EXPLICIT AND IMPLICIT EMOTION PROCESSING: NEURAL BASIS, PERCEPTUAL AND COGNITIVE MECHANISMS

EDITED BY: Giulia Mattavelli, Alessia Celeghin and Noemi Mazzoni
PUBLISHED IN: Frontiers in Psychology

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88966-177-0
DOI 10.3389/978-2-88966-177-0

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: frontiersin.org/about/contact

EXPLICIT AND IMPLICIT EMOTION PROCESSING: NEURAL BASIS, PERCEPTUAL AND COGNITIVE MECHANISMS

Topic Editors:
Giulia Mattavelli, University Institute of Higher Studies in Pavia, Italy
Alessia Celeghin, University of Turin, Italy
Noemi Mazzoni, University of Trento, Italy

Citation: Mattavelli, G., Celeghin, A., Mazzoni, N., eds. (2021). Explicit and Implicit Emotion Processing: Neural Basis, Perceptual and Cognitive Mechanisms. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88966-177-0

Table of Contents

frontiers in Psychology

EDITORIAL
published: 30 September 2020
doi: 10.3389/fpsyg.2020.584469

Editorial: Explicit and Implicit Emotion Processing: Neural Basis, Perceptual and Cognitive Mechanisms

Alessia Celeghin[1]*, *Noemi Mazzoni*[2] *and Giulia Mattavelli*[3]

[1] Department of Psychology, University of Torino, Torino, Italy, [2] Department of Psychology and Cognitive Science, University of Trento, Trento, Italy, [3] School of Advanced Studies Istituto Universitario Studi Superiori Pavia, Pavia, Italy

Keywords: emotion perception, implicit - explicit processing, emotional contagion, autism spectrum disorder (ASD), Parkinson disease, active inference, emotion regulation (ER)

Editorial on the Research Topic

Explicit and Implicit Emotion Processing: Neural Basis, Perceptual and Cognitive Mechanisms

The scientific construct of "emotion" has been characterized in the past century by a plethora of different properties (Mauss and Robinson, 2009; Fox, 2018). Understandably, research has emphasized the systematic decomposition of this complex subject of inquiry into elemental components. This approach made great strides in the understanding of key emotional components, and has been seminal in tackling neural fingerprint of emotion, its different behavioral responses, physiological reactions, motivational boosts, and communicative consequences. The investigation through separate and parallel channels represented the most effective way to build an operational scientific approach capable to provide the rigor, replicability, and proper control over the selected components. Nevertheless, each strategy presents some limitations: for example, neuroimaging studies offer correlational evidence about brain activity and the corresponding emotional response, while lesion studies offer causal evidence on altered mental processing, but scarce statistical reproducibility and issues related to homogeneity among groups of patients (Caramazza, 1986). It follows that the analysis of behavior integrated with more sophisticated methods is crucial to offer new perspectives on the emotion field (Krakauer et al., 2017).

A peculiar property of emotions consists in their capability to modulate human behavior both at explicit and implicit levels (Gyurak et al., 2011; Braunstein et al., 2017). The ongoing debate considers the type of and the way stimuli are presented, the definition of automaticity and its neural mechanism, the proper methods to collect those responses (Pourtois et al., 2013). This debate prompted the launch of the current Research Topic, which collects several articles and offers an updated view on the multiplicity of approaches in affective sciences. For instance, Burra and Kerzel emphasize the role of top-down processes and attentional demands in the perceptual encoding of emotions and the prioritization of threatening stimuli. Oldrati et al. also focus on the interplay between attention and emotion, and identify the contribution of stimulus type and intensity in summoning attentional resources. Besides attention, also social factors and internal states can modulate emotional encoding. The automatic transmission of emotional content from one individual to another one, the so-called emotional contagion, is a prominent example of reflex-like reactions that typifies initial stages of emotional responses, which are likely to have an evolutionary origin. Often described as a motor imitation (i.e., mimicry), it occurs in a way that is highly dependent on the individual's learning history, often insensitive to social context, although some rapid responses seem modulated by social factors (Hatfield et al., 1994). In their study, Norscia et al. investigate the possible relation between emotional contagion and (auditory) yawn contagion,

OPEN ACCESS

Edited and reviewed by:
Antonino Raffone,
Sapienza University of Rome, Italy

****Correspondence:***
Alessia Celeghin
alessia.celeghin@unito.it

Specialty section:
This article was submitted to
Consciousness Research,
a section of the journal
Frontiers in Psychology

Received: *17 July 2020*
Accepted: *25 August 2020*
Published: *30 September 2020*

Citation:
Celeghin A, Mazzoni N and Mattavelli G (2020) Editorial: Explicit and Implicit Emotion Processing: Neural Basis, Perceptual and Cognitive Mechanisms. Front. Psychol. 11:584469. doi: 10.3389/fpsyg.2020.584469

a form of replication of facial display triggered by an auditory signal. The authors address the affiliation-based accounts (Palagi et al., 2020) and provide evidence supporting the contention that familiarity, gender bias, and social bond drive the rate and efficiency of contagion. The role of the contextual state in the judgment of emotional expression is also addressed by Pinilla et al., who investigate how an induced affective state can affect the judgment of the observed affective state in others, as communicated by face. Other relevant contextual elements are the mutual interactions amid visual perception of social cues to derive social categories, such as gender and race, and how these processes modulate emotion perception and guide interpersonal behaviors (Bagnis et al., 2019). These aspects are investigated by Venturoso et al. in their article describing how gender modulates the implicit processing of in-group and out-group faces, and by Vergallito et al., who describe how own vs. others expressions alter emotion perception at the sensorimotor level.

Clinical studies have been invaluable to reveal the causal contribution of specific brain areas in emotion processing. Indeed, deficits in the recognition of distinct emotions result from focal lesions in different brain areas, like the amygdala, anterior insula, basal ganglia and frontal cortex (Calder et al., 2001; Celeghin et al., 2017; Mattavelli et al., 2019). Likewise, patients with neurodegenerative diseases involving the insula and basal ganglia (such as Huntington's and Parkinson's diseases) showed poorer performance in emotion recognition, with possible dissociations related to the stages of processing (Novak et al., 2012; Mattavelli et al., 2020). In their review, Wagenbreth et al. focus on patients with Parkinson and deep brain stimulation (DBS) in the subthalamic nucleus. They conclude that DBS has no consistent impact on facial emotion recognition although improvements were measurable for non-facial signals—like written or spoken semantic stimuli and emotional prosody—and non-facial pictorial stimuli. Moreover, a limited number of studies showed less impairment in implicit processing associated with DBS. This is in line with the view of at least partially distinct pathways for explicit and implicit emotion processing, a perspective proposed also for accounting for emotion regulation dysfunctions in individuals with autism spectrum disorder (ASD). In this regard, Siciliano and Clausi outline the role of the cerebellum in integrating internal state with external environmental information. Two other studies in the Topic address emotional difficulties in ASD: Mazzoni et al. show that the difficulty in recognizing emotional body stimuli in children with ASD may rely on atypical processing of body movement information rather than on emotion itself; Giannotti et al. find that alexithymia rather than ASD predicts perception of attachment to parents.

The hierarchical integration of high and low levels of emotional encoding—and, among these, the interoceptive and exteroceptive embodied cues—is a crucial topic in understanding emotion regulation. In this regard, the perspective discussed by Shalev conceptualizes low emotional clarity as the product of low access to the emotional signal. This results in prediction error associated with a mismatch between the current bodily state and the predicted state. The motivated cue integration theory (MCI) proposed by the author describes the potential transformation of low emotional clarity into the creation of emotion goals. Within this frame, a special emphasis has been observed in deep generative models that afford the capacity to explain multimodal (i.e., interoceptive, proprioceptive, and exteroceptive) sensations that are characteristic of emotion concepts. For instance, Smith et al. introduce a novel model of emotion inference that is sufficiently nuanced to produce synthetic emotional processes. Their computational simulation infers the emotion by considering the feeling in a specific moment and conceptualizing the learning process. Interestingly, they outline an active inference model of emotional states in the brain, which can minimize uncertainty to maximize prior preferences. This model—described by the Bayesian principles—postulates that prior information could influence the interpretation of sensory evidence, especially when this latter is ambiguous or degraded. As it happens, prior probabilities could be represented by automatically driven psychophysiological changes that influence perceptual decisions on emotions. From these pieces of evidence, a new avenue to implement a Bayesian framework in the domain of non-conscious and conscious emotional perception seems an interesting step to follow.

In conclusion, with this Topic, we sought to offer a comprehensive picture of a phenomenon as multifaceted as explicit and implicit emotion processing is. The variety of studies included here well pictures the current research efforts to integrate different points of view. The future challenge is to converge into a comprehensive theoretical model of the emotional brain, taking into account multiple factors that require a more conjoint and complementary form of integration.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

AC was supported by the ERC Consolidator grant LIGHTUP (prot. 772953) and by a PRIN grant 2017 from the Italian Ministry of Education University and Research (MIUR: 2017TBA4KS). NM was funded by the Municipality of Rovereto.

REFERENCES

Bagnis, A., Celeghin, A., Mosso, C. O., and Tamietto, M. (2019). Toward an integrative science of social vision in intergroup bias. *Neurosci. Biobehav. Rev.* 102, 318–326. doi: 10.1016/j.neubiorev.2019.04.020

Braunstein, L. M., Gross, J. J., and Ochsner, K. N. (2017). Explicit and implicit emotion regulation: a multi-level framework. *Soc. Cogn. Affect. Neurosci.* 12, 1545–1557. doi: 10.1093/scan/nsx096

Calder, A. J., Lawrence, A. D., and Young, A. W. (2001). Neuropsychology of fear and loathing. *Nat. Rev. Neurosci.* 2, 352–363. doi: 10.1038/35072584

Caramazza, A. (1986). On drawing inferences about the structure of normal cognitive systems from the analysis of patterns of impaired performance: the case for single-patient studies. *Brain Cogn.* 5, 41–66. doi: 10.1016/0278-2626(86)90061-8

Celeghin, A., Diano, M., Bagnis, A., Viola, M., and Tamietto, M. (2017). Basic emotions in human neuroscience: neuroimaging and beyond. *Front. Psychol.* 8:1432. doi: 10.3389/fpsyg.2017.01432

Fox, E. (2018). Perspectives from affective science on understanding the nature of emotion. *Brain Neurosci. Adv.* 2:2398212818812628. doi: 10.1177/2398212818812628

Gyurak, A., Gross, J. J., and Etkin, A. (2011). Explicit and implicit emotion regulation: a dual-process framework. *Cogn. Emot.* 25, 400–412. doi: 10.1080/02699931.2010.544160

Hatfield, E., Cacioppo, J. T., and Rapson, R. L. (1994). *Emotional Contagion*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139174138

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., and Poeppel, D. (2017). Neuroscience needs behavior: correcting a reductionist bias. *Neuron* 93, 480–490. doi: 10.1016/j.neuron.2016.12.041

Mattavelli, G., Barvas, E., Longo, C., Zappini, F., Ottaviani, D., Malaguti, M. C., et al. (2020). Facial expressions recognition and discrimination in Parkinson' s disease. *J. Neuropsychol.*. doi: 10.1111/jnp.12209

Mattavelli, G., Pisoni, A., Casarotti, A., Comi, A., Sera, G., Riva, M., et al. (2019). Consequences of brain tumour resection on emotion recognition. *J. Neuropsychol.* 13, 1–21. doi: 10.1111/jnp.12130

Mauss, I. B., and Robinson, M. D. (2009). Measures of emotion: a review. *Cogn. Emot.* 23, 209–237. doi: 10.1080/02699930802204677

Novak, M. J. U., Warren, J. D., Henley, S. M. D., Draganski, B., Frackowiak, R. S., and Tabrizi, S. J. (2012). Altered brain mechanisms of emotion processing in pre-manifest Huntington's disease. *Brain* 135, 1165–1179. doi: 10.1093/brain/aws024

Palagi, E., Celeghin, A., Tamietto, M., Winkielman, P., and Norscia, I. (2020). The neuroethology of spontaneous mimicry and emotional contagion in human and non-human animals. *Neurosci. Biobehav. Rev.* 111, 149–165. doi: 10.1016/j.neubiorev.2020.01.020

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

Conflict of Interest: 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.

frontiers
in Psychology

ORIGINAL RESEARCH
published: 24 October 2019
doi: 10.3389/fpsyg.2019.02400

Task Demands Modulate Effects of Threatening Faces on Early Perceptual Encoding

Nicolas Burra and Dirk Kerzel*

Faculté de Psychologie et des Sciences de l'Éducation, Université de Genève, Geneva, Switzerland

The *threat capture hypothesis* states that threatening stimuli are automatically processed with higher priority than non-threatening stimuli, irrespective of observer intentions or focus of attention. We evaluated the threat capture hypothesis with respect to the early perceptual stages of face processing. We focused on an electrophysiological marker of face processing (the lateralized N170) in response to neutral, happy, and angry facial expressions displayed in competition with a non-face stimulus (a house). We evaluated how effects of facial expression on the lateralized N170 were modulated by task demands. In the pixel task, participants were required to identify the gender of the face, which made the face task-relevant and entailed structural encoding of the face stimulus. In the pixel task, participants identified the location of a missing pixel in the fixation cross, which made the face task-irrelevant and placed it outside the focus of attention. When faces were relevant, the lateralized N170 to angry faces was enhanced compared to happy and neutral faces. When faces were irrelevant, facial expression had no effect. These results reveal the critical role of task demands on the preference for threatening faces, indicating that top-down, voluntary processing modulates the prioritization of threat.

Keywords: task demand, N170, lateralized N170, facial expressions, threat, angry faces

Edited by:
Giulia Mattavelli,
University of Milano-Bicocca, Italy

Reviewed by:
Alessandra Vergallito,
University of Milano-Bicocca, Italy
Thomas Straube,
Institute of Medical Psychology and Systems Neuroscience, Germany

****Correspondence:***
Nicolas Burra
Nicolas.Burra@unige.ch

Specialty section:
This article was submitted to Consciousness Research, a section of the journal Frontiers in Psychology

Received: *30 July 2019*
Accepted: *08 October 2019*
Published: *24 October 2019*

Citation:
Burra N and Kerzel D (2019) Task Demands Modulate Effects of Threatening Faces on Early Perceptual Encoding. Front. Psychol. 10:2400. doi: 10.3389/fpsyg.2019.02400

INTRODUCTION

The capacity to detect a threat in the environment is essential for survival. In humans, threatening facial expressions, such as angry or fearful faces, have a high priority. In a social context, they indicate aggressive intentions or potential threats in the environment. The *threat capture hypothesis* posits that threatening signals are prioritized above other visual information (Öhman and Mineka, 2001). In this model, a core assumption is that threatening stimuli are processed even when they are located outside the focus of attention. It is also assumed that they are processed regardless of task demands. In fact, the automatic encoding of emotionally significant events and the rapid orienting toward threatening stimuli outside the focus of attention would be advantageous for survival because it prepares the organism to take appropriate action (i.e., Öhman et al., 2010). However, continuous capture by irrelevant threatening stimuli may also have detrimental consequences for survival because relevant non-threatening stimuli may be missed. Accordingly, it is reasonable to propose that attentional selection of threatening stimuli is context-dependent and only partially driven by automatic pre-attentive encoding of the threat value of the stimulus.

To answer our research question, we used event-related potentials (ERPs) because of their high temporal resolution. The current study is concerned with lateralized ERPs. Lateralized ERPs are triggered by events in the periphery where a target object in the left or right visual field is in competition with a physically balanced non-target object in the opposite visual field. Instead of analyzing ERPs to left and right targets, ERPs are calculated contralateral and ipsilateral to the target stimulus. The advantage of calculating difference waves is that asymmetries between the left and right hemisphere are removed. Accordingly, only target processing is reflected in the difference waves, which allows for very strong conclusions about the timing of the respective neural events (Luck, 2005, 2012). In the past decades, two lateralized ERPs occurring over posterior cortex at electrode sites PO7/8 between 200 and 300 ms have been extensively used to investigate attentional selection. The N2pc is a more negative voltage contralateral than ipsilateral to a target stimulus and is considered a marker of attentional selection (Luck and Hillyard, 1994; Eimer, 1996). The P_D has the opposite polarity and is considered a marker of attentional suppression (Hickey et al., 2009; Sawaki and Luck, 2010). Threatening stimuli have been shown to affect both markers of attentional processing. Changes in the N2pc occur if the threatening stimulus is relevant and attended (i.e., Feldmann-Wustefeld et al., 2011; Weymar et al., 2011; Yao et al., 2013), whereas changes in the P_D and the N2pc occur if the threatening stimulus is irrelevant and suppressed (Burra et al., 2016, 2017; Bretherton et al., 2017).

Before attentional selection of an object is performed, early perceptual encoding discriminates between face and non-face stimuli. However, it is a matter of debate whether the early perceptual encoding of threatening faces is enhanced. In general, the early structural encoding of faces is associated with the N170 (Bentin et al., 1996; George et al., 1996; Eimer, 2000). The N170 is an enhanced negativity for face compared to non-face stimuli, which was originally observed with stimuli presented in the center of the visual field (at fixation), but also occurs with peripheral stimuli. The N170 occurs bilaterally over posterior occipito-temporal electrode sites, but is typically larger over the right hemisphere (Rossion and Jacques, 2008). The larger N170 over the right hemisphere is attributed to the fusiform face area, a brain region selectively activated by faces (Kanwisher et al., 1997). Interestingly, some have argued that the N170 is independent of attention (Cauquil et al., 2000; Carmel and Bentin, 2002). For instance, in Carmel and Bentin (2002), the N170 was not different between conditions where the stimuli were ignored or attended. However, recent studies have shown that top-down factors, such as expectations, instructions and/or spatial attention can modulate the amplitude of the N170 (e.g., Holmes et al., 2003; Sreenivasan et al., 2007, 2009; Aranda et al., 2010; Schinkel et al., 2014), suggesting that the N170 is, in fact, modulated by task demands. Nonetheless, we consider the N170 as pre-attentional because it occurs between 140 and 200 ms post-stimulus and therefore precedes the two markers of attentional processing, the N2pc and P_D.

Bruce and Young (1986) stated in their influential model of face processing that structural encoding and the detection of the emotional expression of faces is achieved in two parallel and independent stages. Typically, the N170 has been associated with the structural encoding of faces and according to Bruce and Young (1986), structural encoding is insensitive to facial expressions. However, their hypothesis is contested because studies have yielded mixed results. In some studies, the N170 was enhanced in response to emotional faces as opposed to neutral faces (Batty and Taylor, 2003; Schyns et al., 2007; Leppanen et al., 2008; Righart and de Gelder, 2008; Smith et al., 2013; Turano et al., 2017), but in other studies the N170 remained insensitive to facial expressions (Pourtois et al., 2005; Rellecke et al., 2012; Calvo and Beltran, 2013; Tamamiya and Hiraki, 2013; Neath-Tavares and Itier, 2016; for review, see Hinojosa et al., 2015). These discrepancies may be related to methodological differences in the research, such as the choice of reference electrodes (Joyce and Rossion, 2005; Rellecke et al., 2013) or stimulus features (daSilva et al., 2016; Schindler et al., 2019). Alternatively, it may be that attention and task demands modulate effects of emotional expression on the N170 (Eimer et al., 2003; Holmes et al., 2003, 2006; Neath-Tavares and Itier, 2016; for review, see Eimer and Holmes, 2007). Importantly, effects of facial expressions on the N170 occurred when participants' attentional focus was spread over the entire face and when they were engaged in a task that required explicit recognition of the emotional expression. Conversely, when attention was directed away from the face stimuli, the effects of emotional expression were absent (Eimer et al., 2003; for review, see Holmes et al., 2003, 2006; Hinojosa et al., 2015). Therefore, the early encoding of facial expression might depend on task demands.

Thus, both the importance of task-demands on the N170 and the absence of systematic difference in the N170 to threatening (fearful, angry) and non-threatening (happy, neutral) faces challenge the threat capture hypothesis (for review, see Hinojosa et al., 2015). Nevertheless, it is likely that researchers did not use the adequate marker to reveal effects of facial expression. In fact, for decades, the N170 was thought to be insensitive to stimulus location. Therefore, the stimulus was commonly presented centrally or, when the stimuli were displayed in the periphery, stimulus location was collapsed in the analysis. However, recent evidence suggested that the N170 is sensitive to the location of the face (Towler and Eimer, 2015). In fact, similar to the N2pc or the P_D component, visually evoked potentials were enhanced over the contralateral hemisphere as compared to the ipsilateral hemisphere. When a face was displayed to the left or right of the fixation point with a competing non-face stimulus on the opposite side, the N170 was larger contralateral to the side of the face stimulus. The reason is that stimuli become increasingly represented by the contralateral hemisphere as stimuli appear at larger retinal eccentricities (Towler and Eimer, 2015; Towler et al., 2016). Critically, the presence of a competing stimulus in the opposite visual field inhibits the transmission of identity-sensitive information between hemispheres and, therefore, reveals a contralateral bias in the visual processing of faces. Similar to the N2pc, this difference between contralateral and ipsilateral removes differences between the left and right hemisphere and isolates early lateralized face encoding. The lateralized N170 increases the ecological validity of the experimental paradigm because competition between visual hemispheres is a common phenomenon in everyday visual processing, as the world mostly contains multiple objects in both visual fields. Accordingly, the contralateral dominance of high-level visual object processing over ipsilateral processing has also been noted in the perception of real-world scenes (Freiwald et al., 2016).

Our review of the literature shows a mixed picture concerning the effect of emotional expression on the N170, but little is known about whether the contralateral bias is sensitive to threat or task demand. Because the lateralized N170 may better reflect everyday processing than the classical N170, we investigated whether the lateralized N170 was sensitive to facial expressions and critically, whether this sensitivity was modulated by task demand. To do so, we used a within-subject design to measure the effect of altering the task with identical stimuli (Neath-Tavares and Itier, 2016; Itier and Neath-Tavares, 2017; Smith and Smith, 2019). The displays were composed of two stimuli, a face on one side and a house on the other. Face stimuli included neutral, angry, or happy expressions. In the gender task, the faces were attended because participants were required to categorize the gender of the face. In the pixel task, the faces were unattended because participants located a missing pixel on the fixation cross. In both tasks, facial expressions were task-irrelevant, which is important because angry and happy expressions are recognized faster than neutral faces. In contrast, differential processing latencies are absent when participants are required to discriminate the gender of the stimuli, which involves an equally demanding encoding across facial expressions (Wronka and Walentowska, 2011). Therefore, the only difference between the gender and pixel tasks was the relevance of the face prior to attentional selection. We expect larger N170 components over the hemisphere contralateral to the visual field in which a face was presented as compared to the hemisphere ipsilateral (Towler and Eimer, 2015; Towler et al., 2016). The threat capture hypothesis suggests that the lateralized N170 to threatening stimuli is enhanced irrespective of the task, i.e., before attentional selection. In contrast, if top-down processing were critical, there should be a larger lateralized N170 to threatening faces as compared to non-threatening stimuli only when the faces are task-relevant, but not when they are irrelevant. Subsequently, an enhanced N2pc to threatening stimuli is expected (Feldmann-Wustefeld et al., 2011; Burra et al., 2016).

EXPERIMENT

Methods

Participants

The participants included 22 students at the University of Geneva (7 male, 15 female, mean age 20.8 ± 1.6 years, all right-handed). All had normal or corrected-to-normal vision. We discarded the data from two participants due to excessive alpha waves, leaving 20 participants in the final sample. The study received clearance from the local ethics committee (Faculty of Psychology and Educational Sciences, Geneva University). All participants gave written informed consent in accordance with the Declaration of Helsinki. Participants received 40 Swiss Francs for their participation.

Apparatus and Stimuli

Stimulus displays consisted of 20 facial identities (10 male and 10 female) images of faces taken from the Emotion Lab at the Karolinska Institute (KDEF) (Lundqvist et al., 1998) and the NimStim database (Tottenham et al., 2009) and 10 pictures of houses were taken from Google (for a similar procedure, see Towler et al., 2016; Framorando et al., 2018). Face stimuli differed regarding valence but not emotional intensity (for more details, see Burra et al., 2017). Each face was shown with a neutral, angry, or happy facial expression, yielding 60 different pictures. All images were matched for luminance with the SHINE package (Willenbockel et al., 2010) and subtended a visual angle of approximately 4° × 3.8°.

Procedure

Participants completed the experiment in a soundproof box with dim lighting. Stimuli were presented using MATLAB and the Psychtoolbox (Kleiner et al., 2007). On each trial, two images were presented simultaneously to the left and right of fixation at a horizontal eccentricity of approximately 4°, measured relative to the center of each image. Each bilateral stimulus array was presented for 200 ms. Then, a black screen was displayed until the participants pressed the keyboard to answer. The inter-trial interval was 1000 ms. Throughout the study, when the two pictures appeared, one pixel of the central fixation cross was removed from its upper or lower branch. Pictures appeared with equal probability and in random order. Each face appeared with the same probability on the left and right side and the pixel was missing with the same probability in the upper and lower part of the fixation cross. Participants saw the same face four times in each block. The experiment consisted of three blocks of the gender discrimination task and three blocks of the pixel discrimination task. The same visual displays were used in both tasks, but the task alternated between blocks. The task in the first block was randomly chosen. Each block consisted of 240 trials, with a 5-s pause after every 48 trials. The experiment lasted about an hour in total. In the gender task, participants had to respond as quickly as possible by pressing one of two keyboard buttons (the "1" and "2" keys of the number pad) to indicate whether the displayed face was male or female. In the pixel task, participants were required to detect whether the missing pixel on the fixation cross was in the upper or lower half of the cross. Participants used two fingers of the right hand to respond and mapping of key to response was counterbalanced across participants.

EEG Recording and Analysis

EEG data were acquired using a 32-channel BioSemi ActiveTwo system (BioSemi, Amsterdam, Netherlands), with electrodes including standard 10–20 system locations as well as six additional reference electrodes[1]. Offline, 0.1–40 Hz filters were applied after EEG data acquisition. Horizontal eye movements (HEOG) and vertical eye movements (VEOG) were measured

---

[1] http://www.biosemi.com/

from two electrodes placed at the outer canthus of each eye and above and below the right eye, respectively. Two additional electrodes, an active common mode sense (CMS) and a passive driven right leg (DRL) electrode, were used in the study. Raw EEG data was recorded relative to CMS. The CMS/DRL electrodes replaced the ground for recordings through a feedback loop that drove the average potential of the subject (i.e., the Common Mode voltage) as close as possible to the "zero" ADC reference voltage in the AD box[2]. Subsequently, the signal was re-referenced to the average voltage across electrodes, which is common in the face processing literature (Joyce and Rossion, 2005). All offline analyses of EEG data were conducted with Brain Vision Analyzer. Ocular artifacts were corrected using Independent Component Analysis. Specifically, components associated with eye blinks were removed from the continuous EEG (Jung et al., 2000). The EEG was epoched into 600 ms segments, from 200 ms before stimulus onset to 400 ms after stimulus onset. A baseline correction of 200 ms was applied. Trials with saccades (voltage exceeding ±30 μV in the HEOG channel), eye blinks (exceeding ±60 μV at VEOG), or muscular artifacts (exceeding ±80 μV at any other electrode) prior to correction by Independent Component Analysis were excluded from analysis. In total, 23% of trials were discarded.

Behavioral Results

We calculated the median response time (RT) for each condition and subject. Trials in which RTs were shorter than 200 ms and longer than 2000 ms were discarded. Mean RTs and accuracy are summarized in **Table 1**.

A repeated-measures 2 (task: gender, pixel) × 3 (facial expression: neutral, angry, and happy) ANOVA was conducted on median RTs of correct responses (accuracy). This revealed a main effect of task, $F(1,19) = 171.33$, $p < 0.001$, $\eta_p^2 = 0.9$, with higher accuracy in the pixel (94%) than in the gender task (85%). There was also a main effect of facial expression, $F(2,38) = 12.91$, $p < 0.001$, $\eta_p^2 = 0.45$, with more correct responses to neutral (91%) compared to angry (89%) and happy faces (89%). The analysis revealed an interaction between task and facial expression, $F(2,38) = 6.97$, $p = 0.003$, $\eta_p^2 = 0.26$. Separate one-way ANOVAs for each task revealed that facial expression modulated accuracy in the gender categorization task, $F(2,38) = 12.9$, $p < 0.001$, $\eta_p^2 = 0.4$, but not in the pixel task, $p = 0.55$. In the gender categorization task, accuracy was higher with neutral (87%) than with angry (84%) or happy faces (83%), $ts(19) > 2.25$, $ps < 0.036$, Cohen's $d_z > 0.68$.

The same 2 × 3 ANOVA was also conducted on median RTs of correct responses. There was a significant main effect of task, $F(1,19) = 1033.69$, $p < 0.001$, $\eta_p^2 = 0.98$, with shorter RTs for pixel localization (422 ms) than gender discrimination (598 ms). The analysis also revealed a main effect of facial expression $F(2,38) = 8.03$, $p < 0.001$, $\eta_p^2 = 0.29$, with longer RTs with happy (513 ms) than neutral (509 ms) or angry faces (507 ms). There was no significant interaction, $F < 0.51$.

Electrophysiological Results

For lateralized components, brain activity at electrode sites ipsilateral to the face stimulus was subtracted from the activity at electrode sites contralateral to the face stimulus. Average ipsilateral and contralateral voltages are shown in **Figure 1A** and differences waves are shown in **Figure 1B** (right panel). The left panel of **Figure 1B** shows that the lateralized N170 was maximal at electrodes P7/P8. Therefore, all components were calculated at these electrodes. We chose a time window from 160 to 200 ms for the lateralized N170, which was around the maximal difference between the contralateral and ipsilateral signal. For the sake of consistency with prior studies (Towler et al., 2016; Framorando et al., 2018; Neumann et al., 2018), we calculated the early N2pc from 200 to 240 ms, the late N2pc from 240 to 280 ms, always at electrodes P7/P8. In order to discard any impact of the lateral eye movement on the lateralized components, we analyzed the magnitudes of residual HEOG deflections in their according time windows. For non-lateralized components, the brain activity was averaged at electrodes P7 and P8, collapsed across left and right locations of the face stimulus. The non-lateralized N170 was calculated around the peak negativity, between 150 and 190 ms (see **Figure 2**).

Lateralized N170

A 2 (task: gender, pixel) × 3 (facial expression: neutral, angry, and happy) repeated-measures ANOVA on the differences between contra- and ipsilateral voltages from 160 to 200 ms revealed a main effect of task, $F(1,19) = 28.95$, $p < 0.001$, $\eta_p^2 = 0.60$. The lateralized N170 was larger in the gender task (−3.91 μV) than in the pixel task (−2.83 μV). The main effect of facial expression was not significant, $p = 0.052$. Critically, we found an interaction between task and facial expression, $F(2,38) = 5.5$, $p = 0.008$, $\eta_p^2 = 0.22$, which motivated separate one-way ANOVAs for each task. There was a significant effect of facial expression in the gender task, $F(2,38) = 4.81$, $p = 0.014$, $\eta_p^2 = 0.2$, which was caused by a larger lateralized N170 for angry faces (−4.29 μV) compared with neutral (−3.72 μV), $t(19) = 2.64$, $p = 0.016$, Cohen's $d_z = 0.59$, and happy faces (−3.70 μV), $t(19) = 2.87$, $p = 0.01$,

TABLE 1 | Behavioral results.

	Gender task			Pixel task		
	Neutral	**Angry**	**Happy**	**Neutral**	**Angry**	**Happy**
Reaction time	596 (46)	595 (45)	602 (47)	422 (55)	420 (57)	424 (58)
Accuracy	87% (5)	84% (4)	83% (5)	94% (3)	94% (3)	94% (3)

Mean reaction times and accuracy in the experimental conditions. Standard deviations are shown in parentheses.

---

[2] http://www.biosemi.com/faq/cms&drl.htm

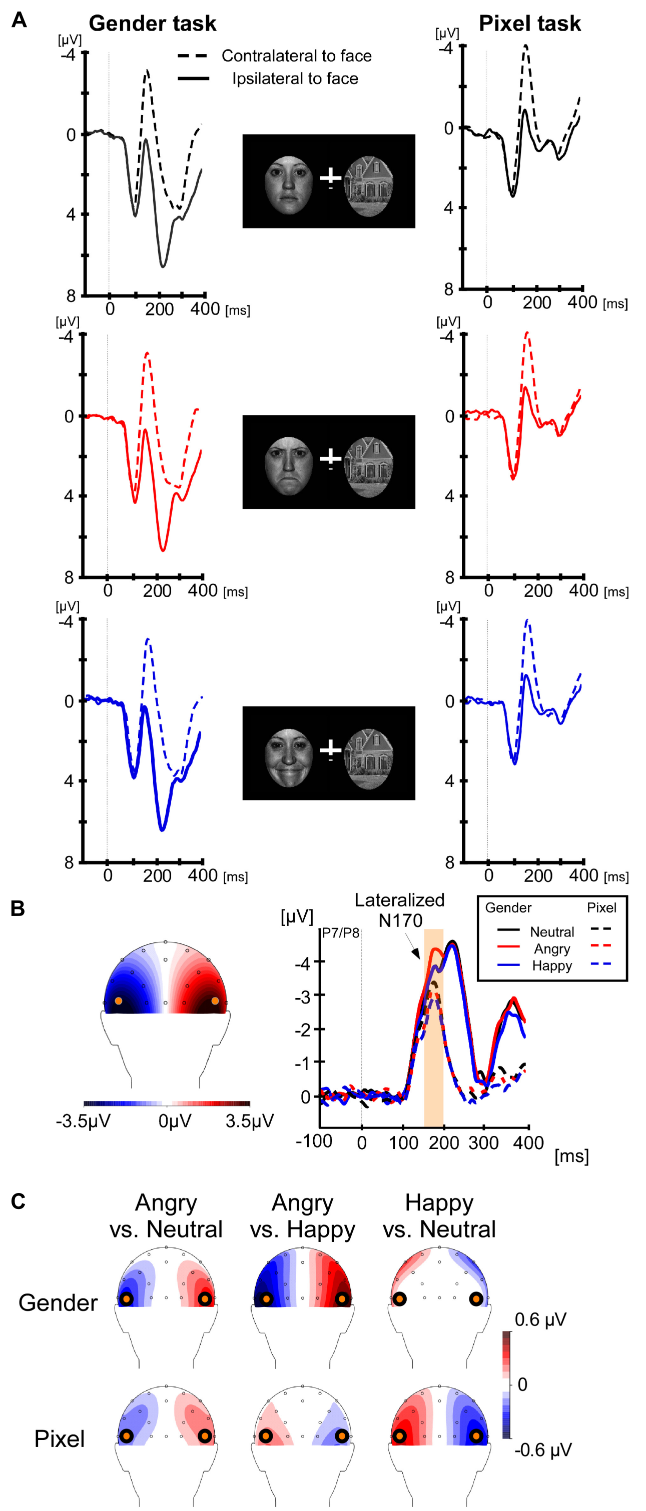


FIGURE 1 | Lateralized event-related components at P7/P8. The interval for baseline correction was from –200 to 0 ms, but only the final 100 ms are shown in the graph to save space. **(A)** Contralateral (dashed line) and ipsilateral (solid line) event-related potentials to the face stimuli are shown for the gender categorization task on the left and for the pixel task on the right.

(Continued)

FIGURE 1 | The top row shows results for neutral facial expressions, the second row for angry and the third row for happy facial expressions. (**B**, left panel) The scalp distribution of the lateralized N170 from 160–200 ms. The contralateral activity was maximal at location P7/P8. (**B**, right panel) The difference waves for each facial expression and task. The orange region highlights the N170 interval where we found a larger amplitude for angry faces in the gender task, but not in the pixel task. **(C)** Scalp topographies of the differences between angry-neutral, angry-happy, and happy-neutral during the N170 time-window, separately for the gender and pixel tasks. The differences between angry-neutral and happy-neutral in the gender task were restricted to the electrodes of interest.

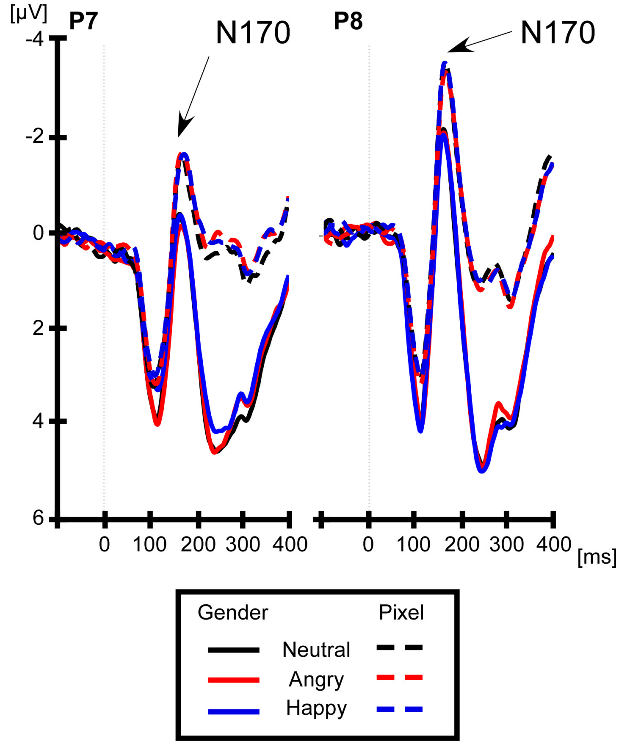


FIGURE 2 | Non-lateralized N170 to facial expressions at P7/P8. The N170 is more negative on the **right** (electrode P8) than on the **left** hemisphere (electrode P7). However, the type of facial expression did not affect the non-lateralized N170, neither in the gender (solid line) nor in the pixel task (dashed line).

Cohen's $d_z = 0.64$. In the pixel task, the main effect of facial expression did not reach significance, $p = 0.07$. The lateralized N170 in the pixel task was −2.83, −3.05, and 2.6 μV for angry, neutral, and happy faces, respectively. Overall, significant lateralized N170 components were present in all six conditions, $ts(19) > 6.73$, $ps < 0.001$, Cohen's d_zs > 1.5.

Early N2pc Interval

The same 2 × 3 ANOVA on the differences between contra- and ipsilateral voltages from 200 to 240 ms revealed a main effect of task, $F(1,19) = 33.43$, $p < 0.001$, $\eta_p^2 = 0.62$. The early N2pc was larger in the gender task (−4.35 μV) than in the pixel task (−0.86 μV). Collapsed across facial expression, the N2pc was significant in both gender and pixel tasks, *t*s(19) > 2.54, *p*s < 0.02, Cohen's d_zs > 0.56. There was no main effect of facial expression and no interaction between task and facial expression, *p*s > 0.68. Note that the differences between angry-neutral and happy-neutral in the gender task were restricted to the electrodes of interest (see **Figure 1C**).

Late N2pc Interval

The same 2 × 3 ANOVA on the differences between contra- and ipsilateral voltages from 240 to 280 ms revealed a main effect of task, $F(1,19) = 17.75$, $p < 0.001$, $\eta_p{}^2 = 0.48$. The late N2pc was larger in the gender task (−2.14 μV) as compared to the pixel task (−0.001 μV). There was no main effect for facial expression and no significant interaction between task and facial expression, *p*s > 0.15. Collapsed across facial expression, a significant N2pc was present in the gender task, $ts(19) = 3.28$, $p = 0.004$, Cohen's $d_z = 0.73$, but not in the pixel task, $p = 0.97$.

HEOG

The same 2 × 3 ANOVA was used to evaluate ocular movements in the lateralized N170, early N2pc, and late N2pc. There was no effect of the HEOG in the time window of the lateralized N170, and late N2pc, suggesting that the voltages measured at electrodes P7/8 were not contaminated by eye movements. However, task had an effect in the early N2pc time window, $F(1,19) = 13.11$, $p = 0.002$, $\eta_p{}^2 = 0.48$, with a larger HEOG signal (and thus more eye movements) in the gender task (−0.7 μV) as compared to the pixel task (0.05 μV). No other main effects or interaction effects reached the level of significance.

Non-lateralized N170

The non-lateralized N170 was the mean voltage from 150 to 190 ms at electrodes P7/8, irrespective of the location of the face stimulus on the screen. We conducted a repeated-measures 2 (hemisphere: left = electrode P7, right = electrode P8) × 2 (task: gender, pixel) × 3 (facial expression: neutral, angry, and happy). The ANOVA returned a main effect of hemisphere, $F(1,19) = 7.69$, $p < 0.0121$, $\eta_p{}^2 = 0.28$, with a more negative N170 on the right (−2.17 μV) than on the left hemisphere (−0.62 μV). The main effect of task was significant, $F(1,19) = 7.69$, $p < 0.012$, $\eta_p{}^2 = 0.28$, with a more negative N170 in the pixel task (−2.21 μV) than in the gender task (−0.66 μV). There was no main effect of facial expression, $p = 0.86$, and critically, there was no interaction of facial expression and task, $p = 0.88$.

DISCUSSION

We investigated whether the effect of threatening faces on early face encoding before attentional selection is modulated by task demands. Past studies have used the non-lateralized N170 as an electrophysiological measure of face encoding and assessed effects of facial expression (Batty and Taylor, 2003; Schyns et al., 2007; Leppanen et al., 2008; Righart and de Gelder, 2008; Smith et al., 2013; Turano et al., 2017). In this study, we did not find changes of the non-lateralized N170 as a function of facial expression, which is in line with previous studies (Pourtois et al., 2005; Rellecke et al., 2012; Calvo and Beltran, 2013; Tamamiya and Hiraki, 2013; Neath-Tavares and Itier, 2016; for review, see Hinojosa et al., 2015). In contrast to prior studies, however, task demands did not modulate the effect of facial expression on the non-lateralized N170 (Eimer et al., 2003; Holmes et al., 2003, 2006; Neath-Tavares and Itier, 2016; for review, see Eimer and Holmes, 2007). Further, we confirm that faces elicited a more negative non-lateralized N170 on the right (electrode P8) than on the left (electrode P7) (Rossion and Jacques, 2008), in line with a predominance of the right hemisphere in face processing (Kanwisher et al., 1997). However, the non-lateralized N170 remained insensitive to facial expressions regardless of task.

Further, we confirmed a contralateral bias in face encoding, the lateralized N170 (Towler et al., 2016; Framorando et al., 2018; Neumann et al., 2018). The amplitude of the lateralized N170 to threat-related stimuli was larger than to non-threatening stimuli when participants performed a gender categorization task. In this task, faces were task-relevant and observers deployed spatial attention to the face location. Conversely, when participants performed a task at central fixation that did not require processing of the structural elements of the faces, the early bias toward threat-related faces was absent. Thus, task demands are critical in the effect of emotional stimuli on the lateralized N170, in line with previous evidence based on the non-lateralized N170 (Eimer et al., 2003; Holmes et al., 2003, 2006; Neath-Tavares and Itier, 2016; for review, see Eimer and Holmes, 2007).

Our research complements studies on the lateralized N170 (Towler et al., 2016; Framorando et al., 2018; Neumann et al., 2018) by investigating the effect of threatening faces. We demonstrated that the lateralized N170 is enhanced in situations where a threatening face is competing with another object, which may be closer to the real world than displays with only a single face (Towler and Eimer, 2015). A possible reason for the dependence on competition is that the presence of a competing stimulus (a house) in the opposite visual field inhibits the interhemispheric communication via the corpus callosum. Consistently, previous research suggests that the lateralized bias originates from brain regions involved in early stages of visual perception. For instance, functional magnetic brain imaging (fMRI) studies suggested that early visual regions, such as the lateral occipital cortex, mediate the contralateral bias in face encoding (Niemeier et al., 2005), rather than the middle fusiform gyrus (Grill-Spector et al., 1999; Hemond et al., 2007) or the posterior superior temporal sulcus (pSTS) (Pitcher et al., 2019). However, the contralateral vs. ipsilateral bias was not present in the amygdala (Pitcher et al., 2019), although this structure is critical in the threat capture account. Possibly, the poor temporal resolution of the fMRI measures in Pitcher et al. (2019) failed to pick up the brief contralateral vs. ipsilateral bias in the amygdala.

Critically, our data shows that task-relevance is critical in the emergence of enhanced perceptual encoding of threatening faces. However, what is the source of this enhancement? Because the enhancement appears before attentional selection, it is difficult to argue that the enhancement results from enhanced attention to the face. Rather, task instructions seem to have increased the sensitivity of face processing, which in turn may have increased the sensitivity to facial expressions. When task demands forced enhanced face processing, faces conveying a threatening content were preferentially processed compared to non-threatening faces. Thus, enhanced processing of threatening faces depends crucially on attention to the faces. Without attention, it may well be that the N170 is insensitive to facial expressions. Nevertheless, we acknowledge that further studies are needed to address the specific role of attention in this effect. For instance, we think that spatial attention was deployed to the location of the face in our study, but it may be interesting to see whether attention to facial features across the display has a similar effect.

More generally, our data reveal the automaticity of the lateralized N170 and at the same time, the influence of task demands. First, face encoding occurs even if faces are task-irrelevant and outside the focus of attention, as indicated by the presence of a lateralized N170 even when attention was focused on the fixation cross (i.e., in the pixel task). Therefore, the N170 is a preattentive marker of the presence of a face, similar to a face detector. However, when the task required more detailed processing of the facial features (i.e., in the gender task), the contralateral bias reflected differences between emotions. Because the same stimuli were used in both gender and pixel tasks, this difference cannot be explained by sensorial differences among faces. It seems likely that the processing required to perform the task increased the sensitivity of early face encoding. While the "automatic" lateralized N170 is elicited solely by the presence of a face in the environment (Neumann et al., 2018), the more "voluntary" lateralized N170 might be elicited by more detailed processing. The two components of the lateralized N170 might correspond to the extraction of two levels of configural information (Rhodes, 1988; Maurer et al., 2002). First-order configural information consists of spatial relations between constituent elements of an object (e.g., the arrangement of the nose above the mouth), which allows for categorization as a face. Second-order information consists of the relative size of these spatial relationships, which may be critical in the gender task (Baudouin and Humphreys, 2006; Zhao and Hayward, 2010). Extraction of second-order features is also necessary to distinguish facial expression. Thus, the larger lateralized N170 to angry faces may reflect the extraction of second-order features.

There are several caveats to this study. First, we found limited behavioral evidence for enhanced processing of threat-related stimuli. In the anger superiority effect (ASE), angry faces are found more rapidly than happy faces in visual search displays (Hansen and Hansen, 1988). While angry faces were processed more quickly than happy faces in our study, there was no difference compared to neutral faces. Moreover, accuracy in the gender task was better with neutral faces compared to angry or happy faces. Thus, there was evidence of enhanced processing of angry faces in the lateralized N170, but none in behavior. Possibly, behavioral markers of the ASE are less sensitive with gender categorization, but may emerge in other tasks, such as face detection. Alternatively, the limited number of items in the display as well as the lack of competing face stimuli might explain the absence of the ASE.

Second, we did not observe an enhanced N2pc to angry faces in the gender task, which is surprising. Previous studies have revealed an enhancement of the N2pc to threat-related objects, which might be seen as the outcome of an attentional bias to threat (Feldmann-Wustefeld et al., 2011; Weymar et al., 2011; Yao et al., 2013). Thus, the larger N2pc to angry faces has been taken as electrophysiological evidence for the ASE. However, angry faces did not produce an enhanced N2pc in the current study, even when they were relevant (i.e., in the gender task). Possibly, the reason is that the target face in our study was not competing with other faces, but with a non-face stimulus. It is plausible that an attentional bias to threat-related stimuli is not only sensitive to faces *per se* but is also sensitive to the distractors competing for selection. Notably, the N2pc is composed of a target-related and a distractor-related component (N_t and P_D components, respectively, Hickey et al., 2009). Possibly, the P_D component of the N2pc was enhanced with face compared to house distractors, which explains the larger N2pc to angry faces in previous studies.

Finally, the lack of attentional bias to angry faces in the pixel task is also inconsistent with evidence of an early attentional bias to threatening faces or objects when the threatening stimulus was task-irrelevant (Burra et al., 2016, 2017, 2019). However, faces in the pixel task of the current experiment were irrelevant *and* entirely outside the focus of attention, whereas in previous studies, the targets were faces competing for selection with irrelevant facial expressions that were within the attentional focus. In fact, salient objects fail to capture attention when presented outside the focus of attention (Belopolsky and Theeuwes, 2010; Kerzel et al., 2012). Because attention in the pixel task was only allocated to the fixation cross and not to the lateral faces, attentional capture by threatening faces may have been absent. Moreover, in previous studies, the target and the irrelevant distractor belonged to the same category, which may have increased attentional priority of the target face. Further research may use a central task with stimuli belonging to the same category as the irrelevant lateralized stimuli (i.e., a face) or requiring the participant to attend to the entire visual display.

