right lesion was found to

provide higher intensity ratings than the patients with left lesions, but only for angry faces ( $z = -4.36, p < 0.001$ ) and auditory joy ( $z = -3.23, p < 0.005$ ). All other significant effects concerned the opposite relationship, namely, the patients with left lesions rated the items as more intense than the patient with a right lesion did. This was the case for visual joy ( $z = 3.19, p < 0.005$ ), auditory fear ( $z = 8.29, p < 0.001$ ), audiovisual fear ( $z = 8.23, p < 0.001$ ), and audiovisual neutral items ( $z = 3.67, p < 0.001$ ).

When we added the normalized lesion size as a covariate and compared the interactions of perceived intensity and modality for a specific emotion on the target scale, we observed a massive effect



in the visual modality across all emotions: the larger the lesion, the less intensely the patients perceived visual anger ( $z = -2.90$ ,  $p < 0.005$ ), visual joy ( $z = -2.79$ ,  $p < 0.005$ ), and visual fear ( $z = -2.96$ ,  $p < 0.005$ ). Neutral visual stimuli, however, failed to reach significance ( $p > 0.15$ ). This relationship also held good for audiovisual joy ( $z = -2.15$ ,  $p < 0.051$ ), but no significant effects were observed either for other audiovisual expressions (angry, fearful or neutral), or for auditory stimuli ( $p > 0.15$ ).

## DISCUSSION

### CATEGORICAL RESPONSES

Our main goal was to investigate the relationship between emotion and modality, comparing patients who had undergone unilateral anterior temporal lobectomy and amydalectomy with a matched control group. Overall, proportional corrected scores revealed that patients detected joy less accurately across all modalities, in contrast to previous studies postulating that impairments are restricted to negatively valenced stimuli (e.g., Brierley et al., 2004). In addition, the patients displayed deficits for auditory fear and visual anger. The massive effect we observed for decoding joy has several possible explanations. First, this effect could be associated with the amount of information needed for accurate decoding. For instance, Graham et al. (2007) reported that patients were impaired in categorizing emotional faces when these were only presented for a limited duration. In the auditory domain, timing is also a crucial feature for prosody decoding. In healthy individuals, researchers have shown that there is a positive correlation between the duration of the sound and the correct recognition of the vocal stimulus (Pollack et al., 1960; Cornew et al., 2010; Pell

and Kotz, 2011). Furthermore, happy prosody needs a duration of at least 1 s to be decoded accurately (Pell and Kotz, 2011), and our stimuli included 0.25 s fade-ins and fade-outs, thus reducing the amount of available information and its actual duration. The second explanation also concerns a lack of information. In the visual items, the lips were animated to match the intensity contour of each vocal stimulus, even in the unimodal visual condition. As a result, this manipulation may have had an impact on emotion recognition because the information needed to detect a smile was masked by the movement of the lips accompanying the vocalization. More specifically, the visual cues in the mouth region that are needed to detect joy (AU 12 – lip corner puller) and anger (AUs 0 – upper lip raise, 17 – chin raise, 23 – lip funnel, 24 – lip press) were less visible, and thus less salient. Although we expected fear perception accuracy to be poorer among patients than among controls across all the modalities, we found that it was only diminished for auditory stimuli, indicating that unilateral amygdala damage is not sufficient to impair fear recognition in the visual domain. Numerical differences in the confusion matrix (Table 3) suggest that the lack of an effect for visual information stemmed from the fact that fearful faces and faces expressing surprise were confused by both patients (62%) and controls (71%). This confusion between fear and surprise at the visual level is easily explained by the proximity of the AUs used to produce these emotional expressions. In actual fact, they differ by only two AUs: one in the brow region (AU 4 – brow lowerer), the other around the mouth (AU 20 – lip stretcher).

Interestingly, the patients were more accurate than controls in their detection of neutral expressions in both the auditory and

**Table 3 | Confusion matrix.**

Responses	Target emotion											
	Auditory information				Audiovisual information				Visual information			
	Anger	Joy	Neutral	Fear	Anger	Joy	Neutral	Fear	Anger	Joy	Neutral	Fear
<b>Controls</b>												
Anger	<b>62</b>	0	1	10	<b>88</b>	1	3	3	<b>81</b>	2	2	1
Disgust	9	1	5	4	4	0	4	5	8	2	3	1
Joy	3	<b>96</b>	2	3	0	<b>94</b>	2	2	0	<b>79</b>	4	2
Neutral	1	1	<b>75</b>	1	1	1	<b>64</b>	2	1	11	<b>59</b>	11
Fear	21	0	1	<b>70</b>	5	1	2	<b>61</b>	1	2	3	<b>13</b>
Surprise	2	1	16	10	1	2	24	23	8	4	28	71
Ambivalent	2	1	0	2	1	1	1	4	1	0	1	1
<b>Patients</b>												
Anger	<b>68</b>	0	0	10	<b>77</b>	3	3	3	<b>53</b>	2	3	1
Disgust	5	0	0	2	7	2	0	0	10	7	2	2
Joy	5	<b>83</b>	0	0	0	<b>79</b>	2	2	0	<b>40</b>	5	0
Neutral	2	5	<b>95</b>	3	2	10	<b>90</b>	8	10	39	<b>64</b>	27
Fear	7	0	2	<b>50</b>	7	0	2	<b>48</b>	15	2	16	<b>3</b>
Surprise	5	12	2	27	0	3	3	37	10	5	7	62
Ambivalent	8	0	1	8	7	3	0	2	2	5	3	5

Percentage of responses for each target emotion in each modality on the six rating scales. Bold values indicate the percentage of correct responses for an emotion on the target scale. The "ambivalent" response category corresponds to high intensity ratings on more than one scale for the same emotion.

audiovisual modalities. In this experiment, controls may have been biased toward emotional stimuli, in that 75% of items contained emotional information. They were therefore more driven to search for emotional cues in the faces. Assuming that emotion detection plays a functional role, we can surmise that it is less detrimental to identify an object as emotional, than to miss information that could indicate a threat. One can also argue that the patients' emotion detection networks were less activated (expressed behaviorally by emotional blunting) by emotional stimuli, meaning that a neutral item was more likely to be perceived of as non-emotional.

### INTENSITY PERCEPTION

First, controls and patients alike provided lower intensity ratings for visual emotional items than for auditory or audiovisual ones. More specifically, the control group rated visual angry, joyful, and fearful items as significantly less intense, while the patient group gave significantly lower intensity ratings for all the visual items (both emotional and neutral), when lesion size was taken into account. This less intense perception of visual stimuli could be explained by the differing nature of the auditory (real human voices) and visual (synthetic faces created with FACSGen) items. Nevertheless, the control group exhibited specific patterns of intensity perception for auditory and audiovisual items, depending on the emotion. In the case of anger, audiovisual items were perceived of as being more intense than unimodal auditory ones. This could be interpreted as an increase in the perceived potential threat, driven by the redundant information in the bimodal

condition, as we are hard-wired to attribute particular importance to threat-related signals in order to avoid danger more effectively (Marsh et al., 2005). For joy, we observed the opposite pattern, in that auditory joyful items were rated as more intense than audiovisual items. Finally, there was no difference between the intensity ratings provided for auditory and audiovisual fear items, either in the control group or in the patient group. It seems, therefore, that anterior temporal lobe lesions disrupt the processing not just of fear-related stimuli, but also of other emotions in the visual modality. An additional analysis comparing emotional and neutral items showed that patients produced higher intensity ratings for neutral items than for emotional ones, regardless of modality. This effect across modalities lends further weight to the assumption of emotional blunting among these patients. When we compared the groups on emotional and neutral items for each modality, we found that differences only showed up in the auditory and audiovisual modalities, with higher ratings for neutral items provided by patients compared with controls. It is not entirely clear whether the lesions alone were responsible for this effect or whether a more general dysfunction of the epileptic brain was to blame, although the correlations between lesion size and emotional judgments suggest that the lesions themselves had an impact, beyond a general epileptic effect.

The present data indicate that the anterior temporal lobe plays a variety of roles, depending on the modality. First, patients exhibited a greater deficit in intensity perception for the visual modality as a linear function of lesion size for all emotional expressions. This

result highlights an important role of this region at the end of the ventral visual pathway, regardless of the nature of the emotional information. Second, modality had an impact on the ratings provided by the controls for specific emotions. Anger, for instance, was perceived of as more intense in the auditory modality than in the audiovisual or visual ones, while joy was perceived of as more intense in the audiovisual modality than in the two unimodal ones. This emotional modality preference has already been flagged up by Bänziger et al. (2009). Until now, however, it has never been observed in patients. It could be linked to the deficit in the visual pathway mentioned earlier, as no differences were observed between the unimodal auditory condition and the audiovisual one, suggesting that the disruption of the visual processing channel meant that the processing focus had to be switched to the auditory modality. We can therefore hypothesize that our patients' audiovisual processing was impaired as a consequence of a lack of input from the visual pathway toward the anterior temporal lobe. Crossmodal integration in the PRC, an associative area in the anterior temporal lobe that has been highlighted in both animal (e.g., Goulet and Murray, 2001) and human (e.g., Taylor et al., 2006) studies, may therefore play a major role in audiovisual integration.

### INTENSITY PERCEPTION AND LESION EFFECT

We expected the patient with right amygdala damage to exhibit a greater deficit than those with left damage, given that emotion perception decoding appears to be right-lateralized (e.g., Adolphs, 2002; Schirmer and Kotz, 2006). Different deficit patterns were observed, however, depending on emotion and modality. The patient with a right temporal lesion displayed a deficit in auditory and audiovisual fear perception, along with a deficit in visual joy perception, while the left-lesion patients rated joyful prosody and angry visual expressions as less intense. These two last emotions can be seen as *approach* emotions, and BAS scores have been shown to correlate with activity in the left hemisphere (Harmon-Jones and Allen, 1997; Coan and Allen, 2003).

In addition to the lateralization effect, results highlighted a major impact of lesion size, mainly for the recognition and intensity ratings of visual emotional items. This massive visual impairment could be explained by the impact of the resection on part of the visual "what" (ventral) pathway: the absence or disruption of this component of the visual pathway system may have had a greater effect because of the reduced cues for determining expressions in the visual stimuli (i.e., masking by lip movements matched with vocalizations). Based on prior research with animals (Ungerleider and Mishkin, 1982), Catani and Thiebaut de Schotten (2008) showed, using diffusion tensor imaging, that the inferior longitudinal fasciculus, a ventral associative bundle, connects the occipital and temporal lobes (more specifically, the visual areas) to the amygdala. Given that lesion size particularly seemed to affect the visual modality in our study, we can surmise that a compensatory mechanism was at work, whereby the lack of discriminating information in a specific modality triggered a shift toward another modality (see, for example, Bavelier and Hirshorn, 2010).

### LIMITATIONS

The first caveat regarding our experiment concerns the small number of patients, and the fact that only one patient had

undergone a right anterior temporal resection, while another had a larger resection. However, the discrepancy between the number of patients and the number of controls did not impede our statistical analysis, owing to our choice of model and the fact that we tested every model excluding Patients JP or CM to see if we observed any change, which was not the case. The more important point to take into consideration is the difference between the visual and auditory information. The sounds were taken from the GEMEP database, which features real human voices. By contrast, the visual stimuli were non-natural faces (i.e., avatars), and this difference could account for the increased difficulty in labeling the expressions, even though they matched the Ekman coding system (see FACSGen; Roesch et al., 2011).