In sum, our results reveal that task demands modulate the preferential encoding of threatening faces prior to attentional selection, which is inconsistent with the hypothesis of automatic threat detection, even outside the focus of attention (Öhman and Mineka, 2001). Notably, our results suggests that top-down control plays a role in the early processing of threatening stimuli, as reflected in the lateralized N170 (Towler and Eimer, 2015; Towler et al., 2016). Therefore, our study contributes to the growing evidence in favor of the critical role of task demands in supposedly automatic effects.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by local ethics committee (Faculty of

Psychology and Educational Sciences, Geneva University). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

NB collected and analyzed the data and wrote the first draft of the manuscript. DK edited the manuscript.

FUNDING

DK was supported by grant no. 100019_182146 from the Swiss National Science Foundation.

ACKNOWLEDGMENTS

We wish to thank Jessica Robinson for help with data acquisition.

REFERENCES

Aranda, C., Madrid, E., Tudela, P., and Ruz, M. (2010). Category expectations: a differential modulation of the N170 potential for faces and words. *Neuropsychologia* 48, 4038–4045. doi: 10.1016/j.neuropsychologia.2010.10.002

Batty, M., and Taylor, M. J. (2003). Early processing of the six basic facial emotional expressions. *Brain Res. Cogn. Brain Res.* 17, 613–620. doi: 10.1016/S0926-6410(03)00174-5

Baudouin, J. Y., and Humphreys, G. W. (2006). Configural information in gender categorisation. *Perception* 35, 531–540. doi: 10.1068/p3403

Belopolsky, A. V., and Theeuwes, J. (2010). No capture outside the attentional window. *Vis. Res.* 50, 2543–2550. doi: 10.1016/j.visres.2010.08.023

Bentin, S., Allison, T., Puce, A., Perez, E., and McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *J. Cogn. Neurosci.* 8, 551–565. doi: 10.1162/jocn.1996.8.6.551

Bretherton, P. M., Eysenck, M. W., Richards, A., and Holmes, A. (2017). Target and distractor processing and the influence of load on the allocation of attention to task-irrelevant threat. *Neuropsychologia* doi: 10.1016/j.neuropsychologia.2017.09.009 [Epub ahead of print].

Bruce, V., and Young, A. (1986). Understanding face recognition. *Br. J. Psychol.* 77(Pt 3), 305–327. doi: 10.1111/j.2044-8295.1986.tb02199.x

Burra, N., Barras, C., Coll, S. Y., and Kerzel, D. (2016). Electrophysiological evidence for attentional capture by irrelevant angry facial expressions. *Biol. Psychol.* 120, 69–80. doi: 10.1016/j.biopsycho.2016.08.008

Burra, N., Coll, S. Y., Barras, C., and Kerzel, D. (2017). Electrophysiological evidence for attentional capture by irrelevant angry facial expressions: naturalistic faces. *Neurosci. Lett.* 637, 44–49. doi: 10.1016/j.neulet.2016.11.055

Burra, N., Pittet, C., Barras, C., and Kerzel, D. (2019). Attentional suppression is delayed for threatening distractors. *Vis. Cogn.* 27, 185–198. doi: 10.1080/13506285.2019.1593272

Calvo, M. G., and Beltran, D. (2013). Recognition advantage of happy faces: tracing the neurocognitive processes. *Neuropsychologia* 51, 2051–2061. doi: 10.1016/j.neuropsychologia.2013.07.010

Carmel, D., and Bentin, S. (2002). Domain specificity versus expertise: factors influencing distinct processing of faces. *Cognition* 83, 1–29. doi: 10.1016/S0010-0277(01)00162-7

Cauquil, A. S., Edmonds, G. E., and Taylor, M. J. (2000). Is the face-sensitive N170 the only ERP not affected by selective attention? *Neuroreport* 11, 2167–2171. doi: 10.1097/00001756-200007140-00021

daSilva, E. B., Crager, K., Geisler, D., Newbern, P., Orem, B., and Puce, A. (2016). Something to sink your teeth into: the presence of teeth augments ERPs to mouth expressions. *Neuroimage* 127, 227–241. doi: 10.1016/j.neuroimage.2015.12.020

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalogr. Clin. Neurophysiol.* 99, 225–234. doi: 10.1016/0013-4694(96)95711-9

Eimer, M. (2000). The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport* 11, 2319–2324. doi: 10.1097/00001756-200007140-00050

Eimer, M., and Holmes, A. (2007). Event-related brain potential correlates of emotional face processing. *Neuropsychologia* 45, 15–31. doi: 10.1016/j.neuropsychologia.2006.04.022

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

Feldmann-Wustefeld, T., Schmidt-Daffy, M., and Schubo, A. (2011). Neural evidence for the threat detection advantage: differential attention allocation to angry and happy faces. *Psychophysiology* 48, 697–707. doi: 10.1111/j.1469-8986.2010.01130.x

Framorando, D., Burra, N., Bapst, M., and Pegna, A. J. (2018). ERP responses greater for faces in the temporal compared to the nasal visual field. *Neurosci. Lett.* 665, 7–12. doi: 10.1016/j.neulet.2017.11.031

Freiwald, W., Duchaine, B., and Yovel, G. (2016). Face processing systems: from neurons to real-world social perception. *Ann. Rev. Neurosci.* 39, 325–346. doi: 10.1146/annurev-neuro-070815-013934

George, N., Evans, J., Fiori, N., Davidoff, J., and Renault, B. (1996). Brain events related to normal and moderately scrambled faces. *Brain Res. Cogn. Brain Res.* 4, 65–76. doi: 10.1016/0926-6410(95)00045-3

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., and Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24, 187–203. doi: 10.1016/S0896-6273(00)80832-6

Hansen, C. H., and Hansen, R. D. (1988). Finding the face in the crowd: an anger superiority effect. *J. Pers. Soc. Psychol.* 54, 917–924. doi: 10.1037/0022-3514.54.6.917

Hemond, C. C., Kanwisher, N. G., and Op de Beeck, H. P. (2007). A preference for contralateral stimuli in human object- and face-selective cortex. *PLoS One* 2:e574. doi: 10.1371/journal.pone.0000574

Hickey, C., Di Lollo, V., and McDonald, J. J. (2009). Electrophysiological indices of target and distractor processing in visual search. *J. Cogn. Neurosci.* 21, 760–775. doi: 10.1162/jocn.2009.21039

Hinojosa, J. A., Mercado, F., and Carretie, L. (2015). N170 sensitivity to facial expression: a meta-analysis. *Neurosci. Biobehav. Rev.* 55, 498–509. doi: 10.1016/j.neubiorev.2015.06.002

Holmes, A., Kiss, M., and Eimer, M. (2006). Attention modulates the processing of emotional expression triggered by foveal faces. *Neurosci. Lett.* 394, 48–52. doi: 10.1016/j.neulet.2005.10.002

Holmes, A., Vuilleumier, P., and Eimer, M. (2003). The processing of emotional facial expression is gated by spatial attention: evidence from event-related brain potentials. *Cogn. Brain Res.* 16, 174–184. doi: 10.1016/S0926-6410(02)00268-9

Itier, R. J., and Neath-Tavares, K. N. (2017). Effects of task demands on the early neural processing of fearful and happy facial expressions. *Brain Res.* 1663, 38–50. doi: 10.1016/j.brainres.2017.03.013

Joyce, C., and Rossion, B. (2005). The face-sensitive N170 and VPP components manifest the same brain processes: the effect of reference electrode site. *Clin. Neurophysiol.* 116, 2613–2631. doi: 10.1016/j.clinph.2005.07.005

Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., McKeown, M. J., Iragui, V., et al. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* 37, 163–178. doi: 10.1111/1469-8986.3720163

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. doi: 10.1523/JNEUROSCI.17-11-04302.1997

Kerzel, D., Born, S., and Schonhammer, J. (2012). Perceptual grouping allows for attention to cover noncontiguous locations and suppress capture from nearby locations. *J. Exp. Psychol.* 38, 1362–1370. doi: 10.1037/a0027780

Kleiner, M., Brainard, D., and Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception* 36, 14–14.

Leppanen, J. M., Hietanen, J. K., and Koskinen, K. (2008). Differential early ERPs to fearful versus neutral facial expressions: a response to the salience of the eyes? *Biol. Psychol.* 78, 150–158. doi: 10.1016/j.biopsycho.2008.02.002

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: The MIT Press.

Luck, S. J. (2012). *Electrophysiological Correlates of the Focusing of Attention within Complex Visual Scenes: N2pc and Related ERP Components*. Oxford: Oxford University Press.

Luck, S. J., and Hillyard, S. A. (1994). Spatial filtering during visual search: evidence from human electrophysiology. *J. Exp. Psychol.* 20, 1000–1014. doi: 10.1037/0096-1523.20.5.1000

Lundqvist, D., Flykt, A., and Öhman, A. (1998). *The Karolinska Directed Emotional Faces (KDEF). CD ROM*. Stockholm: Karolinska Institute, 630.

Maurer, D., Grand, R. L., and Mondloch, C. J. (2002). The many faces of configural processing. *Trends Cogn. Sci.* 6, 255–260. doi: 10.1016/S1364-6613(02)01903-4

Neath-Tavares, K. N., and Itier, R. J. (2016). Neural processing of fearful and happy facial expressions during emotion-relevant and emotion-irrelevant tasks: a fixation-to-feature approach. *Biol. Psychol.* 119, 122–140. doi: 10.1016/j.biopsycho.2016.07.013

Neumann, M. F., Viska, C. G., van Huis, S., and Palermo, R. (2018). Similar distraction, but differential suppression, for faces and non-face objects: evidence from behaviour and event-related potentials. *Biol. Psychol.* 139, 39–46. doi: 10.1016/j.biopsycho.2018.09.011

Niemeier, M., Goltz, H. C., Kuchinad, A., Tweed, D. B., and Vilis, T. (2005). A contralateral preference in the lateral occipital area: sensory and attentional mechanisms. *Cereb. Cortex* 15, 325–331. doi: 10.1093/cercor/bhh134

Öhman, A., Juth, P., and Lundqvist, D. (2010). Finding the face in a crowd: relationships between distractor redundancy, target emotion, and target gender. *Cogn. Emot.* 24, 1216–1228. doi: 10.1080/02699930903166882

Öhman, A., and Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychol. Rev.* 108, 483–522. doi: 10.1037/0033-295X.108.3.483

Pitcher, D., Pilkington, A., Rauth, L., Baker, C., Kravitz, D. J., and Ungerleider, L. G. (2019). The human posterior superior temporal sulcus samples visual space differently from other face-selective regions. *Cereb. Cortex* 00, 1–8. doi: 10.1093/cercor/bhz125

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

Rellecke, J., Sommer, W., and Schacht, A. (2012). Does processing of emotional facial expressions depend on intention? Time-resolved evidence from event-related brain potentials. *Biol. Psychol.* 90, 23–32. doi: 10.1016/j.biopsycho.2012.02.002

Rellecke, J., Sommer, W., and Schacht, A. (2013). Emotion effects on the n170: a question of reference? *Brain Topogr.* 26, 62–71. doi: 10.1007/s10548-012-0261-y

Rhodes, G. (1988). Looking at faces: first-order and second-order features as determinants of facial appearance. *Perception* 17, 43–63. doi: 10.1068/p170043

Righart, R., and de Gelder, B. (2008). Recognition of facial expressions is influenced by emotional scene gist. *Cogn. Affect. Behav. Neurosci.* 8, 264–272. doi: 10.3758/CABN.8.3.264

Rossion, B., and Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *Neuroimage* 39, 1959–1979. doi: 10.1016/j.neuroimage.2007.10.011

Sawaki, R., and Luck, S. J. (2010). Capture versus suppression of attention by salient singletons: electrophysiological evidence for an automatic attend-to-me signal. *Atten. Percept. Psychophys.* 72, 1455–1470. doi: 10.3758/APP.72.6.1455

Schindler, S., Bruchmann, M., Bublatzky, F., and Straube, T. (2019). Modulation of face- and emotion-selective ERPs by the three most common types of face image manipulations. *Soc. Cogn. Affect. Neurosci.* 14, 493–503. doi: 10.1093/scan/nsz027

Schinkel, S., Ivanova, G., Kurths, J., and Sommer, W. (2014). Modulation of the N170 adaptation profile by higher level factors. *Biol. Psychol.* 97, 27–34. doi: 10.1016/j.biopsycho.2014.01.003

Schyns, P. G., Petro, L. S., and Smith, M. L. (2007). Dynamics of visual information integration in the brain for categorizing facial expressions. *Curr. Biol.* 17, 1580–1585. doi: 10.1016/j.cub.2007.08.048

Smith, E., Weinberg, A., Moran, T., and Hajcak, G. (2013). Electrocortical responses to NIMSTIM facial expressions of emotion. *Int. J. Psychophysiol.* 88, 17–25. doi: 10.1016/j.ijpsycho.2012.12.004

Smith, F. W., and Smith, M. L. (2019). Decoding the dynamic representation of facial expressions of emotion in explicit and incidental tasks. *Neuroimage* 195, 261–271. doi: 10.1016/j.neuroimage.2019.03.065

Sreenivasan, K. K., Goldstein, J. M., Lustig, A. G., Rivas, L. R., and Jha, A. P. (2009). Attention to faces modulates early face processing during low but not high face discriminability. *Atten. Percept. Psychophys.* 71, 837–846. doi: 10.3758/APP.71.4.837

Sreenivasan, K. K., Katz, J., and Jha, A. P. (2007). Temporal characteristics of top-down modulations during working memory maintenance: an event-related potential study of the N170 component. *J. Cogn. Neurosci.* 19, 1836–1844. doi: 10.1162/jocn.2007.19.11.1836

Tamamiya, Y., and Hiraki, K. (2013). Individual differences in the recognition of facial expressions: an event-related potentials study. *PLoS One* 8:e57325. doi: 10.1371/journal.pone.0057325

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

Towler, J., and Eimer, M. (2015). Early stages of perceptual face processing are confined to the contralateral hemisphere: evidence from the N170 component. *Cortex* 64, 89–101. doi: 10.1016/j.cortex.2014.09.013

Towler, J., Kelly, M., and Eimer, M. (2016). The focus of spatial attention determines the number and precision of face representations in working memory. *Cereb. Cortex* 26, 2530–2540. doi: 10.1093/cercor/bhv083

Turano, M. T., Lao, J., Richoz, A.-R., Lissa, P. D., Degosciu, S. B., Viggiano, M. P., et al. (2017). Fear boosts the early neural coding of faces. *Soc. Cogn. Affect. Neurosci.* 12, 1959–1971. doi: 10.1093/scan/nsx110

Weymar, M., Low, A., Ohman, A., and Hamm, A. O. (2011). The face is more than its parts–brain dynamics of enhanced spatial attention to schematic threat. *Neuroimage* 58, 946–954. doi: 10.1016/j.neuroimage.2011.06.061

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., and Tanaka, J. W. (2010). Controlling low-level image properties: the SHINE toolbox. *Behav. Res. Methods* 42, 671–684. doi: 10.3758/BRM.42.3.671

Wronka, E., and Walentowska, W. (2011). Attention modulates emotional expression processing. *Psychophysiology* 48, 1047–1056. doi: 10.1111/j.1469-8986.2011.01180.x

Yao, S., Ding, C., Qi, S., and Yang, D. (2013). Value associations of emotional faces can modify the anger superiority effect: behavioral and electrophysiological evidence. *Soc. Cogn. Affect. Neurosci.* 9, 849–856. doi: 10.1093/scan/nst056

Zhao, M., and Hayward, W. G. (2010). Holistic processing underlies gender judgments of faces. *Atten. Percept. Psychophys.* 72, 591–596. doi: 10.3758/APP.72.3.591

Conflict of Interest: 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.

frontiers in Psychology

ORIGINAL RESEARCH
published: 29 November 2019
doi: 10.3389/fpsyg.2019.02649

Effects of Baby Schema and Mere Exposure on Explicit and Implicit Face Processing

Leonardo Venturoso[1], **Giulio Gabrieli**[2], **Anna Truzzi**[3], **Atiqah Azhari**[2], **Peipei Setoh**[2], **Marc H. Bornstein**[4,5] **and Gianluca Esposito**[1,2*]

[1] Department of Psychology and Cognitive Science, University of Trento, Trento, Italy, [2] Psychology Program, School of Social Sciences, Nanyang Technological University, Singapore, Singapore, [3] Trinity College Institute of Neuroscience, Trinity College Dublin, Dublin, Ireland, [4] Institute for Fiscal Studies, London, United Kingdom, [5] National Institute of Child Health and Human Development, Bethesda, MD, United States

Edited by:
Giulia Mattavelli,
University Institute of Higher Studies in Pavia, Italy

Reviewed by:
Michio Nomura,
Kyoto University, Japan
Eirini Mavritsaki,
Birmingham City University,
United Kingdom

***Correspondence:**
Gianluca Esposito
gianluca.esposito@ntu.edu.sg;
gianluca.esposito@unitn.it

Specialty section:
This article was submitted to Consciousness Research, a section of the journal Frontiers in Psychology

Received: 24 June 2019
Accepted: 11 November 2019
Published: 29 November 2019

Citation:
Venturoso L, Gabrieli G, Truzzi A, Azhari A, Setoh P, Bornstein MH and Esposito G (2019) Effects of Baby Schema and Mere Exposure on Explicit and Implicit Face Processing. Front. Psychol. 10:2649. doi: 10.3389/fpsyg.2019.02649

In an increasingly multicultural society, the way people perceive individuals from the same vs different ethnic groups greatly affects their own and societal well-being. Two psychological effects that influence these perceptions are the Mere-Exposure Effect (MRE), wherein familiarity with certain objects or persons suffices for people to develop a preference for them, and the Baby Schema (BS), a set of specific facial features that evokes caregiving behaviors and an affective orientation in adults. In the present study, we aimed to investigate whether these two effects play a role in implicit physiological responses to babies vs. adults faces belonging to participants in-group vs. out-group. In study 1, the pupillary diameter of 62 Caucasian participants (M = 31; F = 31) who observed adult and infant faces of different ethnic groups (Caucasian, Chinese) was measured. In study 2, brain waves of 38 Caucasian participants (M = 19; F = 19), who observed the same set of faces, were recorded using EEG. In both studies, adults explicit preferences (i.e., attitudes) toward faces were assessed using questionnaires. In Study 1, females showed greater attention to infant than adult faces (BS effect) in both pupils, regardless of the ethnic group of the face. By contrast, males attended to infant more than adult faces for out-group faces only (BS effect). In Study 2, greater left posterior-parietal alpha activation toward out-group compared to in-group adult faces was found in males (MRE). Participants with a low BS effect toward in-group baby faces exhibited greater left posterior alpha activation to out-group than in-group baby faces (MRE). These findings reveal how different levels of sensitivity to in-group infants may moderate perceptions of both in-group and out-group baby faces. Questionnaire measures on attitudes showed that males and females preferred in-group to out-group adult faces (MRE). Participants in Study 2 also reported a greater preference for infants than adults faces (BS effect). These findings explicate the roles of gender and the Baby Schema effect in moderating implicit processing of in-group and out-group faces, despite their lack in moderating explicit reports. Contradictory findings at the implicit (physiological) and explicit (self-report) levels suggest that differential processing of faces may occur at a non-conscious level.

Keywords: baby schema, mere exposure effect, face processing, neural oscillations, pupillometry, EEG

1. INTRODUCTION

Faces are central in social cognition. From a person's face, we efficiently extract an immense amount of information (e.g., age and ethnicity) that guide our behavioral responses (Willis and Todorov, 2006). Of particular salience is the infant's face, which is crucial in the context of parenting. It is well-known that rapid and appropriate responses to infants needs enable the wholesome development of social, cognitive, and emotional domains, whereas incoherent responses lead to adverse developmental trajectories (Bowlby, 1969; Waters et al., 2000; Furman and Buhrmester, 2009; Bornstein, 2015). Adults parental responses have been shown to be facilitated by a rapid shift in attention toward infants faces. As compared to faces of adults and non-human infants/adults, faces of human infants command greater levels of attention (Brosch et al., 2007). The automatic orientation to an infant's face suggests that it serves as a cue to trigger a distinct set of brain responses that promotes adults adaptive caregiving (Plutchik, 1987; Seifritz et al., 2003).

Baby Schema refers to the set of physical and behavioral characteristics typical of babies (such as big round eyes, big head, and small face, small ears, short limbs, clumsy gait) that evoke protective caregiving behaviors in adults (Lorenz, 1943; Lorenz and Martin, 1971). The Baby Schema effect is more evident in women than in men, suggesting gender differences in processing infant faces (Baron-Cohen et al., 2003; Glocker et al., 2009a). However, little is known of whether and how the familiarity of ethnic in-group facial cues, as compared to less familiar out-group features, moderates adults responses to infant faces. The Mere Exposure Effect is a phenomenon by which increased exposure to an object or person leads to enhanced familiarity and contributes to developing a preference (Zajonc, 1968). For instance, Bornstein and Tamis-LeMonda (1989) found that a wide range of affective responses increases with exposure (e.g., pleasantness, liking, etc.), and this effect is applicable to a variety of stimuli (e.g., abstract paintings, polygons, line drawings, etc.), including faces. The Mere Exposure Effect suggests that preference for faces would increase to familiar in-group adult and infant faces. However, it remains to be seen whether the protective effect of Baby Schema would lead to an increased preference for infant faces regardless of whether the face belongs to an in-group (familiar) or out-group (unfamiliar). Preferences for facial stimuli can be measured in many ways, such as by recording autonomic physiological responses. Pupillometry assesses pupillary response to emotionally relevant visual stimuli.

Dilation of the pupil indicates physiological arousal, reflecting the activation of the sympathetic branch (i.e., arousal response) of the autonomic nervous system toward visual stimuli (Hess and Polt, 1960; Aktar et al., 2016; Wetzel et al., 2016). In general, larger pupillary dilations are evoked by more pleasant or more aversive stimuli (Steinhauer, 1983; Bradley et al., 2008). On the other hand, pupil constriction happens when the circular muscle contracts, reflecting the activation of the parasympathetic branch (i.e., calming response) of the autonomic nervous system. However, the interpretation of the meaning of pupil constriction is not completely clear and unambiguous. In fact, while Loewenfeld (1966) argued that the only factor causing pupil constriction is the increased light intensity and while Hess (1972) suggested that constriction occurs in response to unpleasant or distasteful visual stimuli, other recent studies (Mathôt et al., 2013; Mathôt and Van der Stigchel, 2015; Landi and Freiwald, 2017; Winn et al., 2018; Zekveld et al., 2018) underlined that distinct cognitive processes, including attentional, emotional and motivational ones, may differently intervene in mediating the strength of pupillary constriction. For example, Landi and Freiwald (2017) found that greater constrictions occur in response to familiar stimuli than unfamiliar ones. The pupillometry has been used in order to study several variables involved in face processing (i.e., age, ethnic group, and gender). Regarding the ethnic group, differential responses to in-group and out-group are reflected in pupil dilation and constriction. For instance, Goldinger et al. (2009) and Wu et al. (2012) found that viewers pupil size was larger when processing out-group than in-group faces, meaning that more attention is given to out-group faces. Moreover, Landi and Freiwald (2017) found that viewers pupil size decrease considerably when processing familiar faces compared to unfamiliar ones, suggesting that familiar faces request fewer cognitive resources in order to be processed. These studies show that the Mere Exposure Effect is reliably observed using measures of pupillometry.

Another variables that have been investigated are gender and age. In accordance with known gender effects of Baby Schema, women display relatively larger pupillary dilation, which means greater attention to certain stimuli, when they view pictures of infants than people of other ages (Hess and Polt, 1960). To our knowledge, no studies have found larger pupillary responses for adult faces compared to baby faces in female participants. Thus, the Baby Schema effect is sensitively captured by pupillometry.

Several studies (Andrew et al., 1982; Hamilton and Vermeire, 1988; Prior et al., 2002; Regolin et al., 2004) have also found that the pupillary response may variate between pupils and that this laterality may indicate differentiated processing of stimuli. For instance, Andrew et al. (1982) found that the left eye of the domestic chick is more reactive to emotional stimuli compared to the right one. Hamilton and Vermeire (1988) also claimed that the domestic chick's right eye is involved in foraging behavior, while the left one is engaged in alert behavior. Moreover, Prior et al. (2002) and Regolin et al. (2004) have proved that the left eye and its contralateral connections to the right hemisphere are relevant in the processing of spatial information.

A more complete understanding of physiological responses requires however measures of both peripheral and central nervous systems. Pupillometry captures the former, but electroencephalography (EEG) records brain oscillatory activity as an index of the latter. Asymmetrical alpha band oscillations (8–12 Hz) indicate facial preferences (Kang et al., 2015). This asymmetrical pattern is elicited if stimuli are especially arousing, regardless of their emotional valence (Balconi and Mazza, 2009). Investigations using infant stimuli have revealed that electrophysiological activity toward infants faces is greater in women than in men (Proverbio et al., 2011). Proverbio and De Gabriele (2017) also showed that there was no observable difference in electrophysiological activity when viewing infant faces of different ethnicities, lending support to the wide application of protective effects of Baby Schema. EEG has also been used to examine the Mere Exposure Effect. Thiruchselvam et al. (2016) showed that increased exposure to infant faces affects attractiveness ratings and posterior neural reactivity. In a study using in-group and out-group faces, Zheng and Segalowitz (2014) learned that group membership influences brain electrical activity specific to faces. Thus, in addition to pupillometry, we used EEG to investigate both Baby Schema and Mere Exposure Effects. Despite past extensive investigation on this topic, no study has considered how responses to in-group baby faces and out-group baby faces are associated, and how the processing of faces in each of these categories may be moderated by participants gender.

The present study aimed to investigate Baby Schema and Mere Exposure Effects on adults peripheral and central nervous system and attitudinal responses to infant and adult faces of ethnic in-group and out-group. We conducted two studies: Study 1 used pupillometry in Caucasians ($N = 62$), and Study 2 recorded EEG in Caucasians ($N = 38$). In both studies, participants were presented with Chinese and Caucasian adults and infants faces. Alongside implicit neurophysiological measures, adults explicit attitudes to faces were assessed using a questionnaire. Several studies, including Glocker et al. (2009a,b) and Caria et al. (2012) focused on how BS activates the neural system of both females and males, while others, such as Esposito et al. (2015) explicitly compared emotional responses in the two genders. However, as far as we know, no study has taken into account an additional variable which is the familiarity of faces (in-group vs. out-group) to investigate how BS influences face processing. The Life History Theory (LHT) highlights that males and females of different species exhibit distinct attitudes toward babies and parental activities (Draper and Harpending, 1982; Mascaro et al., 2013). They adopt distinct strategies to optimize their fitness: parenting and mating. As females place more resources in reproduction (for example for pregnancy and labor), evolution motivates her to engage in parenting/caregiving activities that improve the survival and quality of the offspring. Indeed, it seems that higher rates of reproduction in humans are associated with lower offspring and parental survivorship, especially for mothers (Penn and Smith, 2007). On the contrary, since the investment of males in reproduction is usually lower, evolution motivates them toward mating. That could explain why females, compared to males, generally tend to be more sensitive toward babies. Moreover, in an increasingly multicultural society, the investigation of psychological factors that may affect the way people perceive familiar vs unfamiliar individuals (including infants) become pivotal. Both infant and adult faces of in-group and out-group ethnicity are thus stimuli that are presumed to have salient emotional valence, which are probably different in men and women. This study had four hypotheses. First, we expected females to exhibit the Baby Schema effect to both in-group and out-group infant faces. This result would have Baby Schema prevail over the Mere Exposure Effect. Second, taking into account that men show a greater MRE compared to women, we expected an interaction between BS and MRE to be evident in males. Third, considering that individual differences were found in previous infant and adult face processing studies (e.g., Bornstein and Tamis-LeMonda, 1989; Lehmann et al., 2013), we hypothesized that one's neural activation pattern to in-group baby faces and out-group adult faces could be correlated. Past studies have shown that this correlation is gender-dependent. Generally, women, who are known to have higher levels of Baby Schema Effect (e.g., Hess and Polt, 1960; Glocker et al., 2009a), also display higher levels of activation toward out-group faces compared to in-group faces (Richeson and Trawalter, 2008; Trawalter et al., 2008). Fourth, we expected females, but not males, to report no difference in preference between in-group and out-group infant faces, given that females take a greater interest in infants in general and are more inclined in caregiving activities (Maestripieri and Pelka, 2002).

2. STUDY 1

2.1. Materials and Methods

2.1.1. Participants

A total of 62 Caucasian Italian adults (31 males, 31 females, Mean Age = 22.55 years, SD = 2.57) were recruited through a database of volunteers available through the University of Trento web site. Informed consent was obtained from all participants, and no incentives were provided. Exclusion criteria were parenthood, pregnancy, and non-Caucasian ethnicity. The study was conducted in accordance with the ethical principles stated in the Helsinki declaration and it was approved by the IRB of the Nanyang Technological University.

2.1.2. Stimuli

Used stimuli were drawn from two different available datasets: asian faces have been selected from Yap et al. (2016), while caucasian faces used in this work have been provided by the Computer Vision Laboratory, University of Ljubljana, Slovenia (Solina et al., 2003; Peer, 2005). Permission has been obtained for the usage of the images. Forty pictures of faces were shown to participants. The faces presented belonged to the following categories that allowed for the manipulation of face age and ethnic group: (i) Baby Caucasian, (ii) Adult Caucasian, (iii) Baby Chinese, (iv) Adult Chinese. Pictures represented neutrally expressive infant and adult female faces (13 × 17 cm) and were obtained from public domain databases. Stimuli were presented in black and white and matched for contrast and brightness using iOS Preview's Tools. Each face was circled within a gray frame to exclude possible distracting information such as hair or background (circle: $d = 22$ cm, area = $380\,\text{cm}^2$ and circumference = 69 cm). Thus, the stimuli focused on facial feature information.

2.1.3. Experimental Procedure

Participants arrived in the laboratory for the experimental session and completed informed consent before starting the experiment. The experiment took place in a quiet darkened environment. We recorded pupil change using an eye-tracking device (Tobii T120, screen: 34 × 27 cm) created by Tobii[1]. After eye tracking calibration, the session started. A within-subjects design was

---

[1] https://www.tobii.com

used: each participant saw all pictures in random order. In accordance with previous studies (Bernard et al., 2015; Endendijk et al., 2018), each face was presented for 4 sec followed by a recovery period during which a gray screen, with a central cross served as a fixation point, was displayed for 3 s (**Figure 1**).

At the conclusion of the procedure, participants completed a questionnaire about their attitude toward each face. All faces were presented once again with three questions aimed to assess participants attitudes: (i) How positive is your attitude toward the face? (ii) How close do you feel to this individual? and (iii) How much do you like this individual? Participants answered each question by moving a sliding cursor on a 0–100 scale, where 0 represented the most negative emotional valence, and 100 represented the most positive emotional valence.

2.2. Analysis

2.2.1. Preliminary Analysis

Prior to data analysis, pupil width values were examined for normality, homogeneity of variance, presence of outliers, and influential cases. Outliers, defined as values 2 standard deviations above/below the mean, were log-transformed to make the data conform to normality (Keene, 1995). Average pupil width during the fixation screen preceding each face was considered as a baseline. This baseline measure was subtracted from the average pupil width during the presentation of each face to compute the change in pupil width specific to each face. Change in pupil width for each face was then averaged across the categories of infant and adult in-group and out-group faces (Mathôt, 2018).

Concerning attitudinal responses, participants answers to the three questions for each face category were highly correlated in both genders (females' Attitude and Closeness: Pearson's r = 0.852, p <2.2e-16; Attitude and Pleasantness: Pearson's r = 0.871, p < 2.2e-16; Closeness and Pleasantness: Pearson's r = 0.947, p < 2.2e-16, and males' Attitude and Closeness: Pearson's r = 0.931, p < 2.2e-16; Attitude and Pleasantness: Pearson's r = 0.946, p < 2.2e-16; Closeness and Pleasantness: Pearson's r = 0.960, p < 2.2e-16). Therefore, to take the effect of inter-individual differences on pupil width scores into account, a linear model between pupil width and attitude scores was run. The residuals of the model, representing pupil width changes unexplained by participants' attitude, were used in further analysis.

2.2.1.1. Inferential analysis

Two three-way ANOVA models (one for the left pupil, one for the right pupil) were performed with pupil width residuals as the dependent variable, the two within-subjects factors (face age: baby/adult; ethnic group: Caucasian/Chinese) and one between-subject factor (Sex: Male/Female) as the independent variables. Whenever significant effects emerged, *post-hoc* analyses were carried out by performing Student's t-tests. Effect sizes were evaluated using Cohen's d.

2.3. Results

2.3.1. Physiological Results

Since it is well-documented that the right and left hemispheres process visual stimuli differently and that this lateralization is also reflected in the pupil variation of the two eyes (Andrew et al., 1982; Hamilton and Vermeire, 1988; Prior et al., 2002; Regolin et al., 2004), we decided to analyze the data separately to verify the presence of any differences in our work. To test hypothesis 1, that females exhibit the Baby Schema effect to both in-group and out-group infant faces, we undertook a 2 × 2 × 2 Analysis of Variance (ANOVA) with age and ethnicity of face as within-subjects factors and gender a between-subject factor; this analysis supported hypothesis 1. After application of a Bonferroni correction to take into account multiple tests, a main effect of ethnic group was found on pupil width of both pupils [left pupils: $F_{(1,60)} = 27.49$, $p = 2.157496\text{e-}06$, $\eta^2 = 0.0244$; right pupils: $F_{(1,60)} = 35.53$, $p = 1.430421\text{e-}07$, $\eta^2 = 0.0189$]. *Post-hoc* analysis revealed that only women increased pupil diameter from baseline toward out-group Chinese faces compared to in-group Caucasian faces [left pupils: $t_{(61)} = -3.974$, $p = 0.0002$, $\eta^2 = -0.5047$; right pupils: $t_{(61)} = -3.6825$, $p = 0.0005$, $\eta^2 = -0.4677$]. To test hypothesis 2, that an association between BS and Mere Exposure Effect (MRE) is evident in males, we observed for interaction effects from the ANOVA models; this analysis partially supports our hypothesis. A significant interaction effect between sex (Males vs. Females) and ethnic group (Caucasian vs. Chinese) of the face was found [left pupils: $F_{(1,60)} = 8.33$, $p = 5.400192\text{e-}03$, $\eta^2 = 0.0075$; right pupils : $F_{(1,60)} = 9.82$, $p = 2.673374\text{e-}03$, $\eta^2 = 0.0053$] and in addition a significant interaction effect between age (Baby vs. Adult) and ethnic group (Caucasian vs. Chinese) of the face was found [left pupils: $F_{(1,60)} = 35.62$, $p = 1.385729\text{e-}07$, $\eta^2 = 0.0421$; right pupils : $F_{(1,60)} = 32.53$, $p = 3.815601\text{e-}07$, $\eta^2 = 0.0378$] (**Figure 2**).

Post-hoc analysis revealed that for both pupils of women and men, there was a significant increase in pupil width in response to Chinese babies faces compared to Caucasian babies faces [left pupils (females): $t_{(30)} = -6.7712$, $p = 1.652\text{e-}07$, Cohen's d = −1.0629; right pupils (females): $t_{(30)} = -6.4622$, $p = 3.861\text{e-}07$, Cohen's d = −0.9946; left pupils (males): $t_{(30)} = -4.0984$, $p = 0.0003$; Cohen's d = −0.4358; right pupils (males): $t_{(30)} = -3.4548$, $p = 0.001665$, Cohen's d = −0.3799]. In females and males, both pupil widths were significantly wider in response to Chinese baby faces compared to Chinese Adult faces [left pupils (females): $t_{(30)} = 4.5692$, $p = 7.839\text{e-}05$, Cohen's d = 0.6774; right pupils (females): $t_{(30)} = 4.9375$, $p = 2.779\text{e-}05$, Cohen's d = 0.7047; left pupils (males): $t_{(30)} = 3.3656$, $p = 0.0021$, Cohen's d = 0.3439; right pupils (males): $t_{(30)} = 3.2478$, $p = 0.0029$, Cohen's d = 0.3290]. Only the pupil widths of female participants were significantly reduced in response to Caucasian baby faces compared to Caucasian adult faces [left pupils (females): $t_{(30)} = -2.5445$, $p = 0.016$, Cohen's d = −0.3868; right pupils (females): $t_{(30)} = -2.581$, $p = 0.0149$, Cohen's d = −0.3555]. Pupil widths of female participants significantly increased in response to Chinese baby faces compared to Caucasian Adult faces [left pupils (females): $t_{(30)} = 4.7998$, $p = 4.098\text{e-}05$, Cohen's d = 0.6177; right pupils (females): $t_{(30)} = 4.6619$, $p = 6.042\text{e-}05$, Cohen's d = 0.5829]. However, only females right pupil significantly decreased in response to Caucasian baby faces compared to Chinese Adult faces [right pupils (females): $t_{(30)} = -2.8755$, $p = 0.00735$, Cohen's d = −0.2618].

2.3.2. Attitudinal Results

To test hypothesis 4, that females showed no difference in preference for in-group and out-group infant faces, we undertook

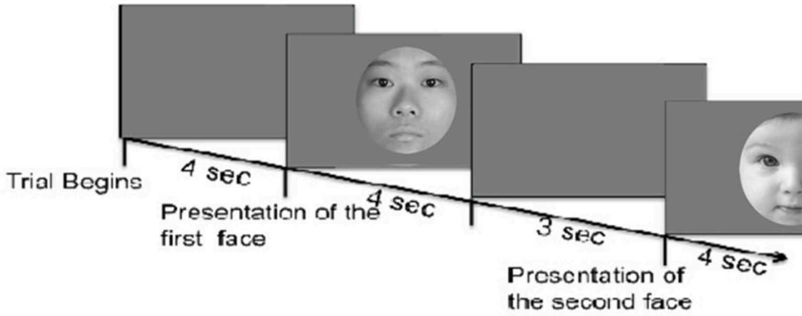


FIGURE 1 | The sketch above illustrates the model representation of the stimuli. Faces were presented in a randomized order across participants. Each face was presented for 4 s followed by an inter-stimulus interval during which a gray screen was shown for 3 s. Face stimuli were selected from Peer (2005) and Yap et al. (2016).

a 2 × 2 × 2 Analysis of Variance (ANOVA) with age and ethnicity of face as within-subjects factors, sex of participants as a between-subjects factor and attitude score as the dependent variable; this analysis partially supported our hypothesis. A significant main effect was found for sex [$F_{(1,60)} = 8.29$, $p = 5.505286\text{e-}03$, $\eta^2 = 4.609139\text{e-}02$], for ethnic group [$F_{(1,60)} = 126.93$, $p = 1.929138\text{e-}16$, $\eta^2 = 2.883455\text{e-}01$] and for age [$F_{(1,60)} = 7.84$, $p = 6.855548\text{e-}03$, $\eta^2 = 3.850743\text{e-}02$]. Moreover, a significant interaction between sex of participants and ethnic group was found on attitude scores [$F_{(1,60)} = 11.06$, $p = 1.509542\text{e-}03$, $\eta^2 = 3.409557\text{e-}02$] (**Figures 3A,B**). *Post-hoc* analysis revealed a significant difference in attitude scores in response to adult faces only [females: $t_{(30)} = 6.06$, $p = 1.1537\text{e-}06.001$, Cohen's d = 1.1277; males: $t_{(30)} = 6.29$, $p = 6.1750\text{e-}07$, Cohen's d = 1.5075], where adult Caucasian faces were rated more positively compared to adult Chinese faces. No significant difference was found between the attitude scores given to Caucasian and Chinese babies faces.

3. STUDY 2

3.1. Materials and Methods

3.1.1. Participants

A total of 40 Caucasian Italian adults (20 males, 20 females, Mean Age = 23.175, SD = 2.84) participated. Exclusion criteria were parenthood, pregnancy, and non-Caucasian ethnicity. Informed consent was obtained from all participants prior to the study, and no incentives were given to the participants. This study was conducted in accordance with the ethical principles stated in the Helsinki declaration. Two participants (1 male and 1 female) were omitted from the final analysis because the collected data was corrupted due to technical malfunctions in EEG recordings.

3.1.2. Stimuli

The same set of faces used in Study 1 was shown to participants (refer to section 2.1.2). Similarly, faces were presented for 4 s each in a random order, with a 3-seconds inter-stimulus interval (ISI) between any two faces, as shown in **Figure 1**. The stimuli were presented using TkInter[2] on a 29 monitor (DELL 29 Ultrasharp U2719 WM Monitor. Resolution was set to 1,920 × 1,080, refresh rate: 60.00 Hz). Participants were asked to sit approximately 50–70 cm away from the screen. At the end of the presentation, participants completed the same questionnaire employed in Study 1 regarding their disposition toward each face.

3.1.3. Data Acquisition

EEG data were collected using a 14 channel setup (Emotiv EPOC) and digitized at 128 Hz. Previous research has demonstrated the reliability of this device for the recording of EEG signals (Debener et al., 2012; Duvinage et al., 2012; Badcock et al., 2013; De Vos et al., 2014; Ries et al., 2014). E8 of the 14 channels were employed in Study 2: four channels for posterior (P7, P8, 01, 02) and four for anterior areas (F3, F4, F7, F8). Data was recorded on an external device (Lenovo ThinkPad Intel Core i5-4210U).

3.1.4. Preprocessing

Preprocessing of EEG data was implemented using the MNE python package (Gramfort et al., 2013). To reduce the impact of external sources of noise and artifacts, a band-pass filter between 0.51 and 45 Hz was applied to the EEG data. Filtered signals were decomposed into 5 frequency bands—theta (4, 8 Hz), alpha (8, 12 Hz), beta (13, 26 Hz), gamma (26, 45 Hz), and delta (0.51, 4 Hz)—by means of Fast Fourier Transformation. Signals were then normalized, using standard scores, to take into account between- and within-subject differences. Finally, for each stimulus presentation (4 s long), the mean activity and its standard deviation was computed and stored for further analysis. Similar to Study 1, participants answers to the three questionnaires were highly correlated (females' Attitude and Closeness: Pearson's $r = 0.852$, $p < 2.2\text{e-}16$; Attitude and Pleasantness: Pearson's $r = 0.870$, $p < 2.2\text{e-}16$; Closeness and Pleasantness:Pearson's $r = 0.947$, $p < 2.2\text{e-}16$ and males' Attitude and Closeness: Pearson's $r = 0.931$, $p < 2.2\text{e-}16$; Attitude and Pleasantness: Pearson's $r = 0.946$, $p < 2.2\text{e-}16$; Closeness and Pleasantness: Pearson's $r = 0.960$, $p < 2.2\text{e-}16$). Therefore, to take into account inter-individual differences, participants attitude toward faces on frequency bands amplitude variations was taken

---

[2] https://wiki.python.org/moin/TkInter

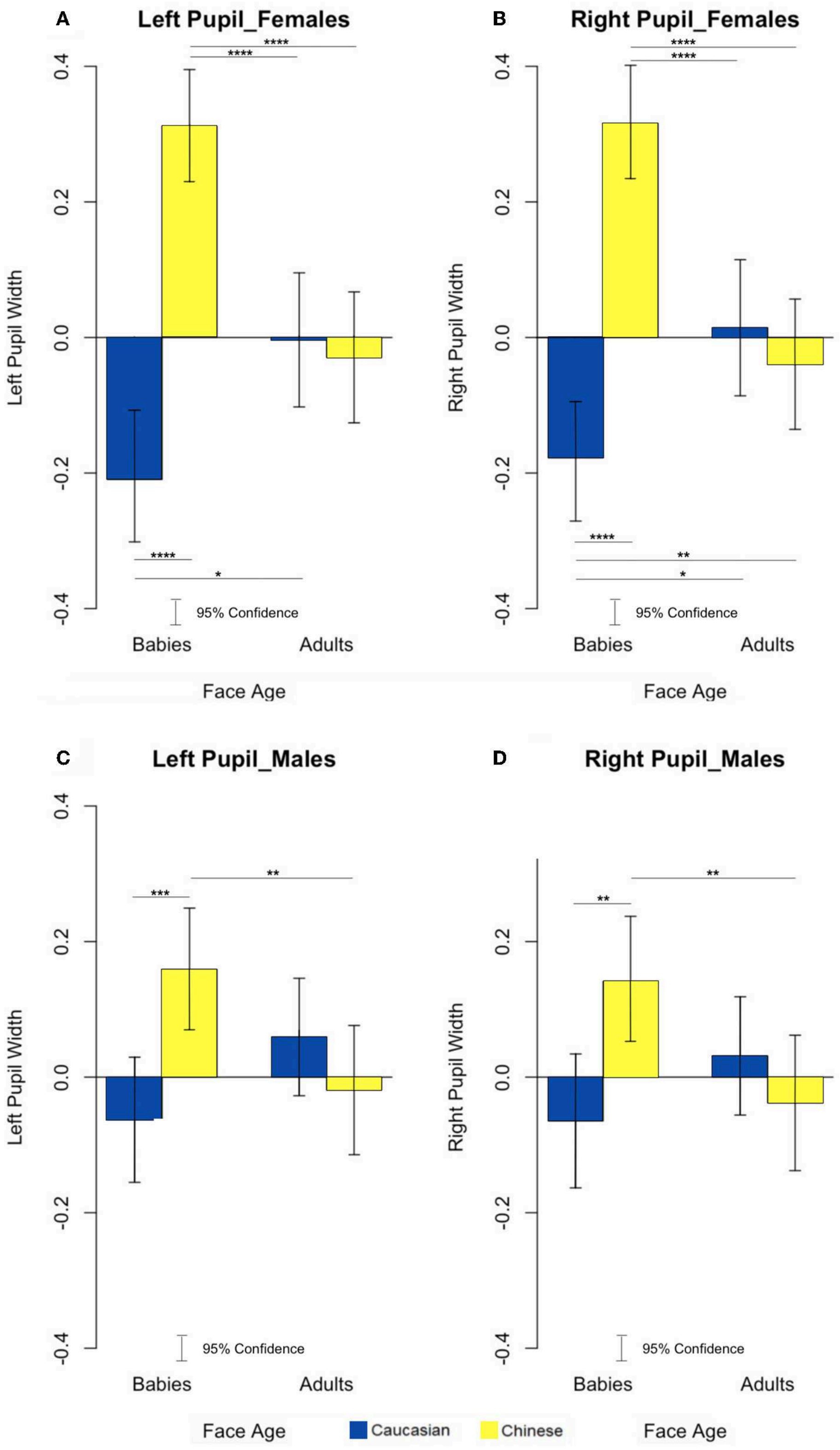


FIGURE 2 | Effects of the interaction between faces age and ethnic group on left and right pupil width in females **(A,B)** and females **(C,D)**. **p* < 0.05, ***p* < 0.01, ****p* < 0.001, *****P* < 0.0001.

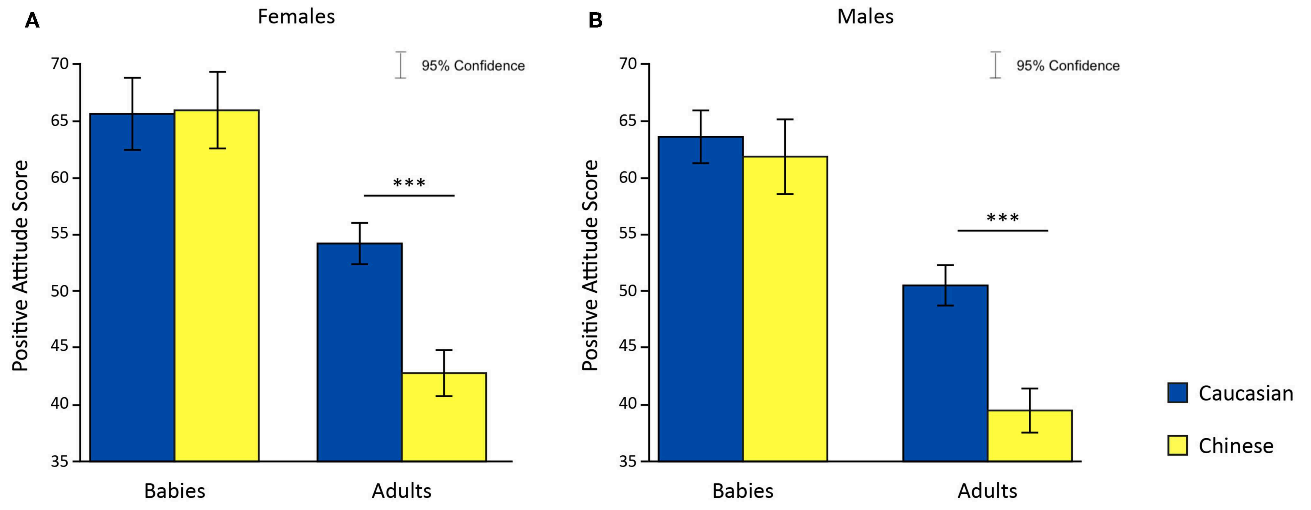


FIGURE 3 | Effects of faces' age and ethnic group on attitude scores in females **(A)** and males **(B)**. $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

out of the linear model between frequencies amplitudes and attitude scores.