### CONCLUSION

The results revealed a visual deficit in the perceived intensity of emotional stimuli. This deficit was explained by lesion size, in that the larger the lesion, the lower the intensity ratings for the visual items. This could be caused by disruption to the visual pathway connecting the occipital lobe and the amygdala, but further investigation is needed to test this hypothesis. Furthermore, emotional blunting effects may also have played a part, given that the neutral expressions were given higher intensity ratings by patients than by controls. It would be useful to determine whether the absence of audiovisual enhancement in the patients' perception can be accounted for solely by the amygdala or whether the absence of the PRC, an area that has already been identified as an integrating area in both animals (Goulet and Murray, 2001) and humans (Taylor et al., 2006), is also an important factor.

### ACKNOWLEDGMENTS

We thank the reviewers for their precious comments, as well as Elizabeth Wiles-Portier for preparing the manuscript. This research was partly funded by the National Center of Competence in Research (NCCR) Affective Sciences, financed by the Swiss National Science Foundation (no. 51NF40-104897 – Didier Grandjean), and hosted by the University of Geneva.

### REFERENCES

- Adolphs, R. (2002). Neural systems for recognizing emotion. *Curr. Opin. Neurobiol.* 12, 169–177. doi: 10.1016/S0959-4388(02)00301-X
- Adolphs, R. (2008). Fear, faces, and the human amygdala. *Curr. Opin. Neurobiol.* 18, 166–172. doi: 10.1016/j.conb.2008.06.006
- Adolphs, R., and Tranel, D. (1999). Intact recognition of emotional prosody following amygdala damage. *Neuropsychologia* 37, 1285–1292. doi: 10.1016/S0028-3932(99)00023-8
- Adolphs, R., and Tranel, D. (2004). Impaired judgments of sadness but not happiness following bilateral amygdala damage. *J. Cogn. Neurosci.* 16, 453–462. doi: 10.1162/089892904322926782
- Adolphs, R., Tranel, D., and Damasio, H. (2001). Emotion recognition from faces and prosody following temporal lobectomy. *Neuropsychology* 15, 396–404. doi: 10.1037/0894-4105.15.3.396
- Adolphs, R., Tranel, D., Damasio, H., and Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature* 372, 669–672. doi: 10.1038/372669a0
- Adolphs, R., Tranel, D., Damasio, H., and Damasio, A. R. (1995). Fear and the human amygdala. *J. Neurosci.* 15, 5879–5891.
- Amaral, D. G., Behniea, H., and Kelly, J. L. (2003). Topographic organization of projections from the amygdala to the visual cortex in the

- macaque monkey. *Neuroscience* 118, 1099–1120. doi: 10.1016/S0306-4522(02)01001-1
- Anderson, A. K., and Phelps, E. A. (1998). Intact recognition of vocal expressions of fear following bilateral lesions of the human amygdala. *Neuroreport* 9, 3607–3613. doi: 10.1093/neucas/6.6.442
- Anderson, A. K., Spencer, D. D., Fulbright, R. K., and Phelps, E. A. (2000). Contribution of the anteromedial temporal lobes to the evaluation of facial emotion. *Neuropsychology* 14, 526–536. doi: 10.1037/0894-4105.14.4.526
- Andersen, S. M., Rapcsak, S. Z., and Beeson, P. M. (2010). Cost function masking during normalization of brains with focal lesions: still a necessity? *Neuroimage* 53, 78–84. doi: 10.1016/j.neuroimage.2010.06.003
- Armony, J. L. (2013). Current emotion research in behavioral neuroscience: the role (s) of the amygdala. *Emot. Rev.* 5, 104–115. doi: 10.1177/1754073912457208
- Armony, J. L., and Ledoux, J. E. (1997). How the brain processes emotional information. *Ann. N. Y. Acad. Sci.* 821, 259–270. doi: 10.1111/j.1749-6632.1997.tb48285.x
- Armony, J. L., and Ledoux, J. E. (1999). “How danger is encoded: towards a systems, cellular, and computational understanding of cognitive-emotional interactions,” in *The New Cognitive Neuroscience*, ed. M. S. Gazzaniga (Cambridge, MA: MIT Press), 1067–1079.
- Ashburner, J., and Friston, K. J. (2005). Unified segmentation. *Neuroimage* 26, 839–851. doi: 10.1016/j.neuroimage.2005.02.018
- Banse, R., and Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *J. Pers. Soc. Psychol.* 70, 614–636. doi: 10.1037/0022-3514.70.3.614
- Bänziger, T., Grandjean, D., and Scherer, K. R. (2009). Emotion recognition from expressions in face, voice, and body: the Multimodal Emotion Recognition Test (MERT). *Emotion* 9, 691–704. doi: 10.1037/a0017088
- Bänziger, T., Mortillaro, M., and Scherer, K. R. (2012). Introducing the Geneva multimodal expression corpus for experimental research on emotion perception. *Emotion* 12, 1161–1179. doi: 10.1037/a0025827
- Batut, A. C., Gounot, D., Namer, I. J., Hirsch, E., Kehrl, P., and Metz-Lutz, M. N. (2006). Neural responses associated with positive and negative emotion processing in patients with left versus right temporal lobe epilepsy. *Epilepsy Behav.* 9, 415–423. doi: 10.1016/j.yebeh.2006.07.013
- Bavelier, D., and Hirshorn, E. A. (2010). I see where you’re hearing: how cross-modal plasticity may exploit homologous brain structures. *Nat. Neurosci.* 13, 1309–1311. doi: 10.1038/nn1110-1309
- Belin, P., Fillion-Bilodeau, S., and Gosselin, F. (2008). The montreal affective voices: a validated set of non-verbal affect bursts for research on auditory affective processing. *Behav. Res. Methods* 40, 531–539. doi: 10.3758/BRM.40.2.531
- Boersma, P., and Weenink, D. (2011). “Praat: doing phonetics by computer [Computer program]” 5.2.37. Available at: <http://www.praat.org/> (accessed September 2, 2011).
- Breiter, H. C., Etcoff, N. L., Whalen, P. J., Kennedy, W. A., Rauch, S. L., Buckner, R. L., et al. (1996). Response and habituation of the human amygdala during visual processing of facial expression. *Neuron* 17, 875–887. doi: 10.1016/S0896-6273(00)80219-6
- Brierley, B., Medford, N., Shaw, P., and David, A. S. (2004). Emotional memory and perception in temporal lobectomy patients with amygdala damage. *J. Neurol. Neurosurg. Psychiatry* 75, 593–599. doi: 10.1136/jnnp.2002.006403
- Broks, P., Young, A. W., Maratos, E. J., Coffey, P. J., Calder, A. J., Isaac, C. L., et al. (1998). Face processing impairments after encephalitis: amygdala damage and recognition of fear. *Neuropsychologia* 36, 59–70. doi: 10.1016/S0028-3932(97)00105-X
- Cahill, L., Weinberger, N. M., Roozendaal, B., and McGaugh, J. L. (1999). Is the amygdala a locus of “conditioned fear?” Some questions and caveats. *Neuron* 23, 227–228. doi: 10.1016/S0896-6273(00)80774-6
- Calder, A. J., and Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nat. Rev. Neurosci.* 6, 641–651. doi: 10.1038/nrn1724
- Calder, A. J., Young, A. W., Rowland, D., Perrett, D. I., Hodges, J. R., and Etcoff, N. L. (1996). Facial emotion recognition after bilateral amygdala damage: differentially severe impairment of fear. *Cogn. Neuropsychol.* 13, 699–745. doi: 10.1080/026432996381890
- Catani, M., and Thiebaut de Schotten, M. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex* 44, 1105–1132. doi: 10.1016/j.cortex.2008.05.004
- Coan, J. A., and Allen, J. J. (2003). Varieties of emotional experience during voluntary emotional facial expressions. *Ann. N. Y. Acad. Sci.* 1000, 375–379. doi: 10.1196/annals.1280.034
- Cornew, L., Carver, L., and Love, T. (2010). There’s more to emotion than meets the eye: a processing bias for neutral content in the domain of emotional prosody. *Cogn. Emot.* 24, 1133–1152. doi: 10.1080/02699930903247492
- Cristinzio, C., Sander, D., and Vuilleumier, P. (2007). Recognition of emotional face expressions and amygdala pathology. *Epileptologie* 24, 130–138.
- De Gelder, B., and Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cogn. Emot.* 14, 289–311. doi: 10.1080/026999300378824
- Dellacherie, D., Hasboun, D., Baulac, M., Belin, P., and Samson, S. (2011). Impaired recognition of fear in voices and reduced anxiety after unilateral temporal lobe resection. *Neuropsychologia* 49, 618–629. doi: 10.1016/j.neuropsychologia.2010.11.008
- Dolan, R. J., Morris, J. S., and De Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proc. Natl. Acad. Sci. U.S.A.* 98, 10006–10010. doi: 10.1073/pnas.171288598
- Ekman, P., and Friesen, W. V. (1978). *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto, CA: Consulting Psychologists Press.
- Etcoff, N. L., and Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition* 44, 227–240. doi: 10.1016/0010-0277(92)90002-Y
- Ethofer, T., Anders, S., Wiethoff, S., Erb, M., Herbert, C., Saur, R., et al. (2006). Effects of prosodic emotional intensity on activation of associative auditory cortex. *Neuroreport* 17, 249–253. doi: 10.1097/01.wnr.0000199466.32036.5d
- Ethofer, T., De Ville, D. V., Scherer, K., and Vuilleumier, P. (2009). Decoding of emotional information in voice-sensitive cortices. *Curr. Biol.* 19, 1028–1033. doi: 10.1016/j.cub.2009.04.054
- Fecteau, S., Belin, P., Joannette, Y., and Armony, J. L. (2007). Amygdala responses to non-linguistic emotional vocalizations. *Neuroimage* 36, 480–487. doi: 10.1016/j.neuroimage.2007.02.043
- Frühholz, S., Ceravolo, L., and Grandjean, D. (2012). Specific brain networks during explicit and implicit decoding of emotional prosody. *Cereb. Cortex* 22, 1107–1117. doi: 10.1093/cercor/bhr184
- Frühholz, S., and Grandjean, D. (2012). Towards a fronto-temporal neural network for the decoding of angry vocal expressions. *Neuroimage* 62, 1658–1666. doi: 10.1016/j.neuroimage.2012.06.015
- Frühholz, S., and Grandjean, D. (2013). Amygdala subregions differentially respond and rapidly adapt to threatening voices. *Cortex* 49, 1394–1403. doi: 10.1016/j.cortex.2012.08.003
- Gosselin, N., Peretz, I., Hasboun, D., Baulac, M., and Samson, S. (2011). Impaired recognition of musical emotions and facial expressions following anteromedial temporal lobe excision. *Cortex* 47, 1116–1125. doi: 10.1016/j.cortex.2011.05.012
- Gosselin, N., Peretz, I., Johnsen, E., and Adolphs, R. (2007). Amygdala damage impairs emotion recognition from music. *Neuropsychologia* 45, 236–244. doi: 10.1016/j.neuropsychologia.2006.07.012
- Goulet, S., and Murray, E. A. (2001). Neural substrates of crossmodal association memory in monkeys: the amygdala versus the anterior rhinal cortex. *Behav. Neurosci.* 115, 271–284. doi: 10.1037/0735-7044.115.2.271
- Graham, R., Devinsky, O., and Labar, K. S. (2007). Quantifying deficits in the perception of fear and anger in morphed facial expressions after bilateral amygdala damage. *Neuropsychologia* 45, 42–54. doi: 10.1016/j.neuropsychologia.2006.04.021
- Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., et al. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nat. Neurosci.* 8, 145–146. doi: 10.1038/nn1392
- Harmon-Jones, E., and Allen, J. J. (1997). Behavioral activation sensitivity and resting frontal EEG asymmetry: covariation of putative indicators related to risk for mood disorders. *J. Abnorm. Psychol.* 106, 159–163. doi: 10.1037/0021-843X.106.1.159
- Haxby, J. V., Hoffman, E. A., and Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends Cogn. Sci.* 4, 223–233. doi: 10.1016/S1364-6613(00)01482-0