3.1.5. Analytical Plan

3.1.5.1. Preliminary analyses

A repeated-measure mixed design Analysis of Variance (ANOVA) was computed on the 5 frequency bands of selected channels. Amplitude residuals were used as dependent variables. Residual scores controlled for homogeneity (Levene's Test of Equality of Error Variances: Residuals Baby Caucasian $p = 0.096693$, Residuals Baby Chinese $p = 0.521395$, Residuals Adult Caucasian $p = 0.904724$, Residuals Adult Chinese $p = 0.128961$) before residuals were used as the dependent variable in the regression model.

3.1.5.2. Inferential analyses

The first analysis of variance (ANOVA) model tested the effects of gender in moderating BS (hypothesis 1) and MRE (hypothesis 2). Two within-subjects factors (face age: baby/adults; ethnic group: Caucasian/ Chinese) and one between-subject factor (participant gender) were used as independent variables. *Post-hoc* analyses were performed using Student's t-tests. Level of significance was corrected using false discovery rate correction. The second ANOVA model was employed to test whether neural responses to in-group baby faces were related to out-group adult faces (hypothesis 3). We created a new variable to represent the in-group Baby Schema Effect (INBSE) by subtracting the corrected mean amplitudes for the stimuli "adult Caucasian" from the corrected mean amplitudes for the stimuli "baby Caucasian." A repeated-measures Analysis of Covariance (ANCOVA) was conducted in which this new variable was fitted as a covariate. Face age and ethnic group were used as within-subject factors, and participant gender was used as between-subjects factor. Finally, a repeated-measure mixed design ANOVA was conducted to investigate the impact of face age, ethnic group, and gender on participants attitudes (hypothesis 4). Attitude scores were dependent variables, while the independent variables were kept as the first ANOVA model.

3.2. Results

3.2.1. Physiological Results

To test our first hypothesis, that females exhibit the Baby Schema (BS) effect to both in-group and out-group infant faces, we employed a 2 × 2 × 2 ANOVA where age (Baby vs. Adult) and ethnic group (Caucasian vs. Chinese) of the faces presented were within-subject factors and participant's gender was a between-subject factor; this analysis did not support the hypothesis. An interaction effect between age (Baby vs. Adults), ethnic group (Caucasian vs. Chinese) and participant's gender was found only in P7 channel for alpha wave [$F_{(1,36)} = 8.66$, $p = 0.0451040$ after Multiple Comparison Correction, $\eta^2 = 7.280281\text{e-}02$]. Moreover, *post-hoc* analysis revealed that males showed a significant increase of P7 alpha's amplitude for Caucasian adult faces and a significant decrease in amplitude for Chinese adult faces ($p = 0.004886$, $\eta^2 = 0.7357$) (**Figures 4A,B**).

To test our third hypothesis, that in-group Baby Schema Effect (INBSE) was related to different activation patterns for out-group stimuli, we undertook a repeated-measures ANCOVA analysis, with INBSE as a covariate, while the age and ethnicity of the face presented were within-subject actors; this analysis supported our hypothesis. A significant three-way interaction between the age of face, the ethnic group of face, and INBSE ($F_{(1,36)} = 31{,}49$, corrected $p = 0.000003$, $\eta^2 = 0.3050$] emerged. We also found a significant two-way interaction effect between age of face and INBSE [$F_{(1,36)} = 28.88$, $p = 0.000005$, after Bonferroni Correction]. *Post-hoc* analysis revealed that participants with higher INBSE showed greater activation toward out-group Adults compared to in-group Adults and that participants with lower INBSE showed greater activation toward in-group Adults compared to out-group Adults ($p = 0.00015$, $\eta^2 =$

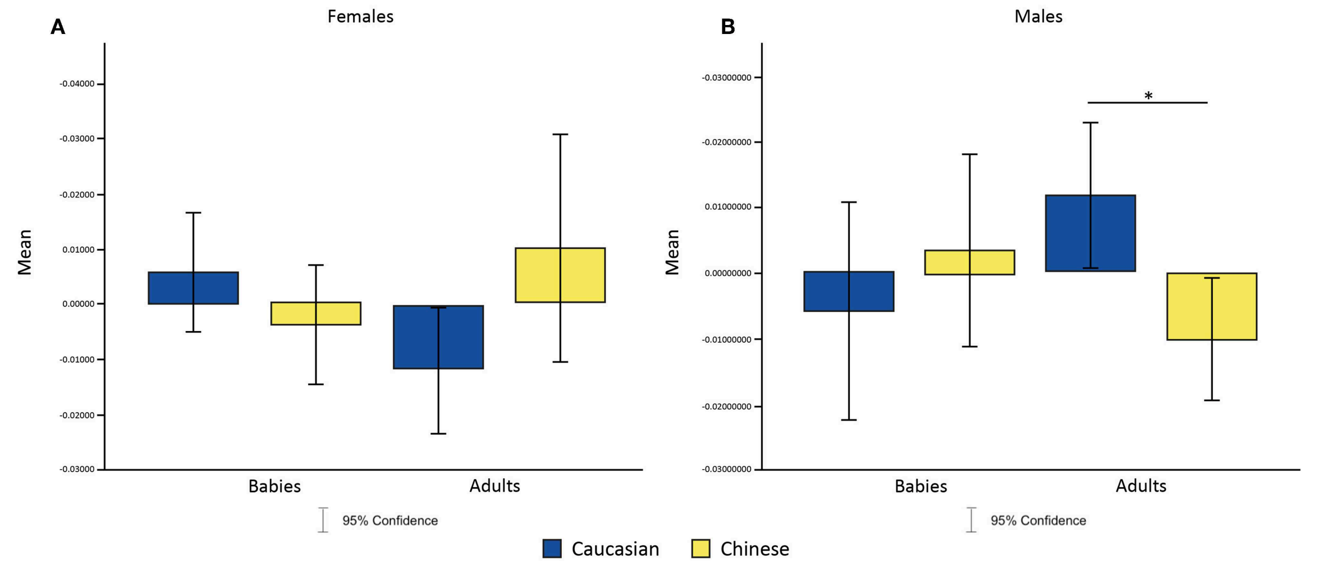


FIGURE 4 | Effects of the interaction between faces' age and ethnic group on P7 alpha amplitude in females **(A)** and males **(B)**. $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

-0.3062; $p = 0.000427$, $\eta^2 = 0.2377$]. Finally, we found a significant effect of ethnic group of baby faces for lower INBSE ($p = 0.00015$, $\eta^2 = -0.4292$). No main effects were found (**Figure 5**).

3.2.2. Attitudinal Results

To test our fourth hypothesis, that preference for in-group and out-group infant faces do not differ for females, we conducted a 2 × 2 × 2 ANOVA with age and ethnicity of face as within-subjects factors, sex of participants as a between-subjects factor and attitude scores as the dependent variable; this analysis partially supported our hypothesis. From the analysis, we found main effects of age and ethnic group of a face [age: $F_{(1,36)} = 43.32$, $p = 1.1683\text{E-}7$, $\eta^2 = 0.1804$; ethnic Group: $F_{(1,36)} = 26.20$, $p = 1.044176\text{e-}05, \eta^2 = 0.0705$]. We also found a significant interaction between age and ethnic group of the face on attitude scores for both males and females [$F_{(1,36)} = 46.03$, $p = 6.3027\text{E-}8$, $\eta^2 = 0.0527$]. Moreover, we found a significant age of face and sex of participant interaction, meaning that males and females treated baby and adult faces differently [$F_{(1,36)} = 7.81$, $p = 0.008$, $\eta^2 = 0.0382$]. *Post-hoc* analysis revealed that Caucasian baby faces evoked greater positive attitudes compared to Chinese adult faces in both genders [females: $t_{(18)} = 7.112$, $p = 0.000001$, Cohen's d $= 1.753$; males: $t_{(18)} = 5.386413$, $p = 0.00004$, Cohen's d $= 1.390$]. Only in females was there a greater preference for Caucasian baby faces compared to Caucasian adult faces [females: $t_{(18)} = 3.604883$, $p = 0.002$, Cohen's d $= 0.8181$]. Both genders showed more positive attitudes toward Caucasian adult faces compared to Chinese adult faces [females: $t_{(18)} = 6.809751$, $p = 0.000002$, Cohen's d $= 1.109$; males: $t_{(18)} = 7.523347$, $p = 5.812\text{E-}7$, Cohen's d $= 1.733$], and Chinese baby faces compared to Chinese adult faces [Females: $t_{(18)} = 6.610146$, $p = 0.000003$, Cohen's d $= 1.323$; Males: $t_{(18)} = 4.834543$, $p = 0.00013$, Cohen's d $= 1.109$]. No significant difference in attitude scores was found between baby faces stimuli (**Figures 6A,B**).

4. DISCUSSION

This study investigates the Baby Schema (BS) and Mere Exposure (MRE) effects on adults implicit and explicit responses to infant and adult in-group and out-group faces. Baby Schema is a collection of infant facial features which has been selected by evolution to inspire caregiving behaviors in adults. MRE dictates that familiarity leads to preference, which begets the question of whether the protective effect of BS applies to both in-group (familiar) and out-group (unfamiliar) infants. We investigated Italian adults implicit physiological responses using pupillometry and electroencephalography while they were presented with in-group (Caucasian) and out-group (Chinese) infant and adult faces. Questionnaire data were also collected to obtain measures of participants explicit self-reported preferences of faces.

Results from Study 1 revealed several key findings. First, females exhibited greater pupillary dilation to out-group infants than out-group adults (Chinese baby >Chinese adult) and greater pupillary constriction to in-group infants than in-group adults (Caucasian baby >Caucasian adult) in both pupils. Two cross-ethnic group comparisons (i.e., Chinese baby >Caucasian adult; Caucasian baby >Chinese adult) were observed in the right pupil while only one was observed in the left pupil (i.e., Chinese baby >Caucasian adult). Males showed a BS effect in both pupils specific to out-group Chinese infants (i.e., Chinese baby >Chinese adult), thus partially supporting hypothesis 1.

From Study 2, we found greater left posterior-parietal alpha activation toward familiar in-group (i.e., Caucasian) adult faces compared to unfamiliar out-group (i.e., Chinese) adult faces in males. Since alpha activation is inversely related to neural

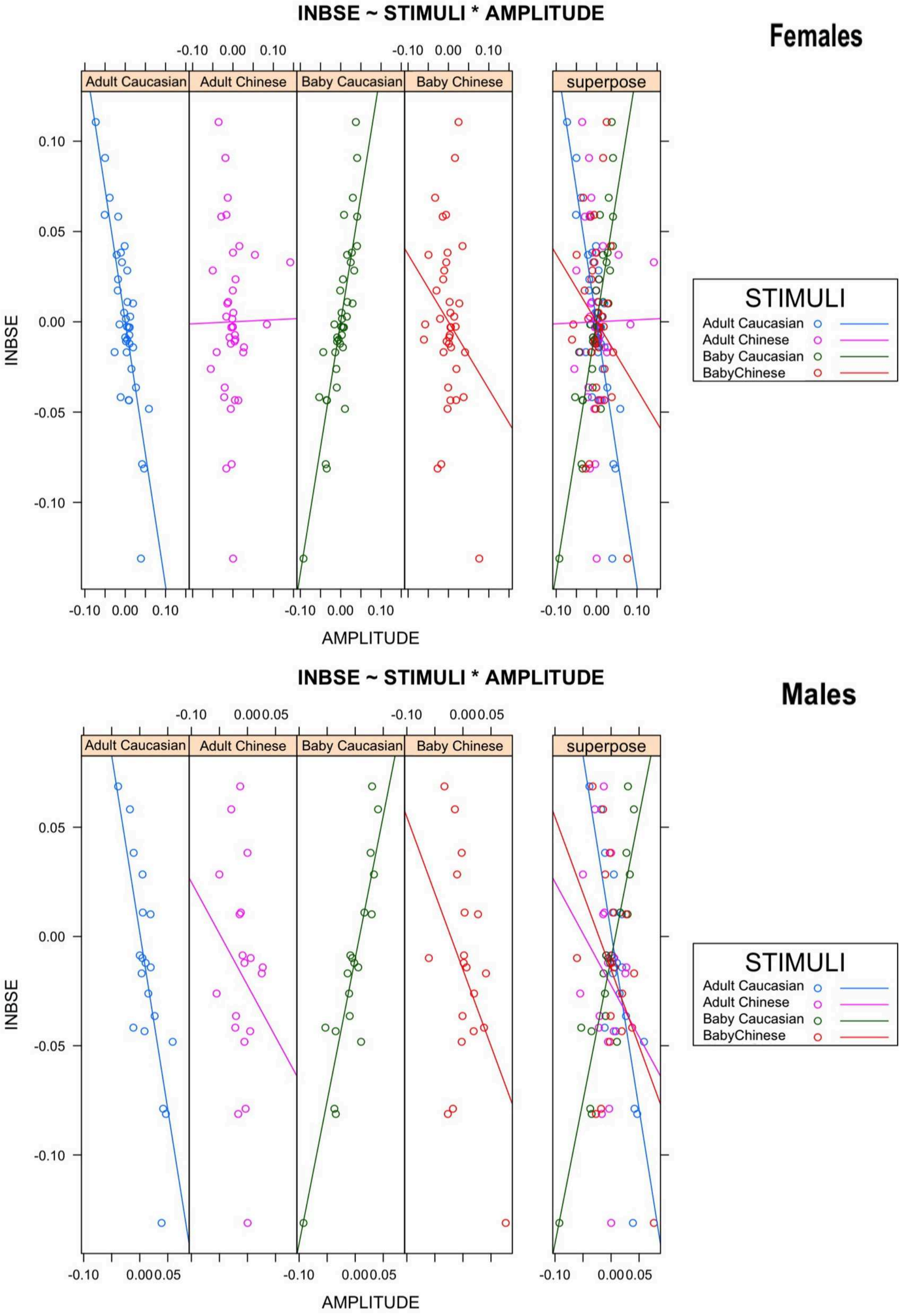


FIGURE 5 | Scatter plots representing the relation between in-group Baby Schema Effect (INBSE) and amplitude of response toward face stimuli in males and females.

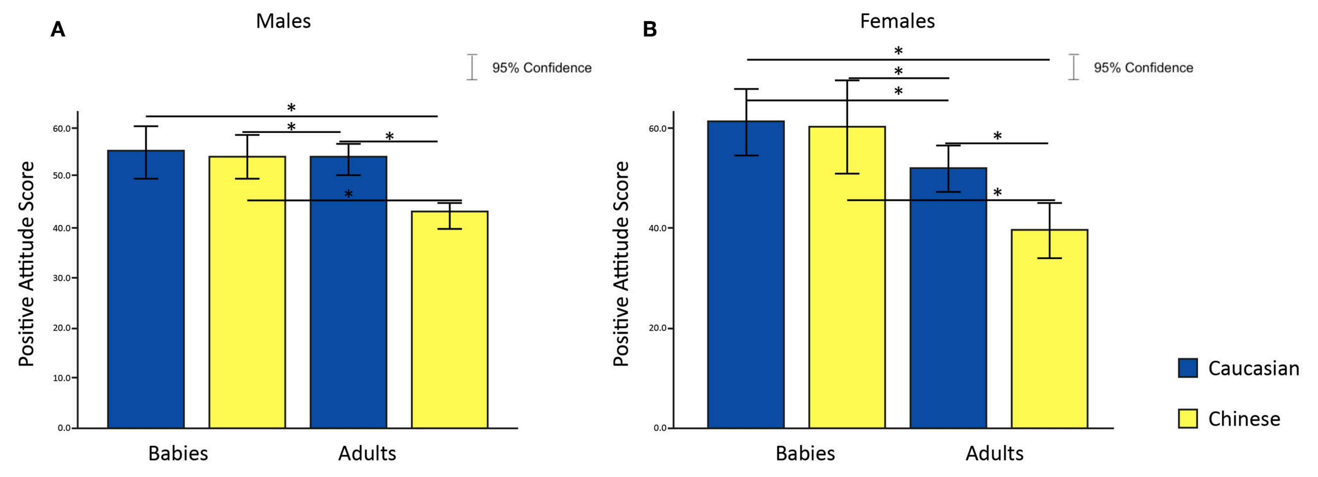


FIGURE 6 | Effects of faces' age and ethnic group on attitude scores in females **(A)** and males **(B)**. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

activity, this shows that males exhibit more activation and thus attention toward out-group faces, which suggests the presence of MRE and supports hypothesis 2. We conducted another analysis by stratifying participants based on their in-group Baby Schema effect (INBSE) score, from which we found that participants lower in INBSE showed significantly different activation patterns between Chinese and Caucasian baby faces, whereas participants higher in INBSE did not. Instead, those with a higher INBSE score exhibited distinct activation between Chinese and Caucasian adult faces, which indicates that those with higher INBSE differentiated between in-group and out-group adults at the early stage of face processing. These findings suggest that level of INBSE moderates MRE for adult and baby faces, which supports hypothesis 3.

Questionnaire responses from participants, which was an explicit measure of self-reported preference of faces, were consistent in both Study 1 and 2. Males and females reported greater preference for in-group adult faces, which suggests an MRE specific to adult faces. Participants in Study 2 showed greater preference for baby over adult faces which reflects an overt BS effect. However, gender was not found to moderate preferences for baby faces as expected in hypothesis 4.

4.1. Gender Moderates Baby Schema and Mere Exposure Effects

Overall, we found in females a greater pupillary response toward children than adults. However, looking specifically at the direction of the variations, we observed that this response goes in two different directions: a pupil dilation toward out-group infants and a pupil constriction toward in-group ones. Regarding pupil dilation, past studies (Glocker et al., 2009a,b), where variables such as in-group and out-group were not taken into account and where the physiological measure was for example the neural response, highlighted that women generally showed greater responses toward infants. Therefore, it is possible to assume that the dilation that we observed in our study, which in general is a synonym for more attention given to stimuli (Aktar et al., 2016), means more attention given to infants than to adults. Hess and Polt (1960) similarly observed larger pupil dilation in women looking at pictures of infants than of other people, lending further support to our findings. Instead, the observed pupil constriction for in-group babies is more challenging to explain, considering the non univocal origin of pupil constriction (Loewenfeld, 1966; Hess, 1972; Mathôt et al., 2013; Mathôt and Van der Stigchel, 2015; Winn et al., 2018; Zekveld et al., 2018). However, in line with this result, Landi and Freiwald (2017) found that a greater pupillary constriction occurs in response to familiar faces compared to unfamiliar ones, suggesting that familiar faces request fewer cognitive resources in order to be processed. Therefore, the picture that emerges is that both BS and MRE are active in females while they process faces of different ages and ethnicity. This co-existence leads them to show more attention to infant faces than adult faces but with opposite directions for Caucasian and Chinese faces due to the familiarity of the stimuli.

By comparison, the BS effect elicited in males was specific to out-group faces only, which is in line with our second hypothesis, that an interaction effect between BS and MRE is likely to be observed in males. Pupil dilation indicates greater preferential attention and has been thought to occur during an increase in processing effort (Kahneman, 1973; Beatty, 1982). As compared to familiar in-group faces, unfamiliar out-group faces are more novel and evoke greater attention. The processing of unfamiliar out-group faces might require the recruitment of more cognitive resources (Goldinger et al., 2009; Hu et al., 2014). Therefore, the BS effect which entices adults to attend to infant faces might have been moderated by the MRE effect which induces enhanced attention to out-group faces, leading to significantly greater pupillary dilation in males toward baby than adult faces for out-groups.

While gender difference in regard to the Baby Schema effect is well-established (Hess and Polt, 1960; Glocker et al., 2009a), we found that gender also moderates MRE. Males in Study 2 exhibited amplitude reduction of the alpha band, corresponding to neural activation, in the left posterior-parietal region toward unfamiliar out-group (Chinese) adult faces when contrasted against familiar in-group (Caucasian) adult faces. From an evolutionary point of view, better processing of new and potentially threatening stimuli in the environment (out-group faces) is pivotal and may guarantee the individual's survival. Amplitude decrease of alpha indicates a state of relatively high excitability and engagement of the brain area, while an amplitude increase reflects inhibition (Hart et al., 2000; Ito and Urland, 2005; He et al., 2009). Posterior parietal areas are involved in attentional processing (Buschman and Miller, 2007) and activity in the alpha system may predict attention toward threatening social stimuli (Yamakawa et al., 2009; Giardina et al., 2012; Grimshaw et al., 2014). We might deduce therefore that males attention is more effectively captured by out-group adults, possibly due to evolutionary-enhanced processing which occurs in response to unfamiliar stimuli that may pose a threat. Behavioral reports from Study 1 and 2 showed greater preference toward in-group adult faces, suggesting that MRE is present at an explicit level in both males and females (Allport et al., 1954; LeVine and Campbell, 1972; Brewer, 1979; Esposito et al., 2014).

Attitudes toward infant and adult faces varied across the two studies. Participants in Study 1 showed no difference in attitude toward infant and adult faces, but both male and female participants in Study 2 reported more positive attitudes to infant faces compared to adult faces. These reports show the existence of conscious perception of the BS effect (Glocker et al., 2009a,b; Parsons et al., 2011). However, we did not find any difference in reports of preference for infant compared to adult faces between male and female participants as we initially hypothesized.

4.2. Baby Schema Moderates Mere Exposure Effect

Individuals who experience the BS effect have been found to possess higher levels of empathy, interpersonal closeness, and need to belong (Lehmann et al., 2013). These personality traits might also influence implicit perceptions toward in-group and out-group faces. To investigate this relation, we stratified our participants according to their Baby Schema score. We found that participants with lower BS scores differed in their processing of in-group and out-group infant faces which fulfilled our third hypothesis. A possible interpretation is that participants with lower BS scores are less susceptible to the BS effect such that they discriminate between in-group and out-group infants. Participants with higher BS scores are generally more prone to find infant faces appealing and are less likely to process in-group and out-group infant faces differently. Similar neural processing of infant faces of different ethnicity also suggests some protective effect of BS when perceiving unfamiliar out-group infants (Proverbio and De Gabriele, 2017).

Higher BS scores were associated with reduced alpha amplitudes (increase in neural activation) in response to out-group adult than in-group adult faces. This suggests that a strong BS effect, in which no difference is detected in the processing of in-group and out-group infants is potentially indicative of substantial MRE effects among adults (Proverbio et al., 2011; Proverbio and De Gabriele, 2017). It is reasonable to assume that participants who are more interested in caregiving activities are also more sensitive to potentially threatening stimuli to infants, such as unfamiliar out-group adults (Lorenz, 1943; Lorenz and Martin, 1971; Glocker et al., 2009a,b).

4.3. Limitations

Here we report some limitations and propose possible future directions of research. First, the main limitation of the present work concerns the absence in the available literature of a unequivocal interpretation of the pupil constriction toward in-group baby faces observed in the pupillometry study. However, this result is in line with (Landi and Freiwald, 2017) who found that familiar faces elicit greater pupillary constrictions compared to unfamiliar ones. The simultaneous presence of MRE and the BS might explain why in females more attention is given to infants than adults but with opposite direction depending on the familiarity (Caucasian vs. Chinese). To confirm this interpretation, however, a future study should try to test the same paradigm in a different population where Chinese faces are more familiar than Caucasian faces. Second, this study utilized neutral faces only, and future studies should assess the effect of emotional valence in facial expressions (i.e., smiling, crying). Third, consistent with other studies (e.g., Esposito et al., 2015), only two ethnic groups were examined. Employing a range of different ethnic groups (e.g., African, Latin Americans, et al.) could better expose relations between social distance from out-groups and face processing. Fourth, only adult female faces were used; it has been well-established that males and females react differently to faces of adult women. This gender bias could be overcome by including adult male faces in the stimuli. Fifth, future studies may combine frequency bands analysis (neural oscillations) of face processing with other existing measures of EEG, like Event Related Potential (ERP). ERP analysis was not conducted in this study as it would have required multiple repetitions of each stimulus that would not have enabled us to be consistent with the pupillometry study. Sixth, despite the fact that we analyzed data collected from a generally individualistic society, it may be also interesting to replicate this study in an more collectivist society to see how culture shapes faces perception. Finally, future investigators should also consider individually held attitudes and experiences with infants as well as out-group members.

5. CONCLUSION

The human face contains tremendous social information and plays a fundamental role in communication with other people. Our study has revealed the rich interplay of Baby Schema (BS) and Mere Exposure (MRE) effects that emerge on merely viewing adult and infant faces. We showed that gender greatly moderates the processing of faces. Given the simultaneous presence of both MRE and BS, females pay more attention to infant than adult faces but the responses are opposite for Caucasian and Chinese faces (constriction vs. dilation) due to the familiarity of the stimuli. By comparison, males attend to out-group infants significantly more than out-group adults, but require more cognitive recruitment to process out-group than in-group adult faces. While gender differences were observed at the implicit processing level, males and females did not differ in their self-reported attitudes, as both reported greater preference for infant over adult faces. Contradictory results that emerge at the implicit (physiological) and explicit (self-report) levels indicate that differential processing of faces varying in age and ethnicity may occur at a non-conscious level. Besides gender, level of Baby Schema effect may reflect empathy and predict differential processing of adult and infant faces belonging to an individual's in-group or out-group. These findings evoke excitement regarding the role of gender and personality in moderating face processing. Ingrained in us is a complex biological attentional mechanism that extracts critical information from the humble face and elicits automatic responses that have been shaped by evolution and society.

DATA AVAILABILITY STATEMENT

The dataset Esposito et al. (2019) generated for this study can be found in the NTU's Data repository DR-NTU Data at the following address: https://doi.org/10.21979/N9/TGTTTR.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the IRB of the Nanyang Technological University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

LV, MB, and GE conceptualized the study. LV collected the data. LV, GG, and AT analyzed the data. LV and AA wrote the original draft. All the authors reviewed and edited the submitted version of the article.

FUNDING

This study was supported by the NAP-SUG program of the Nanyang Technological University (GE), the Singapore Ministry of Education Academic Research Grant - Tier 1 (GE), the Intramural Research Program of the NIH/NICHD, USA, and an International Research Fellowship at the Institute for Fiscal Studies (IFS), London, UK, funded by the European Research Council (ERC) under the Horizon 2020 research and innovation programme (grant agreement No 1334 695300-HKADeC-ERC-2015-AdG).

ACKNOWLEDGMENTS

All participants in this study are gratefully acknowledged.

REFERENCES

Aktar, E., Mandell, D. J., de Vente, W., Majdandžić, M., Raijmakers, M. E. J., and Bögels S. M. (2016). Infants' temperament and mothers', and fathers' depression predict infants' attention to objects paired with emotional faces. *J. Abnorm. Child Psychol.* 44, 975–990. doi: 10.1007/s10802-015-0085-9

Allport, G. W., Clark, K., and Pettigrew, T. (1954). *The Nature of Prejudice*. Reading, MA: Addison-Wesley.

Andrew, R. J., Mench, J., and Rainey, C. (1982). "Right-left asymmetry of response to visual stimuli in the domestic chick," in *Analysis of Visual Behaviour*, eds D. J. Ingle, M. A. Goodale, and R. J. W. Mansfield (Cambridge, MA: MIT Press), 197–209.

Badcock, N. A., Mousikou, P., Mahajan, Y., de Lissa, P., Thie, J., and McArthur, G. (2013). Validation of the emotiv EPOC(®) EEG gaming system for measuring research quality auditory ERPs. *PeerJ* 1:e38. doi: 10.7717/peerj.38

Balconi, M., and Mazza, G. (2009). Brain oscillations and BIS/BAS (behavioral inhibition/activation system) effects on processing masked emotional cues. ERS/ERD and coherence measures of alpha band. *Int. J. Psychophysiol.* 74, 158–165. doi: 10.1016/j.ijpsycho.2009.08.006

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan N., and Wheelwright S. (2003). The systemizing quotient: an investigation of adults with asperger syndrome or high–functioning autism, and normal sex differences. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 361–374. doi: 10.1098/rstb.2002.1206

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychol. Bull.* 91, 276–292. doi: 10.1037/0033-2909.91.2.276

Bernard, K., Simons, R., and Dozier, M. (2015). Effects of an attachment-based intervention on child protective services–referred mothers' event-related potentials to children's emotions. *Child Dev.* 86, 1673–1684. doi: 10.1111/cdev.12418

Bornstein, M. H. (2015). "Children's parents," in *Handbook of Child Psychology and Developmental Science*, ed R. M. Lerner (Hoboken, NJ: John Wiley & Sons), 55–132. doi: 10.1002/9781118963418.childpsy403

Bornstein, M. H., and Tamis-LeMonda, C. S. (1989). Maternal responsiveness and cognitive development in children. *New Dir. Child Adolesc. Dev.* 1989, 49–61. doi: 10.1002/cd.23219894306

Bowlby, J. (1969). *Attachment and Loss v. 3*. Vol. 1. Random House.

Bradley, M. M., Miccoli, L., Escrig, M. A., and Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology* 45, 602–607. doi: 10.1111/j.1469-8986.2008.00654.x

Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: a cognitive-motivational analysis. *Psychol. Bull.* 86:307. doi: 10.1037//0033-2909.86.2.307

Brosch, T., Sander, D., and Scherer, K. R. (2007). That baby caught my eye... attention capture by infant faces. *Emotion* 7, 685–689. doi: 10.1037/1528-3542.7.3.685

Buschman, T. J., and Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315, 1860–1862. doi: 10.1126/science.1138071

Caria, A., Falco, S. D., Venuti, P., Lee, S., Esposito, G., Rigo, P., et al. (2012). Species-specific response to human infant faces in the premotor cortex. *Neuroimage* 60, 884–893. doi: 10.1016/j.neuroimage.2011.12.068

De Vos, M., Kroesen, M., Emkes, R., and Debener, S. (2014). P300 speller BCI with a mobile EEG system: comparison to a traditional amplifier. *J. Neural Eng.* 11:036008. doi: 10.1088/1741-2560/11/3/036008

Debener, S., Minow, F., Emkes, R., Gandras, K., and de Vos, M. (2012). How about taking a low-cost, small, and wireless EEG for a walk?: EEG to go. *Psychophysiology* 49, 1617–1621. doi: 10.1111/j.1469-8986.2012.01471.x

Draper, P., and Harpending, H. (1982). Father absence and reproductive strategy: an evolutionary perspective. *J. Anthropol. Res.* 38, 255–273. doi: 10.1086/jar.38.3.3629848

Duvinage, M., Castermans, T., Dutoit, T., Petieau, M., Hoellinger, T., Saedeleer, C. D., et al. (2012). A P300-based quantitative comparison between the emotiv epoc headset and a medical EEG device. *Biomed. Eng.* 765, 2012–2764. doi: 10.2316/P.2012.764-071

Endendijk, J. J., Spencer, H., van Baar, A. L., and Bos, P. A. (2018). Mothers neural responses to infant faces are associated with activation of the maternal care system and observed intrusiveness with their own child. *Cogn. Affect. Behav. Neurosci.* 18, 609–621. doi: 10.3758/s13415-018-0592-6

Esposito, G., Gabrieli, G., and Venturoso, L. (2019). *Effects of Baby Schema and Mere Exposure on Explicit and Implicit Face Processing.* DR-NTU (Data), V1. doi: 10.21979/N9/TGTTTR

Esposito, G., Nakazawa, J., Ogawa, S., Stival, R., Kawashima, A., Putnick, D. L., et al. (2014). Baby, you light-up my face: culture-general physiological responses to infants and culture-specific cognitive judgements of adults. *PLoS ONE* 9:e106705. doi: 10.1371/journal.pone.0106705

Esposito, G., Nakazawa, J., Ogawa, S., Stival, R., Putnick, D. L., and Bornstein, M. H. (2015). Using infrared thermography to assess emotional responses to infants. *Early Child Dev. Care* 185, 438–447. doi: 10.1080/03004430.2014.932153

Furman, W., and Buhrmester, D. (2009). Methods and measures: the network of relationships inventory: behavioral systems version. *Int. J. Behav. Dev.* 33, 470–478.

Giardina, A., Caltagirone, C., Cipolotti, L., and Oliveri, M. (2012). The role of right and left posterior parietal cortex in the modulation of spatial attentional biases by self and non-self face stimuli. *Soc. Neurosci.* 7, 359–368. doi: 10.1080/17470919.2011.628410

Glocker, M. L., Langleben, D. D., Ruparel, K., Loughead, J. W., Gur, R. C., and Sachser, N. (2009a). Baby schema in infant faces induces cuteness perception and motivation for caretaking in adults. *Ethology* 115, 257–263. doi: 10.1111/j.1439-0310.2008.01603.x

Glocker, M. L., Langleben, D. D., Ruparel, K., Loughead, J. W., Valdez, J. N., Griffin, M. D., et al. (2009b). Baby schema modulates the brain reward system in nulliparous women. *Proc. Natl. Acad. Sci. U.S.A.* 106, 9115–9119. doi: 10.1073/pnas.0811620106

Goldinger, S. D., He, Y., and Papesh, M. H. (2009). Deficits in cross-race face learning: insights from eye movements and pupillometry. *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 1105–1122. doi: 10.1037/a0016548

Gramfort, A., Luessi, M., Larson, E., Engemann, D., Strohmeier, D., Brodbeck, C., et al. (2013). MEG and EEG data analysis with MNE-Python. *Front. Neurosci.* 7:267. doi: 10.3389/fnins.2013.00267

Grimshaw, G. M., Foster, J. J., and Corballis, P. M. (2014). Frontal and parietal EEG asymmetries interact to predict attentional bias to threat. *Brain Cogn.* 90, 76–86. doi: 10.1016/j.bandc.2014.06.008

Hamilton, C. R., and Vermeire, B. A. (1988). Complementary hemispheric specialization in monkeys. *Science* 242, 1691–1694. doi: 10.1126/science.3201258

Hart, A. J., Whalen, P. J., Shin, L. M., McInerney, S. C., Fischer, H., and Rauch, S. L. (2000). Differential response in the human amygdala to racial outgroup vs. ingroup face stimuli. *Neuroreport* 11, 2351–2355. doi: 10.1097/00001756-200008030-00004

He, Y., Johnson, M. K., Dovidio, J. F., and McCarthy, G. (2009). The relation between race-related implicit associations and scalp-recorded neural activity evoked by faces from different races. *Soc. Neurosci.* 4, 426–442. doi: 10.1080/17470910902949184

Hess, E. H. (1972). "Pupillometrics," in *Handbook of Psychophysiology*, eds N. S. Greenfield and R. A. Sternbach (New York, NY: Holt, Richard & Winston), 491–531.

Hess, E. H., and Polt, J. M. (1960). Pupil size as related to interest value of visual stimuli. *Science* 132, 349–350. doi: 10.1126/science.132.3423.349

Hu, C., Wang, Q., Fu, G., Quinn, P. C., and Lee, K. (2014). Both children and adults scan faces of own and other races differently. *Vis. Res.* 102, 1–10. doi: 10.1016/j.visres.2014.05.010

Ito, T. A., and Urland, G. R. (2005). The influence of processing objectives on the perception of faces: an ERP study of race and gender perception. *Cogn. Affect. Behav. Neurosci.* 5, 21–36. doi: 10.3758/cabn.5.1.21

Kahneman, D. (1973). *Attention and Effort*. Vol. 1063. (Englewood Cliffs, NJ: Prentice-Hall).

Kang, J.-H., Kim, S. J., Cho, Y. S., and Kim, S.-P. (2015). Modulation of alpha oscillations in the human EEG with facial preference. *PLoS ONE* 10:e0138153. doi: 10.1371/journal.pone.0138153

Keene, O. N. (1995). The log transformation is special. *Stat. Med.* 14, 811–819. doi: 10.1002/sim.4780140810

Landi, S. M., and Freiwald, W. A. (2017). Two areas for familiar face recognition in the primate brain. *Science* 357, 591–595. doi: 10.1126/science.aan1139

Lehmann, V., Huis in't Veld, E. M. J., and Vingerhoets, A. J. J. M. (2013). The human and animal baby schema effect: correlates of individual differences. *Behav. Process.* 94, 99–108. doi: 10.1016/j.beproc.2013.01.001

LeVine, R. A., and Campbell, D. T. (1972). Ethnocentrism: theories of conflict, ethnic attitudes, and group behavior. *J. Polit.* 35, 1022–1024. doi: 10.2307/2129223

Loewenfeld, I. E. (1966). "Pupillary movements associated with light and near vision," in *Recent Developments in Vision Research*, ed M. A. Whitcomb (Washington, DC: National Academy of Sciences), 17–105.

Lorenz, K. (1943). Die angeborenen formen möglicher erfahrung. *Z. Tierpsychol.* 5, 235–409. doi: 10.1111/j.1439-0310.1943.tb00655.x

Lorenz, K., and Martin, R. (1971). Studies in animal and human behaviour. *Br. J. Philos. Sci.* 22, 81–82. doi: 10.1093/bjps/22.1.81

Maestripieri, D., and Pelka, S. (2002). Sex differences in interest in infants across the lifespan. *Hum. Nat.* 13, 327–344. doi: 10.1007/s12110-002-1018-1

Mascaro, J. S., Hackett, P. D., and Rilling, J. K. (2013). Testicular volume is inversely correlated with nurturing-related brain activity in human fathers. *Proc. Natl. Acad. Sci. U.S.A.* 110, 15746–15751. doi: 10.1073/pnas.1305579110

Mathôt, S. (2018). Pupillometry: psychology, physiology, and function. *J. Cogn.* 1:16. doi: 10.5334/joc.18

Mathôt, S., van der Linden, L., Grainger, J., and Vitu, F. (2013). The pupillary light response reveals the focus of covert visual attention. *PLoS ONE* 8:e78168. doi: 10.1371/journal.pone.0078168

Mathôt, S., and Van der Stigchel, S. (2015). New light on the mind's eye: the pupillary light response as active vision. *Curr. Dir. Psychol. Sci.* 24, 374–378. doi: 10.1177/0963721415593725

Parsons, C. E., Young, K. S., Kumari, N., Stein, A., and Kringelbach, M. L. (2011). The motivational salience of infant faces is similar for men and women. *PLoS ONE* 6:e20632. doi: 10.1371/journal.pone.0020632

Peer, P. (2005). *Cvl face database.* Available online at: http://lrv.fri.uni-lj.si/facedb.html (accessed November 16, 2019).

Penn, D. J., and Smith, K. R. (2007). Differential fitness costs of reproduction between the sexes. *Proc. Natl. Acad. Sci. U.S.A.* 104, 553–558. doi: 10.1073/pnas.0609301103

Plutchik, R. (1987). Evolutionary bases of empathy. *Empathy Dev.* 1, 38–46.

Prior, H., Lingenauber, F., Nitschke, J., and Güntürkün, O. (2002). Orientation and lateralized cue use in pigeons navigating a large indoor environment. *J. Exp. Biol.* 205, 1795–1805.

Proverbio, A. M., and De Gabriele, V. (2017). The other-race effect does not apply to infant faces: an ERP attentional study. *Neuropsychologia* 126, 36–45. doi: 10.1016/j.neuropsychologia.2017.03.028

Proverbio, A. M., Riva, F., Zani, A., and Martin, E. (2011). Is it a baby? Perceived age affects brain processing of faces differently in women and men. *J. Cogn. Neurosci.* 23, 3197–3208. doi: 10.1162/jocn_a_00041

Regolin, L., Marconato, F., and Vallortigara, G. (2004). Hemispheric differences in the recognition of partly occluded objects by newly hatched domestic chicks (gallus gallus). *Anim. Cogn.* 7, 162–170. doi: 10.1007/s10071-004-0208-0

Richeson, J. A., and Trawalter, S. (2008). The threat of appearing prejudiced and race-based attentional biases. *Psychol. Sci.* 19, 98–102. doi: 10.1111/j.1467-9280.2008.02052.x

Ries, A. J., Touryan, J., Vettel, J., McDowell, K., and Hairston, W. D. (2014). A comparison of electroencephalography signals acquired from conventional and mobile systems. *J. Neurosci. Neuroeng.* 3, 10–20. doi: 10.1166/jnsne.2014.1092

Seifritz, E., Esposito, F., Neuhoff, J. G., Lüthi, A., Mustovic, H., Dammann, G., et al. (2003). Differential sex-independent amygdala response to infant crying and laughing in parents versus nonparents. *Biol. Psychiatry* 54, 1367–1375. doi: 10.1016/S0006-3223(03)00697-8

Solina, F., Peer, P., Batagelj, B., Juvan, S., Kovac, J. (2003). "Color-based face detection in the 15 seconds of fame art installation," in *Proceedings of Mirage* (France: INRIA Rocquencourt), 38–47.

Steinhauer, S. (1983). Pupillary dilation to emotional visual stimuli revisited. *Psychophysiology* 20:S472.

Thiruchselvam, R., Harper, J., and Homer, A. L. (2016). Beauty is in the belief of the beholder: cognitive influences on the neural response to facial attractiveness. *Soc. Cogn. Affect. Neurosci.* 11, 1999–2008. doi: 10.1093/scan/nsw115

Trawalter, S., Todd, A. R., Baird, A. A., and Richeson, J. A. (2008). Attending to threat: race-based patterns of selective attention. *J. Exp. Soc. Psychol.* 44, 1322–1327. doi: 10.1016/j.jesp.2008.03.006

Waters, E., Merrick, S., Treboux, D., Crowell, J., and Albersheim, L. (2000). Attachment security in infancy and early adulthood: a twenty-year longitudinal study. *Child Dev.* 71, 684–689. doi: 10.1111/1467-8624.00176

Wetzel, N., Buttelmann, D., Schieler, A., and Widmann, A. (2016). Infant and adult pupil dilation in response to unexpected sounds. *Dev. Psychobiol.* 58, 382–392. doi: 10.1002/dev.21377

Willis, J., and Todorov, A. (2006). First impressions: making up your mind after a 100-ms exposure to a face. *Psychol. Sci.* 17, 592–598. doi: 10.1111/j.1467-9280.2006.01750.x

Winn, M. B., Wendt, D., Koelewijn, T., and Kuchinsky, S. E. (2018). Best practices and advice for using pupillometry to measure listening effort: an introduction for those who want to get started. *Trends Hear.* 22:2331216518800869. doi: 10.1177/2331216518800869

Wu, E. X. W., Laeng, B., and Magnussen, S. (2012). Through the eyes of the own-race bias: eye-tracking and pupillometry during face recognition. *Soc. Neurosci.* 7, 202–216. doi: 10.1080/17470919.2011.596946

Yamakawa, Y., Kanai, R., Matsumura, M., and Naito, E. (2009). Social distance evaluation in human parietal cortex. *PLoS ONE* 4:e4360. doi: 10.1371/journal.pone.0004360

Yap, W. J., Chan, E., and Christopoulos, G. I. (2016). "Nanyang facial emotional expression [N-FEE] database—development and validation," in *Poster presented at the 23rd Congress of the International Association for Cross-Cultural Psychology*. Nagoya.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *J. Pers. Soc. Psychol.* 9:1. doi: 10.1037/h0025848

Zekveld, A. A., Koelewijn, T., and Kramer, S. E. (2018). The pupil dilation response to auditory stimuli: current state of knowledge. *Trends Hear.* 22:2331216518777174. doi: 10.1177/2331216518777174

Zheng, X., and Segalowitz, S. J. (2014). Putting a face in its place: in- and out-group membership alters the N170 response. *Soc. Cogn. Affect. Neurosci.* 9, 961–968. doi: 10.1093/scan/nst069

Conflict of Interest: 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.

REVIEW
published: 03 December 2019
doi: 10.3389/fpsyg.2019.02638

Deep Brain Stimulation of the Subthalamic Nucleus Influences Facial Emotion Recognition in Patients With Parkinson's Disease: A Review

*Caroline Wagenbreth, Maria Kuehne, Hans-Jochen Heinze and Tino Zaehle**

Department of Neurology, Otto-von-Guericke-University Magdeburg, Magdeburg, Germany

OPEN ACCESS

Edited by:
Giulia Mattavelli,
University Institute of Higher Studies in Pavia, Italy

Reviewed by:
Francesca Assogna,
Santa Lucia Foundation (IRCCS), Italy
Carlo Alberto Artusi,
University of Turin, Italy

***Correspondence:**
Tino Zaehle
tino.zaehle@ovgu.de

Specialty section:
This article was submitted to Consciousness Research, a section of the journal Frontiers in Psychology

Received: *26 June 2019*
Accepted: *08 November 2019*
Published: *03 December 2019*

Citation:
Wagenbreth C, Kuehne M, Heinze H-J and Zaehle T (2019) Deep Brain Stimulation of the Subthalamic Nucleus Influences Facial Emotion Recognition in Patients With Parkinson's Disease: A Review. Front. Psychol. 10:2638. doi: 10.3389/fpsyg.2019.02638

Parkinson's disease (PD) is a neurodegenerative disorder characterized by motor symptoms following dopaminergic depletion in the substantia nigra. Besides motor impairments, however, several non-motor detriments can have the potential to considerably impact subjectively perceived quality of life in patients. Particularly emotion recognition of facial expressions has been shown to be affected in PD, and especially the perception of negative emotions like fear, anger, or disgust is impaired. While emotion processing generally refers to automatic implicit as well as conscious explicit processing, the focus of most previous studies in PD was on explicit recognition of emotions only, while largely ignoring implicit processing deficits. Deep brain stimulation of the subthalamic nucleus (STN-DBS) is widely accepted as a therapeutic measure in the treatment of PD and has been shown to advantageously influence motor problems. Among various concomitant non-motor effects of STN-DBS, modulation of facial emotion recognition under subthalamic stimulation has been investigated in previous studies with rather heterogeneous results. Although there seems to be a consensus regarding the processing of disgust, which significantly deteriorates under STN stimulation, findings concerning emotions like fear or happiness report heterogeneous data and seem to depend on various experimental settings and measurements. In the present review, we summarized previous investigations focusing on STN-DBS influence on recognition of facial emotional expressions in patients suffering from PD. In a first step, we provide a synopsis of disturbances and problems in facial emotion processing observed in patients with PD. Second, we present findings of STN-DBS influence on facial emotion recognition and especially highlight different impacts of stimulation on implicit and explicit emotional processing.

Keywords: deep brain stimulation, Parkinson's disease, emotional recognition, facial emotional expression, subthalamic nucleus

FACIAL EMOTION PROCESSING

The ability to recognize and identify emotional cues in other people is a crucial component of human communication and interaction. In fact, deficits in emotion recognition are associated with poor social competence, interpersonal functioning, and reduced quality of life (Ruffman et al., 2008).

Especially faces are complex, concrete, and socially significant stimuli that are linked to emotional reactions (Murphy and Zajonc, 1993). Moreover, human faces hold a natural salience and attract attention more than other visual stimuli (Krebs et al., 2011). Hence, facial stimuli are widely used in studies assessing emotion processing under various conditions and in different clinical disturbances. In contrast to emotional prosodic stimuli, they are independent from language and thus offer the possibility to render comparability across different countries and languages. In fact, Ekman postulated basic emotions that are independent from literate culture and can be identified cross-culturally in facial expressions. Accordingly, pictures of emotional facial expression are a widely used tool in research on affective (neuro)sciences. The Ekman faces (Ekman and Friesen, 1976) are the oldest but still used database, but other face databases have been constituted in research studies, for instance, the Karolinska databases (Lundqvist et al., 1998), PENN Facial Discrimination Test (Erwin et al., 1992), or NimStim (Tottenham et al., 2009).

However, facial stimuli are somehow static and unnatural, as people usually encounter moving and dynamic faces in daily life. Moving facial stimuli are advantageous over static ones because they can be recognized easier and faster, especially for rather seldom and culture-defined emotional expressions as pride or defiance, presumably because photographs contain no dynamic information that helps identifying facial expressions (Kan et al., 2002). Moreover, dynamic faces lead to more widespread neural activation patterns and also arouse cortical regions associated with higher social relevance (Sato et al., 2004; Trautmann et al., 2009).

To a certain degree, the processing of facial stimuli underlies some attentional biases. Fearful, angry, or generally threatening faces are detected faster than neutral ones (Ishai et al., 2004; Susa et al., 2012) in accordance with the evolutionary point of view of a faster processing of threatening or possibly life-endangering stimuli. A facilitation of visual search tasks to identify fear-related pictures among fear-irrelevant ones was demonstrated (Öhman and Mineka, 2001), as well as slower attention disengagement from angry faces compared to neutral or happy ones (Fox et al., 2002). The "automatic vigilance" hypothesis implies that people tend to focus their attention preferentially on negative stimuli and can also rather difficult dissolve it from them (Wentura et al., 2000; Öhman and Mineka, 2001). However, there are also hints that postulate an advantageous processing of positive stimuli, that is, words and facial expressions (Kuchinke et al., 2005; Hofmann et al., 2009; Kissler and Koessler, 2011; Wagenbreth et al., 2014). Thus, decoding and interpreting emotions displayed in facial expressions, especially those with negative valence, play a fundamental role in human interactions.

Another interesting factor when investigating facial emotion processing is the task type used in a study. Emotion processing can be assessed by identification tasks in which participants are requested to select an appropriate label for a given emotion. In contrast, in discrimination or matching tasks, one has to judge whether or not two faces express the same emotion. These different forms of emotion recognition also reflect different and successive processing stages. Both recognition processes occur relatively early after stimulus presentation. But while the perception if two faces are identical happens rather automatic and unconsciously, that is, in an implicit way, identification tasks require additional knowledge and conscious deliberation of facial expression information as they rather involve explicit emotional processing. Hence, implicit emotional processing refers to the automatic and involuntary processing of emotional stimuli that a person is confronted with (e.g., pictures or voices). In contrast, explicit processing is the downstream process afterward that requires cognitive abilities (e.g., concentration, attention) to process and classify these given emotional stimuli.

Both forms of emotion processing may be differentially compromised in patients with Parkinson's disease (PD), as will be shown in *Recognition of Facial Emotion in Parkinson's Patients* of this review. Subthalamic stimulation may further have the potential not only to influence emotional recognition out of facial expressions but also to affect implicit and explicit emotional processing in a distinct way. In the following sections of this review, we will point out findings on facial emotional processing in PD patients. Then, we will focus on how subthalamic stimulation may impact those processes and, finally, discuss methodological differences between studies.

RECOGNITION OF FACIAL EMOTION IN PARKINSON'S PATIENTS

Parkinson's disease is a neurodegenerative disease characterized by increasing motor impairments, such as tremor, rigidity, and bradykinesia due to a depletion of dopaminergic neurons in the striatum. While PD patients occasionally suffer from motor impairments in daily life, a spectrum of concomitant non-motor symptoms including changes in mood, impulsivity, and other neuropsychiatric aspects, as well as deficits in cognitive and emotional processing have the potential to considerably influence subjective well-being and quality of life in patients (Limousin et al., 1998; Romito et al., 2002; Ballanger et al., 2009; Abbes et al., 2018).

In the 1980s, it was first suggested that PD patients also suffer from a selective impairment in facial emotion recognition (Beatty et al., 1989; Blonder et al., 1989). Since then, this observation has been replicated in a vast majority of subsequent studies, showing deficient facial emotion recognition in PD patients when compared to healthy controls (HC) (e.g., Jacobs et al., 1995; Kan et al., 2002; Sprengelmeyer et al., 2003; Yip et al., 2003; Suzuki et al., 2006; Lawrence et al., 2007; Clark et al., 2008). Accordingly, reviews and extended meta-analyses demonstrated deficits in facial emotion processing and outlined potentially biasing and correlated factors between studies like visuospatial or cognitive deficits, disease severity, or mood disorders of patients among others (Assogna et al., 2008; Gray and Tickle-Degnen, 2010; Argaud et al., 2018). In particular, emotion recognition of faces was reported to be affected in all basic emotions, but most deficits concerned the recognition of negative emotions (Argaud et al., 2018). In contrast, however, other studies did not reveal any deterioration in emotion processing in PD (Madeley et al., 1995; Adolphs et al., 1998; Breitenstein et al., 1998; Pell and Leonard, 2005).

It was suggested that emotion recognition deficits in PD were associated with influences of dopaminergic medication, but results are inconsistent. Gray and Tickle-Degnen (2010) reported emotion perception to be largely unaffected by medication status; other studies supported this assumption (Yip et al., 2003; Péron et al., 2009; Cohen et al., 2010; Roca et al., 2010; Enrici et al., 2015). Contrastingly, several publications demonstrated beneficial effects of L-Dopa on emotion recognition (Tessitore et al., 2002; Sprengelmeyer et al., 2003; Delaveau et al., 2010), suggesting that L-Dopa partially restores amygdala response but in dependence of disease progression (Delaveau et al., 2009). In fact, in the early disease stages, mesocorticolimbic pathways are described to be relatively spared compared with the motor pathway (Braak et al., 2004), and L-Dopa needed to improve motor symptoms would at the same time overdose mesolimbic projections to subcortical structures involved in emotion processing like the amygdala, leading to detrimental effects in emotion perception, whereas in later disease stages, these effects would be beneficial.

Interestingly, studies consistently demonstrating deficient emotion recognition in PD patients generally assessed explicit emotional processing by explicit evaluation and naming of the emotional value of facial expressions (Yip et al., 2003; Kan et al., 2004; Schröder et al., 2006; Ibarretxe-Bilbao et al., 2009; Péron et al., 2015; Xi et al., 2015). Here, diminished recognition was demonstrated for specific emotions such as anger (Sprengelmeyer et al., 2003; Dujardin et al., 2004; Lawrence et al., 2007; Clark et al., 2008), surprise (Clark et al., 2008), fear (Kan et al., 2002; Sprengelmeyer et al., 2003), and disgust (Kan et al., 2002; Dujardin et al., 2004; Suzuki et al., 2006; Assogna et al., 2008).

In contrast, implicit emotional recognition abilities in PD have received rather limited attention so far, and thus, implicit processing deficits are less proven. Wieser et al. (2006) tested implicit and explicit emotion processing by means of event-related potentials in response to affective pictures. The authors reported preserved early implicit processing but blunted explicit emotional responses in PD patients, supporting the view that the elementary–implicit–reaction to an emotion is independent from the conscious response to it. Hence, the emotional salience detection from faces may be compromised in PD while early perceptual and structural face processing appears to be intact. Two further studies investigated implicit emotional processing of pictorial and verbal affective stimuli but did not test for emotional facial expressions (Castner et al., 2007; Borg et al., 2012). In accordance, they also reported preserved implicit emotional processing in PD patients. In a previous study, we assessed explicit as well as implicit emotional processing in PD (Wagenbreth et al., 2016). For this purpose, we investigated emotion recognition of circumscribed facial information, for example, emotional states portrayed only in the eye region of a face, in an affective priming paradigm. We found largely preserved implicit (i.e., automatic and unconscious) emotional processing of facial expressions in PD patients, as shown by a preserved sensitivity to emotional priming; however, with a specific altered processing of disgust- and happiness-connoted stimuli. In contrast and consistent with the literature, explicit emotional processing was considerably impaired for facial stimulus material in PD.

Hence, PD compromises explicit emotional processing in general, as was shown for semantic, pictorial, and facial stimuli (Castner et al., 2007; Borg et al., 2012; Wagenbreth et al., 2016), with a particular emphasis on facial emotion recognition due to its significance in social contexts. However, the ability to be implicitly sensitive to emotional content seems to be largely spared by the disease, especially for facial stimuli.

Yet, the presented studies on facial emotional processing in PD differ with respect to several methodological, clinical, and patients' individual factors and thus lack direct comparability. Because of this high variability in studies, several confounding factors need to be considered. In their meta-analysis, Gray and Tickle-Degnen (2010) pointed out seven potential moderators of emotional facial processing in PD that might be associated with inconsistent results of previous studies. Three of these postulated moderators concerned methodological aspects referring to the task (stimulus modality, task type, emotion displayed), and the other four relate to the patients themselves (medication status, motor disability, depression status, executive functions, and visuospatial abilities). The authors also stressed the role of potential working memory constraints during the task. Argaud et al. (2018) further emphasized the relevance of possible interactions between facial emotion impairments and mood disorders.

Functional neuroimaging studies investigating the neural basis of emotion recognition in healthy human subjects have proposed both the ventral striatum and the amygdala to be involved in processing of negative emotions like fear and disgust (Morris et al., 1996; Phillips et al., 1998). These structures consistently receive afferents from the dopaminergic neurons of the mesolimbic ventral tegmental area, which is known to degenerate in PD. The amygdala is supposed to represent an essential factor of the emotional face processing impairments in PD (Braak et al., 1994; Harding et al., 2002; Yoshimura et al., 2005). In a functional imaging study, Kipps et al. (2007) reported that the amygdala volume correlated with the ability to recognize happy facial expressions. In fact, projections from the amygdala reach the striatum, and the amygdala itself undergoes severe pathological changes during the course of PD (Braak et al., 1994). However, PD-related amygdala involvement seems to be unrelated to cognitive impairments (Braak et al., 1994), indicating that emotion processing associated with the amygdala might be spared even with ongoing PD pathology. Other attempts to explain deviant emotion processing in PD thus focus on the involvement of dopamine in emotional processing. Ongoing dopamine depletion of the mesolimbic pathway leads to a dysfunction of the limbic loop, which in turn links the basal ganglia to the orbitofrontal cortex (OFC), which leads to an impaired neuronal processing in the limbic projection area (Alexander et al., 1986; Cummings, 1993). Furthermore, the involvement of the striatum and specifically the insula in the processing of disgust was postulated (Sprengelmeyer et al., 1996; Phillips et al., 1997; Calder et al., 2000). The insula is highly interconnected with the basal ganglia and cortical regions, interacts with multiple brain networks (Chikama et al., 1997; Fudge et al., 2005), and is one of the first cortical regions that is pathologically affected in PD (Braak et al., 2006). Lesions of the insula impair the recognition of facial emotions (Calder et al., 2000), and the loss of normal metabolic activity in insular neurons in PD has been associated with blunted emotions in PD patients (Wieser et al., 2006; Robert et al., 2012). Finally, subcortical activation of the thalamus, putamen, and basal ganglia during emotional processing was observed in several functional MRI (fMRI) studies (Morris et al., 1999; Sander et al., 2005; Cheung et al., 2006), demonstrating involvement of the basal ganglia in emotion processing.

Another approach to explain diminished emotion processing in PD patients has been discussed. According to the *simulation theory of emotions*, some authors assume the contribution of motor impairments to the emotion recognition deficit (Goldman and Sripada, 2005; Niedenthal, 2007). PD-inherent symptoms like facial amimia and dysprosody have been discussed as being partially responsible for the patients' inaccuracy in emotion recognition, but other investigations report contradictory results (see Assogna et al., 2008, for a review).

INFLUENCE OF SUBTHALAMIC STIMULATION ON FACIAL EMOTION PROCESSING IN PD

Method

We conducted a detailed search of literature with the aim of reviewing all of the relevant papers on STN-DBS influence on facial emotion processing in PD patients. We searched PubMed services (April 2019) with the following keywords: *Parkinson's disease, deep brain stimulation, subthalamic stimulation, emotion recognition*, *emotion processing*, *face*, and *facial expression*. We also hand-searched relevant journals and examined the references of retrieved key articles to find possible further publications regarding facial emotion processing. This review particularly focused on reported comparisons between both DBS conditions (ON and OFF) and thus aims to report within-group differences between both measure conditions. Furthermore, only those publications were chosen that explicitly reported patients' performances on emotion recognition out of faces or facial stimuli, at least as a subtask among other tests. Articles were restricted to the English language and were published between 2003 and 2019. Nineteen publications fulfilled the abovementioned criterions and were identified as being relevant to the question of how STN-DBS influences facial emotion recognition in patients (**Table 1**).

Facial Emotion Recognition Under STN-DBS

Anatomical and Structural Correlates of Emotional Processing Under STN-DBS

Subthalamic stimulation has been variously proven to be a helpful therapeutic tool to improve motor disturbances in PD (Frank et al., 2004; Hershey et al., 2004, 2010) but has also been shown to influence cognitive domains, such as executive functions and language (Fasano et al., 2010; Wagenbreth et al., 2015; Zaehle et al., 2017). It further plays a central role in the regulation of limbic and emotional/affective functions (Temel et al., 2005; Mallet et al., 2007). The STN is part of the basal ganglia, which are interconnected to specific motor, cognitive, and limbic cortical regions through partially closed and sequentially arranged circuits (Alexander and Crutcher, 1990; Parent and Hazrati, 1995). The STN is thus closely connected with the ventral striatum and ventral pallidum, as well as to limbic cortical areas such as the anterior cingulated cortex and the OFC; regions that are known to play a major role in the recognition of emotions from faces and voices (Adolphs, 2002; Wildgruber et al., 2006) and from emotional prosody (Sander et al., 2005). DBS of the STN modifies dopaminergic transmission in the basal ganglia and thus affects the limbic, associative, and motor network circuits (Temel et al., 2005; Mallet et al., 2007). Considering the anatomic overlap of the functional areas of the STN (Haynes and Haber, 2013), stimulation of the dorsolateral motor region of the STN may as well have effects on the mesolimbic basal ganglia loops (Castrioto et al., 2014), leading to effects in emotion processing, for instance. In this context, STN-DBS has been described to occasionally cause neuropsychiatric effects like mood disturbances or depression (Berney et al., 2002; Krack et al., 2003). A relation between depression and emotion processing deficits has been suggested, but analyses yielded different results concerning possible interactions, suggesting emotion processing deficits not to be a consequence of depression or mood disturbances (Gray and Tickle-Degnen, 2010). Schneider et al. (2003) investigated the relation between mood induction and emotion recognition under STN-DBS and found a mood-enhancing effect as well as intensified emotional experience under stimulation but no changes in emotion recognition out of faces. The authors concluded that emotional recognition and categorization may involve other brain areas than those involved in the sole emotion memory or emotion experience (Schneider et al., 2003).

Generally, studies investigating the effects of STN-DBS on facial emotion perception in PD demonstrated rather heterogeneous results, and either covered unchanged emotion recognition of facial expressions under DBS (Schneider et al., 2003; Berney et al., 2007; Albuquerque et al., 2014; McIntosh et al., 2015) or worsening of explicit discriminating emotional faces under stimulation (Geday et al., 2006; Péron et al., 2010a,b). Precisely, a tendency for DBS to cause deficits in facial discrimination of especially negative emotions like disgust (Mondillon et al., 2012; Aiello et al., 2014;

TABLE 1 | Influence of STN-DBS on recognition of emotional facial expressions in PD patients.

	N	Disease duration	L-dopa therapy	Experimental design	Task	Emotional stimuli	Results
Schneider et al., 2003	12 PD	17.0 ± 6.3 years	Yes	Med only; Med-off/DBS-OFF; Med-off/DBS-ON	Discrimination	PENN facial discrimination test (Erwin et al., 1992)	• No stimulation effects on emotion discrimination; but mood-enhancing effect and improved emotional memory under DBS
Schroeder et al., 2004	10 PD	16 ± 3.1 years	Yes (except 1)	DBS-OFF; DBS-ON	Identification	FEEST (facial expressions of emotions: stimuli and test) (Young et al., 2002)	• Reduced recognition of angry expressions during DBS • No changes in facial emotion processing for all other emotions
Dujardin et al., 2004	12 PD, 12 HC	13 ± 2.5 years	Yes	DBS-OFF (pre-surgically); DBS-ON (3 months after surgery)	Intensity rating of expressions on emotion rating scales	Emotional facial expressions from Hess and Blairy (1995)	• General impairment in facial emotion decoding after surgery in 9/12 patients • Significant post-operative impairment for sadness, anger, and a trend for disgust
Biseul et al., 2005	15 PD (post-operative group), 15 PD (preoperative group), 15 HC	15 ± 6.2 years	Yes	DBS-OFF; DBS-ON	Identification	Ekman and Friesen, 1976	• Specific impairment to recognize fear in the tested post-operative group compared to preoperative group and HC; but no difference between DBS-ON and DBS-OFF within this group for all facial expressions
Geday et al., 2006	10 PD, 22 HC	13 ± 3.2 years	Yes	DBS-OFF; DBS-ON	PET study: no task, patients were instructed to look at pictures; After 1 week evaluation of "pleasantness"	EPS (Empathy Picture System)	• Inhibited emotional activation of the right fusiform gyrus under DBS • DBS raised emotional activation of the anterior cingulated and lowered activity of the putamen
Berney et al., 2007	15 PD	12.0 ± 6.0 years	Yes	Med-off/DBS-ON; Med-off/DBS-OFF; Med-on/DBS-OFF; Med-on/DBS-ON	Matching	Computerized task consisting of pairs conveying the same or different emotion	• Changes in mood core dimensions under stimulation, but stable emotion discrimination processing
Drapier et al., 2008	17 PD	11.8 ± 2.6 years	Yes	DBS-OFF (3 months pre surgery); DBS-ON (3 months post-surgery)	Identification	Ekman and Friesen, 1976	• Impaired fear and sadness recognition after STN-DBS • Worsened apathy scores after DBS
Le Jeune et al., 2008	13 PD, 30 HC	10.9 ± 2.2 years	Yes	DBS-OFF (3 months pre surgery); DBS-ON (3 months post-surgery)	Identification	Ekman and Friesen, 1976	• Selective reduction of fear recognition under stimulation • Correlation between reduced glucose metabolism in the right orbitofrontal cortex and reduced fear recognition
Péron et al., 2010a	24 PD, 20 untreated PD (pathological control group), 30 HC	11.9 ± 2.5 years	Yes	DBS-OFF (3 months pre surgery); DBS-ON (3 months post-surgery)	Identification	Ekman and Friesen, 1976	• Impaired fear and sadness recognition after STN-DBS
Péron et al., 2010b	13 PD, 13 HC	10.5 ± 3.6 years	Yes	DBS-OFF (3 months pre surgery); DBS-ON (3 months post-surgery)	Identification	RMET (Baron-Cohen et al., 1997)	• Reduced Emotion score for patients under DBS but unchanged results in gender attribution task

(Continued)

TABLE 1 | Continued

	N	Disease duration	L-dopa therapy	Experimental design	Task	Emotional stimuli	Results
Mondillon et al., 2012	14 PD, 14 HC	12.3 ± 0.7 years	Yes	Med-off/DBS-ON; Med-off/DBS-OFF; Med-on/DBS-OFF; Med-on/DBS-ON	Identification/ Categorization	Karolinska emotional faces database (Lundqvist et al., 1998)	• Decreased recognition of disgust under DBS (and Med-off) • No impairment in emotion recognition observed when both therapies (Med and DBS) were "ON" • Combined administration of Med and DBS has more benefit on facial emotion recognition than the separate administration of therapies alone
Aiello et al., 2014	12 PD, 13 HC	10.9 ± 4.1 years	Yes	Med-off/DBS-ON; Med-off/DBS-OFF; Med-on/DBS-OFF; Med-on/DBS-ON	Discrimination, Intensity rating of expressions on emotion rating scales	NimStim set (Tottenham et al., 2009)	• Process of DBS surgery (microlesions) reduced patients' performance on discrimination task independently of stimulus type • No changes in facial emotion recognition, stable performance on facial emotion discrimination after 4 months post-surgery, except for disgust
Albuquerque et al., 2014	30 PD	15.9 ± 7.0 years	Yes	DBS-OFF (before surgery); DBS-ON (1 years after surgery)	Matching, Identification	CATS (Comprehensive Affect Testing System) (Froming et al., 2006)	• No changes in discrimination and naming of emotional faces under DBS
Mermillod et al., 2014	14 PD, 14 HC	12.4 ± 0.7 years	Yes	Med-off/DBS-ON; Med-off/DBS-OFF; Med-on/DBS-OFF; Med-on/DBS-ON	Identification	Ekman and Friesen, 1976	• Lower overall recognition rate for high spatial frequency emotional faces under DBS • No effect of DBS on recognition of either broad or low spatial frequency faces
McIntosh et al., 2015	9 PD, 7 untreated PD (receiving Med), 23 elderly HC, 21 young HC	NA	Yes	Med only; DBS-OFF; DBS-ON	Identification	RMET, TASIT (Awareness of Social Inference Test, face + voice), (McDonald et al., 2003)	• Neither therapy type (Med or DBS + Med) nor therapy state (ON/OFF) changed emotion recognition
Irmen et al., 2017	11 PD, 11 HC	11.5 ± 4.2 years	Yes (except 1)	DBS-OFF; DBS-ON	Emotional Stroop Task	2D Facial Emotional Stimuli dataset (Erwin et al., 1992)	• No conflict-induced slowing under DBS • Valence bias affecting conflict-induced reaction time slowing under DBS-OFF
Enrici et al., 2017*	18 PD, 20 untreated PD (receiving Med), 20 HC	12.6 ± 3.0 years	Yes	Med only; Med-on/DBS-ON	Identification	Ekman and Friesen, 1976; RMET	• Comparisons only for DBS versus HC, no comparison between DBS-ON and DBS-OFF • No impairment in facial emotion recognition and in affective Theory of Mind in DBS patients compared to HC
Martínez-Fernández et al., 2018	16 PD, 16 HC	10.6 ± 3.4 years	Yes	Med-off/DBS-ON; Med-off/DBS-OFF; Med-on/DBS-OFF; Med-on/DBS-ON	Emotional Stroop Task	Karolinska emotional faces database (Lundqvist et al., 1998)	• For fearful faces, emotional Stroop effect was higher under dopaminergic treatment than under DBS • Both treatments did not modulate Stroop effect • EEG: L-Dopa but not DBS increases the amplitude of the event-related potential N170

(Continued)

TABLE 1 | Continued

	N	Disease duration	L-dopa therapy	Experimental design	Task	Emotional stimuli	Results
Wagenbreth et al., 2019	14 PD	11.7 ± 4.4 years	Yes (except 2)	DBS-OFF; DBS-ON	Affective priming paradigm, identification	Ekman and Friesen, 1976	• DBS affected explicit emotional processing more than implicit processing • DBS selectively diminished explicit processing of disgust but had an ameliorating effect on discriminating fear stimuli

**No within-group comparison; PD: Parkinson's disease patients; HC: healthy controls; Med: medication; DBS: deep brain stimulation.*

Wagenbreth et al., 2019), anger (Schroeder et al., 2004), sadness (Drapier et al., 2008), and fear (Drapier et al., 2008; Mondillon et al., 2012) was observed.

Aiello et al. (2014) found diminished disgust discrimination abilities under DBS for facial expressions and emotional prosodic stimuli in PD patients but emphasized that impaired disgust recognition was prominent also before DBS implant in patients, which is in line with other studies (see *Recognition of Facial Emotion in Parkinson's Patients*). Impaired processing of disgust is thus related to the neurodegenerative disease itself rather than just an impact of STN-DBS (Mondillon et al., 2012; Aiello et al., 2014). Still, there seems to be consensus that disgust processing in faces and also in other stimulus modalities (e.g., prosody) even worsened under STN-DBS (Vicente et al., 2009; Mondillon et al., 2012; Aiello et al., 2014). Results concerning other negative emotions are not that conclusive. For instance, while some studies could demonstrate a clear deterioration of fear recognition under stimulation (Drapier et al., 2008; Le Jeune et al., 2008; Péron et al., 2010a), others failed to show any worsening of fear processing under DBS (e.g., Schroeder et al., 2004; Aiello et al., 2014; McIntosh et al., 2015).

There are different attempts to explain findings of reduced facial emotion processing of especially negative emotions. Previous investigations found activity modulations in brain regions associated with emotional processing during STN stimulation. PET studies in patients showed activity changes in non-motor areas of the associative and limbic circuits during STN-DBS and thus contribute to the findings of the central role the STN holds in motor, cognitive, and limbic basal ganglia circuits (Schroeder et al., 2002, 2003; Hilker et al., 2004). Using local field potentials, Kühn et al. (2005) found limbic activation of the STN in response to emotionally arousing pictures and proposed that this might be a reason for altered affect in PD patients. Several authors hence proposed impaired facial emotion recognition to be a result of a limbic dysfunction induced by STN-DBS (Dujardin et al., 2004; Biseul et al., 2005; Drapier et al., 2008). Le Jeune et al. (2008) demonstrated a correlation between impaired fear recognition under STN-DBS and a decrease in glucose metabolism in the right OFC. In their positron emission tomography (PET) study, Geday et al. (2006) proposed that the stimulation of the STN would inhibit regional blood flow rate activity in the right lateral fusiform gyrus, an area that is generally activated by emotional facial expressions, leading to altered emotion perception in faces.

Implicit and Explicit Emotional Processing Under STN-DBS

Most studies investigating DBS influence on emotion recognition clearly concentrated on one form of emotion processing only. Whereas nearly all studies in this review measured the explicit identification of emotional facial expressions, Geday et al. (2006) rather investigated the emotional reaction of patients to the pictures displayed. Patient's general recognition of facial expressions was significantly altered by STN stimulation, whereas emotional assessment *per se* appeared to be unaffected during DBS because patients showed no significant perception changes for DBS-ON or OFF. The authors assumed that patients first needed to identify the emotional valence of a picture to subsequently assess an empathetic reaction to this picture. Geday et al. (2006) hence referred to different aspects of emotional processing, pointing out intact implicit, but impaired explicit emotional processing of facial expressions under DBS. This is in line with a study measuring implicit and explicit emotional lexical-semantic processing, suggesting that basal ganglia-thalamocortical circuits are likely not to be involved in the automatic (implicit) activation of emotion evaluations (Castner et al., 2007). Basal ganglia-thalamocortical activation would be necessary if cognitive-driven decisions are requested.

This notion was seized further in the study by Wagenbreth et al. (2019), in which, for the first time, DBS impact on implicit as well as explicit emotional processing in PD patients was examined. In an affective priming paradigm, the authors used emotional words and emotional circumscribed facial regions, that is, human eyes displaying different emotions. DBS affected explicit emotional processing more than implicit processing and had a considerable diminishing effect on the processing of disgust stimuli, but even improved explicit processing of fear stimuli. This study supported the assumption of the varying involvement of the basal ganglia, depending on demanded conscious or automatic stimuli perception and processing, and further stressed the existence of different neural mechanisms for different emotional expressions. As already described in *Recognition of Facial Emotion in Parkinson's Patients*, referring to studies assessing emotion processing in non-stimulated PD patients, the classification of analyses in implicit and explicit emotional processing is rather neglected in the following presented studies.

Most studies in this review report data of explicit processing only. Hence, a purpose and challenge for future studies lie in the investigation of both aspects combined.

Influence of Medication Status

One possible confounder that may explain differences between the reported studies is additional medication intake despite STN-DBS in patients. In all presented studies, patients received additional dopaminergic medication, which was applied supplementary to DBS (except for single patients). Some publications, however, tried to filter out L-Dopa-associated effects on facial recognition performance under DBS and deployed an experimental setting with alternating testing conditions for medication and DBS: (1) Med OFF/DBS OFF; (2) Med OFF/DBS ON; (3) Med ON/DBS OFF; (4) Med ON/DBS ON (Berney et al., 2007; Mondillon et al., 2012; Aiello et al., 2014; Mermillod et al., 2014; Martínez-Fernández et al., 2018). Two other studies reported data for facial recognition in three different conditions: for DBS ON and OFF and for medicated patients only (Schneider et al., 2003; McIntosh et al., 2015). Whereas Schneider et al. (2003) tested the identical patient group with medication only and with DBS only to ensure direct comparability, the medication group in the study of McIntosh et al. (2015) consisted of different patients from those in the tested DBS patient group.

In general, emotion recognition was described to be unaffected by medication status, and this was even found for patients not yet treated with STN-DBS (Roca et al., 2010; McIntosh et al., 2015). However, different studies demonstrate that the interaction between dopaminergic medication and STN-DBS offers the best results in facial emotion recognition (Mondillon et al., 2012; Martínez-Fernández et al., 2018). To explain this finding, modifications to the non-motor basal ganglia-thalamocortical circuitry and to the emotional functions of the OFC and amygdala through DBS and L-Dopa medication have been proposed. The interaction between L-Dopa and STN-DBS plays a crucial role for patients, since in most cases, dopaminergic medication intake is continued despite DBS implant in patients. In fact, all studies presented in **Table 1** report additional levodopa intake besides STN-DBS. L-Dopa could overdose the mesolimbic projections toward the amygdala and OFC and thus lead to altered amygdala activation in response to emotion perception (Delaveau et al., 2009; Vicente et al., 2009; Aiello et al., 2014). DBS would compensate this overactivation by decreasing OFC activity and thereby restoring the necessary OFC-amygdala interaction (Mondillon et al., 2012). In turn, L-Dopa would compensate for the decreasing effect DBS has on the OFC and amygdala, which may explain facial recognition improvement when both therapeutic measures are "ON." Another explanation was given by Vizcarra et al. (2019), who suggested therapeutic synergism of the effects of L-Dopa and DBS. In their meta-analysis, they found that the combined effect was greater than either treatment alone, while both alone lessened motor severity in patients to a similar effect.

Moreover, neurophysiological studies could show differing regional brain activation of areas associated with facial recognition, depending on therapy means. In their EEG study, Martínez-Fernández et al. (2018) demonstrated that the peak of the event-related potential N170, which is thought to represent facial integration and to be modulated by facial emotional content, was increased by levodopa, but not by DBS. While levodopa induced stronger activity in the right fusiform gyrus that generated the N170, STN-DBS hypoactivated this region. Geday et al. (2006) used PET to compare functional activations of brain regions when processing emotional facial expressions. They found inhibited regional blood flow rates in the right fusiform gyrus under DBS, which was not prominent when stimulation was switched OFF. STN-DBS did not change fusiform reaction to emotional expressions but raised the emotional activation of the anterior cingulate and lowered the activity of the putamen.

Methodological and Experimental Differences Between Studies

In general, it is difficult to draw reliable conclusions for distinct results of the single publications given the methodological and experimental differences between studies. However, one recurring finding throughout studies is that the impairment in emotion recognition in patients could not be attributed to secondary variables, such as depression, anxiety, cognitive declines, or visuospatial deficits (Dujardin et al., 2004; Schroeder et al., 2004; Drapier et al., 2008; Albuquerque et al., 2014). Hence, emotion recognition deficits are to be already associated with disease-specific mechanisms. For instance, the insula and amygdala have been shown to be involved in the processing of facial emotions (Sprengelmeyer et al., 1996; Calder et al., 2000) but to be involved in pathological changes due to PD, as described in *Recognition of Facial Emotion in Parkinson's Patients* of this review.

Moreover, the impairment in facial emotion recognition seems to selectively concern negative emotions, while findings on positive emotions are rather scarce or non-existing. Yet, one has to keep in mind that studies examined far more negative (fear, anger, sadness, disgust) than positively valenced emotions. Actually, only happiness serves as a positive emotion throughout all studies; some publications also included surprise, whose valence is difficult to define as positive or negative without explanatory context due to its ambiguity. Also, happiness can have distinct reasons and thus different implications and consequences. Hence, there exists a bias concerning the frequency and fluency of negative versus positive emotions because negative emotions are far more differentiated than positive ones. From the evolutionary point of view, negative emotions imply greater functional value than positive ones because they help ensure "surviving" and coping in negative situations–and are thus more noticeable than positive emotions in daily life. In accordance, the "*angry face advantage*" postulates that angry faces are detected faster than other emotional and neutral faces (LoBue, 2009). But contrastingly, happiness seems to be easily recognized from facial expressions, which usually leads to ceiling effects in some studies (Kan et al., 2002). The described effect of a PD-inherent specific processing impairment has to be regarded and interpreted under these preliminary aspects.

Methodological differences between the presented studies concern the investigated emotions, the applied tests, as well as calculations. For this reason, it is difficult to determine if STN-DBS selectively impairs the recognition of specific emotions or if it rather leads to an overall deficit. Whereas most studies investigated multiple and different single emotions like disgust, fear, or happiness (e.g., Drapier et al., 2008; Péron et al., 2010a; Mermillod et al., 2014), other studies calculated a general overall emotion score that does not provide information about processing mechanisms for specific emotions (for instance, Schneider et al., 2003; Berney et al., 2007; Péron et al., 2010b; McIntosh et al., 2015). Albuquerque et al. (2014) reported mean values for positive and negative emotions but not for single emotions. Finally, some studies used only a small subset of two facial expressions for their investigation, like happiness and sadness/fear due to experimental settings (Irmen et al., 2017; Martínez-Fernández et al., 2018). Also, Dujardin et al. (2004) decided to eliminate happiness and fear expressions in their setting from the start because of methodological considerations.

Other disparities between studies refer to experimental settings and individual patients' preconditions. For example, some researchers (e.g., Drapier et al., 2008; Péron et al., 2010a; McIntosh et al., 2015) tested early progression PD patients, that is, patients with motor impairments not past Hoehn and Yahr stage II. Other studies provided data of patients with far more advanced disease progressions (Dujardin et al., 2004; Schroeder et al., 2004). All studies except McIntosh et al. (2015) gave information about patients' disease duration [with a range from 10.5 ± 3.6 years (Péron et al., 2010b) to 17.0 ± 6.3 years (Schneider et al., 2003)].

Several studies compared emotion processing of facial expressions in PD patients preoperatively and postoperatively, with DBS-OFF results gained before surgery and DBS-ON results assessed several months after surgery (Dujardin et al., 2004; Drapier et al., 2008; Le Jeune et al., 2008; Péron et al., 2010a,b; Albuquerque et al., 2014). All of these except Albuquerque et al. (2014) reported diminished facial emotion (and prosody) recognition after DBS surgery. Albuquerque et al. (2014) did not find any differences between both measure times.

This experimental setting offers the advantage of precluding learning effects of the stimulus material due to long time intervals between both measures. But this time span of at least several months comes along with further subjective, social, and cognitive influences on patients, which cannot be taken into account when interpreting performances and might thus bias the comparability of both testing results. Plus, changes of medication doses between both testing points are common. Moreover, the DBS surgery itself as well as alterations during the operation can have impacts on the performance outcome under DBS ON, which are not present for the OFF condition. Especially Aiello et al. (2014) emphasized the effects of microlesions on facial recognition. They reported diminished facial emotion recognition in patients even before surgery. According to their results, soon after DBS surgery, before turning the stimulator ON, patients were impaired in facial discrimination and recognition tasks for the emotion sadness only but showed recovered ability to recognize disgust. After 4 months post-surgery, patients' performance in facial emotion recognition remained stable but impaired disgust recognition was prominent again, just like before DBS operation. The authors suggested microlesion effects in the processing of single emotions and stressed the existence of different neural mechanisms for different emotional expressions.

To face the problem of long time intervals and possible diminished comparability of OFF and ON results due to subjective and environmental influences, other studies tested performance only after patients had been operated by switching the stimulator ON and OFF post-operatively. One can assume that possible influences deriving from surgery (microlesions, releasing effects, etc.) have subsided until then and are thus comparable for both testing sessions (DBS-ON and DBS-OFF). However, surgery itself can have a considerable impact on neuropsychological and executive functioning and might provoke alterations in the processing of emotional stimulus material (Brück et al., 2011). Previously, Okun et al. (2009) proposed not STN-DBS itself but rather insertion or lesion effects associated with electrode implantation as a possible underlying mechanism to explain performance differences preoperatively and post-operatively. Reliable comparisons between DBS-treated patients and HC can thus only be made with constraints. Another point that has to be kept in mind when analyzing results of emotion processing under stimulation is the fact that the long-lasting cerebral circuits reorganization following chronic STN-DBS is a long-term procedure and cannot be tackled early after turning the stimulator OFF. Hence, given results in ON/OFF testings rather shortly after DBS surgery do not represent final modulations caused by stimulation; hence, long-term analyses would be desirable in this context. As already mentioned above, general psychiatric anxiety was not associated with impaired emotion perception (Dujardin et al., 2004). However, the sole act of switching the stimulator ON or OFF, respectively, might lead to transient side effects but also subjective inconveniences like bad expectations and worry of experiencing returning (motor) symptoms like tremor in the hands when the stimulation is stopped. This subjectively perceived and testing situation-dependent trouble might also impact patients' performance and might actually contribute to the differing results concerning facial emotion recognition in the presented studies.

Contrary to the remaining studies reported in this review, Enrici et al. (2017) investigated facial emotion recognition while stimulation was ON but did not test (or did not report) results for DBS-OFF and rather demonstrated comparisons to a group of L-Dopa-treated patients without surgery and to a group of HC. It is hence difficult to draw conclusions about the efficacy and impact of STN-DBS on emotional processing in this study.

Two publications in this review adopt a special position within research of facial emotion recognition under DBS. Irmen et al. (2017) and Martínez-Fernández et al. (2018) did not investigate explicit emotion recognition *per se*, but rather recognition of emotional faces during a Stroop task. Both applied a modified version of the classic Stroop test, the so-called facial emotional Stroop test developed by Etkin et al. (2006). Patients were asked to recognize the emotion that was expressed by faces, while they should ignore the emotion of a word written over these faces. Hence, congruent trials would imply the same emotional valence for both face and word. The Stroop effect refers to the difference in reaction times between congruent and incongruent trials and represents a measure of the overall conflict processing. Martínez-Fernández et al. (2018) examined "happiness," "fear," and a neutral condition, whereas Irmen et al. (2017) used "joy" and "grief" in their Stroop test. Irmen et al. (2017) reported no reaction slowing in patients under DBS, thus, demonstrating a defect in within-trial conflict adaptation induced by STN-DBS. Martínez-Fernández et al. (2018) found that STN-DBS, in contrast to levodopa, has no significant impact on emotional conflict processing, but that PD patients suffer from a dysfunction in the early processing of facial emotions, which could be anatomically localized to the inferotemporal cortex and the fusiform gyrus. They further pointed out different regional brain activation, depending on treatment. While levodopa increased activity in areas associated with emotion processing, DBS hypoactivated them. The Stroop effect would hence be modulated by levodopa, but not by STN-DBS, "with this modulation being mainly mediated through the effect of each treatment on the recognition of facial emotion" (Martínez-Fernández et al., 2018).

IMPLICATIONS

This review summarizes studies investigating the influence of STN-DBS on recognition of emotional facial expressions in PD patients. This summary provides double-sided results. The majority of studies either demonstrated worsening in the processing of at least single specific emotions or reported no changes in facial emotion recognition under STN-DBS. No study was able to show any ameliorations of facial emotion processing in patients when DBS was ON. Such improvements were visible only for non-facial material like written or spoken semantic stimuli and emotional prosody or non-facial pictorial stimuli (Castner et al., 2007; Brück et al., 2011; Serranová et al., 2011; Wagenbreth et al., 2019).

However, these insights are of course biased by the fact that facial emotion processing is already deteriorated in PD patients, even before DBS surgery. Hence, it is rather complicated to define a baseline of facial emotion recognition that refers to "normal" performance in PD patients, and studies investigating DBS impact on this performance always underlie this bias. Furthermore, several methodological as well as clinical and individual factors contribute to different findings and exacerbate a final assertion concerning the efficacy of STN-DBS on facial emotion perception. These factors relate to experimental settings, selection criteria and choice of patients, additional dopaminergic medication next to DBS and its simultaneous intake or suspension, or settings regarding switching the stimulator ON or OFF. A general problem of all studies seems to comprise the recruitment and engagement of patients because all reported STN-DBS studies in this review present rather small sample sizes. With the exception of Albuquerque et al. (2014), who reported results of $N = 30$ patients, no other study was able to present data of more than 24 patients. Indeed, most studies showed results of less than 15 patients. Hence, a difficulty and challenge in the research of STN-DBS effects consist of the availability of data of an appropriate sample size to ensure validity and generalizability.

Finally, further possibly confounding or interacting factors have not been a subject of interest over studies, as for instance gender of patients, different cultural circles, living conditions, and other social factors of patients. For example, it is conceivable that single patients who live rather solitarily and secluded may undergo problems in facial emotion recognition due to inexperience and lack of contact with other people. This might concern particularly "socially and culturally defined" emotions, which were subject of studies assessing not (only) basic emotions but rather emotional gradations or social emotions like the Reading the Mind in the Eyes test (Baron-Cohen et al., 1997). This might be taken into consideration in future investigations.

At any rate, studies demonstrated the possibility of deteriorated recognition of negative emotions out of faces, especially disgust or fear, after DBS surgery. This might have important implications for communication and social living for relatives, nurses, or caregivers of patients. Precautionary arrangements and accords should be entered to ensure optimal cooperation and well-being for patients.

Future studies might further aim to give more attention to the combined investigation of implicit and explicit emotional processing and the DBS influence on both. As this review could demonstrate, there is scarcely STN-DBS research investigating implicit emotional processing because most studies concentrated on emotion processing that is associated with thinking, categorizing, and reflecting.

Finally, it was recently postulated that "what is consistently reported as a group effect seems to be mainly driven by a small, but substantial subgroup of DBS-treated patients" (Højlund et al., 2017; Foki et al., 2018). Effects of treatment may be small and specific to certain individuals. Thus, results should be regarded with respect to interindividual characteristics and intensities as well as to possible heterogeneous gains and losses from DBS.

AUTHOR CONTRIBUTIONS

CW, MK, and TZ wrote the main manuscript. H-JH and TZ contributed to conception of the review. All authors reviewed the manuscript, read and approved the final manuscript, and agreed to be accountable for all aspects of the work.

REFERENCES

Abbes, M., Lhommée, E., Thobois, S., Klinger, H., Schmitt, E., Bichon, A., et al. (2018). Subthalamic stimulation and neuropsychiatric symptoms in Parkinson's disease: results from a long-term follow-up cohort study. *J. Neurol. Neurosurg. Psychiatry* 89, 836–843. doi: 10.1136/jnnp-2017-316373

Adolphs, R. (2002). Recognizing emotions from facial expressions: psychological and neurological mechanisms. *Behav. Cogn. Neurosci. Rev.* 1, 21–62. doi: 10.1177/1534582302001001003

Adolphs, R., Schul, R., and Tranel, D. (1998). Intact recognition of facial emotion in Parkinson's disease. *Neuropsychology* 12, 253–258. doi: 10.1037/0894-4105.12.2.253

Aiello, M., Eleopra, R., Lettieri, C., Mondani, M., D'Auria, S., Belgrado, E., et al. (2014). Emotion recognition in Parkinson's disease after subthalamic deep brain stimulation: differential effects of microlesion and STN stimulation. *Cortex* 51, 35–45. doi: 10.1016/j.cortex.2013.11.003

Albuquerque, L., Coelho, M., Martins, M., and Martins, I. P. (2014). STN-DBS does not change emotion recognition in advanced Parkinson's disease. *Parkinson Relat. Disord.* 20, 564–565. doi: 10.1016/j.parkreldis.2014.01.020

Alexander, G. E., and Crutcher, M. D. (1990). Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci.* 13, 266–271. doi: 10.1016/0166-2236(90)90107-l

Alexander, G. E., DeLong, M. R., and Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Ann. Rev. Neurosci.* 9, 357–381. doi: 10.1146/annurev.neuro.9.1.357

Argaud, S., Vérin, M., Sauleau, P., and Grandjean, D. (2018). Facial emotion recognition in Parkinson's disease: a review and new hypotheses. *Mov. Disord.* 33, 554–567. doi: 10.1002/mds.27305

Assogna, F., Pontieri, F., Caltagirone, C., and Spalletta, G. (2008). The recognition of facial emotion expressions in Parkinson's disease. *Eur. Neuropsychopharmacol.* 18, 835–848. doi: 10.1016/j.euroneuro.2008.07.004

Ballanger, B., van Eimeren, T., Moro, E., Lozano, A. M., Hamani, C., Boulinguez, P., et al. (2009). Stimulation of the subthalamic nucleus and impulsivity: release your horses. *Ann. Neurol.* 66, 817–824. doi: 10.1002/ana.21795

Baron-Cohen, S., Jolliffe, T., Mortimore, C., and Robertson, M. (1997). Another advanced test of theory of mind: evidence from very high functioning adults with autism or asperger syndrome. *J. Child Psychol. Psychiatry* 38, 813–822. doi: 10.1111/1469-7610.00715

Beatty, W. W., Staton, R. D., Weir, W. S., Monson, N., and Whitaker, H. A. (1989). Cognitive disturbances in Parkinson's disease. *J. Geriatr. Psychiatry Neurol.* 2, 22–33.

Berney, A., Panisset, M., Sadikot, A. F., Ptito, A., Dagher, A., Fraraccio, M., et al. (2007). Mood stability during acute stimulator challenge in Parkinson's disease patients under long-term treatment with subthalamic deep brainstimulation. *Mov. Dis.* 22, 1093–1096. doi: 10.1002/mds.21245

Berney, A., Vingerhoets, F., Perrin, A., Guex, P., Villemure, J. G., Burkhard, P. R., et al. (2002). Effect on mood of subthalamic DBS for Parkinson's disease: a consecutive series of 24 patients. *Neurology* 59, 1427–1429. doi: 10.1212/01.wnl.0000032756.14298.18

Biseul, I., Sauleau, P., Haegelen, C., Trebon, P., Drapier, D., Raoul, S., et al. (2005). Fear recognition is impaired by subthalamic nucleus stimulation in Parkinson's disease. *Neuropsychologia* 43, 1054–1059. doi: 10.1016/j.neuropsychologia.2004.10.006

Blonder, L. X., Gur, R. E., and Gur, R. C. (1989). The effects of right and left hemiparkinsonism on prosody. *Brain Lang.* 36, 193–207. doi: 10.1016/0093-934x(89)90061-8

Borg, C., Bedoin, N., Bogey, S., Michael, G. A., Poujois, A., Laurent, B., et al. (2012). Implicit and explicit emotional processing in Parkinson's disease. *J. Clin. Exp. Neuropsychol.* 34, 289–296. doi: 10.1080/13803395.2011.639296

Braak, H., Braak, E., Yilmazer, D., de Vos, R. A. I., Jansen, E. N. H., Bohl, J., et al. (1994). Amygdala pathology in Parkinson's disease. *Acta Neuropathol.* 88, 493–500.

Braak, H., Ghebremedhin, E., Rüb, U., Bratzke, H., and Del Tredici, K. (2004). Stages in the development of Parkinson's disease-related pathology. *Cell Tissue Res.* 318, 121–134.

Braak, H., Rüb, U., Schultz, C., and del Tredici, K. (2006). Vulnerability of cortical neurons to Alzheimer's and Parkinson's diseases. *J. Alzheimer Dis.* 9, 35–44.

Breitenstein, C., Daum, I., and Ackermann, H. (1998). Emotional processing following cortical and subcortical brain damage: contribution of the fronto-striatal circuitry. *Behav. Neurol.* 11, 29–42. doi: 10.1155/1998/579029

Brück, C., Wildgruber, D., Kreifelts, B., Krüger, R., and Wächter, T. (2011). Effects of subthalamic nucleus stimulation on emotional prosody comprehension in Parkinson's disease. *Pub. Library Sci. One* 6:e19140. doi: 10.1371/journal.pone 0019140 doi: 10.1371/journal.pone.0019140

Calder, A. J., Keane, J., Manes, F., Antoun, N., and Young, A. W. (2000). Impaired recognition and experience of disgust following brain injury. *Nat. Neurosci.* 3, 1077–1078. doi: 10.1038/80586

Castner, J. E., Chenery, H. J., Copland, D. A., Coyne, T. J., Sinclair, F., and Silburn, P. A. (2007). Semantic and affective priming as a function of stimulation of the subthalamic nucleus in Parkinson's disease. *Brain* 130, 1395–1407. doi: 10.1093/brain/awm059

Castrioto, A., Lhommée, E., Moro, E., and Krack, P. (2014). Mood and behavioural effects of subthalamic stimulation in Parkinson's disease. *Lancet Neurol.* 13, 287–305. doi: 10.1016/S1474-4422(13)70294-1

Cheung, C. C., Lee, T. M., Yip, J. T., King, K. E., and Li, L. S. (2006). The differential effects of thalamus and basal ganglia on facial emotion recognition. *Brain Cogn.* 61, 262–268. doi: 10.1016/j.bandc.2006.01.008

Chikama, M., McFarland, N. R., Amaral, D. G., and Haber, S. N. (1997). Insular cortical projections to functional regions of the striatum correlate with cortical cytoarchitectonic organization in the primate. *J. Neurosci.* 17, 9686–9705. doi: 10.1523/jneurosci.17-24-09686.1997

Clark, U. S., Neargarder, S., and Cronin-Golomb, A. (2008). Specific impairments in the recognition of emotional facial expressions in Parkinson's disease. *Neuropsychologia* 46, 2300–2309. doi: 10.1016/j.neuropsychologia.2008.03.014

Cohen, H., Gagné, M. H., Hess, U., and Pourcher, E. (2010). Emotion and object processing in Parkinson's disease. *Brain Cogn.* 72, 457-463.

Cummings, J. L. (1993). The neuroanatomy of depression. *J. Clin. Psychiatry* 54(Suppl.), 14–20.

Delaveau, P., Salgado-Pineda, P., Fossati, P., Witias, T., Azulay, J. P., and Blin, O. (2010). Dopaminergic modulation of the default mode network in Parkinson's disease. *Eur. Neuropsychopharmacol.* 20, 784–792. doi: 10.1016/j.euroneuro.2010.07.001

Delaveau, P., Salgado-Pineda, P., Witjas, T., Micallef-Roll, J., Fakra, E., Azulay, J. P., et al. (2009). Dopaminergic modulation of amgydala activity during emotion recognition in patients with Parkinson disease. *J. Clin. Psychopharmacol.* 29, 548–554. doi: 10.1097/JCP.0b013e3181bf1c5f

Drapier, D., Péron, J., Leray, E., Sauleau, P., Biseul, I., Drapier, S., et al. (2008). Emotion recognition impairment and apathy after subthalamic nucleus stimulation in Parkinson's disease have separate neural substrates. *Neuropsychologia* 46, 2796–2801. doi: 10.1016/j.neuropsychologia.2008.05.006

Dujardin, K., Blairy, S., Defebvre, L., Krystkowiak, P., Hess, U., Blond, S., et al. (2004). Subthalamic nucleus stimulation induces deficits in decoding emotional facial expressions in Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* 75, 202–208. doi: 10.1136/jnnp.2003.013656

Ekman, P., and Friesen, W. (1976). *Pictures of Facial Affect*. PaloAlto, CA: Consulting Psychologists Press.