- Heberlein, A. S., Adolphs, R., Tranel, D., and Damasio, H. (2004). Cortical regions for judgments of emotions and personality traits from point-light walkers. *J. Cogn. Neurosci.* 16, 1143–1158. doi: 10.1162/0898929041920423
- Kanwisher, N., and Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 361, 2109–2128. doi: 10.1098/rstb.2006.1934
- Klasen, M., Kenworthy, C. A., Mathiak, K. A., Kircher, T. T., and Mathiak, K. (2011). Supramodal representation of emotions. *J. Neurosci.* 31, 13635–13643. doi: 10.1523/JNEUROSCI.2833-11.2011
- Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., and Wildgruber, D. (2007). Audiovisual integration of emotional signals in voice and face: an event-related fMRI study. *Neuroimage* 37, 1445–1456. doi: 10.1016/j.neuroimage.2007.06.020
- Ledoux, J. (2007). The amygdala. *Curr. Biol.* 17, 868–874. doi: 10.1016/j.cub.2007.08.005
- Marsh, A. A., Ambady, N., and Kleck, R. E. (2005). The effects of fear and anger facial expressions on approach- and avoidance-related behaviors. *Emotion* 5, 119–124. doi: 10.1037/1528-3542.5.1.119
- Massaro, D. W., and Egan, P. B. (1996). Perceiving affect from the voice and the face. *Psychon. Bull. Rev.* 3, 215–221. doi: 10.3758/BF03212421
- McClelland, S. III, Garcia, R. E., Peraza, D. M., Shih, T. T., Hirsch, L. J., Hirsch, J., et al. (2006). Facial emotion recognition after curative non-dominant temporal lobectomy in patients with mesial temporal sclerosis. *Epilepsia* 47, 1337–1342. doi: 10.1111/j.1528-1167.2006.00557.x
- Meletti, S., Benuzzi, F., Rubboli, G., Cantalupo, G., Stanzani Maserati, M., Nichelli, P., et al. (2003). Impaired facial emotion recognition in early-onset right mesial temporal lobe epilepsy. *Neurology* 60, 426–431. doi: 10.1212/WNL.60.3.426
- Miller, J. (1982). Divided attention: evidence for coactivation with redundant signals. *Cogn. Psychol.* 14, 247–279. doi: 10.1016/0010-0285(82)9010-X
- Mitchell, R. L. C., Elliott, R., Barry, M., Cruttenden, A., and Woodruff, P. W. R. (2003). The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia* 41, 1410–1421. doi: 10.1016/S0028-3932(03)00017-4
- Morris, J. S., Friston, K. J., Buchel, C., Frith, C. D., Young, A. W., Calder, A. J., et al. (1998). A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain* 121(Pt 1), 47–57. doi: 10.1093/brain/121.1.47
- Morris, J. S., Scott, S. K., and Dolan, R. J. (1999). Saying it with feeling: neural responses to emotional vocalizations. *Neuropsychologia* 37, 1155–1163. doi: 10.1016/S0028-3932(99)00015-9
- Morrison, S. E., and Salzman, C. D. (2010). Re-valuing the amygdala. *Curr. Opin. Neurobiol.* 20, 221–230. doi: 10.1016/j.conb.2010.02.007
- Müller, V. I., Cieslik, E. C., Turetsky, B. I., and Eickhoff, S. B. (2012). Crossmodal interactions in audiovisual emotion processing. *Neuroimage* 60, 553–561. doi: 10.1016/j.neuroimage.2011.12.007
- Murray, E. A. (2007). The amygdala, reward and emotion. *Trends Cogn. Sci.* 11, 489–497. doi: 10.1016/j.tics.2007.08.013
- Palermo, R., Schmalzl, L., Mohamed, A., Bleasel, A., and Miller, L. (2010). The effect of unilateral amygdala removals on detecting fear from briefly presented backward-masked faces. *J. Clin. Exp. Neuropsychol.* 32, 123–131. doi: 10.1080/13803390902821724
- Pell, M. D., and Kotz, S. A. (2011). On the time course of vocal emotion recognition. *PLoS ONE* 6:e27256. doi: 10.1371/journal.pone.0027256
- Phan, K. L., Wager, T., Taylor, S. F., and Liberzon, I. (2002). Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage* 16, 331–348. doi: 10.1006/nimg.2002.1087
- Phelps, E. A., and LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron* 48, 175–187. doi: 10.1016/j.neuron.2005.09.025
- Plichta, M. M., Gerdes, A. B., Alpers, G., Harnisch, W., Brill, S., Wieser, M., et al. (2011). Auditory cortex activation is modulated by emotion: a functional near-infrared spectroscopy (fNIRS) study. *Neuroimage* 55, 1200–1207. doi: 10.1016/j.neuroimage.2011.01.011
- Pollack, I., Rubenstein, H., and Horowitz, A. (1960). Communication of verbal modes of expression. *Lang. Speech* 3, 121–130. doi: 10.1177/002383096000300301
- Pourtois, G., Dan, E. S., Grandjean, D., Sander, D., and Vuilleumier, P. (2005a). Enhanced extrastriate visual response to bandpass spatial frequency filtered fearful faces: time course and topographic evoked-potentials mapping. *Hum. Brain Mapp.* 26, 65–79. doi: 10.1002/hbm.20130
- Pourtois, G., De Gelder, B., Bol, A., and Crommelinck, M. (2005b). Perception of facial expressions and voices and of their combination in the human brain. *Cortex* 41, 49–59.
- Pourtois, G., Schettino, A., and Vuilleumier, P. (2013). Brain mechanisms for emotional influences on perception and attention: what is magic and what is not. *Biol. Psychol.* 92, 492–512. doi: 10.1016/j.biopsycho.2012.02.007
- Pourtois, G., Spinelli, L., Seeck, M., and Vuilleumier, P. (2010). Modulation of face processing by emotional expression and gaze direction during intracranial recordings in right fusiform cortex. *J. Cogn. Neurosci.* 22, 2086–2107. doi: 10.1162/jocn.2009.21404
- Roesch, E. B., Tamarit, L., Reveret, L., Grandjean, D., Sander, D., and Scherer, K. R. (2011). FACSGen: a tool to synthesize emotional facial expressions through systematic manipulation of facial action units. *J. Nonverbal Behav.* 35, 1–16. doi: 10.1007/s10919-010-0095-9
- Sabatinelli, D., Bradley, M. M., Fitzsimmons, J. R., and Lang, P. J. (2005). Parallel amygdala and inferotemporal activation reflect emotional intensity and fear relevance. *Neuroimage* 24, 1265–1270. doi: 10.1016/j.neuroimage.2004.12.015
- Sander, D., Grafman, J., and Zalla, T. (2003). The human amygdala: an evolved system for relevance detection. *Rev. Neurosci.* 14, 303–316. doi: 10.1515/REVNEURO.2003.14.4.303
- Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., et al. (2005). Emotion and attention interactions in social cognition: brain regions involved in processing anger prosody. *Neuroimage* 28, 848–858. doi: 10.1016/j.neuroimage.2005.06.023
- Schirmer, A., Fox, P. M., and Grandjean, D. (2012). On the spatial organization of sound processing in the human temporal lobe: a meta-analysis. *Neuroimage* 63, 137–147. doi: 10.1016/j.neuroimage.2012.06.025
- Schirmer, A., and Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends Cogn. Sci.* 10, 24–30. doi: 10.1016/j.tics.2005.11.009
- Schröger, E., and Widmann, A. (1998). Speeded responses to audiovisual signal changes result from bimodal integration. *Psychophysiology* 35, 755–759. doi: 10.1111/1469-8986.3560755
- Scott, S. K., Young, A. W., Calder, A. J., Hellowell, D. J., Aggleton, J. P., and Johnson, M. (1997). Impaired auditory recognition of fear and anger following bilateral amygdala lesions. *Nature* 385, 254–257. doi: 10.1038/385254a0
- Sprengelmeyer, R., Atkinson, A. P., Sprengelmeyer, A., Mair-Walther, J., Jacobi, C., Wildemann, B., et al. (2010). Disgust and fear recognition in paraneoplastic limbic encephalitis. *Cortex* 46, 650–657. doi: 10.1016/j.cortex.2009.04.007
- Sprengelmeyer, R., Young, A. W., Schroeder, U., Gossenbacher, P. G., Federlein, J., Buttner, T., et al. (1999). Knowing no fear. *Proc. Biol. Sci.* 266, 2451–2456. doi: 10.1098/rspb.1999.0945
- Taylor, K. I., Moss, H. E., Stamatakis, E. A., and Tyler, L. K. (2006). Binding cross-modal object features in perirhinal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 103, 8239–8244. doi: 10.1073/pnas.0509704103
- Ungerleider, L. G., and Mishkin, M. (1982). “Two cortical visual systems,” in *Analysis of Visual Behavior*, eds D. J. Ingle, M. Goodale, and R. J. W. Mansfield (Cambridge, MA: MIT Press), 549–586.
- Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends Cogn. Sci.* 9, 585–594. doi: 10.1016/j.tics.2005.10.011
- Vuilleumier, P., Armony, J. L., Driver, J., and Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron* 30, 829–841. doi: 10.1016/S0896-6273(01)00328-2
- Vuilleumier, P., and Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. *Neuropsychologia* 45, 174–194. doi: 10.1016/j.neuropsychologia.2006.06.003
- Vuilleumier, P., Richardson, M. P., Armony, J. L., Driver, J., and Dolan, R. J. (2004). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nat. Neurosci.* 7, 1271–1278. doi: 10.1038/nn1341

- Whalen, P. J. (1998). Fear, vigilance, and ambiguity: initial neuroimaging studies of the human amygdala. *Curr. Dir. Psychol. Sci.* 7, 177–188.
- Wiethoff, S., Wildgruber, D., Grodd, W., and Ethofer, T. (2009). Response and habituation of the amygdala during processing of emotional prosody. *Neuroreport* 20, 1356–1360. doi: 10.1097/WNR.0b013e328330eb83
- Wildgruber, D., Ackermann, H., Kreifelts, B., and Ethofer, T. (2006). Cerebral processing of linguistic and emotional prosody: fMRI studies. *Prog. Brain Res.* 156, 249–268. doi: 10.1016/S0079-6123(06)56013-3
- Williams, M. A., Moss, S. A., and Bradshaw, J. L. (2004). A unique look at face processing: the impact of masked faces on the processing of facial features. *Cognition* 91, 155–172. doi: 10.1016/j.cognition.2003.08.002
- Young, A. W., Aggleton, J. P., Hellawell, D. J., Johnson, M., Brooks, P., and Hanley, J. R. (1995). Face processing impairments after amygdalotomy. *Brain* 118(Pt 1), 15–24. doi: 10.1093/brain/118.1.15
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 June 2013; accepted: 14 April 2014; published online: 05 May 2014.