Enrici, I., Adenzato, M., Ardito, R. B., Mitkova, A., Cavallo, M., Zibetti, M., et al. (2015). Emotion processing in Parkinson's disease: a three-level study on recognition, representation and regulation. *PLoS One* 10:e0131470. doi: 10.1371/journal.pone.0131470

Enrici, I., Mitkova, A., Castelli, L., Lanotte, M., Lopiano, L., and Adenzato, M. (2017). Deep brain stimulation of the subthalamic nucleus does not negatively affect social cognitive abilities of patients with Parkinson's disease. *Sci. Rep.* 7:9413. doi: 10.1038/s41598-017-09737-6

Erwin, R. J., Gur, R. C., Gur, R. E., Skolnick, B., Mawhinney-Hee, M., and Smailies, J. (1992). Facial emotion discrimination: I. task construction and behavioral findings in normal subjects. *Psychol. Res.* 42, 231–240. doi: 10.1016/0165-1781(92)90115-j

Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., and Hirsch, J. (2006). Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron* 51, 871–882. doi: 10.1016/j.neuron.2006.07.029

Fasano, A., Romito, L. M., Daniele, A., Piano, C., Zinno, M., Bentivoglio, A. R., et al. (2010). Motor and cognitive outcome in patients with Parkinson's disease 8 years after subthalamic implants. *Brain* 133, 2664–2676. doi: 10.1093/brain/awq221

Foki, T., Hitzl, D., Pirker, W., Novak, K., Pusswald, G., and Lehrner, J. (2018). Individual cognitive change after DBS-surgery in Parkinson's disease patients using reliable change index methodology. *Neuropsychiatry* 32, 149–158. doi: 10.1007/s40211-018-0271-4

Fox, E., Russo, R., and Dutton, K. (2002). Attentional bias for threat: evidence for delayed disengagement from emotional faces. *Cogn. Emot.* 16, 355–379. doi: 10.1080/02699930143000527

Frank, M. J., Seeberger, L. C., and O'Reilly, R. C. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306, 1940–1943. doi: 10.1126/science.1102941

Froming, K., Levy, M., Schaffer, S., and Ekman, P. (2006). *The Comprehensive Affect Testing System*. Sharpsburg, PA: Psychology Software, Inc.

Fudge, J. L., Breitbart, M. A., Danish, M., and Pannoni, V. (2005). Insular and gustatory inputs to the caudal ventral striatum in primates. *J. Comp. Neurol.* 490, 101–118. doi: 10.1002/cne.20660

Geday, J., Ostergaard, K., and Gjedde, A. (2006). Stimulation of subthalamic nucleus inhibits emotional activation of fusiform gyrus. *Neuroimage* 33, 706–714. doi: 10.1016/j.neuroimage.2006.06.056

Goldman, A. I., and Sripada, C. S. (2005). Stimulationist models of face-based emotion recognition. *Cognition* 94, 193–213. doi: 10.1016/j.cognition.2004.01.005

Gray, H. M., and Tickle-Degnen, L. (2010). A meta-analysis of performance on emotion recognition tasks in Parkinson's disease. *Neuropsychology* 24, 176–191. doi: 10.1037/a0018104

Harding, A. J., Stimson, E., Henderson, J. M., and Halliday, G. M. (2002). Clinical correlates of selective pathology in the amygdala of patients with Parkinson's disease. *Brain* 125, 2431–2445. doi: 10.1093/brain/awf251

Haynes, W. I., and Haber, S. N. (2013). The organization of prefrontal-subthalamic inputs in primates provides an anatomical substrate for both functional specificity and integration: implications for Basal Ganglia models and deep brain stimulation. *J. Neurosci.* 33, 4804–4814. doi: 10.1523/JNEUROSCI.4674-12.2013

Hershey, T., Campbell, M. C., Videen, T. O., Lugar, H. M., Weaver, P. M., Hartlein, J., et al. (2010). Mapping Go – No-Go performance within the subthalamic nucleus region. *Brain* 133, 3625–3634. doi: 10.1093/brain/awq256

Hershey, T., Revilla, F. J., Wernle, A., Gibson, P. S., Dowling, J. L., and Perlmutter, J. S. (2004). Stimulation of STN impairs aspects of cognitive control in PD. *Neurology* 62, 1110–1114. doi: 10.1212/01.wnl.0000118202.19098.10

Hess, U., and Blairy, S. (1995). *Set of Emotional Facial Stimuli*. Montreal: University of Quebec at Montreal.

Hilker, R., Voges, J., Weisenbach, S., Kalbe, E., Burghaus, L., Ghaemi, M., et al. (2004). Subthalamic nucleus stimulation restores glucose metabolism in associative and limbic cortices and in cerebellum: evidence from a FDG-PET study in advanced Parkinson's disease. *J. Cereb. Blood Flow Metab.* 24, 7–16. doi: 10.1097/01.WCB.0000092831.44769.09

Hofmann, M. J., Kuchinke, L., Tamm, S., Võ, M. L.-H., and Jacobs, A. M. (2009). Affective processing within 1/10th of a second: high arousal is necessary for early facilitative processing of negative but not positive words. *Cogn. Affect. Behav. Neurosci.* 9, 389–397. doi: 10.3758/9.4.389

Højlund, A., Petersen, M. V., Sridharan, K. S., and Østergaard, K. (2017). Worsening of verbal fluency after deep brain stimulation in Parkinson's disease: a focused review. *Comp. Struct. Biotechnol. J.* 15, 68–74. doi: 10.1016/j.csbj.2016.11.003

Ibarretxe-Bilbao, N., Junqué, C., Tolosa, E., Marti, M. J., Valldeoriola, F., Bargallo, N., et al. (2009). Neuroanatomical correlates of impaired decision-making and facial emotion recognition in early Parkinson's disease. *Eur. J. Neurosci.* 30, 1162–1171. doi: 10.1111/j.1460-9568.2009.06892.x

Irmen, F., Huebl, J., Schroll, H., Brücke, C., Schneider, G. H., Hamker, F. H., et al. (2017). Subthalamic nucleus stimulation impairs emotional conflict adaptation in Parkinson's disease. *Soc. Cogn. Affect. Neurosci.* 12, 1594–1604. doi: 10.1093/scan/nsx090

Ishai, A., Pessoa, L., Bikle, P. C., and Ungerleider, L. G. (2004). Repetition suppression of faces is modulated by emotion. *Proc. Natl. Acad. Sci. U.S.A.* 101, 9827–9832. doi: 10.1073/pnas.0403559101

Jacobs, D. H., Shuren, J., Bowers, D., and Heilman, K. M. (1995). Emotional facial imagery, perception and expression in Parkinson's disease. *Neurology* 45, 1696–1702. doi: 10.1212/wnl.45.9.1696

Kan, Y., Kawamura, M., Hasegawa, Y., Mochizuki, S., and Nakamura, K. (2002). Recognition of emotion from facial, prosodic and written verbal stimuli in Parkinson's disease. *Cortex* 38, 623–630. doi: 10.1016/s0010-9452(08)70026-1

Kan, Y., Mimura, M., Kamijima, K., and Kawamura, M. (2004). Recognition of emotion from moving facial and prosodic stimuli in depressed patients. *J. Neurol. Neurosurg. Psychiatry* 75, 1667–1671. doi: 10.1136/jnnp.2004.036079

Kipps, C. M., Duggins, A. J., McCusker, E. A., and Calder, A. J. (2007). Disgust and happiness recognition correlate with anteroventral insula and amygdala volume respectively in preclinical Huntington's disease. *J. Cogn. Neurosci.* 19, 1206–1217. doi: 10.1162/jocn.2007.19.7.1206

Kissler, J., and Koessler, S. (2011). Emotionally positive stimuli facilitate lexical decisions-an ERP study. *Bio. Psychol.* 86, 254–264. doi: 10.1016/j.biopsycho.2010.12.006

Krack, P., Batiz, A., Van Blercom, N., Chabardes, S., Fraix, V., Ardouin, C., et al. (2003). Five-year follow-up of bilateral stimulation of the subthalamic nucleus in advanced Parkinson's disease. *N. Engl. J. Med.* 349, 1925–1934. doi: 10.1056/nejmoa035275

Krebs, J. F., Biswas, A., Pascalis, O., Kamp-Becker, I., Remschmidt, H., and Schwarzer, G. (2011). Face processing in children with autism spectrum disorder: independent or interactive processing of facial identity and facial expression? *J. Autism. Dev. Disord.* 41, 796–804. doi: 10.1007/s10803-010-1098-4

Kuchinke, L., Jacobs, A. M., Grubich, C., Võ, M. L.-H., Conrad, M., and Herrmann, M. (2005). Incidental effects of emotional valence in single word processing: an fMRI study. *NeuroImage* 28, 1022–1032. doi: 10.1016/j.neuroimage.2005.06.050

Kühn, A. A., Hariz, M. I., Silberstein, P., Tisch, S., Kupsch, A., Schneider, G. H., et al. (2005). Activation of the subthalamic region during emotional processing in Parkinson disease. *Neurology* 65, 707–713. doi: 10.1212/01.wnl.0000174438.78399.bc

Lawrence, A. D., Goerendt, I. K., and Brooks, D. J. (2007). Impaired recognition of facial expressions of anger in Parkinson's disease patients acutely withdrawn from dopamine replacement therapy. *Neuropsychologia* 45, 65–74. doi: 10.1016/j.neuropsychologia.2006.04.016

Le Jeune, F., Péron, J., Biseul, I., Fournier, S., Sauleau, P., Drapier, S., et al. (2008). Subthalamic nucleus stimulation affects orbitofrontal conrtex in facial emotion recognition: a PET study. *Brain* 131, 1599–1608. doi: 10.1093/brain/awn084

Limousin, P., Krack, P., Pollak, P., Benazzouz, A., Ardouin, C., Hoffmann, D., et al. (1998). Electrical stimulation of the subthalamic nucleus in advanced Parkinson's disease. *New Engl. J. Med.* 339, 1105–1111.

LoBue, V. (2009). More than just a face in the crowd: detection of emotional facial expressions in young children and adults. *Dev. Sci.* 12, 305–313. doi: 10.1111/j.1467-7687.2008.00767.x

Lundqvist, D., Flykt, A., and Öhman, A. (1998). *The Karolinska Directed Emotional Faces – KDEF*. Stockholm: Karolinska Institutet.

Madeley, P., Ellis, A. W., and Mindham, R. H. (1995). Facial expressions and Parkinson's disease. *Behav. Neurol.* 8, 115–119. doi: 10.3233/BEN-1995-8207

Mallet, L., Schüpbach, M., N'Diaye, K., Remy, P., Bardinet, E., Czernecki, V., et al. (2007). Stimulation of subterritories of the subthalamic nucleus reveals its role in the integration of the emotional and motor aspects of behavior. *Proc. Natl. Acad. Sci. U.S.A.* 104, 10661–10666. doi: 10.1073/pnas.0610849104

Martínez-Fernández, R., Kibleur, A., Chabardès, S., Fraix, V., Castrioto, A., Lhommée, E., et al. (2018). Different effects of levodopa and subthalamic stimulation on emotional conflict in Parkinson's disease. *Hum. Brain Mapp.* 39, 5014–5027. doi: 10.1002/hbm.24341

McDonald, S., Flanagan, S., Rollins, J., and Kinch, J. (2003). TASIT: a new clinical tool for assessing social perception after traumatic brain injury. *J. Head Trauma Rehabil.* 18, 219–238. doi: 10.1097/00001199-200305000-00001

McIntosh, L. G., Mannavy, S., Camalier, C. R., Folley, B. S., Albritton, A., Konrad, P. E., et al. (2015). Emotion recognition in early Parkinson's disease patients undergoing deep brain stimulation or dopaminergic therapy: a comparison to healthy participants. *Front. Aging Neurosci.* 6:349. doi: 10.3389/fnagi.2014.00349

Mermillod, M., Mondillon, L., Rieu, I., Devaux, D., Chambres, P., Auxiette, C., et al. (2014). Dopamine replacement therapy and deep brain stimulation of the subthalamic nuclei induce modulation of emotional processes at different spatial frequencies in Parkinson's disease. *J. Parkinsons Dis.* 4, 97–110. doi: 10.3233/JPD-130256

Mondillon, L., Mermillod, M., Musca, S. C., Rieu, I., Vidal, T., Chambres, P., et al. (2012). The combined effect of subthalamic nuclei deep brain stimulation and L-dopa increases emotion recognition in Parkinson's disease. *Neuropsychologia* 50, 2868–2879. doi: 10.1016/j.neuropsychologia.2012.08.016

Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., et al. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature* 383, 812–815. doi: 10.1038/383812a0

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

Murphy, S. T., and Zajonc, R. B. (1993). Affect, cognition and awareness: affective priming with optimal and suboptimal stimulus exposures. *J. Pers. Soc. Psychol.* 64, 723–739. doi: 10.1037/0022-3514.64.5.723

Niedenthal, P. M. (2007). Embodying emotion. *Science* 316, 1002–1005. doi: 10.1126/science.1136930

Öhman, A., and Mineka, S. (2001). Fears, phobias and preparedness: toward an evolved module of fear and fear learning. *Psychol. Rev.* 108, 483–522. doi: 10.1037/0033-295x.108.3.483

Okun, M. S., Fernandez, H. H., Wu, S. S., Kirsch-Darrow, L., Bowers, D., Bova, F., et al. (2009). Cognition and mood in Parkinson's disease in subthalamic nucleus versus globus pallidus interna deep brain stimulation: the COMPARE trial. *Ann. Neurol.* 65, 586–595. doi: 10.1002/ana.21596

Parent, A., and Hazrati, L. N. (1995). Functional anatomy of the basal ganglia. I. the cortico-basal ganglia-thalamo-cortical loop. *Brain Res. Rev.* 20, 91–127. doi: 10.1016/0165-0173(94)00007-c

Pell, M. D., and Leonard, C. L. (2005). Facial expression decoding in early Parkinson's disease. *Cogn. Brain Res.* 23, 327–340. doi: 10.1016/j.cogbrainres.2004.11.004

Péron, J., Biseul, I., Leray, E., Vicente, S., Le Jeune, F., Drapier, S., et al. (2010a). Subthalamic nucleus stimulation affects fear and sadness recognition in Parkinson's disease. *Neuropsychology* 24, 1–8. doi: 10.1037/a0017433

Péron, J., Le Jeune, F., Haegelen, C., Dondaine, T., Drapier, D., and Sauleau, P. (2010b). Subthalamic nucleus stimulation affects theory of mind network: a PET study in Parkinson's disease. *PLoS One* 5:e9919. doi: 10.1371/journal.pone.0009919

Péron, J., Cekic, S., Haegelen, C., Sauleau, P., Patel, S., Drapier, D., et al. (2015). Sensory contribution to vocal emotion deficit in Parkinson's disease after subthalamic stimulation. *Cortex* 63, 172–183. doi: 10.1016/j.cortex.2014.08.023

Péron, J., Vicente, S., Leray, E., Drapier, S., Drapier, D., Cohen, R., et al. (2009). Are dopaminergic pathways involved in theory of mind? a study in Parkinson's disease. *Neuropsychologia* 47, 406–414. doi: 10.1016/j.neuropsychologia.2008.09.008

Phillips, M. L., Young, A. W., Scott, S. K., Calder, A. J., Andrew, C., Giampietro, V., et al. (1998). Neural responses to facial and vocal expressions of fear and disgust. *Proc. R. Soc. Bio. Sci.* 265, 1809–1817. doi: 10.1098/rspb.1998.0506

Phillips, M. L., Young, A. W., Senior, C., Brammer, M., Andrew, C., Calder, A. J., et al. (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature* 389, 495–498. doi: 10.1038/39051

Robert, G., Le Jeune, F., Lozachmeur, C., Drapier, S., Dondaine, T., Péron, J., et al. (2012). Apathy in patients with Parkinson disease without dementia or depression: a PET study. *Neurology* 79, 1155–1160. doi: 10.1212/WNL.0b013e3182698c75

Roca, M., Torralva, T., Gleichgerrcht, E., Chade, A., Arévalo, G. G., Gershanik, O., et al. (2010). Impairments in social cognition in early medicated and unmedicated Parkinson disease. *Cogn. Behav. Neurol.* 23, 152–158. doi: 10.1097/wnn.0b013e3181e078de

Romito, L., Raja, M., Daniele, A., Contarino, M. F., Bentivoglio, A. R., Barbier, A., et al. (2002). Transient mania with hypersexuality after surgery for high frequency stimulation of the subthalamic nucleus in Parkinson's disease. *Mov. Disord.* 17, 1371–1374. doi: 10.1002/mds.10265

Ruffman, T., Henry, J. D., Livingstone, V., and Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: implications for neuropsychological models of aging. *Neurosci. Biobehav. Rev.* 32, 863–881. doi: 10.1016/j.neubiorev.2008.01.001

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

Sato, W., Yoshikawa, S., Kochiyama, T., and Matsumura, M. (2004). The amygdala processes the emotional significance of facial expressions: an fMRI investigation using the interaction between expression and face direction. *Neuroimage* 22, 1006–1013. doi: 10.1016/j.neuroimage.2004.02.030

Schneider, F., Habel, U., Volkmann, J., Regel, S., Kornischka, J., Sturm, V., et al. (2003). Deep brain stimulation of the subthalamic nucleus enhances emotional processing in Parkinson disease. *Arch. Gen. Psychiatry* 60, 296–302. doi: 10 1001/archpsyc.60.3.296

Schröder, C., Möbes, J., Schütze, M., Szymanowski, F., Nager, W., Bangert, M., et al. (2006). Perception of emotional speech in Parkinson's disease. *Mov. Disord.* 21, 1774–1778. doi: 10.1002/mds.21038

Schroeder, U., Kuehler, A., Haslinger, B., Erhard, P., Fogel, W., Tronnier, V. M., et al. (2002). Subthalamic nucleus stimulation affects striato-anterior cingulate cortex circuit in a response conflict task: a PET study. *Brain* 125, 1995–2004. doi: 10.1093/brain/awf199

Schroeder, U., Kuehler, A., Hennenlotter, A., Haslinger, B., Tronnier, V. M., Krause, M., et al. (2004). Facial expression recognition and subthalamic nucleus stimulation. *J. Neurol. Neurosurg. Psychiatry* 75, 648–650. doi: 10.1136/jnnp.2003.019794

Schroeder, U., Kuehler, A., Lange, K. W., Haslinger, B., Tronnier, V. M., Krause, M., et al. (2003). Subthalamic nucleus stimulation affects a frontotemporal network: a PET study. *Ann. Neurol.* 54, 445–450. doi: 10.1002/ana.10683

Serranová, T., Jech, R., Dusek, P., Sieger, T., Ruzicka, F., Urgosik, D., et al. (2011). Subthalamic nucleus stimulation affects incentive salience attribution in Parkinson's disease. *Mov. Disord.* 26, 2260–2266. doi: 10.1002/mds.23880

Sprengelmeyer, R., Young, A. W., Calder, A. J., Karnat, A., Lange, H., and Hömberg, V. (1996). Loss of disgust. perception of faces and emotions in huntington's disease. *Brain* 119, 1647–1665. doi: 10.1093/brain/119.5.1647

Sprengelmeyer, R., Young, A. W., Mahn, K., Schroeder, U., Woitalla, D., Büttner, T., et al. (2003). Facial expression recognition in people with medicated and unmedicated Parkinson's disease. *Neuropsychologia* 41, 1047–1057. doi: 10.1016/s0028-3932(02)00295-6

Susa, G., Pitică, I., Benga, O., and Miclea, M. (2012). The self regulatory effect of attentional control in modulating the relationship between attentional biases toward threat and anxiety symptoms in children. *Cogn. Emot.* 26, 1069–1083. doi: 10.1080/02699931.2011.638910

Suzuki, A., Hoshini, T., Shigemasu, K., and Kawamura, M. (2006). Disgust specific impairment of facial expression recognition in Parkinson's disease. *Brain* 129, 707–717. doi: 10.1093/brain/awl011

Temel, Y., Blokland, A., Steinbusch, H. W., and Visser-Vandewalle, V. (2005). The functional role of the subthalamic nucleus in cognitive and limbic circuits. *Prog. Neurobiol.* 76, 393–413. doi: 10.1016/j.pneurobio.2005.09.005

Tessitore, A., Hariri, A. R., Fera, F., Smith, W. G., Chase, T. N., Hyde, T. M., et al. (2002). Dopamine modulates the response of the human amgydala: a study in Parkinson's disease. *J. Neurosci.* 22, 9099–9103. doi: 10.1523/jneurosci.22-20-09099.2002

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

Trautmann, S. A., Fehr, T., and Herrmann, M. (2009). Emotions in motion: dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion-specific activations. *Brain Res.* 1284, 100–115. doi: 10.1016/j.brainres.2009.05.075

Vicente, S., Biseul, I., Péron, J., Philippot, P., Drapier, S., Drapier, D., et al. (2009). Subthalamic nucleus stimulation affects subjective emotional experience in Parkinson's disease patients. *Neuropsychologia* 47, 1928–1937. doi: 10.1016/j.neuropsychologia

Vizcarra, J. A., Situ-Kcomt, M., Artusi, C. A., Duker, A. P., Lopiano, L., Okun, M. S., et al. (2019). Subthalamic deep brain stimulation and levodopa in Parkinson's disease: a meta-analysis of combined effects. *J. Neurol.* 266, 289–297. doi: 10.1007/s00415-018-8936-2

Wagenbreth, C., Kuehne, M., Voges, J., Heinze, H. J., and Zaehle, T. (2019). Deep brain stimulation of the subthalamic nucleus selectively modulates emotion recognition of facial stimuli in Parkinson's patients. *J. Clin. Med.* 8:E1335. doi: 10.3390/jcm8091335

Wagenbreth, C., Rieger, J., Heinze, H. J., and Zaehle, T. (2014). Seeing emotions in the eyes – inverse priming effects induced by eyes expressing mental states. *Front. Psychol.* 5:1039. doi: 10.3389/fpsyg.2014.01039

Wagenbreth, C., Wattenberg, L., Heinze, H. J., and Zaehle, T. (2016). Implicit and explicit processing of emotional facial expressions in Parkinson's disease. *Behav. Brain Res.* 303, 182–190. doi: 10.1016/j.bbr.2016.01.059

Wagenbreth, C., Zaehle, T., Galazky, I., Voges, J., Guitart-Masip, M., Heinze, H. J., et al. (2015). Deep brain stimulation of the subthalamic nucleus modulates reward processing and action selection in Parkinson patients. *J. Neurol.* 262, 1541–1547. doi: 10.1007/s00415-015-7749-9

Wentura, D., Rothermund, K., and Bak, P. (2000). Automatic vigilance: the attention-grabbing power of approach- and avoidance-related social information. *J. Pers. Soc. Psychol.* 78, 1024–1037. doi: 10.1037//0022-3514.78.6.1024

Wieser, M. J., Muhlberger, A., Alpers, G. W., Macht, M., Ellgring, H., and Pauli, P. (2006). Emotion processing in Parkinson's disease: dissociation between early neuronal processing and explicit ratings. *Clin. Neurophysiol.* 117, 94–102. doi: 10.1016/j.clinph.2005.09.009

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

Xi, C., Zhu, Y., Mu, Y., Chen, B., Dong, B., Cheng, H., et al. (2015). Theory of mind and decision-making processes are impaired in Parkinson's disease. *Behav. Brain Res.* 279, 226–233. doi: 10.1016/j.bbr.2014.11.035

Yip, J. T. H., Lee, T. M. C., Ho, S. L., Tsang, K. L., and Li, L. S. W. (2003). Emotion recognition in patients with Parkinson's Disease. *Mov. Disord.* 18, 1115–1122. doi: 10.1002/mds.10497

Yoshimura, N., Kawamura, M., Masaoka, Y., and Homma, I. (2005). The amygdala of patients with Parkinson's disease is silent in response to fearful facial expressions. *Neuroscience* 131, 523–534. doi: 10.1016/j.neuroscience.2004.09.054

Young, A. W., Perrett, D., Calder, A., Sprengelmeyer, R. H., and Ekman, P. (2002). *Facial Expressions of Emotions: Stimuli and Test (FEEST)*. Thurstone: Thames Valley Test Company.

Zaehle, T., Wagenbreth, C., Voges, J., Heinze, H. J., and Galazky, I. (2017). Effects of deep brain stimulation of the subthalamic nucleus on perceptual decision making. *Neuroscience* 343, 140–146. doi: 10.1016/j.neuroscience.2016.11.044

Conflict of Interest: 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.

frontiers in Psychology

HYPOTHESIS AND THEORY
published: 19 December 2019
doi: 10.3389/fpsyg.2019.02844

Simulating Emotions: An Active Inference Model of Emotional State Inference and Emotion Concept Learning

Ryan Smith[1]*, Thomas Parr[2] and Karl J. Friston[2]

[1] Laureate Institute for Brain Research, Tulsa, OK, United States, [2] Wellcome Centre for Human Neuroimaging, Institute of Neurology, University College London, London, United Kingdom

The ability to conceptualize and understand one's own affective states and responses – or "Emotional awareness" (EA) – is reduced in multiple psychiatric populations; it is also positively correlated with a range of adaptive cognitive and emotional traits. While a growing body of work has investigated the neurocognitive basis of EA, the neurocomputational processes underlying this ability have received limited attention. Here, we present a formal Active Inference (AI) model of emotion conceptualization that can simulate the neurocomputational (Bayesian) processes associated with learning about emotion concepts and inferring the emotions one is feeling in a given moment. We validate the model and inherent constructs by showing (i) it can successfully acquire a repertoire of emotion concepts in its "childhood", as well as (ii) acquire new emotion concepts in synthetic "adulthood," and (iii) that these learning processes depend on early experiences, environmental stability, and habitual patterns of selective attention. These results offer a proof of principle that cognitive-emotional processes can be modeled formally, and highlight the potential for both theoretical and empirical extensions of this line of research on emotion and emotional disorders.

Keywords: emotion concepts, trait emotional awareness, learning, computational neuroscience, active inference

OPEN ACCESS

Edited by:
Alessia Celeghin,
University of Turin, Italy

Reviewed by:
Mauro Ursino,
University of Bologna, Italy
Carlos Andrés Méndez,
University of Turin, Italy

****Correspondence:***
Ryan Smith
rsmith@laureateinstitute.org

Specialty section:
This article was submitted to Consciousness Research, a section of the journal Frontiers in Psychology

Received: *02 August 2019*
Accepted: *02 December 2019*
Published: *19 December 2019*

Citation:
Smith R, Parr T and Friston KJ (2019) Simulating Emotions: An Active Inference Model of Emotional State Inference and Emotion Concept Learning. Front. Psychol. 10:2844. doi: 10.3389/fpsyg.2019.02844

INTRODUCTION

The ability to conceptualize and understand one's affective responses has become the topic of a growing body of empirical work (McRae et al., 2008; Smith et al., 2015, 2017b,c, 2018c,d,e, 2019a,c; Wright et al., 2017). This body of work has also given rise to theoretical models of its underlying cognitive and neural basis (Wilson-Mendenhall et al., 2011; Lane et al., 2015; Smith and Lane, 2015, 2016; Barrett, 2017; Kleckner et al., 2017; Panksepp et al., 2017; Smith et al., 2018b). Attempts to operationalize this cognitive-emotional ability have led to a range of overlapping constructs, including trait emotional awareness (Lane and Schwartz, 1987), emotion differentiation or granularity (Kashdan and Farmer, 2014; Kashdan et al., 2015), and alexithymia (Bagby et al., 1994a,b).

This work is motivated to a large degree by the clinical relevance of emotion conceptualization abilities. In the literature on the construct of emotional awareness, for example, lower levels of conceptualization ability have been associated with several psychiatric disorders as well as poorer physical health (Levine et al., 1997; Berthoz et al., 2000; Bydlowski et al., 2005; Donges et al., 2005; Lackner, 2005; Subic-Wrana et al., 2005, 2007; Frewen et al., 2008; Baslet et al., 2009; Consoli et al., 2010; Moeller et al., 2014); conversely, higher ability levels have been associated with a range of adaptive emotion-related traits and abilities (Lane et al., 1990, 1996, 2000; Ciarrochi et al., 2003; Barchard and Hakstian, 2004; Bréjard et al., 2012). Multiple evidence-based psychotherapeutic modalities also aim to improve emotion understanding as a central part of psycho-education in psychotherapy (Hayes and Smith, 2005; Barlow et al., 2016).

While there are a number of competing views on the nature of emotions, most (if not all) accept that emotion concepts must be acquired through experience. For example, "basic emotions" theories hold that emotion categories like sadness and fear each have distinct neural circuitry, but do not deny that knowledge about these emotions must be learned (Panksepp and Biven, 2012). Constructivist views instead hold that emotion categories do not have a 1-to-1 relationship to distinct neural circuitry, and that emotion concept acquisition is necessary for emotional experience (Barrett, 2017). While these views focus on understanding the nature of emotions themselves, we have recently proposed a neurocognitive model – termed the "three-process model" (TPM; Smith et al., 2018b, 2019a; Smith, 2019) of emotion episodes – with a primary focus on accounting for individual differences in emotional awareness. This model characterizes a range of emotion-related processes that could contribute to trait differences in both the learning and deployment of emotion concepts in order to understand one's own affective responses (and in the subsequent use of these concepts to guide adaptive decision-making). The TPM distinguishes the following three broadly defined processes (see **Figure 1**):

1. Affective response generation: a process in which somatovisceral and cognitive states are automatically modulated in response to an affective stimulus (whether real, remembered, or imagined) in a context-dependent manner, based on an (often implicit) appraisal of the significance of that stimulus for the survival and goal-achievement of the individual (i.e., predictions about the cognitive, metabolic, and behavioral demands of the situation).
2. Affective response representation: a process in which the somatovisceral component of an affective response is subsequently perceived via afferent sensory processing, and then conceptualized as a particular emotion (e.g., sadness, anger, etc.) in consideration of all other available sources of information (e.g., stimulus/context information, current thoughts/beliefs about the situation, etc.).
3. Conscious access: a process in which the representations of somatovisceral percepts and emotion concept representations may or may not enter and be held in working memory – constraining the use of this information in goal-directed decision-making (e.g., verbal reporting, selection of voluntary emotion regulation strategies, etc.).

The TPM has also proposed a tentative mapping to the brain in terms of interactions between large-scale neural networks serving domain-general cognitive functions. Some support for this proposal has been found within recent neuroimaging studies (Smith et al., 2017b,c, 2018a,c,d,e). However, the neurocomputational implementation of these processes has not been thoroughly considered. The computational level of description offers the promise of providing more specific and mechanistic insights, which could potentially be exploited to inform and improve pharmacological and psychotherapeutic interventions. While previous theoretical work has applied active inference concepts to emotional phenomena (Joffily and Coricelli, 2013; Seth, 2013; Barrett and Simmons, 2015; Seth and Friston, 2016; Smith et al., 2017d, 2019e; Clark et al., 2018), no formal modeling of emotion concept learning has yet been performed. In this manuscript, we aim to take the first steps in constructing an explicit computational model of the acquisition and deployment of emotion concept knowledge (i.e., affective response representation) as described within the TPM (subsequent work will focus on affective response generation and conscious access processes; see Smith et al., 2019b). Specifically, we present a simple Active Inference model (Friston et al., 2016, 2017a) of emotion conceptualization, formulated as a Markov Decision Process. We then outline some initial insights afforded by simulations using this model.

In what follows, we first provide a brief review of active inference. We will place a special emphasis on deep generative models that afford the capacity to explain multimodal (i.e., interoceptive, proprioceptive, and exteroceptive) sensations that are characteristic of emotional experience. We then introduce a particular model of emotion inference that is sufficiently nuanced to produce synthetic emotional processes but sufficiently simple to be understood from a "first principles" account. We then establish the validity of this model using numerical analyses of emotion concept learning during (synthetic) neurodevelopment. We conclude with a brief discussion of the implications of this work; particularly for future applications.

AN ACTIVE INFERENCE MODEL OF EMOTION CONCEPTUALIZATION

A Primer on Active Inference

Active Inference (AI) starts from the assumption that the brain is an inference machine that approximates optimal probabilistic (Bayesian) belief updating across all biopsychological domains (e.g., perception, decision-making, motor control, etc.). AI postulates that the brain embodies an internal model of the world (including the body) that is "generative" in the sense that it is able to simulate the sensory data that it should receive if its model of the world is correct. This simulated (predicted) sensory data

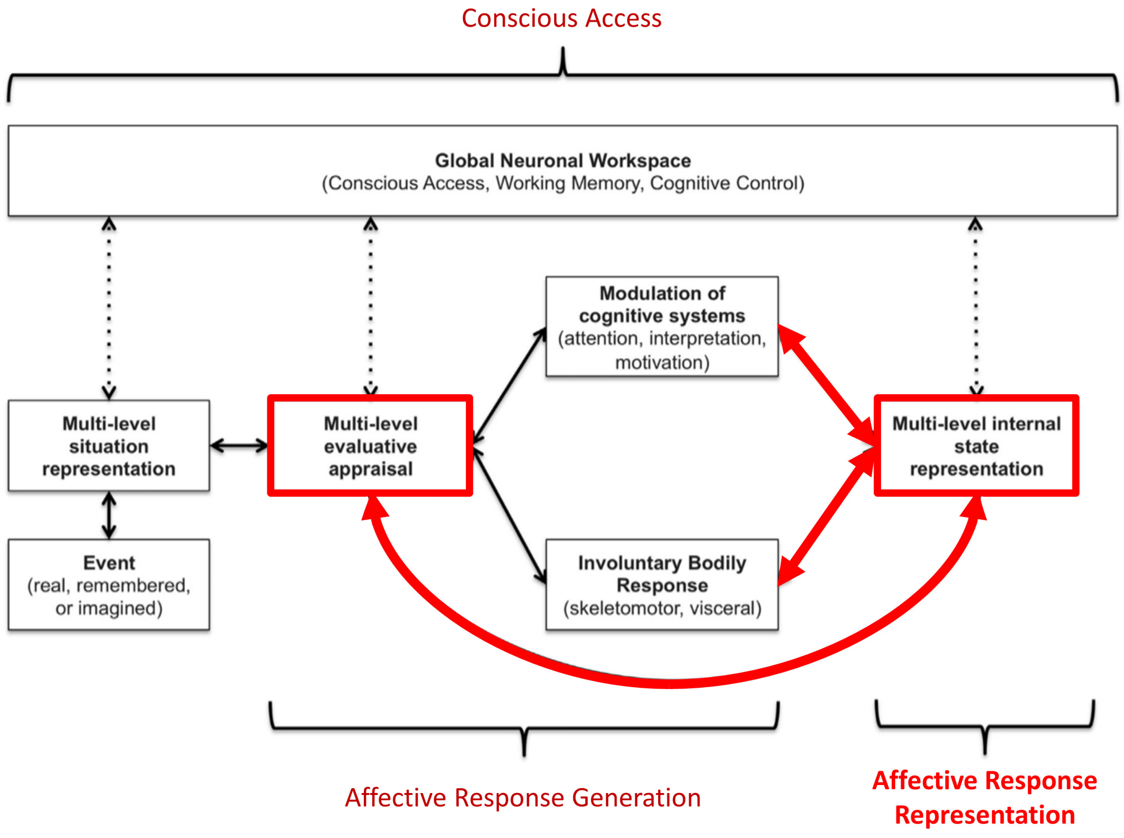


FIGURE 1 | Graphical illustration of the three-process model of emotion episodes (Smith et al., 2018b, 2019a,b; Smith, 2019). Here an event (either internal or external, and either real, remembered, or imagined) is represented in the brain at various levels of abstraction. This complex set of representations is evaluated in terms of its significance to the organism at both implicit (e.g., low-level conditioned or unconditioned responses) and explicit (e.g., needs, goals, values) levels, leading to predictions about the cognitive, metabolic, and behavioral demands of the situation. These predictions initiate an "affective response" with both peripheral and central components. This includes quick, involuntary autonomic and skeletomotor responses (e.g., heart rate changes, facial expression changes) as well as involuntary shifts in the direction of cognitive resources (e.g., attention and interpretation biases, action selection biases). This multicomponent affective response is represented centrally at a perceptual and conceptual level (e.g., valenced bodily sensations and interpretations of those responses as corresponding to particular emotions). Representations of events, evaluations, bodily sensations, emotion concepts (etc.) can then be attended to and gated into working memory ("conscious access"). If gated into and held in working memory, these representations can be integrated with explicit goals, reflected upon, and used to guide deliberative, goal-directed decision-making. The thick red arrows highlight the processes we explicitly model in this paper.

can be compared to actual observations, and deviations between predicted and observed sensations can then be used to update the model. On short timescales (e.g., a single trial in a perceptual decision-making task) this updating corresponds to perception, whereas on longer timescales it corresponds to learning (i.e., updating expectations about what will be observed on subsequent trials). One can see these processes as ensuring the generative model (embodied by the brain) remains an accurate model of the world (Conant and Ashbey, 1970).

Action (be it skeletomotor, visceromotor, or cognitive action) can be cast in similar terms. For example, actions can be chosen to resolve uncertainty about variables within a generative model (i.e., sampling from domains in which the model does not make precise predictions). This can prevent future deviations from predicted outcomes. In addition, the brain must continue to make certain predictions simply in order to survive. For example, if the brain did not in some sense continue to "expect" to observe certain amounts of food, water, shelter, social support, and a range of other quantities, then it would cease to exist (McKay and Dennett, 2009); as it would not pursue those behaviors leading to the realization of these expectations [c.f. "the optimism bias" (Sharot, 2011)]. Thus, there is a deep sense in which the brain must continually seek out observations that support – or are internally consistent with – its own continued existence. As a result, decision-making can be cast as a process in which the brain infers the sets of actions (policies) that would lead to observations most consistent with its own survival-related expectations (i.e., its "prior preferences"). Mathematically, this can be described as selecting policies that maximize a quantity called "Bayesian model evidence" – that is, the probability that sensory data would be observed under a given model. In other words, because the brain is itself a model of the world, action can be understood as a process by which the brain seeks out evidence for itself – sometimes known as self-evidencing (Hohwy, 2016).

In a real-world setting, directly computing model evidence becomes mathematically intractable. Thus, the brain must use some approximation. AI proposes that the brain instead computes a statistical quantity called free energy. Unlike model evidence, computing free energy is mathematically tractable. Crucially, this quantity provides a bound on model evidence, such that minimization of free energy is equivalent to maximizing model evidence. By extension, in decision-making an agent can evaluate the *expected* free energy of the alternative policies she could select – that is, the free energy of future trajectories under each policy (i.e., based on predicted future outcomes, given the future states that would be expected under each policy). Therefore, decision-making will be approximately (Bayes) optimal if it operates by inferring (and enacting) the policy that minimizes expected free energy – and thereby maximizes evidence for the brain's internal model. Interestingly, expected free energy can be decomposed into terms reflecting uncertainty and prior preferences, respectively. This decomposition explains why agents that minimize expected free energy will first select exploratory policies that minimize uncertainty in a new environment (often called the "epistemic value" component of expected free energy). Once uncertainty is resolved, the agent then selects policies that exploit that environment to maximize her prior preferences (often called the "pragmatic value" component of expected free energy). The formal mathematical basis for AI has been detailed elsewhere (Friston et al., 2017a), and the reader is referred there for a full mathematical treatment (also see **Figure 2** for some additional detail).

When a generative model is formulated as a partially observable Markov decision process, active inference takes a particular form. Specifically, specifying a generative model in this context requires specifying the allowable policies, hidden states of the world (that the brain cannot directly observe but must infer), and observable outcomes, as well as a number of matrices that define the probabilistic relationships between these quantities (see **Figure 2**). The "A" matrix specifies which outcomes are generated by each combination of hidden states (i.e., a likelihood mapping indicating the probability that a particular set of outcomes would be observed given a particular hidden state). The "B" matrix encodes state transitions, specifying the probability that one hidden state will evolve into another over time. Some of these transitions are controlled by the agent, according to the policy that has been selected. The "D" matrix encodes prior expectations about the initial hidden state of the world. The "E" matrix encodes prior expectations about which policies will be chosen (e.g., frequently repeated habitual behaviors will have higher prior expectation values). Finally, the "C" matrix encodes prior preferences over outcomes. Outcomes and hidden states are generally factorized into multiple outcome *modalities* and hidden state *factors*. This means that the likelihood mapping (the "A" matrix) plays an important role in modeling the interactions among different hidden states at each level of a hierarchical model when generating the outcomes at the level below. One can think of each factor and modality as an independent group of competing states or observations within a given category. For example, one hidden state factor could be "birds," which includes competing interpretations of sensory input as corresponding to either hawks, parrots, or pigeons, whereas a separate factor could be "location," with competing representations of where a bird is in the sky. Similarly, one outcome modality could be size (e.g., is the bird big or small?) whereas another could be color (is the bird black, white, green, etc?).

As shown in the middle and right panels of **Figure 2**, active inference is also equipped with a neural process theory – a proposed manner in which neuronal circuits and their dynamics can invert generative models via a set of linked update equations that minimize prediction errors. In this neuronal implementation, the probability of neuronal firing in specific populations is associated with the expected probability of a state, whereas postsynaptic membrane potentials are associated with the logarithm of this probability. A softmax function acts as an activation function – transforming membrane potentials into firing rates. With this setup, postsynaptic depolarizations (driven by ascending signals) can be understood as prediction errors (free energy gradients) about hidden states – arising from linear mixtures in the firing rates of other neural populations. These prediction errors (postsynaptic currents) in turn drive membrane potential changes (and resulting firing rates). When predictions errors are minimized, postsynaptic influences no longer drive changes in activity (depolarizations and firing rates), corresponding to minimum free energy.

Via similar dynamics, predictions errors about outcomes (i.e., the deviation between preferred outcomes and those predicted under each policy) can also be computed and integrated (i.e., averaged) to evaluate the expected free energy (value) of each policy (i.e., underwriting selection of the policy that best minimizes these prediction errors). Dopamine dynamics also modulate policy selection, by encoding estimates of the expected uncertainty over policies – where greater expected uncertainty promotes less deterministic policy selection. Phasic dopamine responses correspond to updates in expected uncertainty over policies – which occur when there is a prediction error about expected free energy; that is, when there is a difference between the expected free energy of policies before and after a new observation.

Finally, and most centrally for the simulations we report below, learning in this theory corresponds to a form of synaptic plasticity remarkably similar to Hebbian coincidence-based learning mechanisms associated with empirically observed synaptic long-term potentiation and depression (LTP and LTD) processes (Brown et al., 2009). Here one can think of the strength of each synaptic connection as a parameter in one of the matrices described above. For example, the strength of one synapse could encode the amount of evidence a given observation provides for a given hidden state (i.e., an entry in the "A" matrix), whereas another synapse could encode the probability of a state at a later time given a state at an earlier time (i.e., an entry in the "B" matrix). Mathematically, the synaptic strengths correspond to Dirichlet parameters that increase in value in response to new observations. One can think of this process as adding counts to each matrix entry based on coincidences in pre- and post-synaptic activity. For example, if beliefs favor one hidden state, and this co-occurs with a specific observation, then the strength of the value in the "A" matrix encoding

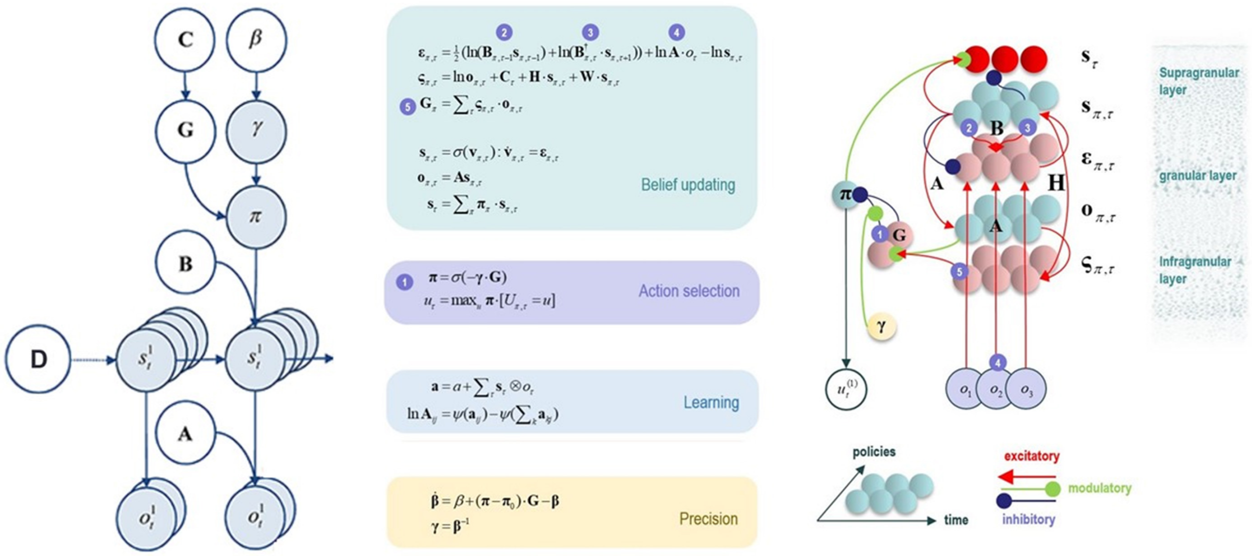


FIGURE 2 | This figure illustrates the mathematical framework of active inference and associated neural process theory used in the simulations described in this paper. **(Left)** Illustration of the Markov decision process formulation of active inference. The generative model is here depicted graphically, such that arrows indicate dependencies between variables. Here observations (o) depend on hidden states (s), as specified by the A matrix, and those states depend on both previous states (as specified by the B matrix, or the initial states specified by the D matrix) and the policies (π) selected by the agent. The probability of selecting a particular policy in turn depends on the expected free energy (G) of each policy with respect to the prior preferences (C) of the agent. The degree to which expected free energy influences policy selection is also modulated by a prior policy precision parameter (γ), which is in turn dependent on beta (β) – where higher values of beta promote more randomness in policy selection (i.e., less influence of the differences in expected free energy across policies). **(Middle/Right)** The differential equations in the **middle panel** approximate Bayesian belief updating within the graphical model depicted on the **left** via a gradient descent on free energy (F). The **right panel** also illustrates the proposed neural basis by which neurons making up cortical columns could implement these equations. The equations have been expressed in terms of two types of prediction errors. State prediction errors (ε) signal the difference between the (logarithms of) expected states (s) under each policy and time point and the corresponding predictions based upon outcomes/observations (A matrix) and the (preceding and subsequent) hidden states (B matrix, and, although not written, the D matrix for the initial hidden states at the first time point). These represent prior and likelihood terms, respectively – also marked as messages 2, 3, and 4, which are depicted as being passed between neural populations (colored balls) via particular synaptic connections in the **right panel** (note: the dot notation indicates transposed matrix multiplication within the update equations for prediction errors). These (prediction error) signals drive depolarization (v) in those neurons encoding hidden states (s), where the probability distribution over hidden states is then obtained via a softmax (normalized exponential) function (σ). Outcome prediction errors (ς) instead signal the difference between the (logarithms of) expected observations (o) and those predicted under prior preferences (C). This term additionally considers the expected ambiguity or conditional entropy (H) between states and outcomes as well as a novelty term (W) reflecting the degree to which beliefs about how states generate outcomes would change upon observing different possible state-outcome mappings (computed from the A matrix). This prediction error is weighted by the expected observations to evaluate the expected free energy (G) for each policy (π), conveyed via message 5. These policy-specific free energies are then integrated to give the policy expectations via a softmax function, conveyed through message 1. Actions at each time point (u) are then chosen out of the possible actions under each policy (U) weighted by the value (negative expected free energy) of each policy. In our simulations, the agent learned associations between hidden states and observations (A) via a process in which counts were accumulated (a) reflecting the number of times the agent observed a particular outcome when she believed that she occupied each possible hidden state. Although not displayed explicitly, learning prior expectations over initial hidden states (D) is similarly accomplished via accumulation of concentration parameters (d). These prior expectations reflect counts of how many times the agent believes she previously occupied each possible initial state. Concentration parameters are converted into expected log probabilities using digamma functions (ψ). As already stated, the **right panel** illustrates a possible neural implementation of the update equations in the **middle panel**. In this implementation, probability estimates have been associated with neuronal populations that are arranged to reproduce known intrinsic (within cortical area) connections. Red connections are excitatory, blue connections are inhibitory, and green connections are modulatory (i.e., involve a multiplication or weighting). These connections mediate the message passing associated with the equations in the **middle panel**. Cyan units correspond to expectations about hidden states and (future) outcomes under each policy, while red states indicate their Bayesian model averages (i.e., a "best guess" based on the average of the probability estimates for the states and outcomes across policies, weighted by the probability estimates for their associated policies). Pink units correspond to (state and outcome) prediction errors that are averaged to evaluate expected free energy and subsequent policy expectations (in the lower part of the network). This (neural) network formulation of belief updating means that connection strengths correspond to the parameters of the generative model described in the text. Learning then corresponds to changes in the synaptic connection strengths. Only exemplar connections are shown to avoid visual clutter. Furthermore, we have just shown neuronal populations encoding hidden states under two policies over three time points (i.e., two transitions), whereas in the task described in this paper there are greater number of allowable policies. For more information regarding the mathematics and processes illustrated in this figure, see Friston et al. (2016, 2017a,b,c).

the relationship between that state and that observation will increase. Counts also increase in similar fashion in the "D" matrix encoding prior beliefs about initial states (whenever a given hidden state is inferred at the start of a trial) as well as in the "B" matrix encoding beliefs about transition probabilities (whenever one specific state is followed by another). For a more detailed discussion, please see the legend for **Figure 2** and associated references.

In what follows, we describe how this type of generative model was specified to perform emotional state inference and emotion concept learning. We also present simulated neural responses based on the neural process theory described above.

A Model of Emotion Inference and Concept Learning

In this paper we focus on the second process in the TPM – affective response representation – in which a multifaceted affective response is generated and the ensuing (exteroceptive, proprioceptive, and interoceptive) outcomes are used to infer or represent the current emotional state. The basic idea is to equip the generative model with a space of emotion concepts (i.e., latent or hidden states) that generate the interoceptive, exteroceptive and proprioceptive consequences (at various levels of abstraction) of being in a particular emotional state. Inference under this model then corresponds to inferring that one of several possible emotion concepts is the best explanation for the data at hand (e.g., "my unpleasant feeling of increased heart rate and urge to run away must indicate that I am afraid to give this speech"). Crucially, to endow emotion concept inference with a form of mental action (Metzinger, 2017; Limanowski and Friston, 2018), we also included a state factor corresponding to *selective attention*. Transitions between attentional states were under control of the agent (i.e., "B" matrices were specified for all possible transitions between these states). The "A" matrix mapping emotion concepts to lower-level observations differed in each attentional state, such that precise information about each type of lower-level information was only available in one attentional state (e.g., the agent needed to transition into the "attention to valence" state to gain precise information regarding whether she was feeling pleasant or unpleasant, and so forth; see **Figure 3C**).

The incorporation of selective attention in emotional state inference and learning within our model was motivated by several factors. First, multiple psychotherapeutic modalities improve clients' understanding of their own emotions in just this way; that is, by having them selectively attend to and record the contexts, bodily sensations, thoughts, action tendencies, and behaviors during emotion-episodes (e.g., Hayes and Smith, 2005; Barlow et al., 2016). Second, low emotional awareness has been linked to biased attention in some clinical contexts (Lane et al., 2018). Third, related personality factors (e.g., biases toward "externally oriented thinking") are included in leading self-report measures of the related construct of alexithymia (Parker et al., 2003). Finally, emotion learning in childhood appears to involve parent-child interactions in which parents draw attention to (and label) bodily feelings and behaviors during a child's affective responses [e.g., see work on attunement, social referencing, and related aspects of emotional development (Mumme et al., 1996; Licata et al., 2016; Smith et al., 2018b)] – and the lack of such interactions hinders emotion learning (and mental state learning more generally; Colvert et al., 2008).

In our model, we used relatively high level "outcomes" (i.e., themselves standing in for lower-level representations) to summarize the products of belief updating at lower levels of a hierarchical model. These outcomes were domain-specific, covering interoceptive, proprioceptive and exteroceptive modalities. A full hierarchical model would consider lower levels, unpacked in terms of sensory modalities; however, the current model, comprising just two levels, is sufficient for our purposes. The bottom portion of **Figure 3A** (in gray) acknowledges the broad form that these lower-level outcomes would be expected to take. The full three-process model would also contain a higher level corresponding to conscious accessibility (for an explicit model and simulations of this higher level, see Smith et al., 2019b). This is indicated by the gray arrows at the top of **Figure 3A**.

Crucially, as mentioned above, attentional focus was treated as a (mental) action that determines the outcome modality or domain to which attention was selectively allocated. Effectively, the agent had to decide which lower-level representations to selectively attend to (i.e., which sequential attention policy to select) in order to figure out what emotional state she was in. Mathematically, this was implemented via interactions in the likelihood mapping – such that being in a particular attentional state selected one and only one precise mapping between the emotional state factor and the outcome information in question (see **Figure 3C**). Formally, this implementation of mental action or attentional focus is exactly the same used to model the exploration of a visual scene using overt eye movements (Mirza et al., 2016). However, on our interpretation, this epistemic foraging was entirely covert; hence mental action (c.f., the premotor theory of attention; (Rizzolatti et al., 1987; Smith and Schenk, 2012; Posner, 2016).

Figure 3 illustrates the resulting model. The first hidden state factor was a space of (exemplar) emotion concepts (SAD, AFRAID, ANGRY, and HAPPY). The second hidden state factor was attentional focus, and the "B" matrix for this second factor allowed state transitions to be controlled by the agent. The agent could choose to attend to three sources of bodily (interoceptive/proprioceptive) information, corresponding to affective valence (pleasant or unpleasant sensations), autonomic arousal (e.g., high or low heart rate), and motivated proprioceptive action tendencies (approach or avoid). The agent could also attend to two sources of exteroceptive information, including the perceived situation (involving social rejection or a crowded event) and subsequent beliefs about responsibility (attributing agency/blame to self or another). These different sources of information are based on a large literature within emotion research, indicating that they are jointly predictive of self-reported emotions and/or are important factors in affective processing (Russell, 2003; Siemer et al., 2007; Lindquist and Barrett, 2008; Scherer, 2009; Harmon-Jones et al., 2010; Barrett et al., 2011; Barrett, 2017).

Our choice of including valence in particular reflects the fact that our model deals with high levels of hierarchical processing (this choice also enables us to connect more fluently with current literature on emotion concept categories). In this paper, we are using labels like "unpleasant" as pre-emotional constructs. In other words, although affective in nature, we take concepts like "unpleasant" as contributing to elaborated emotional constructs during inference. Technically, valenced states provide evidence for emotional state inference at a higher level (e.g., pleasant sensations provide evidence that one is

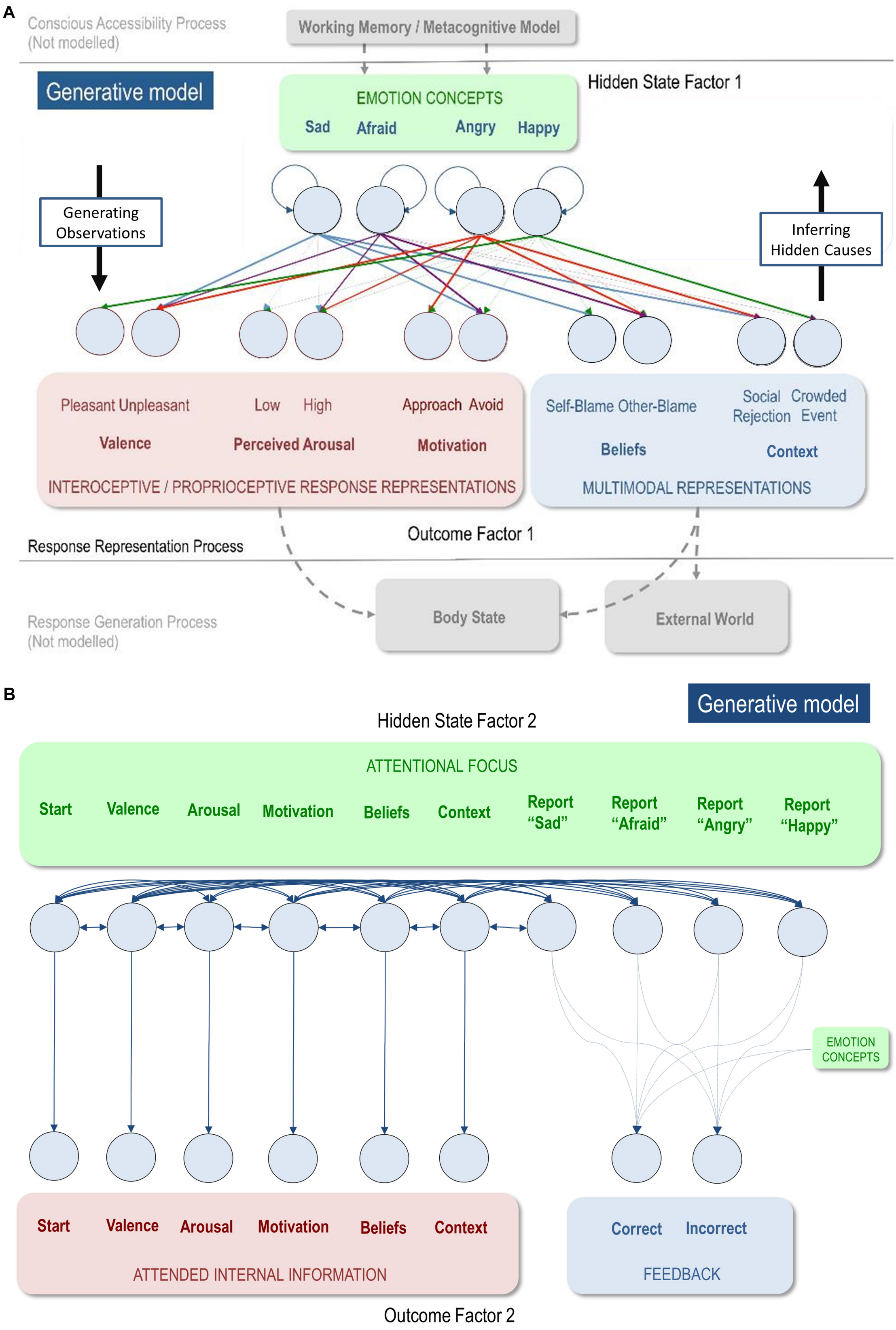


FIGURE 3 | Continued

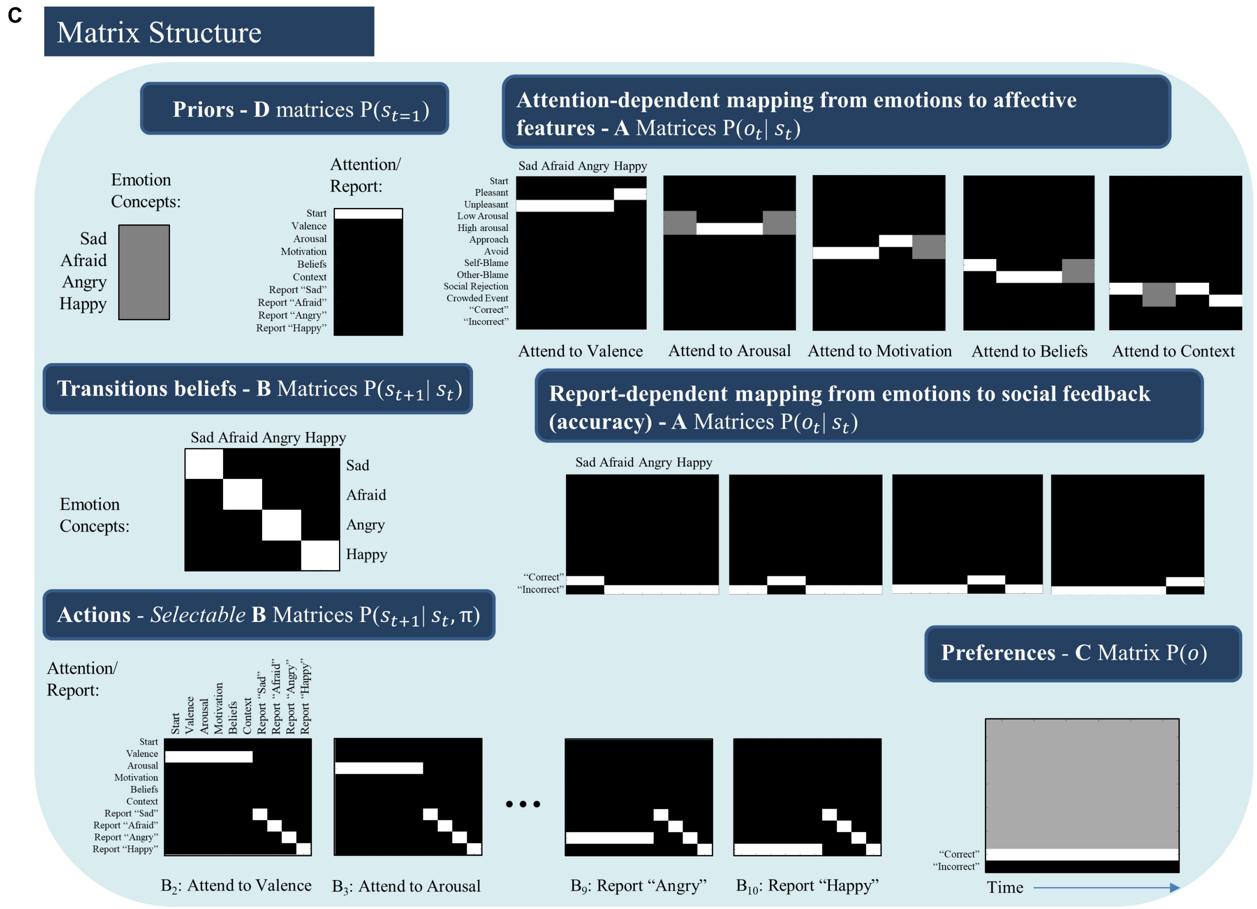


FIGURE 3 | (A) Displays the levels of hidden state factor 1 (emotion concepts) and their mapping to different lower-level representational contents (here modeled as outcomes). Each emotion concept generated different outcome patterns (see text for details), although some were more specific than others (e.g., HAPPY generated high and low arousal equally; denoted by thin dotted connections). The A-matrix encoding these mappings is shown in **(C)**. The B-matrix, also shown in **(C)**, was an identity mapping between emotion states, such that emotions were stable within trials. The precision of this matrix (i.e., implicit beliefs about emotional stability) could be adjusted via passing this matrix through a softmax function with different temperature parameter values. This model simulates the affective response representation process within the three-process model of emotion episodes (Smith et al., 2018b, 2019a; Smith, 2019). Black arrows on the right and left indicate the direction by which hidden causes generate observations (the generative process) and the direction of inference in which observations are used to infer their hidden causes (emotion concepts) using the agent's generative model. The gray arrows/boxes at the bottom and top of the figure denote other processes within the three-process model (i.e., affective response generation and conscious accessibility) that are not explicitly modeled in the current work (for simulations of the affective response generation process, see Hesp et al., 2019, and for simulations of the conscious accessibility process, see Smith et al., 2019b). **(B)** Displays the levels of hidden state factor 2 (focus of attention) and its mapping to representational outcomes. Each focus of attention mapped deterministically (the A-matrix was a fully precise identity matrix) to a "location" (i.e., an internal source of information) at which different representational outcomes could be observed. Multiple B-matrices, depicted in **Figure 3C**, provided controllable transitions (i.e., actions) such that the agent could choose to shift her attention from one internal representation to another to facilitate inference. The agent always began a trial in the "start" attentional state, which provided no informative observations. The final attentional shift in the trial was toward a (proprioceptive) motor response to report an emotion (i.e., at whatever point in the trial the agent became sufficiently confident, at which point the state could not change until the end of the trial), which was either correct or incorrect. The agent preferred (expected) to be correct and not to be incorrect. Because policies (i.e., sequences of implicit attentional shifts and subsequent explicit reports) were selected to minimize expected free energy, emotional state inference under this model entails a sampling of salient representational outcomes and subsequent report – under the prior preference that the report would elicit the outcome of "correct" social feedback. In other words, policy selection was initially dominated by the epistemic value part of expected free energy (driving the agent to gather information about her emotional state); then, as certainty increased, the pragmatic value part of expected free energy gradually began to dominate (driving the agent to report her emotional state). **(C)** Displays select matrices defining the generative model (lighter colors indicate higher probabilities). "D" matrices indicate a flat distribution over initial emotional states and a strong belief that the attentional state will begin in the "start" state. "A" matrices indicate the observations (rows) that would be generated by each emotional state (columns) depending on the current state of attention, as well as the social feedback (accuracy information) that would be generated under each emotional state depending on chosen self-reports. "B" matrices indicate that emotional states are stable within a trial (i.e., states transition only to themselves such that the state transition matrix is an identity matrix) and that the agent can choose to shift attention to each modality of lower-level information (i.e., by selecting the transitions encoded by "B" matrices 2–6), or report her emotional state (i.e., by selecting the transitions entailed by "B" matrices 7–10), at which point she could no longer leave that state. Only the first and last two possible attentional shifting actions ("B" matrices) are shown due to space constraints (note: the first "B" matrix for this factor corresponds to remaining in the starting state and is not shown); "B" matrices 4–8 take identical form, but with the row vector (1 1 1 1 1 1) spanning the first six columns progressively shifted downward with each matrix, indicating the ability to shift from each attentional state to the attentional state corresponding to that row. The "C" matrix indicates a preference for "correct" social feedback and an aversion to "incorrect" social feedback.

feeling a positive emotion like excitement, joy, or contentment, whereas unpleasant sensations provide evidence that one may be feeling a negative emotion such as sadness, fear, or anger). Based on previous work (Joffily and Coricelli, 2013; Clark et al., 2018; Hesp et al., 2019), we might expect valence to correspond to changes in the precision/confidence associated with lower-level visceromotor and skeletomotor policy selection, or to related internal estimates that can act as indicators of success in uncertainty resolution; see Joffily and Coricelli (2013), de Berker et al. (2016), Peters et al. (2017), Clark et al. (2018). Put another way, feeling good may correspond to high confidence in one's model of how to act, whereas feeling bad may reflect the opposite. Explicitly modeling these lower-level dynamics in a deep temporal model will be the focus of future work.

In the simulations we report here, there were 6 time points in each epoch or trial of emotion inference. At the first time point, the agent always began in an uninformative initial state of attentional focus (the "start" state). The agent's task was to choose what to attend to, and in which order, to infer her most likely emotional state. When she became sufficiently confident, she could choose to respond (i.e., reporting that she felt sad, afraid, angry, or happy). In these simulations the agent selected "shallow" one-step policies, such that she could choose what to attend to next – to gain the most information. Given the number of time points, the agent could choose to attend to up to four of the five possible sources of lower-level information before reporting her beliefs about her emotional state. The "A" matrix mapping attentional focus to attended outcomes was an identity matrix, such that the agent always knew which lower-level information she was currently attending to. This may be thought of as analogous to the proprioceptive feedback consequent on a motor action.

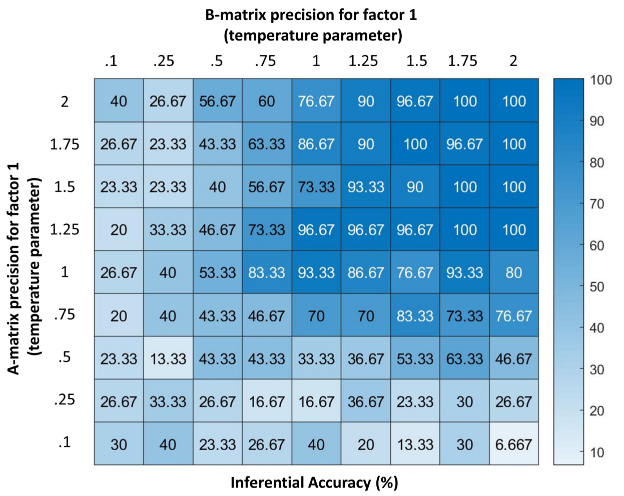
The "B" matrix for hidden emotional states was also an identity matrix, reflecting the belief that emotional states are stable within a trial (i.e., if you start out feeling sad, then you will remain sad throughout the trial). This sort of probability transition matrix in the generative model allows evidence to be accumulated for one state or another over time; here, the emotion concept that provides the best explanation for actively attended evidence in the outcome modalities. The "A" matrix – mapping emotion concepts to outcomes – was constructed such that certain outcome combinations were more consistent with certain emotional states than others: SAD was probabilistically associated with unpleasant valence, either low or high arousal (e.g., lying in bed lethargically vs. intensely crying), avoidance motivation, social rejection, and self-attribution (i.e., self-blame). AFRAID generated unpleasant valence, high arousal, avoidance, other-blame (c.f., fear often being associated with its perceived external cause), and either social rejection or a crowded event (e.g., fear of a life without friends vs. panic in crowded spaces). ANGRY generated unpleasant valence, high arousal, approach, social rejection, and other-blame outcomes. Finally, HAPPY generated pleasant valence, either low or high arousal, either approach or avoidance (e.g., feeling excited to wake up and go to work vs. feeling content in bed and not wanting to go to work), and a crowded event (e.g., having fun at a concert). Because HAPPY does not have strong conceptual links to blame, we defined a flat mapping between HAPPY and blame, such that either type of blame provided no evidence for or against being happy. Although this mapping from emotional states to outcomes has some face validity, it should not be taken too seriously. It was chosen primarily to capture the ambiguous and overlapping correlates of emotion concepts, and to highlight why adaptive emotional state inference and emotion concept learning can represent difficult problems.

If the "A" matrix encoding state-outcome relationships was completely precise (i.e., if the contingencies above were deterministic as opposed to probabilistic), sufficient information could be gathered through (at most) three attentional shifts; but this becomes more difficult when probabilistic mappings are imprecise (i.e., as they more plausibly are in the real world). **Figure 4** illustrates this by showing how the synthetic subject's confidence about her state decreases as the precision of the mapping between emotional states and outcomes decreases (we measured confidence here in terms of the accuracy of responding in relation to the same setup with infinite precision). Changes in precision were implemented via a temperature parameter of a softmax function applied to a fully precise version of likelihood mappings between emotion concepts and the 5 types of lower-level information that the agent could attend to (where a higher value indicates higher precision). For a more technical account of this type of manipulation, please see Parr and Friston (2017b).

Figure 4 additionally demonstrates how reporting confidence decreases with decreasing precision of the "B" matrix encoding emotional state transitions, where low precision corresponds to the belief that emotional states are unstable over time. Interestingly, these results suggest that expectations about emotional instability would reduce the ability to understand or infer one's own emotions. From a Bayesian perspective, this result is very sensible: if we are unable to use past beliefs to contextualize the present, it is much harder to accumulate evidence in favor of one hypothesis about emotional state relative to another. Under moderate levels of precision, our numerical analysis demonstrates that the model can conceptualize the multimodal affective responses it perceives with high accuracy.

Figure 5 illustrates a range of simulation results from an example trial under moderately high levels of "A" and "B" matrix precision (temperature parameter = 2 for each). The upper left plot shows the sequence of (inferred) attentional shifts (note: darker colors indicate higher probability beliefs of the agent, and cyan dots indicate the true states). In this trial, the agent chose a policy in which it attended to valence (observing "unpleasant"), then beliefs (observing "other-blame"), then action (observing "approach"), at which time she became sufficiently confident and chose to report that she was angry. The lower left plot displays the agent's *posterior* beliefs at the end of the trial about her emotional state at each timepoint in the trial, in this case inferring that she had been (and still was) angry. Note that this reflects retrospective inference, and not the agent's beliefs at each timepoint. The lower right and upper right plots display simulated neural responses (based on the neural process theory that accompanies this form of active inference; Friston et al., 2017a), in terms of single-neuron firing rates (raster plots) and local field potentials, respectively. The simulated firing rates in the lower right plot illustrate that the agent's confidence that she was angry increased gradually with each new observation.

The simulations presented in **Figures 4**, **5** make some cardinal points. First, it is fairly straightforward to simulate emotion processing in terms of emotional state inference. This rests upon a particular sort of generative model that can generate outcomes in multiple modalities. The recognition of an emotional state corresponds to the inversion of such models – and therefore necessarily entails multimodal integration. In other words, successfully disambiguating the most likely emotional state here requires consideration of the specific multimodal patterns of experience (i.e., incorporating interoceptive, exteroceptive, and proprioceptive sensations) that would be expected under each emotional state. We have also seen that this form of belief updating – or evidence accumulation – depends sensitively on what sort of evidence is actively attended. This equips the model of emotion concept representation with a form of mental action, which speaks to a tight link between emotion processing and attention to various sources of evidence from within the body – and beyond. Choices to shift attention vs. to self-report are, respectively, driven by the epistemic and pragmatic value of each allowable policy, such that pragmatic value gradually comes to drive the selection of reporting policies as the expected information gain of further attentional shifts decreases. The physiological plausibility of this emotion inference process has been briefly considered in terms of simulated responses. In the next section, we turn to a more specific construct validation, using empirical phenomenology from neurodevelopmental studies of emotion.



	.1	.25	.5	.75	1	1.25	1.5	1.75	2
2	40	26.67	56.67	60	76.67	90	96.67	100	100
1.75	26.67	23.33	43.33	63.33	86.67	90	100	96.67	100
1.5	23.33	23.33	40	56.67	73.33	93.33	90	100	100
1.25	20	33.33	46.67	73.33	96.67	96.67	96.67	100	100
1	26.67	40	53.33	83.33	93.33	86.67	76.67	93.33	80
.75	20	40	43.33	46.67	70	70	83.33	73.33	76.67
.5	23.33	13.33	43.33	43.33	33.33	36.67	53.33	63.33	46.67
.25	26.67	33.33	26.67	16.67	16.67	36.67	23.33	30	26.67
.1	30	40	23.33	26.67	40	20	13.33	30	6.667

FIGURE 4 | Displays the accuracy of the model (percentage of correct inferences over 30 trials) under different levels of precision for two parameters (denoted by temperature values for a softmax function controlling the specificity of the A and B matrices for hidden state factor 1; higher values indicate higher precision). As can be seen, the model performs with high accuracy at moderate levels of precision. However, its ability to infer its own emotions becomes very poor if the precision of either matrix becomes highly imprecise. Accuracy here is defined in relation to the response obtained from an agent with infinite precision – and can be taken as a behavioral measure of the quality of belief updating about emotional states. These results illustrate how emotion concepts could be successfully inferred despite variability in lower-level observations (e.g., contexts, arousal levels), as would be expected under constructivist theories of emotion (Barrett, 2017); however, they also demonstrate limits in variability, beyond which self-focused emotion recognition would begin to fail.

SIMULATING THE INFLUENCE OF EARLY EXPERIENCE ON EMOTIONAL STATE INFERENCE AND EMOTION CONCEPT LEARNING

Having confirmed that our model could successfully infer emotional states – if equipped with emotion concepts – we are now in a position to examine emotion concept learning. Specifically, we investigated the conditions under which emotion concepts could be acquired successfully and the conditions under which this type of emotion learning and inference fails.

Can Emotion Concepts Be Learned in Childhood?

The first question we asked was whether our model could learn about emotions, if it started out with no prior beliefs about how emotions structure its experience. To answer this question, we first ran the model's "A" matrix (mapping emotion concepts to attended outcome information) through a softmax function with a temperature parameter of 0, creating a fully imprecise likelihood mapping. This means that each hidden emotional state predicted all outcomes equally (effectively, none of the hidden states within the emotion factor had any conceptual content). Then we generated 200 sets of observations (i.e., 50 for each emotion concept, evenly interleaved) based on the probabilistic state-outcome mappings encoded in the model described above (i.e., the "generative process"). That is, 50 interleaved learning trials for each emotion were generated by probabilistically sampling from a moderately precise version of the "A" matrix distribution depicted in **Figure 3C** (i.e., temperature parameter = 2). This resulted in 50 sets of observations consistent with the probabilistic mappings for each emotion (e.g., this entailed that roughly 50% of HAPPY trials involved observations of low vs. high arousal, whereas only roughly 1% of HAPPY trials involved the observation of social rejection, etc.). After the 200 learning trials, we then examined the changes in the model's reporting accuracy over time. This meant that the agent, who began with no emotion knowledge (i.e., a fully uninformative "A" matrix), observed patterns of observations consistent with each emotion (as specified above) at 50 timepoints spread out across the 200 trials and needed to learn these associations (i.e., learn the appropriate "A" matrix mapping). This analysis was repeated at several levels of outcome ("A" matrix) and transition ("B" matrix) precision in the generative process – to explore how changes in the predictability or consistency of observed outcome patterns affected the model's ability to learn. In this model, learning was implemented through updating (concentration) parameters for the model's "A" matrix after each trial. The model could also learn prior expectations

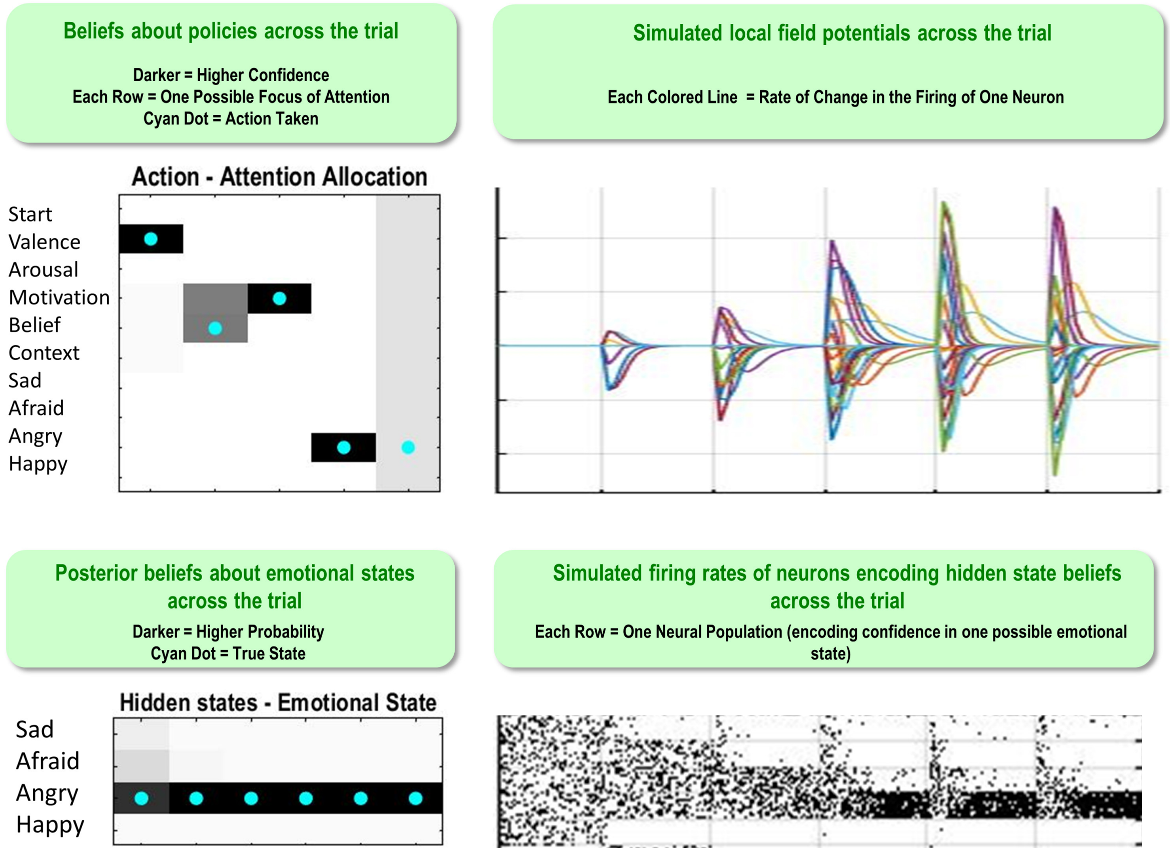


FIGURE 5 | This figure illustrates a range of simulation results from one representative trial under high levels of A- and B-matrix precision (temperature parameter = 2 for each matrix). **(Upper left)** Displays a sequence of chosen attentional shifts. The agent here chose a policy in which it attended to valence, then beliefs, then action motivation, and then chose to report that it was angry. Darker colors indicate higher confidence (probability estimates) in the model about its actions, whereas the cyan dot indicates the true action. **(Bottom left)** Displays the agent's *posterior* beliefs about her emotional state across the trial. These posterior beliefs indicate that, at the end of the trial, the agent retrospectively inferred (correctly) that she was angry throughout the whole trial (i.e., despite not being aware of this until the fourth timestep, as indicated by the firing rates encoding confidence updates over time in the **bottom right**). **(Lower right/Upper right)** Plots simulated neural responses in terms of single-neuron firing rates (raster plots) and local field potentials (rates of change in neural firing), respectively. Here each neuron's activity encodes the probability of occupying a particular hidden state at a particular point in time during the trial (based on the neural process theory depicted in **Figure 2**; see Friston et al., 2017a).

for being in different emotional states, based on updating concentration parameters for its "D" matrix after each trial (i.e., the emotional state it started in). For details of these free energy minimizing learning processes, please see Friston et al. (2016).

We observed that the model could successfully reach 100% accuracy (with minor fluctuation) when both outcome and transition precisions in the generative process were moderately high (i.e., when the temperature parameters for the "A" and "B" matrix of the generative process were 2). The top panel in **Figure 6** illustrates this by plotting the percentage accuracy across all emotions during learning over 200 trials (in bins of 10 trials), and for each emotion (in bins of 5). As can be seen, the model steadily approaches 100% accuracy across trials. The middle and lower panels of **Figure 6** illustrate the analogous results when outcome precision and transition precision were lowered, respectively. The precision values chosen for these illustrations ("A" precision = 1.5, "B" precision = 0.5) represent observed "tipping points" at which learning began to fail (i.e., at progressively lower precision values learning performance steadily approached 0% accuracy). As can be seen, lower precision in either the stability of emotions over time or the consistency between observations and emotional states confounded learning. Overall, these findings provide a proof of principle that this sort of model can learn emotion concepts, if provided with a representative and fairly consistent sample of experiences in its "childhood."

Can a New Emotion Concept Be Learned in Adulthood?

We then asked whether a new emotion could be learned later, after others had already been acquired (e.g., as in adulthood). To answer this question, we again initialized the model with a fully imprecise "A" matrix (temperature parameter = 0) and set the precision of the "A" and "B" matrices of the generative process to the levels at which "childhood" learning was successful (i.e.,

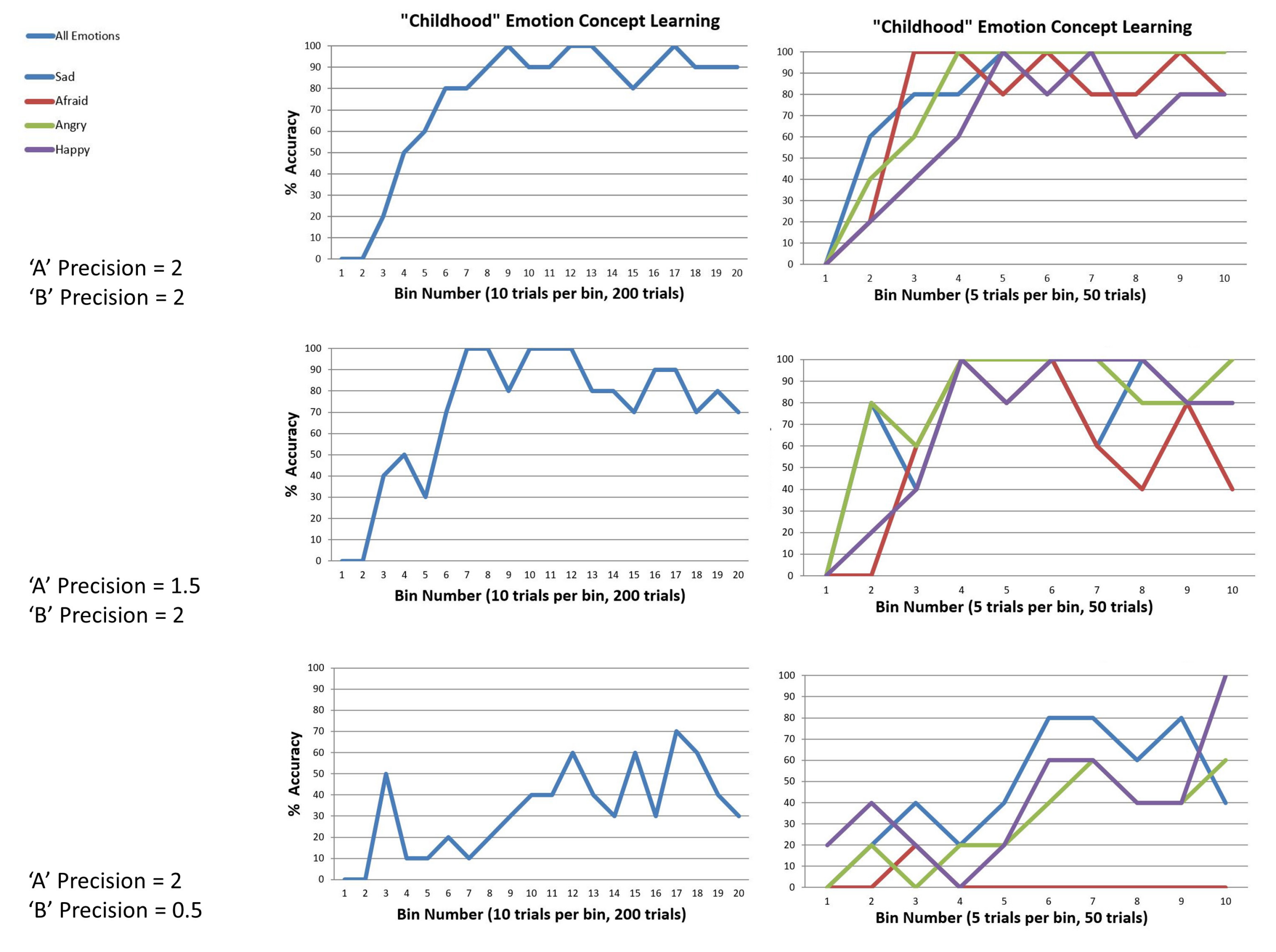


FIGURE 6 | This figure illustrates simulated "childhood" emotion concept learning, in which the agent started out with no emotion concept knowledge (a uniform likelihood mapping from emotional states to outcomes) and needed to learn the correct likelihood mapping over 50 interleaved observations of the outcome patterns associated with each emotion concept (200 trials total). **Left panels** show changes in accuracy over time in 10-trial bins, at different levels of outcome pattern precision (i.e., "A" and "B" matrix precision, reflecting the consistency in the mapping between emotions and outcomes and the stability of emotional states over time, respectively). **Right panels** show the corresponding results for each emotion separately (5-trial bins). **(Top)** with moderately high precision (temperature parameter = 2 for both "A" and "B" matrices), learning was successful. (**Middle** and **Bottom**) with reduced "A" precision or "B" precision (respectively), learning began to fail.

temperature parameter = 2 for each). We then exposed the model to 150 observations that only contained the outcome patterns associated three of the four emotions (50 for each emotion, evenly interleaved). We again allowed the model to accumulate experience in the form of concentration parameters for its "D" matrix – allowing it to learn strong expectations for the emotional states it repeatedly inferred it was in. After these initial 150 trials, we then exposed the model to 200 further trials – using the outcome patterns under all four emotions (50 for each emotion, evenly interleaved). We then asked whether the emotion that was not initially necessary to explain outcomes could be acquired later, when circumstances change.

We first observed that, irrespective of which three emotions were initially presented, accuracy was high by the end of the initial 150 trials (i.e., between 80–100% accuracy for each of the three emotion concepts learned). The upper and middle panels of **Figure 7** illustrate the accuracy over the subsequent 200 trials as the new emotion was learned. As can be seen in the upper left and right sections of **Figure 7**, ANGRY and SAD were both successfully learned. Interestingly, performance for the other emotions appeared to temporarily drop and then increase again as the new emotion concept was acquired (a type of temporary retroactive interference).

The middle left section of **Figure 7** demonstrates that HAPPY could also be successfully learned; however, it appeared to interfere with prior learning for SAD. Upon further inspection, it appeared that SAD may not have been fully acquired in the first 150 trials (only reaching 80% accuracy near the end). We therefore chose to examine whether an "extended" or "emotionally enriched" childhood might prevent this interference, by increasing the initial learning trial number from 150 to 225 (75 interleaved exposures to SAD, ANGRY, and AFRAID outcome patterns). As can be seen in the lower left panel of **Figure 7**, HAPPY was quickly acquired in the

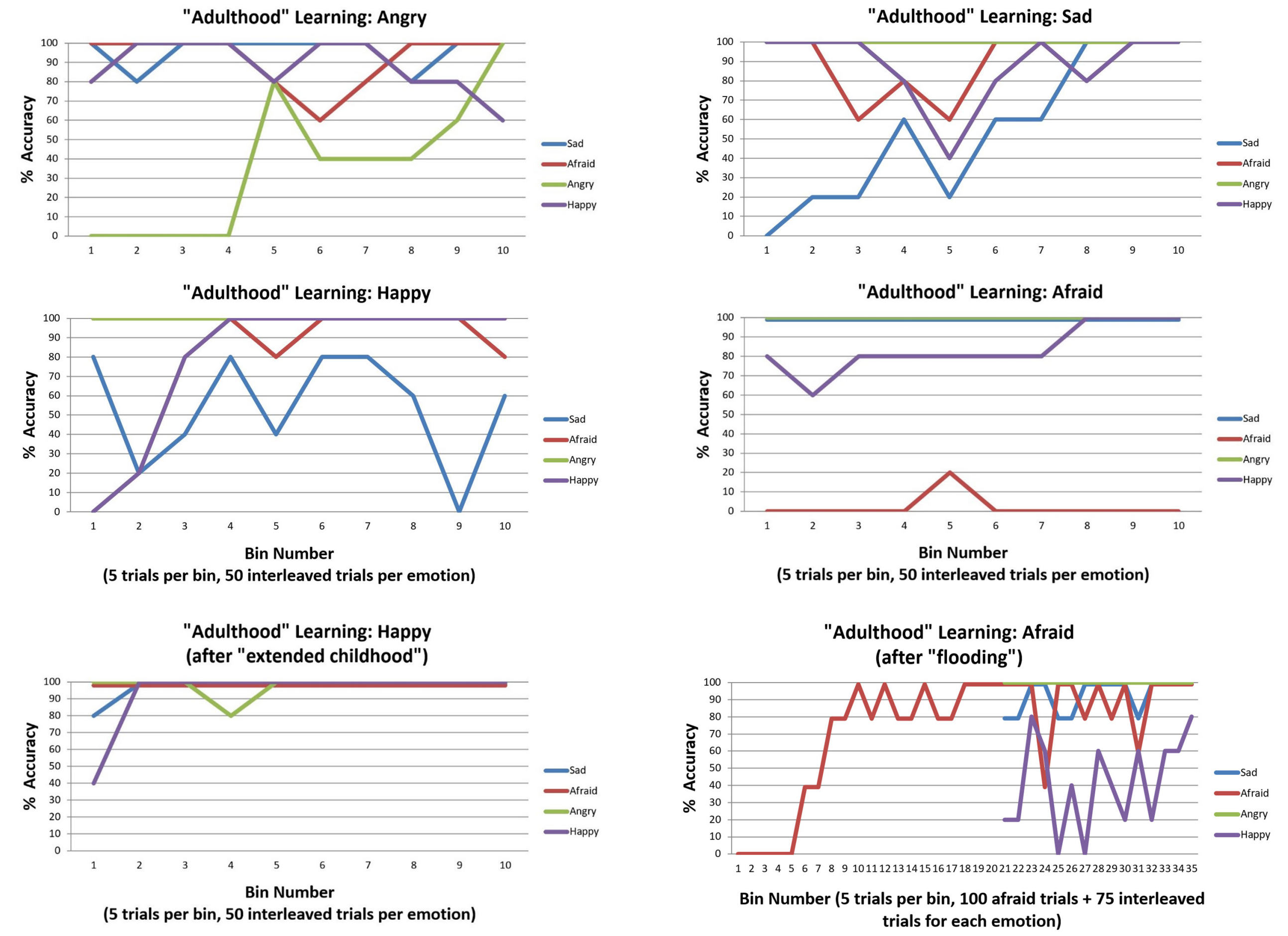


FIGURE 7 | This figure illustrates emotion concept learning in "adulthood", where three emotion concepts had already been learned (in 150 previous trials not illustrated) and a fourth was now needed to explain patterns of outcomes (over 200 trials, 50 of which involved the new emotion in an interleaved sequence). The **top** four plots illustrate learning for each of the four emotion concepts. ANGRY and SAD were successfully acquired. HAPPY was also successfully acquired, but interfered with previous SAD concept learning. As shown in the **bottom left**, this was ameliorated by providing the model with more "childhood" learning trials before the new emotion was introduced. AFRAID was not successfully acquired. This was (partially) ameliorated by first flooding the model with repeated observations of AFRAID-consistent outcomes before interleaving it with the other emotions, as displayed in the **bottom right**. All simulations were carried out with moderately high levels of precision (temperature parameter = 2 for both "A" and "B" matrices).

subsequent 200 trials under these conditions, without interfering with previous learning.

Somewhat surprisingly, the model was unable to acquire the AFRAID concept in its "adulthood" (**Figure 7**, middle right). To better understand this, for each trial we computed the expected evidence for each state, under the distribution of outcomes expected under the generative process (using bar notation to distinguish the process from the model) given a particular state, $(E_{\bar{P}(o|s)}[P(o|s)])$. This was based on the reasoning that, if we treat the different emotional states as alternative models to explain the data, then the likelihood of data given states is equivalent to the evidence for a given state. **Figure 8A** plots the log transform of this expected evidence for each established emotion concept expected under the distribution of outcomes that would be generated if the "real" emotional state were AFRAID (right panel) and contrasts it with the analogous plot for HAPPY (left panel), which was a more easily acquired concept [we took the logarithm of this expected evidence (or likelihood) to emphasize the lower evidence values; note that higher (less negative) values correspond to greater evidence in these plots].