Citation: Milesi V, Cekic S, Péron J, Frühholz S, Cristinzio C, Seeck M and Grandjean D (2014) Multimodal emotion perception after anterior temporal lobectomy (ATL). *Front. Hum. Neurosci.* 8:275. doi: 10.3389/fnhum.2014.00275

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2014 Milesi, Cekic, Péron, Frühholz, Cristinzio, Seeck and Grandjean. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Emotional valence and spatial congruency differentially modulate crossmodal processing: an fMRI study

Dhana Wolf<sup>1,2,3</sup> \*†, Lisa Schock<sup>1,2,3†</sup>, Saurabh Bhavsar<sup>1,2,3</sup>, Liliana R. Demenescu<sup>1,2,3</sup>, Walter Sturm<sup>2,4</sup> and Klaus Mathiak<sup>1,2,3</sup>

<sup>1</sup> Department of Psychiatry, Psychotherapy and Psychosomatics, Medical School, RWTH Aachen University, Aachen, Germany

<sup>2</sup> Interdisciplinary Centre for Clinical Research, Medical School, RWTH Aachen University, Aachen, Germany

<sup>3</sup> JARA–Translational Brain Medicine, Research Centre Jülich, Jülich, Aachen, Germany

<sup>4</sup> Department of Neurology, Clinical Neuropsychology, Medical School, RWTH Aachen University, Aachen, Germany

## Edited by:

Benjamin Kreifelts, University of Tübingen, Germany

## Reviewed by:

Marianne Latinus, Aix-Marseille Université, France

Ralf Veit, Institute of Medical Psychology, Germany

## \*Correspondence:

Dhana Wolf, Department of Psychiatry, Psychotherapy and Psychosomatics, Medical School, RWTH Aachen University, Pauwelsstr. 30, D-52074 Aachen, Germany  
e-mail: dhwolf@ukaachen.de

† These authors have contributed equally to this work.

Salient exogenous stimuli modulate attentional processes and lead to attention shifts—even across modalities and at a pre-attentive level. Stimulus properties such as hemispheric laterality and emotional valence influence processing, but their specific interaction in audio-visual attention paradigms remains ambiguous. We conducted an fMRI experiment to investigate the interaction of supramodal spatial congruency, emotional salience, and stimulus presentation side on neural processes of attention modulation. Emotionally neutral auditory deviants were presented in a dichotic listening oddball design. Simultaneously, visual target stimuli (schematic faces) were presented, which displayed either a negative or a positive emotion. These targets were presented in the left or in the right visual field and were either spatially congruent (valid) or incongruent (invalid) with the concurrent deviant auditory stimuli. According to our expectation we observed that deviant stimuli serve as attention-directing cues for visual target stimuli. Region-of-interest (ROI) analyses suggested differential effects of stimulus valence and spatial presentation on the hemodynamic response in bilateral auditory cortices. These results underline the importance of valence and presentation side for attention guidance by deviant sound events and may hint at a hemispheric specialization for valence and attention processing.

**Keywords: hemodynamic mismatch response, spatial congruency, crossmodal spatial cueing, attention, valence, auditory cortex**

## INTRODUCTION

For an efficient interaction with the environment, information from various sensory modalities is integrated into a unified spatial representation. Since the processing capacity of perceptual input is limited, spatial attention needs to be selectively allocated to relevant stimuli. The importance of such a mechanism is reflected, for instance, in the huge impact of impairments of attention in the left visual field—unilateral neglect—on simple activities of daily living in patients with right-hemisphere damage (for a review see Danckert and Ferber, 2006). The allocation of attention is guided by the processing and evaluation of information from various sensory modalities as well as top-down control mechanisms. The underlying neural mechanisms of attention allocation are modulated by different stimulus properties such as presentation side and emotional content.

## SPATIAL CUEING AND SPATIAL CONGRUENCY

Across modalities, attention networks interact at a supramodal level to guide attention to the most important stimulus in the environment. Attention distribution mechanisms modulate processing in the target-modality sensory cortices and also in other, usually to be ignored modalities. This was demonstrated for audio-visual, tactile-visual, and tactile-auditory interactions (for

a review, see Eimer and Driver, 2001). Thus, attention shifts following exogenous cues influence target detection even across modalities (Eimer and Driver, 2001; McDonald et al., 2003; Menning et al., 2005).

When a sensory stimulation is presented on the same side as the target stimulus, this valid cue reduces reaction time. When the stimulation is presented on the opposite side (invalid cue), attention is directed away from the target location, leading to increased reaction time (Posner, 1980). The auditory oddball paradigm is a powerful mean to investigate this cueing effect at a pre-attentive level. The paradigm comprises deviant auditory stimuli presented in a rapid sequence of frequent standard stimuli. Deviants of any type implemented in a series of standard events constitute a violation of the established pattern and elicit a mismatch response in the auditory cortex, triggering attentional shifts to the side of the deviant (Schröger, 1996; Schock et al., 2013). Schröger et al. combined an auditory oddball design presented on a to-be-ignored ear with an auditory GoNogo-task on the other ear. They found a prolonged reaction time to the task when it was preceded by a deviant sound on the ignored ear. This was interpreted as a shift of attention induced by the deviant sound.

This oddball-induced mismatch response is a well-established tool in the investigation of neural responses (EEG: Näätänen et al.,

1989; Schröger, 1996; Näätänen et al., 2004; MEG: Mathiak et al., 2005; fMRI: Mathiak et al., 2002; Schock et al., 2012). These studies reported that even non-attended changes in the auditory stream induce increased activation of the auditory cortex. Further, this effect is modulated by attention (Alho et al., 1999; Zvyagintsev et al., 2011), cross-modal information (Calvert and Campbell, 2003; Kayser et al., 2005, 2007; Zvyagintsev et al., 2009) and top-down regulation (van Atteveldt et al., 2007; Zvyagintsev et al., 2009; Hsieh et al., 2012).

### THE INFLUENCE OF VALENCE ON ATTENTION DISTRIBUTION

Emotional valence, an important factor contributing to the salience of stimuli, influences attention-related processing. Negative valence draws attention effectively and enhances stimulus processing (Eastwood et al., 2001; Fenske and Eastwood, 2003; Rowe et al., 2007). A negative stimulus can narrow the attentional focus accompanied by enhanced processing. In visual search, negative stimuli are found faster and with fewer errors than positive stimuli (Eastwood et al., 2001). In a preceding behavioral study, Schock et al. (2013) combined schematic faces—representing emotionally salient target stimuli—with dichotic syllables presented in an oddball paradigm. Visual and auditory stimuli were presented on the same side (spatially congruent) or on opposite sides (spatially incongruent). Prolonged response time was found to left-lateralized spatially incongruent stimuli if they were of positive valence, while emotionally negative targets were not affected by spatial incongruency. This behavioral effect may be induced by pre-attentive auditory processing.

### AIMS AND HYPOTHESES

Although neural mechanisms of attention have been extensively studied, the influence of emotional valence, hemispheric lateralization, and their specific interactions on early auditory processing, remain ambiguous. The goal of the present study was to investigate the interaction of emotional valence, visual field, and spatial congruency on auditory cortex activation. For this aim, we conducted an fMRI experiment combining a dichotic listening oddball design with emotional visual targets. We hypothesized that deviant sounds would trigger a shift of attention towards the side of the deviant and thereby accelerate response time to spatially congruent visual stimuli, while response time to spatially incongruent stimuli would be increased. Emotional salience was expected to interact with spatial congruency, in that negative salience would reduce this reaction time costs in spatially incongruent trials. Regarding the neurophysiological data, we hypothesized that auditory cortex responses to deviant sounds would be modulated by spatial congruency with, and emotion of, the target faces. Furthermore, attention and salience-related networks were expected to respond stronger to spatially incongruent than to congruent stimuli.

### MATERIALS AND METHODS

We conducted an fMRI experiment combining a dichotic listening oddball design with emotional visual targets. A deviant stimulus would induce an attention shift to the side of the deviant. Visual target stimuli were schematic faces that exclusively displayed the feature of positive and negative face expressions. They were

presented in the left or right visual field either spatially congruent (valid) or incongruent (invalid) with the auditory deviants. Participants were instructed to ignore the auditory modality and to indicate the detection of a visual target stimulus via button press. We have successfully applied this paradigm in a previous behavioral study and demonstrated significant interactions of valence and visual field and of valence and spatial congruency on reaction time (Schock et al., 2013).

### SUBJECTS

Sixteen healthy volunteers (8 females, age 21–36 years, mean age  $24.9 \pm 4.5$  years) participated in the study. Fifteen subjects were right-handed, as indicated by the laterality quotient of the Edinburgh Inventory (Oldfield, 1971) and one subject was ambidextrous. Subjects were screened with the Structured Clinical Interview (SCID-I; Wittchen et al., 1997) for the diagnostic and Statistical Manual of Mental Disorders (DSM-IV) to exclude subjects with psychiatric disorders. The study was approved by the local Ethics Committee of the Medical School of the RWTH Aachen University and was conducted in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Informed consent was obtained from all subjects.