As can be seen, in the case of HAPPY the three previously acquired emotion concepts had a relatively low ability to account for all observations, and so HAPPY was a useful construct in providing more accurate explanations of observed outcomes. In contrast, when learning AFRAID the model was already confident in its ability to explain its observations (i.e., the other concepts already had much higher evidence than in the case of HAPPY), and the "AFRAID" outcome pattern also provided moderate evidence for ANGRY and SAD (i.e., the outcome patterns between AFRAID and these other emotion concepts had considerable overlap). In **Figure 8B**, we illustrate the "A" and "D" matrix values the agent had learned after the total 350 trials (i.e., 150 + 200, as described above), and an exemplar trial in which the agent mistook fear for anger (which was

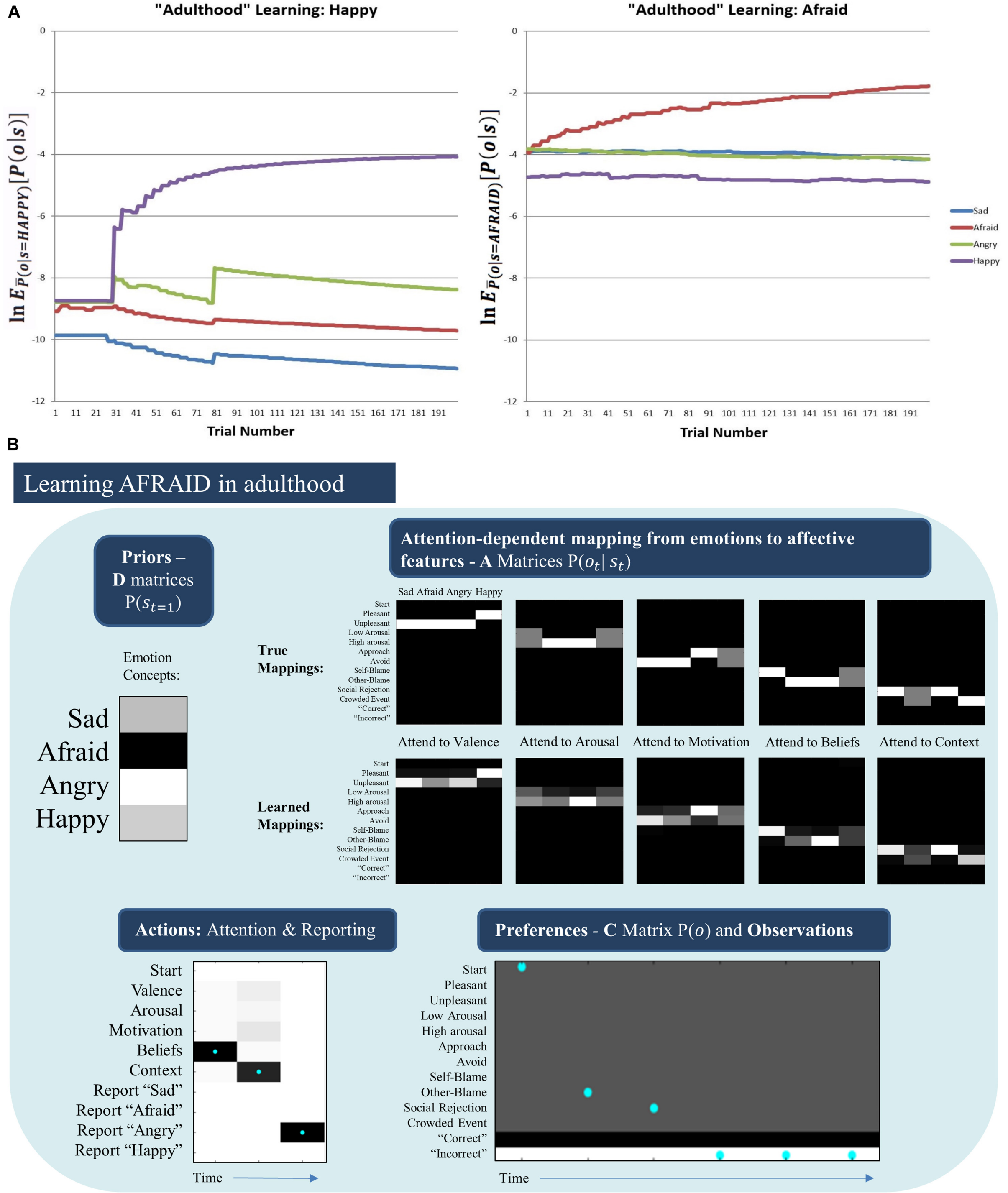


FIGURE 8 | (A) (Left) The log expected probability of outcomes under the HAPPY outcome distribution, given each of the four emotion concepts in the agent's model (i.e., given the "A" matrix it had learned) at each of 200 learning trials, $E_{\bar{P}(o|s=HAPPY)}[P(o|s)]$ (note: the bar over P indicates the generative process distribution). (Right) The analogous results for the AFRAID outcome pattern. These plots illustrate that HAPPY may have been more easily acquired than AFRAID because the agent was less confident in the explanatory power of its current conceptual repertoire when it began to observe the HAPPY outcome pattern than when it began to observe the AFRAID outcome pattern. This also shows that the AFRAID outcome pattern provided some evidence for SAD and ANGRY in the agent's model, likely due to outcome pattern overlap between AFRAID and these other two concepts. **(B)** Illustrates the priors (D) and likelihoods (A) learned after 50 observations of SAD, ANGRY, and HAPPY (150 trials total) followed by 200 trials in which SAD, AFRAID, ANGRY, and HAPPY were each presented 50 times. As can be seen, while the "A" matrices are fairly well learned, the agent acquired a strong prior expectation for ANGRY. As shown in the example trial in the bottom left, this led the agent to "jump to conclusions" and report ANGER on AFRAID trials after making 2 observations consistent with both ANGRY and AFRAID (other-blame and social rejection, as shown in the bottom right). In this case, the agent would have needed to attend to her motivated actions (avoid) to correctly infer AFRAID.

the most common confusion). As can be seen, while the "A" matrix mappings were learned fairly well, the agent had learned a strong prior expectation for ANGRY in comparison to its expectation for AFRAID. In the example trial the agent first attends to beliefs (observing other-blame) and then to the context (observing social rejection). These observations are consistent with both ANGRY and AFRAID; however, social rejection is more uniquely associated with ANGRY (i.e., AFRAID is also associated with crowded events, while ANGER is not). Combined with the higher prior expectation for ANGRY, the agent "jumps to conclusions" and becomes sufficiently confident to report ANGER (at which point she receives "incorrect" social feedback). Here, correct inference of AFRAID (i.e., disambiguating AFRAID from ANGRY) would have required that the agent also attend to her action tendencies (where she would have observed avoidance motivation) before deciding which emotion to report.

In this context, greater evidence for an unexplained outcome pattern would be required to "convince" the agent that her currently acquired concepts were not sufficient and that further information gathering (i.e., a greater number of attentional shifts) was necessary before becoming sufficiently confident to report her emotions. Based on this insight, we examined ways in which the model could be given stronger evidence that its current conceptual repertoire was insufficient to account for its observations. We first observed that we could improve model performance by "flooding" the model with an extended pattern of only AFRAID-consistent outcomes (i.e., 100 trials in a row), prior to reintroducing the other emotions in an interleaved fashion. As can be seen in **Figure 7** (bottom right), this led to successful acquisition of AFRAID. However, it temporarily interfered with previous learning of the HAPPY concept. We also observed that by instead increasing the number of AFRAID learning trials from 200 to 600, the model eventually increased its accuracy to between 40 and 80% across the last 10 bins (last 50 trials) – indicating that learning could occur, but at a much slower rate.

Overall, these results confirmed that a new emotion concept could be learned in synthetic "adulthood," as may occur, for example, in psycho-educational interventions during psychotherapy. However, these results also demonstrate that this type of learning can be more difficult. These results therefore suggest a kind of "sensitive period" early in life where emotion concepts may be more easily acquired.

Can Maladaptive Early Experiences Bias Emotion Conceptualization?

The final question we asked was whether unfortunate early experiences could hinder our agent's ability to adaptively infer and/or learn about emotions. Based on the three-process model (Smith et al., 2018b), it has previously been suggested that at least two mechanisms could bring this about:

1. Impoverished early experiences (i.e., not being exposed to the different patterns of observations that would facilitate emotion concept learning).
2. Having early experiences that reinforce maladaptive cognitive habits (e.g., selective attention biases), which can hinder adaptive inference (if the concepts have been acquired) and learning (if the concepts have not yet been acquired).

We chose to examine both of these possibilities below.

Non-representative Early Emotional Experiences

To examine the first mechanism (involving the maladaptive influence of "unrepresentative" early experiences), we used the same learning procedure and parameters described in the previous sections. In this case, however, we exposed the agent to 200 outcomes generated by a generative process where one emotion was experienced 50 times more often than others. Specifically, we examined the cases of a childhood filled with either chronic fear/threat or chronic sadness, as a potential means of simulating the effects of continual childhood abuse or neglect (the sadness simulations might also be relevant to chronic depression over several years). We then examined the model's ability to learn to infer new emotions in a subsequent 200 trials.

In general, we observed that primarily experiencing fear or sadness during childhood (which could also be thought of as undifferentiated in the sense that they could not be contrasted with other emotions) led the agent to have notable difficulties in learning new emotions later in life. These results were variable upon repeated simulations with different emotions (e.g., verbal reporting continually fluctuated between high and low levels of accuracy for some emotions, while accuracy remained near 0% for others, while yet others were well acquired). For example, in one representative simulation, in which the agent primarily experienced fear during childhood, reporting accuracy continually varied for HAPPY (45% accuracy in the final 20 trials), remained at 0% for ANGRY, remained at 100% for AFRAID, and was stable at 100% for SAD). Whereas primarily experiencing sadness in childhood during a representative simulation led to 0% accuracy for HAPPY, continually varying accuracy for ANGRY and AFRAID (55% accuracy in the final 20 trials for each), and stable high accuracy for SAD (95% accuracy in final 20 trials). Similar patterns of (highly variable) results were observed when performing the same simulations with the other two emotions.

Unlike the results shown in **Figure 8B** – in which likelihood mappings were fairly well acquired (and precise prior expectations for specific emotions hindered correct inference) – poor performance was here explained primarily by poorly acquired likelihood mappings (i.e., the content of the other emotions concepts was often not learned). **Figure 9** illustrates this by presenting the "A" matrices learned by the agent after childhoods dominated by either fear or sadness. As can be seen there, the likelihood mappings do not strongly resemble the true mappings within the generative process. These results in general support the notion that having unrepresentative or insufficiently diverse early emotional experiences could hinder later learning.

Maladaptive Attention Biases

To examine the second mechanism proposed by the three-process model (involving maladaptive patterns in habitual attention allocation), we equipped the model's "E" matrix with high prior expectations over specific policies, which meant that it was 50 times more likely to attend to some information and

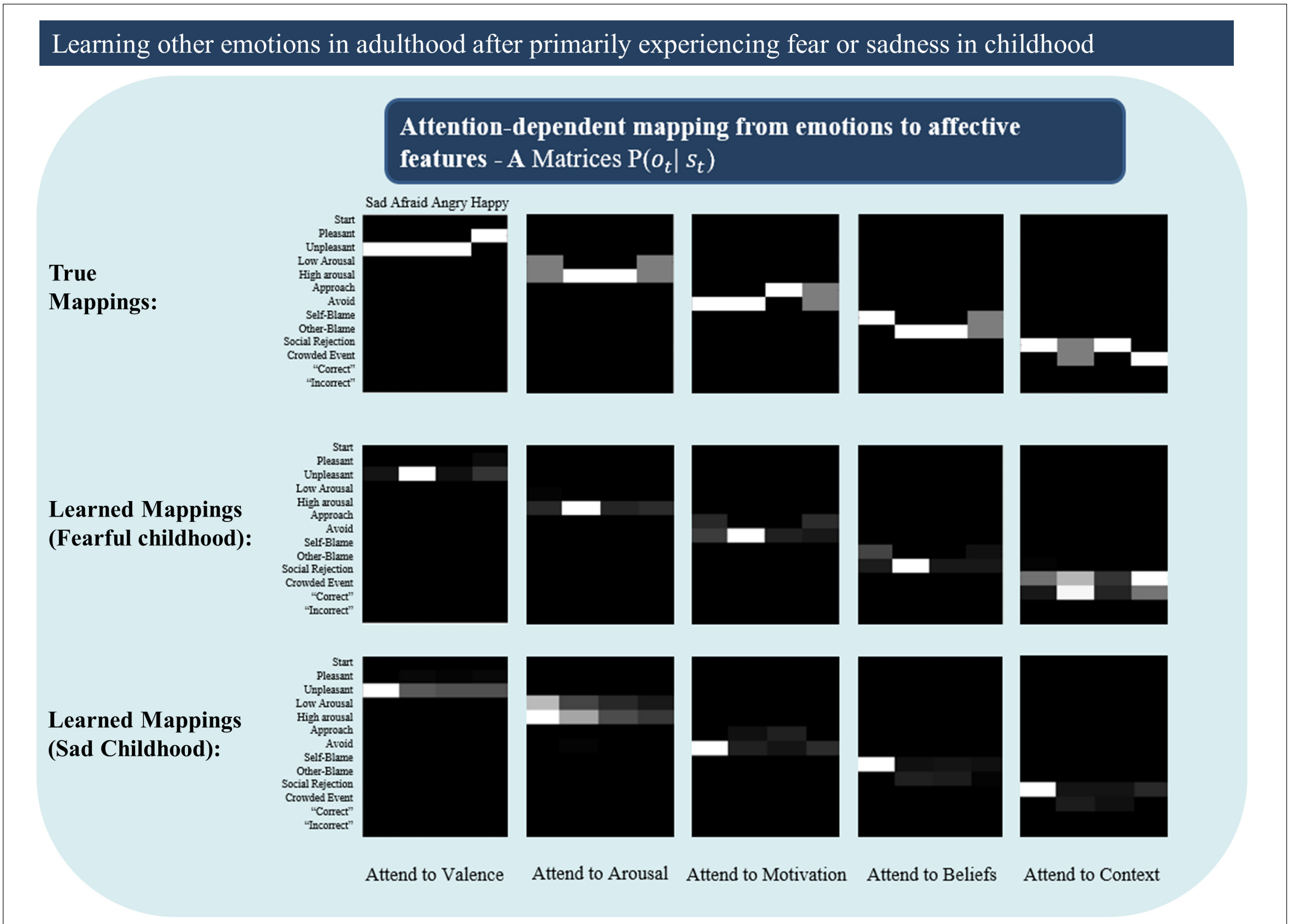


FIGURE 9 | Illustrates poorly learned emotion concepts (likelihood mappings or "A" matrices) in adulthood due to a childhood primarily characterized by either fear or sadness. This could be thought of as simulating early adversity involving continual abuse or neglect, or perhaps cases of chronic depression. See text for more details.

not to other information. This included: (i) an "external attention bias," where the agent had a strong habit of focusing on external stimuli (context) and its beliefs about self- and other-blame; (ii) an "internal attention bias," where the agent had a strong habit of only attending to valence and arousal; and (iii) a "somatic attention bias," where the agent had a strong habit to attend only to its arousal level and the approach vs. avoid modality.

Figure 10 shows how these different attentional biases promote false inference. On the left, the true state is AFRAID, and the externally focused agent first attends to the stimulus/context (social rejection) and then to her beliefs (other-blame); however, without paying attention to her motivated action (avoid), she falsely reports feeling ANGRY instead of AFRAID (note that, following feedback, there is a retrospective inference that afraid was more probable; similar retrospective inferences after feedback are also shown in the other two examples in **Figure 10**). In the middle, the true state is ANGRY, and the internally focused agent first attends to valence (unpleasant) and then to arousal (high); however, without paying attention to her action tendency (approach), she falsely reports feeling AFRAID. On the right, the true state is SAD, and the somatically focused agent attends to her motivated action (avoid) and to arousal (high); however, without attending to beliefs (self-blame) she falsely reports feeling AFRAID instead of SAD.

Importantly, these false reports occur in an agent that has already acquired very precise emotion concepts. Thus, this does not represent a failure to learn about emotions, but simply the effect of having learned poor habits for mental action.

The results of these examples were confirmed in simulations of 40 interleaved emotion trials (10 per emotion) in an agent who had already acquired precise emotion concepts (temperature parameter = 2 for both "A" and "B" matrices; no learning). In these simulations, we observed that the externally focused agent had 100% accuracy for SAD, 10% accuracy for AFRAID, 100% accuracy for ANGRY, and 60% accuracy for HAPPY. The internally focused agent had 100% accuracy for SAD, 50% accuracy for AFRAID, 60% accuracy for ANGRY, and 100% accuracy for HAPPY. The somatically focused agent had 10% accuracy for SAD, 100% accuracy for AFRAID, 100% accuracy for ANGRY, and 0% accuracy for HAPPY. Thus,

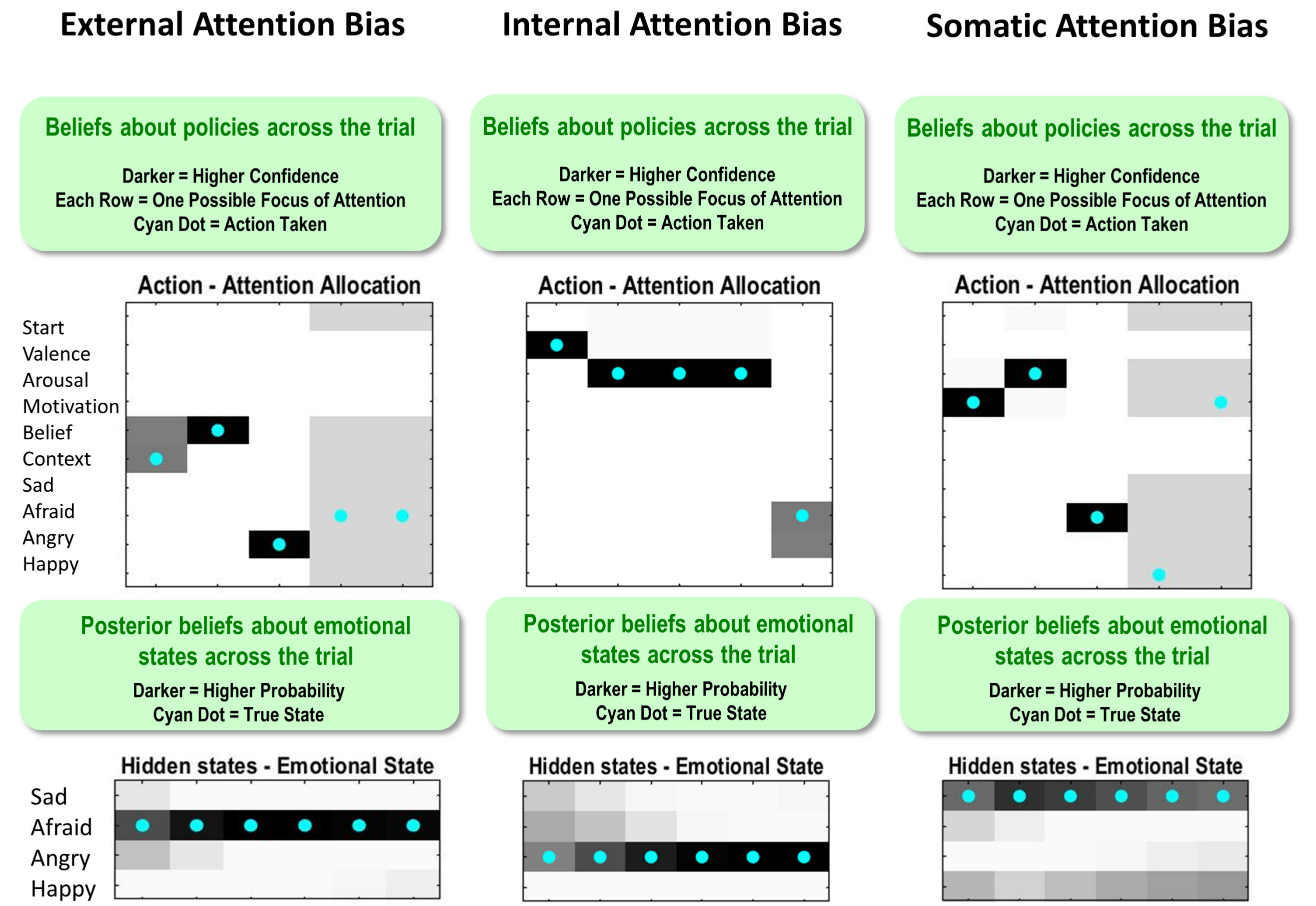
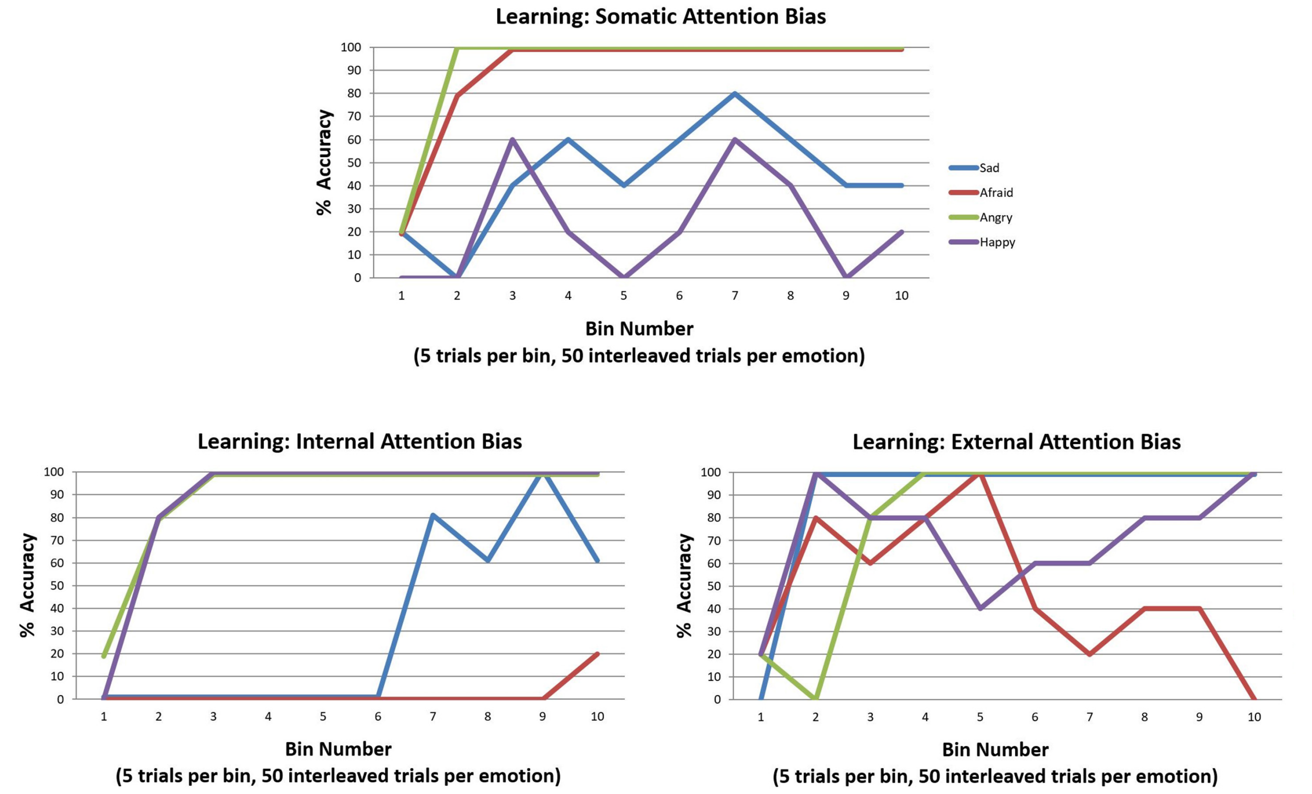


FIGURE 10 | This figure illustrates single trial effects of different attentional biases, each promoting false inference in a model that has already acquired precise emotion concepts (i.e., temperature parameter = 2 for both "A" and "B" matrices). External attention bias was implemented by giving the agent high prior expectations ("E" matrix values) that she would attend to the context and to her beliefs. Internal attention bias was implemented by giving the agent high prior expectations that she would attend to valence and arousal. Somatic attention bias was implemented by giving the agent high prior expectations that she would attend to arousal and action motivation. See the main text for a detailed description of the false inferences displayed for each bias. Note: the **bottom panels** indicate *posterior* beliefs, illustrating the agent's *retrospective* beliefs at the end of the trial about her emotional state throughout the trial. Also note that the agent's actions in the **upper left** and **upper right** after reporting should be ignored (i.e., they are simply an artifact of the trial continuing after the agent's report, when there is no goal and therefore no confidence in what to do).

without adaptive attentional habits, the agent was prone to misrepresent her emotions.

In our final simulations, we examined how learning these kinds of attentional biases in childhood could hinder emotion concept learning. To do so, we used the same learning procedure described in the previous section "Can Emotion Concepts Be Learned in Childhood?". However, in this case we simply equipped the model with the three different attentional biases ("E" matrix prior distributions over policies) and assessed its ability to learn emotion concepts over the 200 trials. The results of these simulations are provided in **Figure 11**. A somatic attention bias primarily allowed the agent to learn two emotion concepts, which corresponded to ANGER and FEAR. However, it is worth highlighting that the low accuracy for the other emotions means that their respective patterns were subsumed under the first two. Thus, it is more accurate to say that the agent learned two affective concepts, which largely predicted only approach vs. avoidance. In contrast, the internally biased agent easily acquired the distinction between pleasant (HAPPY) and unpleasant emotions (all lumped into ANGER) and began to learn a third concept that distinguished between low vs. high arousal (i.e., SAD vs. ANGRY). However, it did not conceptualize the distinction between approach and avoidance (i.e., ANGRY vs. AFRAID). Lastly, the externally focused agent was somewhat labile in concept acquisition; by the end, she could not predict approach vs. avoidance, but she possessed externally focused concepts with content along the lines of "the state I am in when I'm socially rejected and think it's my fault vs. someone else's fault" (SAD vs. ANGRY) and "the state I am in when at a crowded event" (HAPPY). These results provide strong support for the potential role of attentional biases in subverting emotional awareness.



Learning: Somatic Attention Bias

Learning: Internal Attention Bias

Learning: External Attention Bias

FIGURE 11 | Illustration of the simulated effects of each of the three attention biases (described in the legend for **Figure 10**) on emotion concept learning over 200 trials (initialized with a model that possessed a completely flat "A" matrix likelihood distribution). Each bias induces aberrant learning in a different way, often leading to the acquisition of two more coarse-grained emotion concepts as opposed to the four fine-grained concepts distinguished in the generative process. See the main text for a detailed description and interpretation.

DISCUSSION

The active inference formulation of emotional processing we have presented represents a first step toward the goal of building quantitative computational models of the ability to learn, recognize, and understand (be "aware" of) one's own emotions. Although this is clearly a toy model, it does appear to offer some insights, conceptual advances, and possible predictions.

First, in simulating differences in the precision (specificity) of emotion concepts, some intuitive but interesting phenomena emerged. As would be expected, differences in the specificity of the content of emotion concepts – here captured by the precision of the likelihood mapping from states to outcomes (i.e., the precision of what pattern of outcomes each emotion concept predicted) – led to differences in inferential accuracy. This suggests that, as would be expected, those with more precise emotion concepts would show greater understanding of their own affective responses. Perhaps less intuitively, beliefs about the stability of emotion concepts – here captured by the precision of expected state transitions – also influenced inferential accuracy. This predicts that a belief that emotional states are more stable (less labile) over time would also facilitate one's ability to correctly infer what they are feeling. This appears consistent with the low levels of emotional awareness or granularity observed in borderline personality disorder, which is characterized by emotional instability (Levine et al., 1997; Suvak et al., 2011).

Next, in simulating emotion concept learning, a few interesting insights emerged. Our simulations first confirmed that emotion concepts could successfully be learned, even when their content was cast (as done here) as complex, probabilistic, and highly overlapping response patterns across interoceptive, proprioceptive, exteroceptive, and cognitive domains. This was true when all emotion concepts needed to be learned simultaneously (as in childhood; see Widen and Russell, 2008; Hietanen et al., 2016), and was also true when a single new emotion concept was learned after others had already been acquired (as in adulthood during psycho-educational therapeutic interventions; e.g., see Hayes and Smith, 2005; Barlow et al., 2016; Burger et al., 2016; Lumley et al., 2017).

These results depended on whether the observed outcomes during learning were sufficiently precise and consistent. One finding worth highlighting was that emotion concept learning was hindered when the precision of transitions among emotional states was too low. This result may be relevant to previous empirical results in populations known to show reduced understanding of emotional states, such as those with autism

(Silani et al., 2008; Erbas et al., 2013) and those who grow up in socially impoverished or otherwise adverse (unpredictable) environments (Colvert et al., 2008; Lane et al., 2018). In autism, it has been suggested that overly imprecise beliefs about state transitions may hinder mental state learning, because such states require tracking abstract behavioral patterns over long timescales (Lawson et al., 2014, 2017; Haker et al., 2016). Children who grow up in impoverished environments may not have the opportunity to interact with others to observe stable patterns in other's affective responses; or receive feedback about their own (Pears and Fisher, 2005; Lane et al., 2018; Smith et al., 2018b). Our results successfully reproduce these phenomena – which represent important examples of mental state learning that may depend on consistently observed outcome patterns that are relatively stable over time.

As emotion concepts are known to differ in different cultures (Russell, 1991), our model and results may also relate to the learning mechanisms allowing for this type of culture-specific emotion categorization learning. Specifically, the "correct" and "incorrect" social feedback in our model could be understood as linguistic feedback from others in one's culture (e.g., a parent labeling emotional reactions for a child using culture-specific categories). If this feedback is sufficiently precise, then emotion concept learning could proceed effectively – even if the probabilistic mapping from emotion categories to other perceptual outcomes is fairly imprecise (i.e., which appears to be the case empirically; Barrett, 2006, 2017).

Another insight worth highlighting was that learning was more difficult when the agent had already acquired previous concepts but entertained the possible existence of new emotions she had not already learned. An interesting observation was that, in some cases (e.g., learning SAD), new learning temporarily interfered with old learning before being fully integrated into the agent's cognitive repertoire (an effect – termed retroactive interference – that has been extensively studied empirically within learning and memory research; Martínez et al., 2014; Darby and Sloutsky, 2015). In the case of learning HAPPY, we found that more extensive learning in the model's "childhood" was necessary to prevent this type of interference with respect to previous acquisition of the concept SAD. A second interesting observation was that AFRAID was very difficult to learn after the other concepts were fully acquired. This appeared to be because the agent had already learned to be highly confident about the explanatory power of her current conceptual repertoire (combined with the fact that AFRAID had considerable outcome overlap with SAD and ANGRY). It was necessary to provide the model with a persistent "flooding" of observations consistent with the new emotion to reduce its confidence sufficiently to acquire a new concept. It is not clear that this type of flooding is realistic, but perhaps resembles the extended periods of fear that occur during exposure-based behavioral therapies (Cooper et al., 2017). We should also emphasize that, due to the oversimplified nature of the mappings between emotion concepts and affective response features in our simulations, the difficulties observed in learning these specific emotions should not be taken too seriously. However, these overall results do predict that (and illustrate why) emotion concept learning in general should be more difficult in adulthood, and that emotion learning may have a kind of "sensitive period" in childhood (as supported by previous empirical findings; e.g., Pears and Fisher, 2005; Colvert et al., 2008).

The manner in which concept learning was implemented in these simulations may also have more general implications (for considerations of this approach to concept learning more generally, see Smith et al., 2019d). Typically, in Active Inference simulations the state space structure of a model is specified in advance (e.g., Schwartenbeck et al., 2015; Mirza et al., 2016; Parr and Friston, 2017a). Our model was instead equipped with "blank" hidden states devoid of content (i.e., these states started out predicting all outcomes with equal probability in the simulated learning). Over multiple exposures to the observed outcomes, these blank hidden states came to acquire conceptual content that captured distinct statistical patterns in the lower-level affective response components of the model. In some current neural process theories (Bastos et al., 2012; Friston et al., 2017a, 2018; Parr and Friston, 2018), distinct cortical columns are suggested to represent distinct hidden states. Under such theories, our learning model would suggest that the brain might contain "reserve" cortical columns available to capture new patterns of lower-level covariance if/when they begin to be observed in interaction with the world. To our knowledge, no direct evidence of such "reserve neurons" has been observed, although the generation of new neurons (with new synaptic connections) is known to occur in the hippocampus (Chancey et al., 2013). There is also the well-known phenomenon of "silent synapses" in the brain, which can persist into adulthood and become activated when new learning becomes necessary (e.g., see Kerchner and Nicoll, 2008; Chancey et al., 2013; Funahashi et al., 2013). Another interesting consideration is that, during sleep, it appears that many (but not all) synaptic strength increases acquired in the previous day are attenuated (Tononi and Cirelli, 2014). This has been suggested to correspond to a process of Bayesian model reduction, in which redundant model parameters are identified and removed to prevent model over-fitting and promote selection of the most parsimonious model that can successfully account for previous observations (Hobson and Friston, 2012). This also suggests that increases in "reserve" representational resources available for state space expansion (as in concept learning) could perhaps occur after sleep. In short, the acquisition of new concepts, emotion-related or otherwise, speaks to important issues in structure learning. The approach used here offers one solution to the question of how to expand a model, which could complement work on strategies for reducing a model (Friston et al., 2017b).

Although the neural process theory associated with active inference is cast at the level of canonical microcircuits and message passing, and therefore does not make *a priori* predictions about the brain regions that implement the emotion-related processes in our model, it nonetheless can afford empirical testing of macro-anatomical correlates. That is, this process theory can be used to generate predicted neural response time courses during emotional state inference and emotion concept learning in our model, and the macro-anatomical correlates of these time courses can then be established using neuroimaging methods. At present, the three-process model (Smith et al., 2018b), and supporting evidence (McRae et al., 2008; Smith et al., 2015, 2017b,c, 2018c,d,e, 2019a,b), has identified a number of large-scale networks that plausibly implement the processes we have simulated – and could therefore provide *a priori* hypotheses for future studies along these lines (Yeo et al., 2011; Barrett and Satpute, 2013). For example, "limbic network" regions (including orbitofrontal cortex and amygdala, among others), "salience network" regions (including the anterior insula and dorsal anterior cingulate, among others) and somatomotor/posterior insula regions all appear to be involved in generating affective responses and representing either visceral, somatic, or proprioceptive states at a perceptual level. Regions of the paralimbic cortex (e.g., "default mode network," with major hubs in the medial prefrontal cortex and posterior cingulate) are in turn most strongly implicated in conceptual inference (Binder et al., 2009) – such as the emotion concept representation processes simulated here. Thus, activity in a number of distinct brain regions/networks would be expected to show associations with distinct belief updating processes in our model.

A final insight offered by our model pertains to the possibility that maladaptive emotional state inference could be due to early experience. We demonstrated this in two ways. First, we simulated exposure to a large number of single emotion-provoking situations in childhood, promoting precise and highly engrained prior expectations for being in a single emotional state, as well as preferential learning of the respective outcome patterns for that state over others. We found that different kinds of "unrepresentative" (single-emotion) outcome patterns in early experience (e.g., chronic fear or sadness in childhood abuse/neglect or severe chronic depression) prevented learning other emotion concepts in somewhat inconsistent ways in repeated simulations. Overall, however, these results supported the idea that later emotional state inferences and emotion concept learning could be compromised by this type of maladaptive early experience. This could potentially relate to cognitive bias learning, such as the negative interpretation biases characteristic of mood and anxiety disorders (which have been interpreted within computational frameworks; Mogg and Bradley, 2005; Smith et al., 2017a).

Second, we examined the possibility that maladaptive cognitive habits could hinder emotional awareness. Here, we demonstrated that such habits can promote false emotional state inference and can hinder emotion concept learning. Specifically, we found that different types of external, internal, and somatic biases led to the acquisition of coarser-grained emotion concepts that failed to distinguish between various elements of affective responses. Aside from its relevance to cognitive biases more generally, these results could also explain certain empirical phenomena in emotional awareness research, such as the finding that males tend to score lower on emotional awareness measures than females (Wright et al., 2017). Specifically, while a genetic contribution to such findings is possible, it is also known that many cultures reinforce emotion avoidance in boys more than in girls in childhood (Fivush et al., 2000; Diener and Lucas, 2004; Chaplin et al., 2005), and can promote beliefs that paying attention to emotions is a sign of weakness or that emotional information simply carries little practical value. This type of learning could plausibly reinforce biased patterns of attention similar to those simulated here. Thus, our simulations suggest an interesting, testable mechanism by which such (potentially socialization-based) differences may arise.

In closing, it is important to note that this model is deliberately simple and is meant only to represent a proof of principle that emotion inference and learning can be modeled within a neurocomputational framework from first principles. We chose a particular pattern of state-outcome mappings to simulate the content of emotion concepts, but this is unlikely to represent a fully accurate depiction of human emotion concepts or the outcomes they predict. Human emotion concepts likely draw on much higher-dimensional patterns of somatic and visceral sensations, behavioral motivations, and cognitive appraisal patterns. There are also "secondary" emotion concepts like jealousy or embarrassment, which may require including more specific context and appraisal observations in a model (e.g., observing a lover with a competing suitor, observing oneself committing actions that break social norms, etc.). Further, human agents (or at least some of them) are likely to have a much richer space of both emotion and non-emotion concepts available for explaining their patterns of internal experience in conjunction with other beliefs and exteroceptive evidence (e.g., a pattern of low arousal, unpleasant valence, and avoidance in many contexts could also be explained by the concept of sickness rather than sadness; see Smith et al., 2019b). A more complete model would take into account many different possible conceptual interpretations of this sort. In addition, our simulations only attempted to capture the second process within the three-process model (i.e., affective response representation; Smith et al., 2018b). Incorporating the other two processes (affective response generation and conscious access) would undoubtedly induce additional dynamics (including explicit brain-body interactions) that could alter or nuance the simulation results we have provided. Modeling these additional processes will be an important goal of future work (see Smith et al., 2019b).

A final more general limitation with this type of modeling is that, in its current form, there are limited means of evaluating how well it represents the true form of emotional state inference and emotion concept learning implemented in the human brain. Here, we have focused on reproducing and validating a minimal model that evinces emotional state inference and learning within the active inference framework. Crucially, this model has – by construction – a construct validity with the three-process model and associated empirical evidence. As noted above, external validation of the model's ability to capture human brain processes will be an important next step, and can be done, for example, by examining whether the simulated neural responses we have presented are observable within particular brain regions during future neuroimaging studies of attending to – and reporting – one's own emotions (e.g., Lane et al., 1997; Gusnard et al., 2001; Smith et al., 2014, 2018c,d).

With these limitations in mind, however, this approach to computationally modeling emotion-related processes appears promising with respect to the initial insights it can offer. It can illustrate selective information integration in the service of conceptual inference, it can successfully simulate concept learning and some of its known vulnerabilities, and it can highlight maladaptive interactions between cognitive habits, early experience, and the ability to understand and be aware of one's own emotions later in life, all of which may play important roles in the development of emotional pathology. Finally, it highlights the potential for future empirical work in which tasks could be adapted to the broad structure of such models, which would allow investigation of individual differences in emotion processing as well as its neural basis. In other words, once we have a validated model of these emotion-related processes – at the subjective and neuronal level – we can, in principle, fit the model to observed responses and thereby phenotype subjects in terms of their emotion-related beliefs states (Schwartenbeck and Friston, 2016).

Software Note

Although the generative model – specified by the various matrices described in this paper – changes from application to application, the belief updates are generic and can be implemented using standard routines (here spm_MDP_VB_X.m). These routines are available as Matlab code in the latest version of SPM academic software[1]. The simulations in this paper can be reproduced (and customized) via running the Matlab code included here in the **Supplementary Material** (Emotion_learning_model.m).

AUTHOR CONTRIBUTIONS

RS took the lead in writing the manuscript and constructing the model. TP assisted with programming the model and running simulations, and also assisted in writing and editing the manuscript. KF provided guidance in constructing the model and also contributed to the writing and editing of the manuscript.

FUNDING

TP is supported by the Rosetrees Trust (Award Number 173346). KF is a Wellcome Trust Principal Research Fellow (Ref: 088130/Z/09/Z).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02844/full#supplementary-material

---

[1] http://www.fil.ion.ucl.ac.uk/spm/

REFERENCES

Bagby, R., Parker, J., and Taylor, G. (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., Parker, J., and Taylor, G. (1994b). The twenty-item toronto alexithymia scale—II. Convergent, discriminant, and concurrent validity. *J. Psychosom. Res.* 38, 33–40. doi: 10.1016/0022-3999(94)90006-x

Barchard, K., and Hakstian, A. (2004). The nature and measurement of emotional intelligence abilities; basic dimensions and their relationships with other cognitive abilities and personality variables. *Educ. Psychol. Measure.* 64, 437–462. doi: 10.1177/0013164403261762

Barlow, D., Allen, L., and Choate, M. (2016). Toward a Unified Treatment for Emotional Disorders - Republished Article. *Behav. Ther.* 47, 838–853. doi: 10.1016/j.beth.2016.11.005

Barrett, L. (2006). Are emotions natural kinds? *Perspect. Psychol. Sci.* 1, 28–58. doi: 10.1111/j.1745-6916.2006.00003.x

Barrett, L. (2017). *How Emotions are Made: The Secret Life of the Brain*. New York, NY: Houghton Mifflin Harcourt.

Barrett, L., Mesquita, B., and Gendron, M. (2011). Context in emotion perception. *Curr. Direct. Psychol. Sci.* 20, 286–290. doi: 10.1177/0963721411422522

Barrett, L., and Satpute, A. (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., and Simmons, W. (2015). Interoceptive predictions in the brain. *Nat. Rev. Neurosci.* 16, 419–429. doi: 10.1038/nrn3950

Baslet, G., Termini, L., and Herbener, E. (2009). Deficits in emotional awareness in schizophrenia and their relationship with other measures of functioning. *J. Nerv. Mental Dis.* 197, 655–660. doi: 10.1097/NMD.0b013e3181b3b20f

Bastos, A., Usrey, W., Adams, R., Mangun, G., Fries, P., and Friston, K. (2012). Canonical microcircuits for predictive coding. *Neuron* 76, 695–711. doi: 10.1016/j.neuron.2012.10.038

Berthoz, S., Ouhayoun, B., and Parage, N. (2000). Etude preliminaire des niveaux de conscience emotionnelle chez des patients deprimes et des controles. (Preliminary study of the levels of emotional awareness in depressed patients and controls.). *Ann. Med. Psychol.* 158, 665–672.

Binder, J., Desai, R., Graves, W., and Conant, 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

Bréjard, V., Bonnet, A., and Pedinielli, J. (2012). The role of temperament and emotional awareness in risk taking in adolescents. *L'Encéphale* 38, 1–9. doi: 10.1016/j.encep.2011.04.005

Brown, T. H., Zhao, Y., and Leung, V. (2009). "Hebbian plasticity A2–squire," in *Encyclopedia of Neuroscience*, ed. R. Larry (Oxford: Academic Press), 1049–1056.

Burger, A., Lumley, M., Carty, J., Latsch, D., Thakur, E., Hyde-Nolan, M., et al. (2016). The effects of a novel psychological attribution and emotional awareness and expression therapy for chronic musculoskeletal pain: a preliminary, uncontrolled trial. *J. Psychosom. Res.* 81, 1–8. doi: 10.1016/j.jpsychores.2015.12.003

Bydlowski, S., Corcos, M., Jeammet, P., Paterniti, S., Berthoz, S., Laurier, C., et al. (2005). Emotion-processing deficits in eating disorders. *Int. J. Eat. Disord.* 37, 321–329. doi: 10.1002/eat.20132

Chancey, J., Adlaf, E., Sapp, M., Pugh, P., Wadiche, J., and Overstreet-Wadiche, L. (2013). GABA depolarization is required for experience-dependent synapse unsilencing in adult-born neurons. *J. Neurosci.* 33, 6614–6622. doi: 10.1523/JNEUROSCI.0781-13.2013

Chaplin, T., Cole, P., and Zahn-Waxler, C. (2005). Parental socialization of emotion expression: gender differences and relations to child adjustment. *Emotion* 5, 80–88. doi: 10.1037/1528-3542.5.1.80

Ciarrochi, J., Caputi, P., and Mayer, J. (2003). The distinctiveness and utility of a measure of trait emotional awareness. *Pers. Individ. Differ.* 34, 1477–1490. doi: 10.1016/s0191-8869(02)00129-0

Clark, J., Watson, S., and Friston, K. (2018). What is mood? A computational perspective. *Psychol. Med.* 48, 2277–2284. doi: 10.1017/S0033291718000430

Colvert, E., Rutter, M., Kreppner, J., Beckett, C., Castle, J., Groothues, C., et al. (2008). Do theory of mind and executive function deficits underlie the adverse outcomes associated with profound early deprivation?: findings from the english and romanian adoptees study. *J. Abnorm. Child Psychol.* 36, 1057–1068. doi: 10.1007/s10802-008-9232-x

Conant, C., and Ashbey, W. (1970). Every good regulator of a system must be a model of that system. *Int. J. Syst. Sci.* 1, 89–97. doi: 10.1080/00207727008920220

Consoli, S., Lemogne, C., Roch, B., Laurent, S., Plouin, P., and Lane, R. (2010). Differences in emotion processing in patients with essential and secondary hypertension. *Am. J. Hyperten.* 23, 515–521. doi: 10.1038/ajh.2010.9

Cooper, A. A., Clifton, E. G., and Feeny, N. C. (2017). An empirical review of potential mediators and mechanisms of prolonged exposure therapy. *Clin. Psychol. Rev.* 56, 106–121. doi: 10.1016/j.cpr.2017.07.003

Darby, K., and Sloutsky, V. (2015). The cost of learning: interference effects in memory development. *J. Exp. Psychol. Gen.* 144, 410–431. doi: 10.1037/xge0000051

de Berker, A., Rutledge, R., Mathys, C., Marshall, L., Cross, G., Dolan, R., et al. (2016). Computations of uncertainty mediate acute stress responses in humans. *Nat. Commun.* 7:10996. doi: 10.1038/ncomms10996

Diener, M., and Lucas, R. (2004). Adults desires for childrens emotions across 48 countries: $ssociations with individual and national characteristics. *J. Cross Cult. Psychol.* 35, 525–547. doi: 10.1177/0022022104268387

Donges, U., Kersting, A., Dannlowski, U., Lalee-Mentzel, J., Arolt, V., and Suslow, T. (2005). Reduced awareness of others' emotions in unipolar depressed patients. *J. Nerv. Mental Dis.* 193, 331–337. doi: 10.1097/01.nmd.0000161683.02482.19

Erbas, Y., Ceulemans, E., Boonen, J., Noens, I., and Kuppens, P. (2013). Emotion differentiation in autism spectrum disorder. *Res. Autism Spectr. Disord.* 7, 1221–1227. doi: 10.1016/j.rasd.2013.07.007

Fivush, R., Brotman, M., Buckner, J., and Goodman, S. (2000). Gender differences in parent–child emotion narratives. *Sex Roles* 42, 233–253. doi: 10.1023/A:1007091207068

Frewen, P., Lane, R., Neufeld, R., Densmore, M., Stevens, T., and Lanius, R. (2008). Neural correlates of levels of emotional awareness during trauma script-imagery in posttraumatic stress disorder. *Psychosom. Med.* 70, 27–31. doi: 10.1097/psy.0b013e31815f66d4

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., O Doherty, J., and Pezzulo, G. (2016). Active inference and learning. *Neurosci. Biobehav. Rev.* 68, 862–879. doi: 10.1016/j.neubiorev.2016.06.022

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., and Pezzulo, G. (2017a). Active inference: a process theory. *Neural Comput.* 29, 1–49. doi: 10.1162/NECO_a_00912

Friston, K., Lin, M., Frith, C., Pezzulo, G., Hobson, J., and Ondobaka, S. (2017b). Active inference. Curiosity and insight. *Neural Comput.* 29, 2633–2683. doi: 10.1162/neco_a_00999

Friston, K., Parr, T., and de Vries, B. (2017c). The graphical brain: belief propagation and active inference. *Netw. Neurosci.* 1, 381–414. doi: 10.1162/NETN_a_00018

Friston, K., Rosch, R., Parr, T., Price, C., and Bowman, H. (2018). Deep temporal models and active inference. *Neurosci. Biobehav. Rev.* 90, 486–501. doi: 10.1016/J.NEUBIOREV.2018.04.004

Funahashi, R., Maruyama, T., Yoshimura, Y., and Komatsu, Y. (2013). Silent synapses persist into adulthood in layer 2/3 pyramidal neurons of visual cortex in dark-reared mice. *J. Neurophysiol.* 109, 2064–2076. doi: 10.1152/jn.00912.2012

Gusnard, D., Akbudak, E., Shulman, G., and Raichle, M. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 4259–4264. doi: 10.1073/pnas.071043098

Haker, H., Schneebeli, M., and Stephan, K. (2016). Can bayesian theories of autism spectrum disorder help improve clinical practice? *Front. Psychiatry* 7:107. doi: 10.3389/fpsyt.2016.00107

Harmon-Jones, E., Gable, P., and Peterson, C. (2010). The role of asymmetric frontal cortical activity in emotion-related phenomena: a review and update. *Biol. Psychol.* 84, 451–462. doi: 10.1016/J.BIOPSYCHO.2009.08.010

Hayes, S., and Smith, S. (2005). *Get Out of Your Mind and Into Your Life: The New Acceptance and Commitment Therapy*. Oakland, CA: New Harbinger Publications.

Hesp, C., Smith, R., Allen, M., Friston, K., and Ramstead, M. (2019). Deeply felt affect: the emergence of valence in deep active inference. *PsyArXiv* [Preprint]. doi: 10.31234/osf.io/62pfd

Hietanen, J., Glerean, E., Hari, R., and Nummenmaa, L. (2016). Bodily maps of emotions across child development. *Develop. Sci.* 19, 1111–1118. doi: 10.1111/desc.12389

Hobson, J., and Friston, K. (2012). Waking and dreaming consciousness: neurobiological and functional considerations. *Prog. Neurobiol.* 98, 82–98. doi: 10.1016/j.pneurobio.2012.05.003

Hohwy, J. (2016). The self-evidencing brain. *Noûs* 50, 259–285. doi: 10.1111/nous.12062

Joffily, M., and Coricelli, G. (2013). Emotional valence and the free-energy principle. *PLoS Comput. Biol.* 9:e1003094. doi: 10.1371/journal.pcbi.1003094

Kashdan, T., Barrett, L., and McKnight, P. (2015). Unpacking emotion differentiation: transforming unpleasant experience by perceiving distinctions in negativity. *Curr. Direct. Psychol. Sci.* 24, 10–16. doi: 10.1177/0963721414550708

Kashdan, T., and Farmer, A. (2014). Differentiating emotions across contexts: comparing adults with and without social anxiety disorder using random, social interaction, and daily experience sampling. *Emotion* 14, 629–638. doi: 10.1037/a0035796

Kerchner, G., and Nicoll, R. (2008). Silent synapses and the emergence of a postsynaptic mechanism for LTP. *Nat. Rev. Neurosci.* 9, 813–825. doi: 10.1038/nrn2501

Kleckner, I., Zhang, J., Touroutoglou, A., Chanes, L., Xia, C., Simmons, W., et al. (2017). Evidence for a large-scale brain system supporting allostasis and interoception in humans. *Nature Hum. Behav.* 1:0069. doi: 10.1038/s41562-017-0069

Lackner, J. (2005). "Is IBS a problem of emotion dysregulation? Testing the levels of emotional awareness model," in *Proceeding of the Presented at the Annual Meeting of the American Psychosomatic Society*, McLean, VA.