### STIMULI

#### *Auditory stimuli*

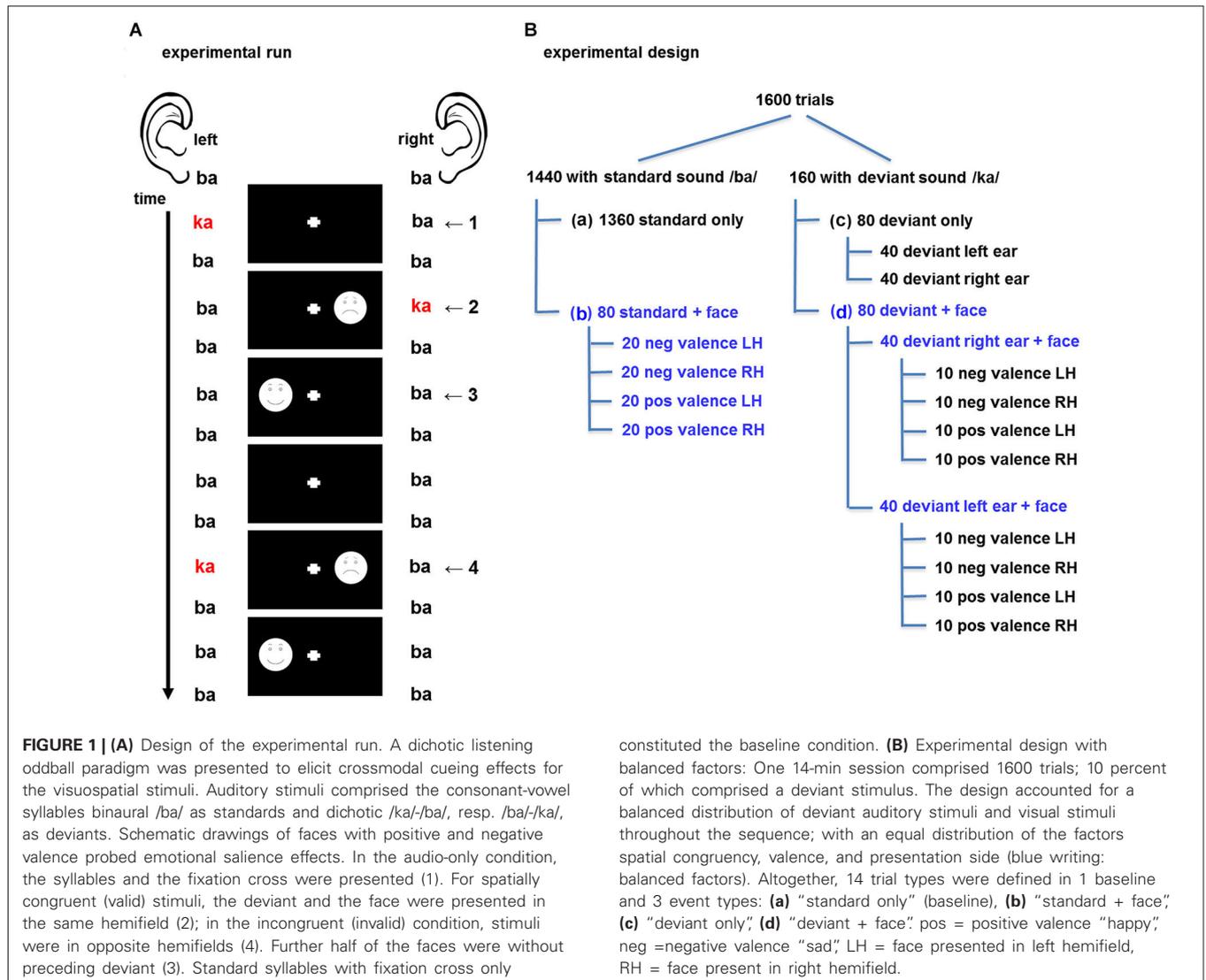
The stimuli comprised two consonant-vowel syllables (/ba/ and /ka/) presented in a dichotic listening oddball design. In 90% of the presentations the standard /ba/ was presented to both ears. In the remaining 10% the deviant /ka/ was presented to one ear (5% to the left ear and 5% to the right ear) and the standard /ba/ to the other ear (left/ba/-right/ka/ and left/ka/-right/ba/). The syllables were delivered synchronous over both ears, with equal sound level and duration (300 ms), and with a constant stimulus onset asynchrony (SOA) of 500 ms.

#### *Visual stimuli*

Visual stimuli were composed of schematic drawings of faces with either a positive or negative emotional expression. The positive expression included a half-elliptic mouth shape and round-shaped eyebrows. The negative expression was achieved by inverting the mouth downwards and orienting the inner eyebrow ends upwards. These expressions were validated beforehand and were rated as either happy or sad (see Schock et al., 2013). For trials including a face, the visual stimuli were presented 175 ms after auditory stimulus onset and lasted for 150 ms.

### DESIGN

The experiment comprised three sessions of 14 min each. In each session, 1600 trials were presented in a rapid, event-related design. A trial was defined as the presentation of one auditory stimulus with or without a face stimulus, so that the length of each trial corresponded to the auditory stimuli onset asynchrony of 500 ms (Figure 1A). Audio-visual trial types were specified by the factors *valence* (positive vs. negative facial expression), *visual field* (left vs. right visual field), and *ear* (left- vs. right-ear deviant). In total 14 events of interest emerged: 4 times



face and standard sound (face left/right, positive/negative), 2 distinct deviants (left/right), and 8 events comprising a face and a deviant (combination of the previous factors). These trial types are displayed in **Figure 1B**. Each session comprised the baseline condition with 1440 trials with “standard only” event types (standard auditory event without visual stimulus; **Figure 1B-a**), 80 trials with the four “standard + face” event types (20 each, **Figure 1B-b**), 80 trials with two “deviant only” event types (40 each, **Figure 1B-c**) and 80 trials with eight “deviant + face” event types (10 each, **Figure 1B-d**). In 50% of the cases the auditory and visual stimulus were presented on the same side (spatially congruent, “valid”) and in the other 50% the stimuli were presented on the opposite side (spatially incongruent, “invalid”; see **Figure 1A**).

The design accounted for a balanced distribution of auditory deviants and visual stimuli throughout the oddball sequence. A minimum of two standards (left/ba/-right/ba/) were placed between two deviants (left/ba/-right/ka/ or left/ka/-right/ba/).

Likewise, at least two trials without visual stimulation were presented between two trials with face presentation to prevent a button press overlap. Each session was individually randomized: first the deviant distribution was randomized. Then, left-ear and right-ear deviant events were randomly selected and assigned to the “deviant + face” events and in a third step the standard events were randomly selected and assigned to the “standard + face” events.

Participants were instructed to ignore the sound and report detection of the face stimulus via button pressing. Visual stimuli were displayed via MR-compatible video goggles (Visua-StimDigital, Resonance Technology, RT, Northridge, CA, USA) and sounds were delivered with MR-compatible headphones. Sound levels were adjusted to comfortable hearing level and good audibility during scanning. Stimulus presentation and response time recording was performed using the software *Presentation* (Version 10.0; Neurobehavioral Systems, Inc., Albany, CA).

## MR IMAGING

Functional imaging was conducted on a 3T Magnetom Trio MR scanner (Siemens Medical Systems, Erlangen, Germany) in the department of Psychiatry, Psychotherapy and Psychosomatics at the Medical School of RWTH Aachen University. Functional images were collected with echo planar imaging (EPI) sensitive to blood oxygenation level dependent (BOLD) contrast (interleaved acquisition of 34 slices; repetition time [TR] = 2000 ms; echo time [TE] = 28 ms; flip angle [FA] = 77°; slice thickness = 3 mm; gap 0.75 mm; matrix size = 64 × 64; field of view [FOV] = 192 × 192 mm<sup>2</sup>; voxel size = 3 × 3 × 3 mm<sup>3</sup>). Slices covered the entire cerebral cortex and were positioned oblique-transversally to achieve maximal brain coverage. 420 volumes were collected per session, of which the first seven were discarded to remove the influence of T1 saturation effects. Head movement was minimized with the use of foam wedges to securely hold the head in the 12-channel head coil. Structural images were obtained using a high-resolution T1-weighted 3D-sequence (TR = 1900 ms; inversion time TI = 900 ms; TE = 2.52 ms; FA = 9°; FOV = 256 × 256 mm<sup>2</sup>; 176 3D-partitions with an isotropic resolution of 1 mm<sup>2</sup>).

## BEHAVIORAL DATA ANALYSIS

Button presses in response to visual stimuli were analyzed by subtracting the onset of the visual stimulus from the onset of button press. The mean reaction time (in ms) was calculated for each event type and each participant. A 2 × 2 × 2 repeated-measures ANOVA with the factors *valence* (positive, negative emotion), *visual field* (left, right), and *spatial congruency* (auditory and visual presentation on the same/opposite side) was conducted. Significance level was set at  $p < 0.05$  after Bonferroni correction. Paired  $t$ -tests disentangled the effects *post-hoc*.

## IMAGING DATA ANALYSIS

Functional MRI data analysis was conducted with the Statistical Parametric Mapping software (SPM8<sup>1</sup>; implemented in MATLAB) (MathWorks, Natick, MA, USA). After discarding the first seven volumes, 413 volumes per session from each participant were spatially realigned to the mean image to correct for head movement, normalized to the stereotaxic anatomical MNI (Montreal Neurological Institute) space with 2 mm isotropic voxels, and spatially smoothed with an 8 mm (FWHM) isotropic Gaussian kernel to account for inter-subject variability in brain anatomy and to increase signal-to-noise ratio. A rapid event-related design was chosen to model the experimental conditions (14 audio-visual events, see Section Design) with a general linear model (GLM).

### Whole brain analysis

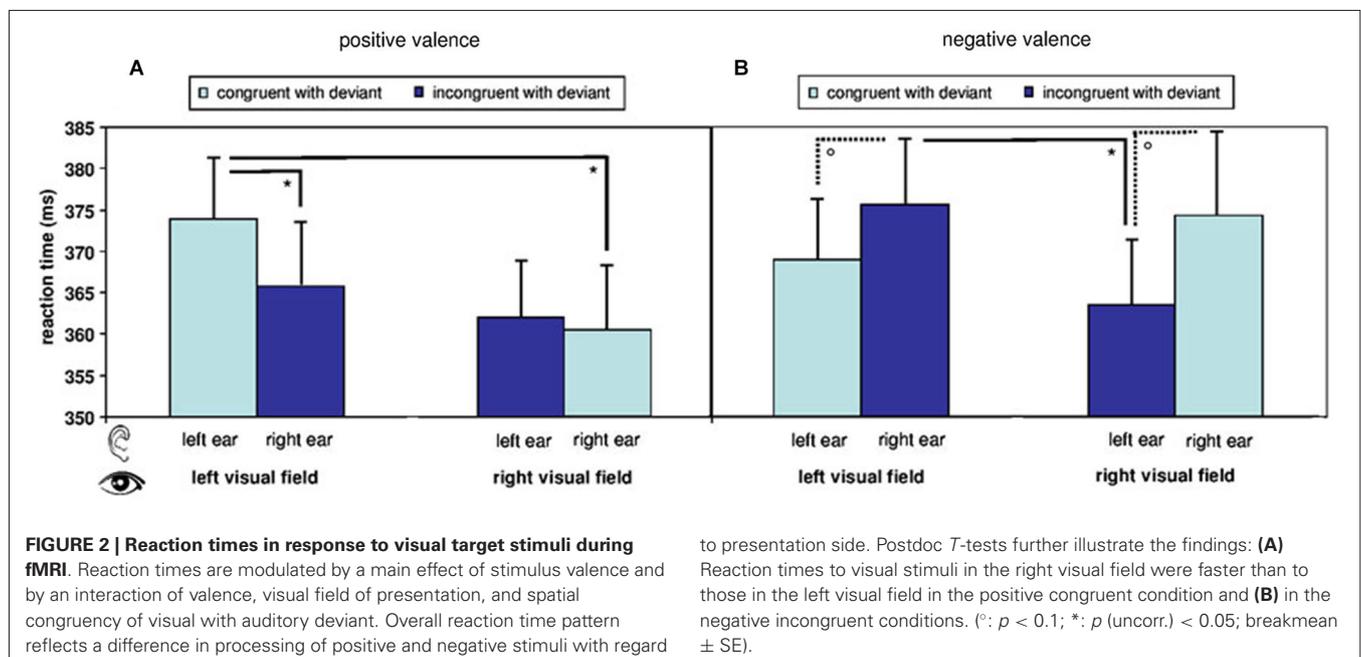
The contrast “deviant without face” vs. “standard without face” (baseline) was built to document auditory cortex activation in response to deviant stimuli with a  $t$ -test. Results were corrected with a family-wise-error (FWE) rate of  $p < 0.05$  to account for multiple testing.

To investigate the neural mechanisms underlying spatial congruency effects, a  $t$ -contrast was designed, which compared spatially incongruent trials (“left deviant + right face”, and “right face + left deviant”) with spatially congruent trials (“left deviant + left face”, and “right deviant + right face”).

### Region-of-interest (ROI) analysis

The contrast estimates for each subject were extracted from bilateral auditory activation cluster peak voxel of the contrast “deviant without face” vs. “standard without face” (baseline; **Figure 2**,

<sup>1</sup>www.fil.ion.ucl.ac.uk



**Table 1 | Deviance responses.**

Contrast	Anatomical area	Hemisphere	BA	MNI coordinates			Peak T-values	Cluster size [voxel]
				X	Y	Z		
“deviant only” vs. “standard only”	Superior temporal gyrus	R	42	68	-26	8	6.80	229
	Superior temporal gyrus	L	22	-62	-24	4	5.77	56

Threshold  $p < 0.05$ , FWE-corr. and cluster extent  $\geq 50$  voxels, BA: Brodmann Area, MNI: Montreal Neurological Institute.

**Table 1).** The response amplitudes at bilateral Auditory cortex (AC) were assessed by a  $2 \times 2 \times 2$  repeated-measures ANOVA, with the factors stimulus *valence* (positive, negative), *visual field* (left, right) and *ear* of deviant presentation (left, right). For this analysis, only trials comprising a deviant and a face were considered.

## RESULTS

### BEHAVIORAL RESULTS

The here applied, uninformative cueing paradigm yielded no speeding of reaction times after auditory deviants ( $368.10 \pm 28.98$  ms) as compared to standards ( $367.13 \pm 25.39$  ms;  $T_{15} = 0.575$ ,  $p = 0.574$ , n.s.). A repeated-measure ANOVA including the factors *spatial congruency*, *visual field*, and *valence* of the visual stimulus analyzed the reaction times in response to stimuli after deviants. A significant main effect emerged for *valence* ( $F_{[1,15]} = 7.820$ ,  $p = 0.014$ ). Mean differences ( $\pm$ SE) yielded significantly faster reactions to positive as compared to negative targets ( $-5.010 \pm 1.792$ ). The factor *spatial congruency* yielded a trend-level effect ( $F_{[1,15]} = 3.476$ ,  $p = 0.082$ ), while *visual field* did not yield significant effects ( $F_{[1,15]} = 2.305$ ,  $p = 0.150$ , n.s.). However, a significant result emerged for the triple interaction between *visual field*, *valence*, and *spatial congruency* ( $F_{[1,15]} = 9.262$ ,  $p = 0.008$ ). In the *post-hoc t*-tests with Bonferroni-correction for multiple comparisons, no contrast between the eight audio-visual conditions yielded significance (see descriptive analysis in **Figure 2**).

### fMRI RESULTS

#### Whole brain analysis

Deviant auditory stimuli yielded large cluster in bilateral AC ( $p < 0.05$ , FWE corr., **Figure 3**, **Table 1**). The incongruency contrast, comparing spatially incongruent with congruent trials, did not reveal significant effects in a whole-brain contrast.

#### ROI analysis

The contrast estimates of the contrast “deviant” vs. “standard” were extracted from the cluster peak voxel in bilateral AC (MNI coordinates [68, -26, 8] and [-62, -23, 4]; **Figure 4**). *Visual field* yielded a significant main effect (left AC:  $F_{[1,15]} = 6.64$ ,  $p = 0.0112$ ; right AC:  $F_{[1,15]} = 10.04$ ,  $p = 0.0019$ ), and *ear* (deviant presentation side) yielded a main effect in the right AC ( $F_{[1,15]} = 9.94$ ,  $p = 0.002$ ). *Valence* trended toward significance in the right AC ( $F_{[1,15]} = 3.16$ ,  $p = 0.078$ ). Furthermore, a significant interaction of *visual field* and *valence* emerged in both ACs (left AC:  $F_{[1,15]} = 4.29$ ,  $p = 0.0405$ ; right AC:  $F_{[1,15]} = 16.4$ ,  $p < 0.001$ ). *Spatial congruency* (interaction of *ear* and *visual field*) failed a significant effect (left AC:  $F_{[1,15]} = 0.01$ ,  $p = 0.918$ , n.s.; right AC:  $F_{[1,15]} = 0.13$ ,  $p = 0.717$ , n.s.). Likewise, significance failed for the

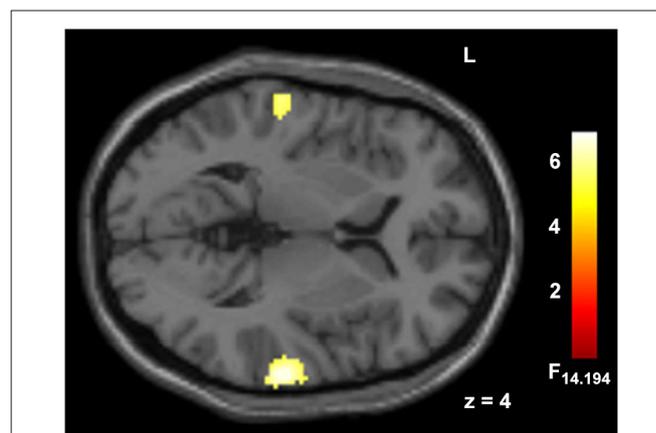
interaction of *ear* and *valence* (left AC:  $F_{[1,15]} = 0.57$ ,  $p = 0.451$ ; right AC:  $F_{[1,15]} = 0.82$ ,  $p = 0.367$ ), as well as a triple interaction of *ear*, *visual field*, and *valence* (left AC:  $F_{[1,15]} = 0.01$ ,  $p = 0.908$ ; right AC:  $F_{[1,15]} = 0.45$ ,  $p = 0.503$ ).

## DISCUSSION

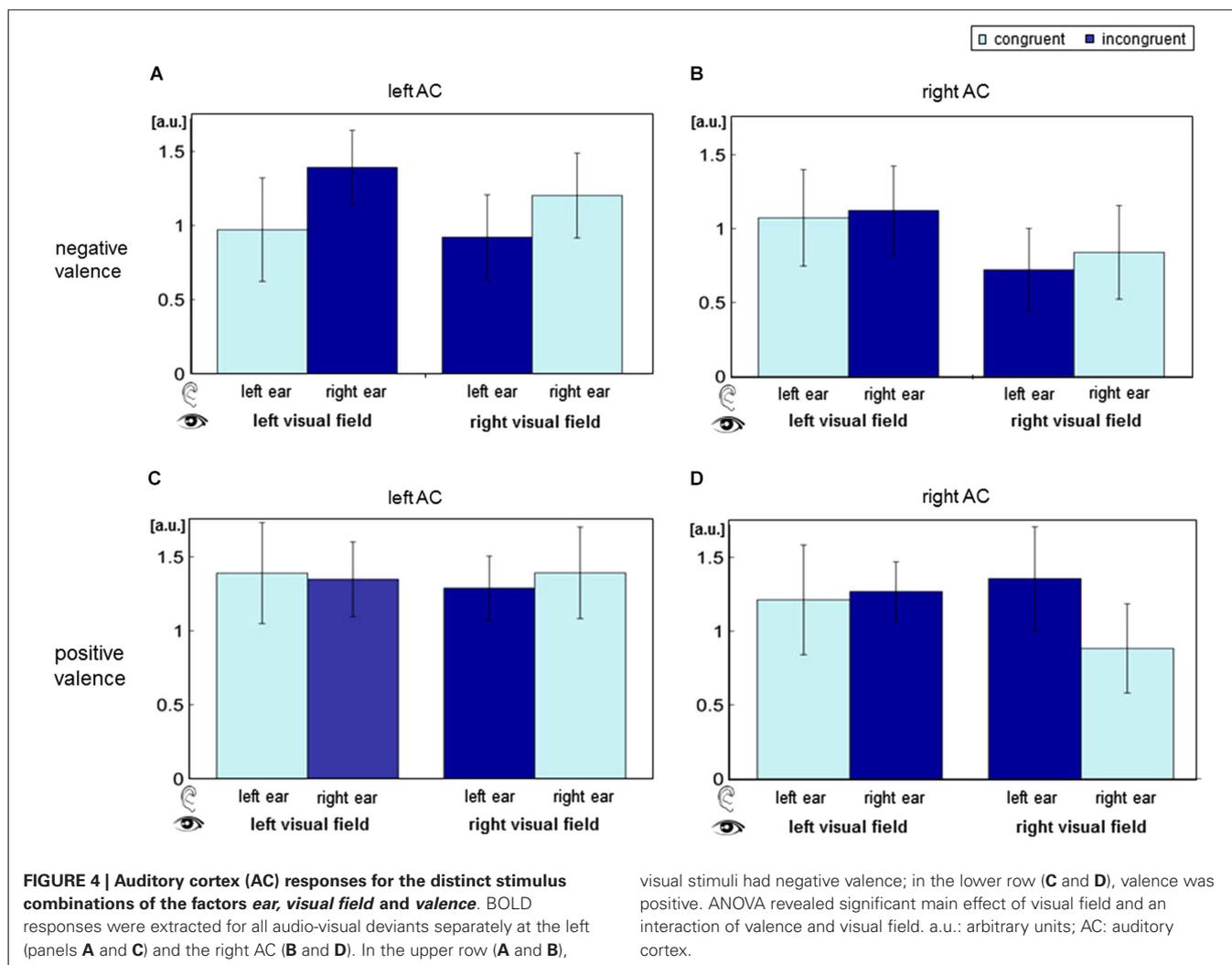
In the present study neural correlates of attention processes in a dichotic listening oddball paradigm were investigated. An auditory stimulus (standard or deviant) preceded visual target presentation (positive and negative schematic faces in the left or right visual field). Events comprising a deviant auditory stimulus were equally distributed between spatially congruent and incongruent trials, thus constituting uninformative spatial cues with a 50% validity ratio. In analogy to the cueing effect of Posner (1980), but applying uninformative crossmodal cues (Spence and Driver, 1997; Ward et al., 2000; Mazza et al., 2007), the auditory deviants were expected to serve as crossmodal spatial cues for visual target stimuli. Indeed, presentation side and valence of visual stimuli significantly affected AC response.

### BEHAVIORAL DATA

To our knowledge, this is the first study to combine schematic faces as emotional stimuli with a spatial cueing (mismatch auditory oddball) design in fMRI. The effectiveness of schematic faces for realistic emotion display (Dyck et al., 2008) and for affective facial stimuli in modulating the allocation of attention in behavioral studies has been demonstrated previously.



**FIGURE 3 | Hemodynamic responses to auditory deviants.** Deviant syllables (/ka/) as compared to ongoing standards (/ba/) elicited responses in auditory cortex. Peak voxel MNI coordinates of right AC: [68, -26, 8] and left AC: [-62, -23, 4];  $p < 0.05$ , FWE corr., extend threshold 5 voxels.



Fox et al. (2001) used schematic faces as spatially incongruent cues in a visual paradigm and demonstrated that negative faces draw and hold attention effectively. In a similar vein, other studies have demonstrated comparable valence effects for the presentation of schematic, or cartoon faces (Eastwood et al., 2001; Fenske and Eastwood, 2003; Santos et al., 2011).

In our previous behavioral study employing the here reported paradigm, we observed that reaction time in response to a visual stimulus following a deviant auditory stimulus was facilitated. Furthermore, reaction time costs emerged in response to spatially incongruent stimuli only in the positive valence condition, whereas negative valence overcame this incongruency effect (Schock et al., 2013). The interaction between spatial presentation and emotional valence of target stimuli is a well-documented phenomenon (Eastwood et al., 2001; Fenske and Eastwood, 2003). The current findings partially reflect the previously observed effects (Schock et al., 2013). Deviant sounds preceding a face did not reduce reaction times compared to faces preceded by a standard sound, but a main effect of *valence* and an interaction

of spatial *congruency*, *visual field* and *valence* was replicated. In summary, alerting may not have been efficient in the noisy MR-environment, but the overall reaction time pattern reflects a difference in processing of positive and negative stimuli with spatial congruency constituting a modulating factor.

#### EFFECT OF DEVIANTS AND SPATIAL INCONGRUENCY

Contrasting the presentation of a deviant sound with a standard, deviance yielded increased hemodynamic responses in bilateral AC. Increased AC activation in response to non-attended deviants within an oddball listening design is well established (EEG: Sams et al., 1985; Näätänen et al., 1989, 2004; MEG: Alho et al., 1998; Mathiak et al., 2000; Phillips et al., 2000; fMRI: Mathiak et al., 2002; Schock et al., 2012). This effect on very early processing steps may underlie the redirection of attention toward changing features in the environment (Näätänen, 1995). Deviant stimuli were presented either spatially congruent or incongruent with a visual target. Contrasting our hypothesis, no spatial congruency effect emerged at the whole-brain level. Our design, comprising uninformative cues (equal probability

of presentation with or without a face) and a wide range of stimulus combinations (valence, visual field, and spatial congruency) may have been inadequate for this question. Until now, only few fMRI studies have investigated whole-brain activation changes in response to spatially incongruent vs. congruent audiovisual stimuli. Sestieri et al. (2006) reported increased response of superior temporal sulcus to spatially congruent audiovisual stimuli, as compared to spatially incongruent stimuli. However, in contrast to our paradigm, participants were explicitly asked to pay attention to both the visual and the auditory stimuli.

#### EFFECTS OF DEVIANCE AND VISUAL FIELD ON AC PROCESSING

ANOVA analyses of AC peak voxel responses revealed a main effect of *visual field* and an interaction effect of *visual field* and *valence* in both AC. Furthermore, a trend-level effect of *valence* emerged in the right AC.

Up until now, the direct comparison of AC responses to spatially incongruent vs. congruent audiovisual stimuli has not been investigated with an auditory oddball design and fMRI. In an EEG study, Teder-Sälejärvi et al. (2005) compared the ERP of bimodal audiovisual stimuli in spatially congruent and incongruent conditions. For incongruent stimulus pairings, the authors reported a shift in phase and amplitude of activity at 100–400 ms, which were allocated to ventral occipito-temporal cortex. Spatially congruent pairings yielded an amplitude modulation of activity at 260–280 ms, localized to superior temporal regions. Though this result hints at an effect of spatial congruency on early auditory processing, we did not find an interaction of *ear* of presentation and *visual field* (i.e., spatial congruency). However, the significant effect of deviance and of visual field of face presentation on AC response suggests a similar notion.

In line with our results, several studies reported that visual input can modulate auditory processing (Kayser et al., 2007; Zvyagintsev et al., 2009; Hsieh et al., 2012). Direct connections between primary auditory and visual cortex have additionally been identified by anatomic studies in primates and fMRI connectivity analysis (Eckert et al., 2008).

#### EFFECTS OF EMOTIONAL VALENCE ON AC PROCESSING

The influence of emotion and valence of visual stimuli on early auditory processing has been reported in a wide range of electrophysiological studies. In an EEG-experiment (Alexandrov et al., 2007) emotional context (monetary reward or punishment) led to significantly larger auditory cortex event-related potentials in response to negative as compared to positive trials. Visually induced emotional states with positive or negative pictures have also been reported to modulate event-related potentials of auditory stimuli (Surakka et al., 1998; Yamashita et al., 2005; Sugimoto et al., 2007; Domínguez-Borràs et al., 2009; Wang et al., 2009). Functional MRI studies complement this picture: Schock et al. (2012) observed right lateralized prefrontal cortex (PFC) activation and enhanced processing of right-ear deviants in the sad mood condition. Fear conditioning, as conveyed by visual stimuli, also affected AC activity (Armony and Dolan, 2001). AC response to fear-conditioned stimuli was modulated by the

presence of a visual context for the likelihood of aversive stimuli appearance.

Complementing these studies, we were the first to investigate face valence effect on AC response in the context of spatial congruency (induced by deviating syllables in an auditory oddball design) and visual field of presentation. The significant interaction of valence and visual field indicates that emotional valence of visual stimuli may modulate auditory processing at an early stage. This finding may support the hypotheses of hemispheric specialization for multimodal integration and emotion processing. According to the valence hypothesis, emotional contexts, especially negative stimuli, are primarily lateralized to the right hemisphere, while positive stimuli are dominantly processed in the left hemisphere (for a review of the valence hypothesis, see Demaree et al., 2005). Similarly, Schock et al. reported enhanced ipsilateral processing of right-ear deviants in the right AC during sad mood, but not during happy or neutral mood. Further, Petit et al. report a right hemispheric enhancement in frontal and temporal areas in response to unattended deviant tones (Petit et al., 2007). Although recent meta-analyses of neuroimaging studies investigating laterality effects on emotional face processing were not able to confirm the valence hypothesis (Wager et al., 2003; Fusar-Poli et al., 2009), our results suggest that within multimodal settings, processing of deviant tones may be affected by laterality and valence—at least at an early auditory stage.

#### LIMITATIONS

To our knowledge, this study is the first to combine fMRI measurement with a dichotic listening oddball design and emotional faces as target stimuli. The resulting high number of conditions and parameters are challenging to take into account. Due to limitations in paradigm length, visual stimuli with neutral emotion were not implemented. Furthermore, deviant oddball paradigms yield comparably small number of events for subsequent analyses since a subdivision into the large number of varying conditions is necessary. However, increasing the experiment duration may induce unwanted effects on attention and arousal. This design-intrinsic small event number as well as a relatively low number of participants may account for the missing results of the whole brain contrasts. The neural response to deviants was clearly detectable in the AC, yet this observation was not fully reflected in the behavioral responses. Our inability to fully reproduce the behavioral effect of spatially congruent stimuli may be explained by the fMRI measurement. The difficulty to reproduce behavioral data within the scanner environment and concurrent reaction time slowing is frequently reported (Plank et al., 2012). Participants in our fMRI experiment responded on average 40 ms slower and with a larger deviation than participants in the behavioral study utilizing the same paradigm (Schock et al., 2013). The fMRI environment may create increased arousal and ongoing distraction due to scanner noise. Indeed, Skouras et al. (2013) recently reported an effect of scanner noise on affective brain processes. Therefore, the fMRI environment may have prevented any salience-driven response acceleration in our study. Nevertheless, although the visual target stimuli exhibit clear valence, the stimuli's arousal value is rather low. Future studies may address the question of stimulus arousal and valence

differences. Higher arousal values may increase the observed effects of valence and visual presentation side. Furthermore, a neutral condition could be implemented to further disentangle emotion and attention effects and laterality of hemispheric activation.

## CONCLUSION

Deviant stimuli of a dichotic listening oddball design were shown to serve as attention-directing cues for visual target stimuli during fMRI measurement. Deviant events and visual targets of different valence induced supramodal modulation of auditory cortices. This modulation was significantly dependent on visual field of presentation and on the interaction of valence and visual field. The present results underline the importance of valence and presentation side for attention guidance by deviant sound events and support the notion of intermodal effects on auditory processing, which may be modulated by emotions.

## ACKNOWLEDGMENTS

This research project was supported by the “Interdisciplinary Centre for Clinical Research (IZKF) Aachen” (N4.2) of the Faculty of Medicine at RWTH Aachen University and the German Research Foundation (Deutsche Forschungsgemeinschaft DFG, IRTG 1328 and MA2631/4-1).

## REFERENCES

- Alexandrov, Y. I., Klucharev, V., and Sams, M. (2007). Effect of emotional context in auditory-cortex processing. *Int. J. Psychophysiol.* 65, 261–271. doi: 10.1016/j.ijpsycho.2007.05.004
- Alho, K., Medvedev, S. V., Pakhomov, S. V., Roudas, M. S., Tervaniemi, M., Reinikainen, K., et al. (1999). Selective tuning of the left and right auditory cortices during spatially directed attention. *Brain Res. Cogn. Brain Res.* 7, 335–341. doi: 10.1016/s0926-6410(98)00036-6
- Alho, K., Winkler, I., Escera, C., Huottilainen, M., Virtanen, J., Jääskeläinen, I. P., et al. (1998). Processing of novel sounds and frequency changes in the human auditory cortex: magnetoencephalographic recordings. *Psychophysiology* 35, 211–224. doi: 10.1111/1469-8986.3520211
- Armory, J. L., and Dolan, R. J. (2001). Modulation of auditory neural responses by a visual context in human fear conditioning. *Neuroreport* 12, 3407–3411. doi: 10.1097/00001756-200110290-00051
- Calvert, G. A., and Campbell, R. (2003). Reading speech from still and moving faces: the neural substrates of visible speech. *J. Cogn. Neurosci.* 15, 57–70. doi: 10.1162/089892903321107828
- Danckert, J., and Ferber, S. (2006). Revisiting unilateral neglect. *Neuropsychologia* 44, 987–1006. doi: 10.1016/j.neuropsychologia.2005.09.004
- Demaree, H. A., Everhart, D. E., Youngstrom, E. A., and Harrison, D. W. (2005). Brain lateralization of emotional processing: historical roots and a future incorporating “dominance”. *Behav. Cogn. Neurosci. Rev.* 4, 3–20. doi: 10.1177/1534528305276837
- Domínguez-Borrás, J., Trautmann, S.-A., Erhard, P., Fehr, T., Herrmann, M., and Escera, C. (2009). Emotional context enhances auditory novelty processing in superior temporal gyrus. *Cereb. Cortex* 19, 1521–1529. doi: 10.1093/cercor/bhn188
- Dyck, M., Winbeck, M., Leiberg, S., Chen, Y., Gur, R. C., and Mathiak, K. (2008). Recognition profile of emotions in natural and virtual faces. *PLoS One* 3:e3628. doi: 10.1371/journal.pone.0003628
- Eastwood, J. D., Smilek, D., and Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Percept. Psychophys.* 63, 1004–1013. doi: 10.3758/bf03194519
- Eckert, M. A., Kamdar, N. V., Chang, C. E., Beckmann, C. F., Greicius, D., and Menon, V. (2008). A cross-modal system linking primary auditory and visual cortices: evidence from intrinsic fMRI connectivity analysis. *Hum. Brain Mapp.* 29, 848–857. doi: 10.1002/hbm.20560
- Eimer, M., and Driver, J. (2001). Crossmodal links in endogenous and exogenous spatial attention: evidence from event-related brain potential studies. *Neurosci. Biobehav. Rev.* 25, 497–511. doi: 10.1016/s0149-7634(01)00029-x
- Fenske, M. J., and Eastwood, J. D. (2003). Modulation of focused attention by faces expressing emotion: evidence from flanker tasks. *Emotion* 3, 327–343. doi: 10.1037/1528-3542.3.4.327
- Fox, E., Russo, R., Bowles, R., and Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *J. Exp. Psychol. Gen.* 130, 681–700. doi: 10.1037/0096-3445.130.4.681
- Fusar-Poli, P., Placentino, A., Carletti, E., Allen, P., Landi, P., Abbamonte, M., et al. (2009). Laterality effect on emotional faces processing: ALE meta-analysis of evidence. *Neurosci. Lett.* 452, 262–267. doi: 10.1016/j.neulet.2009.01.065
- Hsieh, P.-J., Colas, J. T., and Kanwisher, N. (2012). Spatial pattern of BOLD fMRI activation reveals cross-modal information in auditory cortex. *J. Neurophysiol.* 107, 3428–3432. doi: 10.1152/jn.01094.2010
- Kayser, C., Petkov, C. I., Augath, M., and Logothetis, N. K. (2005). Integration of touch and sound in auditory cortex. *Neuron* 48, 373–384. doi: 10.1016/j.neuron.2005.09.018
- Kayser, C., Petkov, C. I., Augath, M., and Logothetis, N. K. (2007). Functional imaging reveals visual modulation of specific fields in auditory cortex. *J. Neurosci.* 27, 1824–1835. doi: 10.1523/jneurosci.4737-06.2007
- Mathiak, K., Hertrich, I., Lutzenberger, W., and Ackermann, H. (2000). Encoding of temporal speech features (formant transients) during binaural and dichotic stimulus application: a whole-head magnetoencephalography study. *Brain Res. Cogn. Brain Res.* 10, 125–131. doi: 10.1016/s0926-6410(00)00035-5
- Mathiak, K., Hertrich, I., Zvyagintsev, M., Lutzenberger, W., and Ackermann, H. (2005). Selective influences of cross-modal spatial-cues on preattentive auditory processing: a whole-head magnetoencephalography study. *Neuroimage* 28, 627–634. doi: 10.1016/j.neuroimage.2005.06.030
- Mathiak, K., Rapp, A., Kircher, T. T. J., Grodd, W., Hertrich, I., Weiskopf, N., et al. (2002). Mismatch responses to randomized gradient switching noise as reflected by fMRI and whole-head magnetoencephalography. *Hum. Brain Mapp.* 16, 190–195. doi: 10.1002/hbm.10041
- Mazza, V., Turatto, M., Rossi, M., and Umiltà, C. (2007). How automatic are audiovisual links in exogenous spatial attention? *Neuropsychologia* 45, 514–522. doi: 10.1016/j.neuropsychologia.2006.02.010
- McDonald, J. J., Teder-Sälejärvi, W. A., Di Russo, F., and Hillyard, S. A. (2003). Neural substrates of perceptual enhancement by cross-modal spatial attention. *J. Cogn. Neurosci.* 15, 10–19. doi: 10.1162/089892903321107783
- Menning, H., Ackermann, H., Hertrich, I., and Mathiak, K. (2005). Spatial auditory attention is modulated by tactile priming. *Exp. Brain Res.* 164, 41–47. doi: 10.1007/s00221-004-2212-7
- Näätänen, R. (1995). The mismatch negativity: a powerful tool for cognitive neuroscience. *Ear Hear.* 16, 6–18. doi: 10.1097/00003446-199502000-00002
- Näätänen, R., Paavilainen, P., Alho, K., Reinikainen, K., and Sams, M. (1989). Do event-related potentials reveal the mechanism of the auditory sensory memory in the human brain? *Neurosci. Lett.* 98, 217–221. doi: 10.1016/0304-3940(89)90513-2
- Näätänen, R., Pakarinen, S., Teemu, R., and Rika, T. (2004). The mismatch negativity (MMN): towards the optimal paradigm. *Clin. Neurophysiol.* 115, 140–144. doi: 10.1016/j.clinph.2003.04.001
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- Petit, L., Simon, G., Joliot, M., Andersson, F., Bertin, T., Zago, L., et al. (2007). Right hemisphere dominance for auditory attention and its modulation by eye position: an event related fMRI study. *Restor. Neurol. Neurosci.* 25, 211–225.
- Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., et al. (2000). Auditory cortex accesses phonological categories: an MEG mismatch study. *J. Cogn. Neurosci.* 12, 1038–1055. doi: 10.1162/08989290051137567
- Plank, T., Rosengarth, K., Song, W., Ellermeier, W., and Greenlee, M. W. (2012). Neural correlates of audio-visual object recognition: effects of implicit spatial congruency. *Hum. Brain Mapp.* 33, 797–811. doi: 10.1002/hbm.21254
- Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25. doi: 10.1080/00335558008248231

- Rowe, G., Hirsh, J. B., and Anderson, A. K. (2007). Positive affect increases the breadth of attentional selection. *Proc. Natl. Acad. Sci. U S A* 104, 383–388. doi: 10.1073/pnas.0605198104
- Sams, M., Paavilainen, P., Alho, K., and Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalogr. Clin. Neurophysiol.* 62, 437–448. doi: 10.1016/0168-5597(85)90054-1
- Santos, A., Mier, D., Kirsch, P., and Meyer-Lindenberg, A. (2011). Evidence for a general face salience signal in human amygdala. *Neuroimage* 54, 3111–3116. doi: 10.1016/j.neuroimage.2010.11.024
- Schock, L., Bhavsar, S., Demenescu, L. R., Sturm, W., and Mathiak, K. (2013). Does valence in the visual domain influence the spatial attention after auditory deviants? Exploratory data. *Front. Behav. Neurosci.* 7:6. doi: 10.3389/fnbeh.2013.00006
- Schock, L., Dyck, M., Demenescu, L. R., Edgar, J. C., Hertrich, I., Sturm, W., et al. (2012). Mood modulates auditory laterality of hemodynamic mismatch responses during dichotic listening. *PLoS One* 7:e31936. doi: 10.1371/journal.pone.0031936
- Schröger, E. (1996). A neural mechanism for involuntary attention shifts to changes in auditory stimulation. *J. Cogn. Neurosci.* 8, 527–539. doi: 10.1162/jocn.1996.8.6.527
- Sestieri, C., Di Matteo, R., Ferretti, A., Del Gratta, C., Caulo, M., Tartaro, A., et al. (2006). “What” versus “where” in the audiovisual domain: an fMRI study. *Neuroimage* 33, 672–680. doi: 10.1016/j.neuroimage.2006.06.045
- Skouras, S., Gray, M., Critchley, H., and Koelsch, S. (2013). fMRI scanner noise interaction with affective neural processes. *PLoS One* 8:e80564. doi: 10.1371/journal.pone.0080564
- Spence, C., and Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Percept. Psychophys.* 59, 1–22. doi: 10.3758/bf03206843
- Sugimoto, S., Nittono, H., and Hori, T. (2007). Visual emotional context modulates brain potentials elicited by unattended tones. *Int. J. Psychophysiol.* 66, 1–9. doi: 10.1016/j.ijpsycho.2007.05.007
- Surakka, V., Tenhunen-Eskelinen, M., Hietanen, J. K., and Sams, M. (1998). Modulation of human auditory information processing by emotional visual stimuli. *Brain Res. Cogn. Brain Res.* 7, 159–163. doi: 10.1016/s0926-6410(98)00021-4
- Teder-Sälejärvi, W. A., Di Russo, F., McDonald, J. J., and Hillyard, S. A. (2005). Effects of spatial congruity on audio-visual multimodal integration. *J. Cogn. Neurosci.* 17, 1396–1409. doi: 10.1162/0898929054985383
- van Atteveldt, N. M., Formisano, E., Goebel, R., and Blomert, L. (2007). Top-down task effects overrule automatic multisensory responses to letter-sound pairs in auditory association cortex. *Neuroimage* 36, 1345–1360. doi: 10.1016/j.neuroimage.2007.03.065
- Wager, T. D., Phan, K. L., Liberzon, I., and Taylor, S. F. (2003). Valence, gender and lateralization of functional brain anatomy in emotion: a meta-analysis of findings from neuroimaging. *Neuroimage* 19, 513–531. doi: 10.1016/s1053-8119(03)00078-8
- Wang, J., Nicol, T., Skoe, E., Sams, M., and Kraus, N. (2009). Emotion modulates early auditory response to speech. *J. Cogn. Neurosci.* 21, 2121–2128. doi: 10.1162/jocn.2008.21147
- Ward, L. M., McDonald, J. J., and Lin, D. (2000). On asymmetries in cross-modal spatial attention orienting. *Percept. Psychophys.* 62, 1258–1264. doi: 10.3758/bf03212127
- Wittchen, H. U., Wunderlich, U., Gruschwitz, S., and Zaudig, M. (1997). *SKID-I. Strukturiertes Klinisches Interview für DSM-IV. Achse I: Psychische Störungen. Interviewheft.* Göttingen: Hogrefe.
- Yamashita, H., Okamoto, Y., Morinobu, S., Yamawaki, S., and Kähkönen, S. (2005). Visual emotional stimuli modulation of auditory sensory gating studied by magnetic P50 suppression. *Eur. Arch. Psychiatry Clin. Neurosci.* 255, 99–103. doi: 10.1007/s00406-004-0538-6
- Zvyagintsev, M., Nikolaev, A. R., Sachs, O., and Mathiak, K. (2011). Early attention modulates perceptual interpretation of multisensory stimuli. *Neuroreport* 22, 586–591. doi: 10.1097/WNR.0b013e32834901f3
- Zvyagintsev, M., Nikolaev, A. R., Thönnessen, H., Sachs, O., Dammers, J., and Mathiak, K. (2009). Spatially congruent visual motion modulates activity of the primary auditory cortex. *Exp. Brain Res.* 198, 391–402. doi: 10.1007/s00221-009-1830-5

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 04 June 2013; accepted: 08 August 2014; published online: 27 August 2014.

Citation: Wolf D, Schock L, Bhavsar S, Demenescu LR, Sturm W and Mathiak K (2014) Emotional valence and spatial congruency differentially modulate crossmodal processing: an fMRI study. *Front. Hum. Neurosci.* 8:659. doi: 10.3389/fnhum.2014.00659

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2014 Wolf, Schock, Bhavsar, Demenescu, Sturm and Mathiak. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.