Lane, R., Anderson, F., and Smith, R. (2018). Biased competition favoring physical over emotional pain: a possible explanation for the link between early adversity and chronic pain. *Psychosom. Med.* 80, 880–890. doi: 10.1097/PSY.0000000000000640

Lane, R., Fink, G., Chua, P., and Dolan, R. (1997). Neural activation during selective attention to subjective emotional responses. *Neuroreport* 8, 3969–3972. doi: 10.1097/00001756-199712220-00024

Lane, R., Quinlan, D., Schwartz, G., Walker, P., and Zeitlin, S. (1990). The levels of emotional awareness scale: a cognitive-developmental measure of emotion. *J. Pers. Assess.* 55, 124–134. doi: 10.1080/00223891.1990.9674052

Lane, R., and Schwartz, G. (1987). Levels of emotional awareness: a cognitive-developmental theory and its application to psychopathology. *Am. J. Psychiatry* 144, 133–143. doi: 10.1176/ajp.144.2.133

Lane, R., Sechrest, L., Reidel, R., Weldon, V., Kaszniak, A., and Schwartz, G. (1996). Impaired verbal and nonverbal emotion recognition in alexithymia. *Psychosom. Med.* 58, 203–210. doi: 10.1097/00006842-199605000-00002

Lane, R., Sechrest, L., Riedel, R., Shapiro, D., and Kaszniak, A. (2000). Pervasive emotion recognition deficit common to alexithymia and the repressive coping style. *Psychosom. Med.* 62, 492–501. doi: 10.1097/00006842-200007000-00007

Lane, R., Weihs, K., Herring, A., Hishaw, A., and Smith, R. (2015). Affective agnosia: expansion of the alexithymia construct and a new opportunity to integrate and extend Freud's legacy. *Neurosci. Biobehav. Rev.* 55, 594–611. doi: 10.1016/j.neubiorev.2015.06.007

Lawson, R., Mathys, C., and Rees, G. (2017). Adults with autism overestimate the volatility of the sensory environment. *Nat. Neurosci.* 20, 1293–1299. doi: 10.1038/nn.4615

Lawson, R., Rees, G., and Friston, K. (2014). An aberrant precision account of autism. *Front. Hum. Neurosci.* 8:302. doi: 10.3389/fnhum.2014.00302

Levine, D., Marziali, E., and Hood, J. (1997). Emotion processing in borderline personality disorders. *J. Nerv. Mental Dis.* 185, 240–246.

Licata, M., Kristen, S., and Sodian, B. (2016). Mother-child interaction as a cradle of theory of mind: the role of maternal emotional availability. *Soc. Develop.* 25, 139–156. doi: 10.1111/sode.12131

Limanowski, J., and Friston, K. (2018). "Seeing the Dark": grounding phenomenal transparency and opacity in precision estimation for active inference. *Front. Psychol.* 9:643. doi: 10.3389/fpsyg.2018.00643

Lindquist, K., and Barrett, L. (2008). Constructing emotion: the experience of fear as a conceptual act. *Psychol. Sci.* 19, 898–903. doi: 10.1111/j.1467-9280.2008.02174.x

Lumley, M., Schubiner, H., Lockhart, N., Kidwell, K., Harte, S., Clauw, D., et al. (2017). Emotional awareness and expression therapy, cognitive behavioral therapy, and education for fibromyalgia: a cluster-randomized controlled trial. *Pain* 158, 2354–2363. doi: 10.1097/j.pain.0000000000001036

Martínez, M., Villar, M. E., Ballarini, F., and Viola, H. (2014). Retroactive interference of object-in-context long-term memory: role of dorsal hippocampus and medial prefrontal cortex. *Hippocampus* 24, 1482–1492. doi: 10.1002/hipo.22328

McKay, R., and Dennett, D. (2009). The evolution of misbelief. *Behav. Brain Sci.* 32, 493–510. doi: 10.1017/S0140525X09990975

McRae, K., Reiman, E., Fort, C., Chen, K., and Lane, R. (2008). Association between trait emotional awareness and dorsal anterior cingulate activity during emotion is arousal-dependent. *Neuroimage* 41, 648–655. doi: 10.1016/j.neuroimage.2008.02.030

Metzinger, T. (2017). "The problem of mental action predictive control without sensory sheets," in *Philosophy and Predicitive Processing*, eds T. Metzinger, and W. Wiese. (Hong Kong: MIND Group).

Mirza, M., Adams, R., Mathys, C., and Friston, K. (2016). Scene construction, visual foraging, and active inference. *Front. Comput. Neurosci.* 10:56. doi: 10.3389/fncom.2016.00056

Moeller, S., Konova, A., Parvaz, M., Tomasi, D., Lane, R., Fort, C., et al. (2014). Functional, structural, and emotional correlates of impaired insight in cocaine addiction. *JAMA Psychiatry* 71, 61–70. doi: 10.1001/jamapsychiatry.2013.2833

Mogg, K., and Bradley, B. (2005). Attentional bias in generalized anxiety disorder versus depressive disorder. *Cogn. Ther. Res.* 29, 29–45. doi: 10.1007/s10608-005-1646-y

Mumme, D., Fernald, A., and Herrera, C. (1996). Infants' responses to facial and vocal emotional signals in a social referencing paradigm. *Child Develop.* 67, 3219–3237. doi: 10.1111/j.1467-8624.1996.tb01910.x

Panksepp, J., and Biven, L. (2012). *The Archaeology of Mind: Neuroevolutionary Origins of Human Emotions*. New York, NY: W.W. Norton & Company.

Panksepp, J., Lane, R., Solms, M., and Smith, R. (2017). Reconciling cognitive and affective neuroscience perspectives on the brain basis of emotional experience. *Neurosci. Biobehav. Rev.* 76, 187–215. doi: 10.1016/j.neubiorev.2016.09.010

Parker, J., Taylor, G., and Bagby, R. (2003). The 20-item Toronto alexithymia scale: III. Reliability and factorial validity in a community population. *J. Psychosom. Res.* 55, 269–275. doi: 10.1016/S0022-3999(02)00578-0

Parr, T., and Friston, K. (2017a). Working memory, attention, and salience in active inference. *Sci. Rep.* 7:14678. doi: 10.1038/s41598-017-15249-0

Parr, T., and Friston, K. (2017b). Uncertainty, epistemics and active inference. *J. R. Soc. Interf.* 14:20170376. doi: 10.1098/rsif.2017.0376

Parr, T., and Friston, K. (2018). The anatomy of inference: generative models and brain structure. *Front. Comput. Neurosci.* 12:90. doi: 10.3389/fncom.2018.00090

Pears, K., and Fisher, P. (2005). Emotion understanding and theory of mind among maltreated children in foster care: evidence of deficits. *Develop. Psychopathol.* 17, 47–65. doi: 10.1017/S0954579405050030

Peters, A., McEwen, B. S., and Friston, K. (2017). Uncertainty and stress: why it causes diseases and how it is mastered by the brain. *Prog. Neurobiol.* 156, 164–188. doi: 10.1016/J.PNEUROBIO.2017.05.004

Posner, M. (2016). Orienting of attention: then and now. *Q. J. Exp. Psychol.* 69, 1864–1875. doi: 10.1080/17470218.2014.937446

Rizzolatti, G., Riggio, L., Dascola, I., and Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia* 25, 31–40. doi: 10.1016/0028-3932(87)90041-8

Russell, J. (1991). Culture and the categorization of emotions. *Psychol. Bull.* 110, 426–450. doi: 10.1037/0033-2909.110.3.426

Russell, J. (2003). Core affect and the psychological construction of emotion. *Psychol. Rev.* 110, 145–172. doi: 10.1037/0033-295x.110.1.145

Scherer, K. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cogn. Emot.* 23, 1307–1351. doi: 10.1080/02699930902928969

Schwartenbeck, P., FitzGerald, T., Mathys, C., Dolan, R., and Friston, K. (2015). The dopaminergic midbrain encodes the expected certainty about desired outcomes. *Cereb. Cortex* 25, 3434–3445. doi: 10.1093/cercor/bhu159

Schwartenbeck, P., and Friston, K. (2016). Computational phenotyping in psychiatry: a worked example. *ENeuro* 3: ENEURO.0049-16.2016. doi: 10.1523/ENEURO.0049-16.2016

Seth, A. (2013). Interoceptive inference, emotion, and the embodied self. *Trends Cogn. Sci.* 17, 565–573. doi: 10.1016/j.tics.2013.09.007

Seth, A., and Friston, K. (2016). Active interoceptive inference and the emotional brain. *Philos. Trans. R. Soc. Lond. B* 371:20160007. doi: 10.1098/rstb.2016.0007

Sharot, T. (2011). The optimism bias. *Curr. Biol.* 21, R941–R945. doi: 10.1016/J.CUB.2011.10.030

Siemer, M., Mauss, I., and Gross, J. (2007). Same situation–different emotions: how appraisals shape our emotions. *Emotion* 7, 592–600. doi: 10.1037/1528-3542.7.3.592

Silani, G., Bird, G., Brindley, R., Singer, T., Frith, C., and Frith, U. (2008). Levels of emotional awareness and autism: an fMRI study Levels of emotional awareness and autism: an fMRI study. *Psychology* 3, 97–112. doi: 10.1080/17470910701577020

Smith, D., and Schenk, T. (2012). The premotor theory of attention: time to move on? *Neuropsychologia* 50, 1104–1114. doi: 10.1016/J.NEUROPSYCHOLOGIA.2012.01.025

Smith, R. (2019). "The three-process model of implicit and explicit emotion," in *Neuroscience of Enduring Change: Implications for Psychotherapy*, eds R. Lane, and L. Nadel (Oxford: Oxford University Press).

Smith, R., Akozei, A., Killgore, W., and Lane, R. (2017a). Nested positive feedback loops in the maintenance of major depression: an integration and extension of previous models. *Brain Behav. Immun.* 67, 374–397. doi: 10.1016/j.bbi.2017.09.011

Smith, R., Alkozei, A., Bao, J., Smith, C., Lane, R., and Killgore, W. (2017b). Resting state functional connectivity correlates of emotional awareness. *NeuroImage* 159, 99–106. doi: 10.1016/j.neuroimage.2017.07.044

Smith, R., Lane, R., Alkozei, A., Bao, J., Smith, C., Sanova, A., et al. (2017c). Maintaining the feelings of others in working memory is associated with activation of the left anterior insula and left frontal-parietal control network. *Soc. Cogn. Affect. Neurosci.* 12, 848–860. doi: 10.1093/scan/nsx011

Smith, R., Thayer, J., Khalsa, S., and Lane, R. (2017d). The hierarchical basis of neurovisceral integration. *Neurosci. Biobehav. Rev.* 75, 274–296. doi: 10.1016/j.neubiorev.2017.02.003

Smith, R., Bajaj, S., Dailey, N., Alkozei, A., Smith, C., Sanova, A., et al. (2018a). Greater cortical thickness within the limbic visceromotor network predicts higher levels of trait emotional awareness. *Conscious. Cogn.* 57, 54–61. doi: 10.1016/j.concog.2017.11.004

Smith, R., Killgore, W., and Lane, R. (2018b). The structure of emotional experience and its relation to trait emotional awareness: a theoretical review. *Emotion* 18, 670–692. doi: 10.1037/emo0000376

Smith, R., Lane, R., Alkozei, A., Bao, J., Smith, C., Sanova, A., et al. (2018c). The role of medial prefrontal cortex in the working memory maintenance of one's own emotional responses. *Sci. Rep.* 8:3460. doi: 10.1038/s41598-018-21896-8

Smith, R., Lane, R., Sanova, A., Alkozei, A., Smith, C., and Killgore, W. W. D. (2018d). Common and unique neural systems underlying the working memory maintenance of emotional vs. bodily reactions to affective stimuli: the moderating role of trait emotional awareness. *Front. Hum. Neurosci.* 12:370. doi: 10.3389/fnhum.2018.00370

Smith, R., Sanova, A., Alkozei, A., Lane, R., and Killgore, W. (2018e). Higher levels of trait emotional awareness are associated with more efficient global information integration throughout the brain: a graph-theoretic analysis of resting state functional connectivity. *Soc. Cogn. Affect. Neurosci.* 13, 665–675. doi: 10.1093/scan/nsy047

Smith, R., Braden, B., Chen, K., Ponce, F., Lane, R., and Baxter, L. (2015). The neural basis of attaining conscious awareness of sad mood. *Brain Imaging Behav.* 9, 574–587. doi: 10.1007/s11682-014-9318-8

Smith, R., Fass, H., and Lane, R. (2014). Role of medial prefrontal cortex in representing one's own subjective emotional responses: a preliminary study. *Conscious. Cogn.* 29, 117–130. doi: 10.1016/j.concog.2014.08.002

Smith, R., Kaszniak, A., Katsanis, J., Lane, R., and Nielsen, L. (2019a). The importance of identifying underlying process abnormalities in alexithymia: implications of the three-process model and a single case study illustration. *Conscious. Cogn.* 68, 33–46. doi: 10.1016/J.CONCOG.2018.12.004

Smith, R., Lane, R., Parr, T., and Friston, K. (2019b). Neurocomputational mechanisms underlying emotional awareness: insights afforded by deep active inference and their potential clinical relevance. *Neurosci. Biobehav. Rev.* 107, 473–491. doi: 10.1101/681288

Smith, R., Quinlan, D., Schwartz, G. E., Sanova, A., Alkozei, A., and Lane, R. D. (2019c). Developmental contributions to emotional awareness. *J. Pers. Assess.* 101, 150–158. doi: 10.1080/00223891.2017.1411917

Smith, R., Schwartenbeck, P., Parr, T., and Friston, K. (2019d). An active inference approach to modeling structure learning: concept learning as an example case. *BioRxiv* [Preprint]. doi: 10.1101/633677

Smith, R., Weihs, K., Alkozei, A., Killgore, W., and Lane, R. (2019e). An embodied neurocomputational framework for organically integrating biopsychosocial processes: an application to the role of social support in health and disease. *Psychosom. Med.* 81, 125–145. doi: 10.1097/PSY.0000000000000661

Smith, R., and Lane, R. (2015). The neural basis of one's own conscious and unconscious emotional states. *Neurosci. Biobehav. Rev.* 57, 1–29. doi: 10.1016/j.neubiorev.2015.08.003

Smith, R., and Lane, R. (2016). Unconscious emotion: a cognitive neuroscientific perspective. *Neurosci. Biobehav. Rev.* 69, 216–238. doi: 10.1016/j.neubiorev.2016.08.013

Subic-Wrana, A., Beetz, M., Paulussen, J., Wiltnik, J., and Beutel, M. (2007). "Relations between attachment, childhood trauma, and emotional awareness in psychosomatic inpatients," *in Proceedings of the Presented at the Annual Meeting of the American Psychosomatic Society*. Budapest.

Subic-Wrana, C., Bruder, S., Thomas, W., Lane, R., and Köhle, K. (2005). Emotional awareness deficits in inpatients of a psychosomatic ward: a comparison of two different measures of alexithymia. *Psychosom. Med.* 67, 483–489. doi: 10.1097/01.psy.0000160461.19239.13

Suvak, M., Litz, B., Sloan, D., Zanarini, M., Barrett, L., and Hofmann, S. (2011). Emotional granularity and borderline personality disorder. *J. Abnorm. Psychol.* 120, 414–426. doi: 10.1037/a0021808

Tononi, G., and Cirelli, C. (2014). Sleep and the price of plasticity: from synaptic and cellular homeostasis to memory consolidation and integration. *Neuron* 81, 12–34. doi: 10.1016/J.NEURON.2013.12.025

Widen, S., and Russell, J. (2008). Children acquire emotion categories gradually. *Cogn. Develop.* 23, 291–312. doi: 10.1016/j.cogdev.2008.01.002

Wilson-Mendenhall, C., Barrett, L., Simmons, W., and Barsalou, L. (2011). Grounding emotion in situated conceptualization. *Neuropsychologia* 49, 1105–1127. doi: 10.1016/j.neuropsychologia.2010.12.032

Wright, R., Riedel, R., Sechrest, L., Lane, R., and Smith, R. (2017). Sex differences in emotion recognition ability: the mediating role of trait emotional awareness. *Motiv. Emot.* 42, 149–160. doi: 10.1007/s11031-017-9648-0

Yeo, B., Krienen, F., Sepulcre, J., Sabuncu, M., 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

Conflict of Interest: 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.

frontiers in Psychology

ORIGINAL RESEARCH
published: 21 January 2020
doi: 10.3389/fpsyg.2019.02971

Influence of Attention Control on Implicit and Explicit Emotion Processing of Face and Body: Evidence From Flanker and Same-or-Different Paradigms

Viola Oldrati[1]*, *Alessandra Bardoni*[1], *Geraldina Poggi*[1] *and Cosimo Urgesi*[1,2,3]

[1] *Scientific Institute, IRCCS E. Medea, Bosisio Parini, Lecco, Italy,* [2] *Laboratory of Cognitive Neuroscience, Department of Languages and Literatures, Communication, Education and Society, University of Udine, Udine, Italy,* [3] *Scientific Institute, IRCCS E. Medea, San Vito al Tagliamento, Pordenone, Italy*

Edited by:
Giulia Mattavelli,
Scuola Universitaria Superiore IUSS Pavia, Italy

Reviewed by:
Dirk Kerzel,
Université de Genève, Switzerland
Khatereh Borhani,
Shahid Beheshti University, Iran

***Correspondence:**
Viola Oldrati
viola.oldrati@lanostrafamiglia.it

Specialty section:
This article was submitted to Consciousness Research, a section of the journal Frontiers in Psychology

Received: *17 September 2019*
Accepted: *16 December 2019*
Published: *21 January 2020*

Citation:
Oldrati V, Bardoni A, Poggi G and Urgesi C (2020) Influence of Attention Control on Implicit and Explicit Emotion Processing of Face and Body: Evidence From Flanker and Same-or-Different Paradigms. Front. Psychol. 10:2971. doi: 10.3389/fpsyg.2019.02971

Many existing findings indicate that processing of emotional information is pre-attentive, largely immune from attentional control. Nevertheless, inconsistent evidence on the interference of emotional cues on cognitive processing suggests that this influence may be a highly conditional phenomenon. The aim of the present study was twofold: (1) to examine the modulation of attention control on emotion processing using facial expressions (2) explore the very same effect for emotional body expressions. In Experiment 1, participants performed a Flanker task in which they had to indicate either the emotion (happy/fearful) or the gender of the target stimulus while ignoring the distracting stimuli at the side. We found evidence for intrusion of the emotional dimension of a stimulus in both the emotion and gender discrimination performance, thus when either task-relevant or task-irrelevant. To further explore the influence of attention control mechanisms, in Experiment 2 participants performed a same-or-different judgment task in which they were asked to pay attention to both the central and lateral stimuli and indicated whether the central stimulus matched the lateral for emotion or gender. Results showed that emotional features exerted an influence at an implicit level (i.e., during gender judgments) for bodies only. Gender features did not affect emotional processing in either experiments. To rule out the possibility that this effect was driven by postural rather than emotional features of fearful vs. happy stimuli, a control experiment was conducted. In Experiment 3, bodies with an opening/up-ward or closing/down-ward posture but with no emotional valence were presented. Results revealed that the body posture did not influence gender discrimination. Findings suggest that the emotional valence of a face or body stimulus can overpass attention filtering mechanisms, independently from the level of attentional modulation (Experiment 1). However, broadening the focus of attention to include the lateral stimuli led emotional information to intrude on the main task, exerting an implicit, bottom–up influence on gender processing, only when conveyed by bodies (Experiment 2). Results point to different mechanisms for the implicit processing of face and body emotional expressions, with the latter likely having role on action preparation processes.

Keywords: face expression, body expression, emotion, implicit processing, top–down control, bottom–up interference

INTRODUCTION

Adapting in a complex and dynamic environment requires the ability to remain oriented to an ongoing task and, at the same time, to direct attention to salient incoming stimuli even when they are not relevant for that task. While proactive mechanisms of attention, commonly called top–down, are deployed to maintain cognitive resources on ongoing tasks, reactive, bottom–up processes are driven by input characterized by novelty or saliency, such as emotional stimuli. The interplay between top–down and bottom–up mechanisms has been investigated with different methods (Park and Thayer, 2014; John et al., 2016; Müsch et al., 2017; Shechner et al., 2017) and in different populations, including healthy individuals and individuals in different life span with neuropsychiatric disorders (Macnamara and Proudfit, 2014; Klein et al., 2017; Mackiewicz et al., 2017; Abend et al., 2018). These studies have approached the conclusion that top–down and bottom–up mechanisms mutually influence each other in a highly conditional way (Kim et al., 2004; Pessoa et al., 2006).

One goal of the study of the interaction between attention control and emotional saliency is to weight the modulatory impact of the main ongoing task (i.e., the task at hand, which recruits the majority of cognitive resources), the affective nature of the distracting stimuli (e.g., valence, arousal and category) and the individual (clinical) characteristics of the participants (e.g., affective disorders). Overall, extensive lines of research confirmed that emotional stimuli capture attention to a greater extent than neutral stimuli at both behavioral and neural levels. At a behavioral level, stimuli charged with emotional salience have been demonstrated to elicit quicker responses in visual search paradigms using pictures of snakes/spiders among flowers/mushrooms (Ohman et al., 2001), angry/frightened faces among neutral or happy faces (Lobue, 2009; Haas et al., 2016), as well as unpleasant scenarios among neutral or pleasant scenarios (Astudillo et al., 2018). Many of these studies also indicate that stimuli with a negative valence may tap attention more than positive stimuli, suggesting that emotional valence, more than stimulus relevance, is crucial for the automatic capture of attention. The expression "negativity bias" originates from a consistent body of results stressing the importance of threat-related valence over non-threat-related contents (Carretié et al., 2009; Carretié, 2014). Nevertheless, this notion needs to be reconciled with results questioning the prioritized processing of threat-related stimuli. There is evidence stressing that the allocation of attention to emotionally salient stimuli is not automatic, but it is rather conditional to the relevance of the emotional dimension for the top–down instructions of the ongoing task (Barratt and Bundesen, 2012; Victeur et al., 2019). Other studies have argued that the "priority" dedicated to threat-related stimuli may derive from (low-level) perceptual characteristics of stimuli (e.g., visible teeth or direct eye gaze in case of angry faces) rather than from (high-level) configurational properties of the emotional expression (Horstmann and Bauland, 2006; Cooper et al., 2013). In a similar vein, evidence has been gathered around the modulatory role of the level of arousal in emotion information processing, where stimulus intensity rather than valence is considered as the key factor that drives attentional resources (Schimmack, 2005; Padmala et al., 2018). In sum, the bottom–up processing of the emotional salience, threat-related valence and arousal of stimuli seems to automatically capture attention at different extent according to the top–down requirements of the main task.

At a neural level, experimental research kept confirming that the interplay between attentional control and processing of emotional information is indeed a conditional phenomenon. The attentional competition elicited by top–down prioritization of task-related stimulus processing and automatic stimulus saliency is thought to recruit separate yet complementary brain circuits according to context demands. These circuits involve prefrontal regions, associated with attentional and complex cognitive operations, and a subcortical thalamo-amygdala pathway, which is thought as a privileged neural candidate for threat detection mechanism (Bishop, 2008). Functional magnetic resonance imaging (fMRI) studies showed that the differential amygdala response to emotional vs. neutral stimuli disappears when the cognitive load of the main task increases (Pessoa et al., 2002, 2005). In a similar vein, while attended emotional stimuli elicit greater amygdala activation as compared to attended neutral stimuli, the differential amygdala response to the two types of stimuli disappears when they are not attended to. This indicates that attentional modulation may exert a substantial influence on the pre-attentive mechanisms that underlie the automatic processing of salient stimuli (Dolan and Vuilleumier, 2003).

Among the different environmental stimuli, the processing of social-cues like faces or bodies is fundamental for our survival. Indeed, despite a heterogeneity in research methods, the automatic processing of facial (and, less extensively, bodily) emotion expressions has been the focus of several previous studies. Overall, these studies have pointed to the automatic allocation of attention to emotional face (or body) expressions even when they were presented as distractors in tasks that focused attentional resources to non-emotional target stimuli, such as letter or color detection tasks (Nordström and Wiens, 2012; Sussman et al., 2013). However, since the interplay between top–down and bottom–up mechanisms is modulated by the requirements of the ongoing tasks and by the type of stimulus saliency, investigating how the processing of facial or bodily emotional expressivity interacts with the processing of other features of the same stimuli seems relevant to dissociate the role of stimulus category (e.g., faces vs. objects) and stimulus features (e.g., emotion vs. gender) in attention allocation. Indeed, visual attention can be allocated to an object falling into a specific region of space (space-based attention) as well as to a specific feature of that object (feature-based attention; Maunsell and Treue, 2006). In this vein, studying the interaction between emotion and gender may represent an interesting opportunity, since both dimensions shape the way we interact in social environments (Harris et al., 2016).

The human visual system is highly sensitive to both emotion (Meeren et al., 2005) and gender cues (Bruce et al., 1993; Johnson and Tassinary, 2005), either conveyed by faces or bodies. Furthermore, the perception of these two salient cues seems highly interdependent, as shown by the interference exerted by variations of one dimension (i.e., gender) while recognizing the other (e.g., emotion; the so-called Garner interference effect; Gilboa-Schechtman et al., 2004; Atkinson et al., 2005; Becker, 2017; but see Gandolfo and Downing, 2019). This reveals the difficulties of the attentional control system in filtering out gender or emotional cues when the main task is focused on emotion or gender, respectively. In a recent behavioral study, the authors adopted a flanker paradigm using computer-generated faces and asked participants to indicate the emotion or the gender of the target stimulus, which could either match or not two flanker stimuli appearing at the side (Zhou and Liu, 2013). Crucially, the congruency between the target and the flanker could refer to the task-relevant (e.g., emotion congruency in the emotion task and gender congruency in the gender task) or the task-irrelevant (e.g., emotion congruency in the gender task and gender congruency in the emotion task) dimension, thus allowing the authors to directly asses the effect of emotional conflict at different levels of attention modulations. Results indicated that the incongruence of emotional (or gender) features between target and flanker interfered with emotion (or gender) discrimination especially when task-relevant, suggesting the modulation of top–down regulation mechanisms on both emotional and non-emotional feature incongruences. The researchers then applied the same paradigm coupled with the recording of ERPs and found that emotional conflict, accompanied by the peaking of the N200 component, was more pronounced when task-relevant than when task-irrelevant (Zhou et al., 2016). Nevertheless, the conflict signaled by the cortical component was observed to rise earlier for emotional incongruence than for gender incongruence. In this regard, the authors interpreted this anticipation as an index of a pre-attentive processing bias in favor of emotional over non-emotional features. Going toward a similar direction, a recent study observed an enhanced N170 component for emotional as compared to neutral face expressions when participants were asked to recognize the gender of laterally presented faces. Interestingly, this effect disappeared when participants were asked, rather than responding on a dimension of the faces, to detect a missing pixel of the central fixation cross, providing further evidence of a modulation of task demands on emotional features interference (Burra and Kerzel, 2019).

In light of the latest literature, the present work was aimed at replicating previous findings (Zhou and Liu, 2013) that either emotion or gender information is modulated by attentional control when task-relevant applying a similar flanker paradigm (Experiment 1). However, we also tested whether the same effects were maintained when using a same-or-different comparison task (Experiment 2). This second paradigm was selected to explore the modulatory role of visual search strategy on stimuli processing and distractor filtering and to test the modulatory role of space-based and feature-based attention allocation (Maunsell and Treue, 2006). The flanker task invites to focus the attention on the central target of a set of stimuli, which may favor space-based filtering of distractors. The same-or-different judgment task requires to expand the focus of attention to the entire set. This could hinder space-based filtering (i.e., suppressing the processing of an object based on its location) and possibly favor feature-based strategy (i.e., suppressing the processing of a specific feature of an object). This fundamental difference between the two tasks is reflected in a more pronounced center-to-periphery gradient of visual search in the flanker than in the same-or-different paradigm (Wendt et al., 2017).

A second aim was to explore the modulatory effect of attention on emotion and gender processing when conveyed not only by faces, but also by bodies. In fact, both faces and bodies provide important cues for the adaptation to social environments. Existing evidence suggests that face and body processing share similar cognitive mechanisms and neural correlates (Minnebusch and Daum, 2009 for a detailed review). For example, there is evidence that both faces and bodies trigger configural processing, defined as the perception of relations among the features of a stimulus (Reed et al., 2006). Moreover, faces and body stimuli specifically activate distinct yet adjacent and overlapping brain regions within the lateral and medial occipito-temporal cortex (Peelen and Downing, 2019). Another indication of a certain degree of similarity between the two comes from clinical studies. Patients diagnosed with prosopagnosia, an acquired disorder of face recognition, have also difficulties with the configural processing of bodies (Righart and de Gelder, 2007; Moro et al., 2012). Nevertheless, analogies from faces to bodies cannot ignore the perceptual and affective differences existing between the two classes of stimuli. As to their functional role, theoretical models suggest that faces would trigger more emphatic-related processes, while bodies processing would involve systems deputed to action observation and preparation (van de Riet et al., 2009).

To these aims, we run three experiments in which we presented a central face or body sided by two lateral faces or bodies that could match or mismatch the gender or emotion of the central stimulus. In Experiment 1, we used a typical flanker paradigm, in which attention was directed only to the central target stimulus and the influences of task-relevant and task-irrelevant features of the lateral stimuli were tested. In Experiment 2, we used a same-or-different paradigm, in which attention was directed to both central and lateral stimuli and participants had to match the central with the lateral stimuli for either gender or emotion. Finally, a third control experiment tested the specificity of the interference effects for bodily emotions as compared to neutral body postures.

EXPERIMENT 1

In this experiment, we used a typical Flanker paradigm to test whether the attentional modulation (i.e., task relevance) of the relative influence of gender and emotional cues was comparable for face and body stimuli. We presented a central face or body target sided by two flanker faces or bodies that could match or not the central target in gender or emotion. In separate blocks, we asked participants to discriminate the gender (i.e., male or female) or the emotion (i.e., positive or negative) of the central target. In keeping with previous studies with faces (Zhou and Liu, 2013), we expected that the presentation of the two flanker stimuli should interfere with the discrimination of the gender or emotion of the central target more when they are incongruent for the task-relevant (i.e., incongruent gender for the gender task and incongruent emotion for the emotion task) than for the task-irrelevant (i.e., incongruent emotion for the gender task and incongruent gender for the emotion task) dimension. This modulation should be comparable for faces and bodies as well for gender and emotional tasks, even if emotional incongruence was expected to exert a marginal effect also on gender discrimination, in keeping with a greater resilience of emotion processing to top–down attentional control (Zhou and Liu, 2013).

Participants

Twenty-four healthy volunteers (7 men, *M* = 28.04 *SD* = 4.15 years old) participated in Experiment 1. All participants had normal or corrected-to-normal vision. Prior to the beginning of the experiment, written informed consent was obtained from all participants. The experiment was approved by the local ethical committee and conducted in accordance with the declaration of Helsinki.

Stimuli and Task

Experimental stimuli consisted of a total of 24 pictures of Caucasian faces and bodies. Faces were taken from the NimStim dataset (Tottenham et al., 2009). Two emotion expressions (happy and fearful) were chosen for each face, so that in total six female and six male faces (three happy and three fearful) were used. Faces whose emotion expression was recognized with an accuracy level equal or superior to 80% were selected (Tottenham et al., 2009). Both happy and fearful faces were with the mouth open, as these have been shown to be better recognized than closed-mouth expressions (Tottenham et al., 2009).

Bodies pictures were taken from a validated pool of static images depicting bodies with blurred faces in emotional whole-body movements (see Borgomaneri et al., 2012 for details). Two emotion expressions (happy and fearful) were chosen for each body image, so that in total six female and six male bodies (three happy and three fearful) were used.

Participants were presented with two Eriksen and Flanker tasks (Eriksen and Eriksen, 1974) focusing either on the emotion expression or on the gender of the stimuli. Each task consisted of two blocks, one presenting face stimuli and the other one presenting body stimuli (**Figure 1**). The order of presentation of both tasks and types of stimuli were counterbalanced between participants. Face and body stimuli were displayed on a dark background. A trial displayed an array of three stimuli, each one subtending a visual angle of 2.5° × 3.5° for faces and 2.5° × 5.6° on average for bodies, considering a viewing distance of about 60 cm from the computer screen. A visual angle of 3° wat set between the center of the target and the center of each flanker. Faces and bodies differ in terms of structure, which is closed for faces and open for bodies. Moreover, while emotion expressions do change the disposition of bodily parts within space (e.g., opening/up-ward vs. closing/down-ward posture for expressing happy vs. fearful emotions, respectively), this is not the case for faces. Given this structural difference, the width of body images including the portion of space covered by legs and arms (2.7 cm on average) was manipulated to match the width of face images (2.6 cm). Participants were asked to indicate the emotion (emotion task) or the gender (gender task) of the face/body displayed in the middle of the array (i.e., the target) while ignoring the ones displayed at the side (i.e., the flanker). Emotional and gender features across the array of stimuli were presented in four different combinations: emotion different/gender different, emotion different/gender same, emotion same/gender different, emotion same/gender same. Thus, in the emotion task, emotional features were task-relevant and gender features were task-irrelevant; conversely, in the gender task, emotional features were task-irrelevant and gender features were task-relevant. Participants provided their response with a left/right mouse click using their thumbs. Response key assignment was counterbalanced among participants. The array of stimuli appeared on the screen for 500 ms, followed by a blank screen displayed for 2,000 ms. A time-window of 500 ms for stimuli presentation has been proven sufficient to elicit an emotional congruency effect in a flanker paradigm with facial expressions (Holthausen et al., 2016). The arrival of a new trial was signaled by a fixation cross of 400–600 ms appearing in the center of the screen. Participants could provide their response from the beginning of the trial till the presentation of the fixation cross.

Statistical Analysis and Sample-Size Calculation

Inverse efficiency scores (IES), computed by dividing mean response times by proportion of correct responses per subject and per condition, were selected as the main dependent measure. Integrating mean response times and correct responses into a single dependent variable represents a suitable option to appropriately weighs the impact of speed and accuracy, particularly in case of a negative correlation between the two measures (Bruyer and Brysbaert, 2011). For all the experiments, trials with response times below 100 ms were excluded from further analysis. A 2 × 2 × 2 × 2 ANOVA on IES was conducted with TASK (emotion vs. gender recognition), STIMULI (face vs. body) EMOTION CONGRUENCY (congruent vs. incongruent emotion) and GENDER CONGRUENCY (congruent vs. incongruent gender) as within-subjects variables.

In case of significant main or interaction effects, follow-up pairwise comparisons were executed adopting the Duncan's

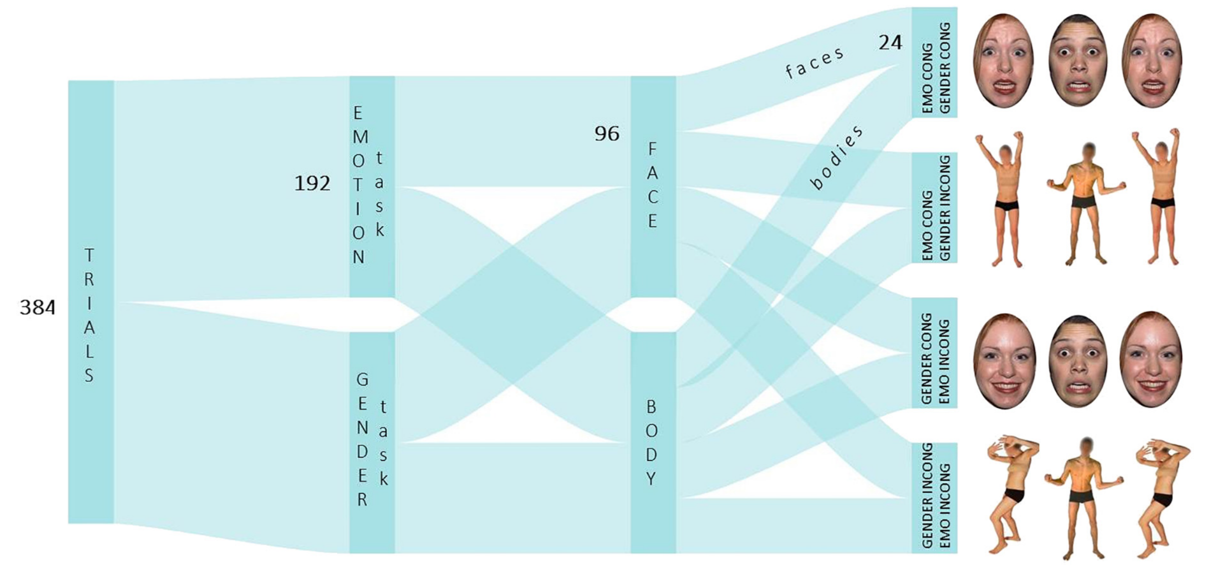


FIGURE 1 | Schematic depiction of the task structure in Experiment 1 and Experiment 2.

Multiple Range Test. The alpha level of significance was set at 0.05 (two-tailed).

On the basis of a 2 × 2 × 2 × 2 repeated measure ANOVA on IES, considering a medium-low effect size of 0.25 – as reported in a published study using a similar paradigm (Zhou and Liu, 2013) – and applying an alpha level of 0.05, a sample of 24 participants was defined as adequate to achieve a power of 0.80 (1-Beta).

Results

The four-way repeated-measure ANOVA on IES yielded significant main effects of TASK [$F(1,23) = 7.8$, $p < 0.02$, $\eta^2_p = 0.25$], STIMULI [$F(1,23) = 25.5$, $p < 0.0001$, $\eta^2_p = 0.53$] and a significant effect of EMOTION CONGRUENCY [$F(1,23) = 4.8$, $p = 0.04$, $\eta^2_p = 0.17$]. Overall, participants were less efficient in emotion discrimination than gender discrimination and in processing body than face stimuli. Furthermore, they showed a tendency to be less efficient to respond to the central target when the lateral flanker had an incongruent vs. congruent emotion (mean IES in emotion congruent = 626.4 ms, SEM = 10.7 vs. incongruent = 636.5 ms, SEM = 10.8). This occurred in both the emotion-discrimination and the gender-discrimination tasks (interaction TASK × EMOTION CONGRUENCY: [$F(1,23) < 1$]), thus when the (in)congruency was either relevant or irrelevant to the task, and for both the body and the face stimuli (interaction STIMULUS × EMOTION CONGRUENCY: [$F(1,23) < 1$]). The main effect of GENDER CONGRUENCY [$F(1,23) < 1$] and any other interactions were non-significant [for all $F(1,23) < 3$, $p > 0.1$]. Accuracy and RTs among conditions are reported in **Table 1**.

Discussion

In this flanker task, participants were asked to recognize either the emotion (happy/fearful) or the gender of the target stimulus while ignoring the distractors, which could match or not either the task-relevant features (e.g., emotional features in emotion discrimination) or the task-irrelevant ones (e.g., gender features in emotion discrimination). The results suggested that, despite the participants had to focus on the central target and ignore the lateral flanker, the congruency between the emotional valence of the target and flanker stimuli impacted both the emotion and gender discrimination performance (**Figure 2**). Thus, in contrast with Zhou and Liu (2013)'s results, we found evidence for intrusion of the emotional dimension of a stimulus independently from attentional modulation. Conversely, non-emotional feature conflict (i.e., gender incongruence) did not affect performance when either relevant (i.e., gender discrimination) or irrelevant for the task at hand (**Figure 2**). This result is inconsistent with previous findings showing, firstly, the presence of both emotional and gender congruency effects – with participants being slower in processing incongruent trials – when features are task-relevant and, secondly, that both these effects are attenuated when congruency became task-irrelevant (Zhou and Liu, 2013). It should be noted that the lack of interaction effects between task and congruency may be attributed to a statistical power issue. In fact, one can fairly assume that increasing the number of trials from 192 to 256, as in Zhou and Liu (2013)'s work, may result not only in a significant main effect of emotion but also in a significant interaction between emotional congruency and task, which would indicate the presence of attention modulation. This notwithstanding, the authors also observed that the attenuation of the congruency effect in the task-irrelevant condition was less pronounced for emotion than gender. This finding drove them to the conclusion that, even though attentional control can exert a top–down modulation on emotion processing, yet emotional saliency can be modulated by attention resources to a lesser degree than non-emotional contents.

TABLE 1 | Mean RTs and Accuracy for each experimental condition in Experiment 1.

		Emotion Recognition				Gender Recognition			
		Face		**Body**		**Face**		**Body**	
		RT (ms)	***Acc %***	***RT (ms)***	***Acc %***	***RT (ms)***	***Acc %***	***RT (ms)***	***Acc %***
Emotion features	**Gender features**								
Congruent	Congruent	594.88	0.97	642.88	0.94	554.06	0.98	609.72	0.95
	Incongruent	597.14	0.97	622.66	0.94	555.14	0.98	599.31	0.93
Incongruent	Congruent	603.01	0.96	637.74	0.95	555.00	0.98	622.47	0.92
	Incongruent	600.20	0.95	635.10	0.93	546.85	0.97	608.75	0.94

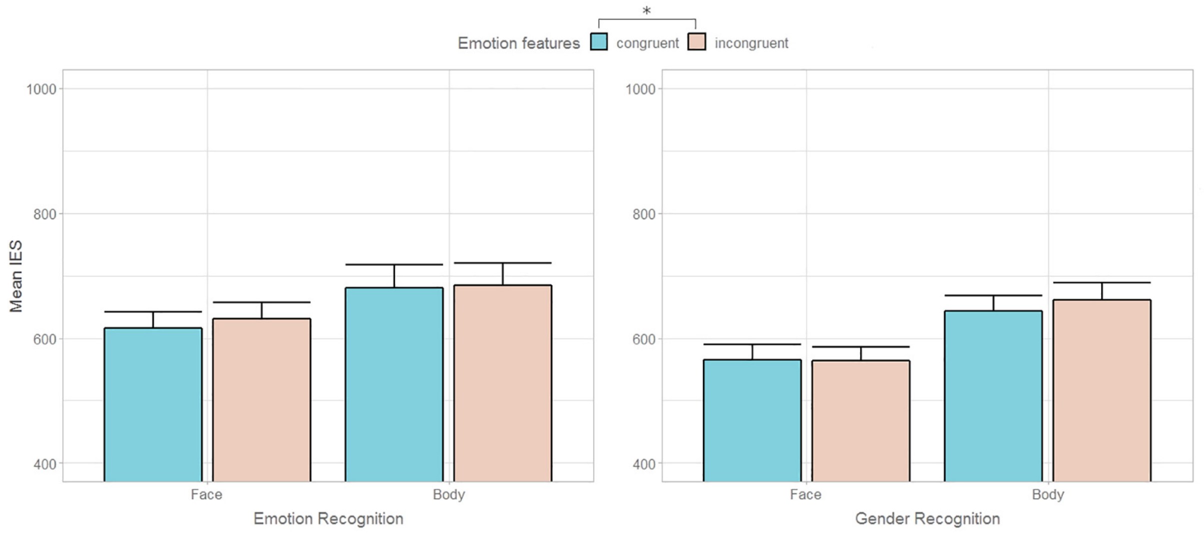


FIGURE 2 | Mean inverse efficiency scores (IES) during Emotion and Gender recognition according to task-relevant **(left graph)** and task-irrelevant **(right graph)** emotional congruency between target and flanker in Experiment 1. Error bars represent + 1 SEM. The asterisk ∗ indicates the main effect of emotion congruency ($p < 0.05$).

Then, how can the discrepancy between our findings and previous evidence be explained? In both Zhou and Liu (2013)'s and our data, gender recognition was easier than emotion recognition, thus arguing against the possibility that the asymmetry between emotion and gender conflict could be due to a mismatch in task difficulty. Furthermore, interference from automatic processing is expected to be higher for more easily encoded and, thus, more salient features (Atkinson et al., 2005; Gandolfo and Downing, 2019). In contrast, we found that the more difficult-to-encode dimension (i.e., emotion) interfered with the easier one (i.e., gender), thus making it unlikely, albeit not ruling out, that the asymmetry between emotion and gender conflict was due to differential speed of processing. A possible answer to the discrepancy between our results and those of Zhou and Liu (2013) may be found in the choice of the type of stimuli. While in the present experiment real face stimuli (and real bodies) were applied, Zhou and Liu used computer-generated faces. Despite little evidence is available on the resemblance between computer-generated and real faces, the emotion recognition of virtual (Dyck et al., 2008) or iconic (Kendall et al., 2016) and of real or photographic face images may be different. Moreover, the use of multiple identities in the present study, as opposed to the adoption of a single face identity in Zhou and Liu (2013)'s study, may explain the discrepancy between the two findings. In fact, displaying different identities presumably increased the variability of low-level features of the stimuli and may have challenged the processing of their gender.

Another possible answer may call into play a temporal factor. Zhou and Liu displayed target and flanker for 1,400 ms, while we presented them for 500 ms. Even though the mean response times reported by the researchers seem comparable to the response times measured in the present experiment, having posed a more strict time-constraint, as in this case, may have allowed only emotional conflict to be processed enough for exerting bottom-up interference on the main task. This is in keeping with the finding of an earlier modulation of ERP cortical components exerted by emotional than by non-emotional feature conflict (Zhou et al., 2016). Based on these considerations, it can be hypothesized that reducing the exposure time to stimuli and focusing on an early decoding time window may have increased the chance to observe emotional features over non-emotional features conflict, at least when participants had to focus their attention on the central target and completely ignore the objects at the side. However, the different weight of emotional or non-emotional conflict on attentional control may change when participants have to focus on both the central target and the lateral flanker to perform a gender or emotion comparison, while ignoring the irrelevant dimension (i.e., emotion or gender, respectively), thus tapping on feature-based rather the space-based attention allocation mechanisms.

EXPERIMENT 2

Experiment 2 was aimed at examining whether expanding the focus of attention across the visual field may modulate the interplay between attentional control and emotional and non-emotional feature processing. To this aim, we used the same stimuli from Experiment 1 in a same-or-different judgment paradigm in which participants had to match the emotion or the gender of the central and lateral stimuli. Thus, while the participants' main task focused on the task-relevant congruency of the stimuli (i.e., emotion congruency in the emotion task and gender congruence in the gender task), we tested the effects of the task-irrelevant dimension (i.e., gender congruency in the emotion task and emotion congruency in the gender task). We expected that, different from Experiment 1, both gender and emotional task-irrelevant incongruence between the lateral and central stimuli should affect the main task when task relevant. Furthermore, in keeping with Experiment 1, the same effects should be obtained for body and face stimuli.

Participants

Twenty-four healthy volunteers (10 men, $M = 28.6$ $SD = 5$ years old) participated in Experiment 2. One participant was excluded from analyses due to below-chance accuracy in one of the tasks. All participants had normal or corrected-to-normal vision. Prior to the beginning of the experiment, written informed consent was obtained from all participants. The experiment was approved by the local ethical committee and conducted in accordance with the declaration of Helsinki.

Stimuli, Task and Analysis

The same stimuli, task structure, procedure and data handling as in Experiment 1 were used in Experiment 2. However, in this experiment participants were asked to perform two same-or-different judgment tasks focusing either on the emotion expression or on the gender of the stimuli. Participants indicated whether the emotion (emotion task) or the gender (gender task) of the stimulus displayed in the middle of the array matched or not that of the stimuli at the sides.

Results

A 2 (TASK, emotion vs. gender comparison) × 2 (STIMULI, face vs. body) × 2 (task-irrelevant CONGRUENCY) repeated measure ANOVA on IES yielded significant main effect of STIMULI [$F(1,22) = 20.95$, $p < 0.001$, $\eta^2_p = 0.49$], significant interaction effects of TASK*STIMULI [$F(1,22) = 10.52$, $p < 0.005$, $\eta^2_p = 0.32$] and of TASK*STIMULI*CONGRUENCY [$F(1,22) = 8.09$, $p < 0.01$, $\eta^2_p = 0.27$]. No other significant main nor interaction effect emerged. The analysis showed that participants were less efficient in processing bodily stimuli, similarly to what emerged in Experiment 1, and that this effect emerged during the gender congruency comparison (mean IES of face = 1101.0 vs. body = 1385.3 vs.; $p < 0.001$). Indeed, during the emotion congruency comparison no difference between type of stimuli emerged (face = 1215.3 body = 1252.6; $p < 0.41$). The three-way interaction revealed that participants were less efficient in processing body gender congruency in emotional (i.e., task-irrelevant) incongruent trials (emotion congruent = 1345.9 vs. emotion incongruent = 1424.6; $p < 0.02$). Conversely, no difference emerged for face gender congruency processing according to emotional features congruency (emotion congruent = 1103.6 vs. emotion incongruent = 1098.4; $p < 0.77$). No other significant difference emerged during emotion congruency comparison according to gender feature congruency for face expressions (gender congruent = 1201.5 vs. gender incongruent = 1229.1; $p < 0.33$) nor body expressions (gender congruent = 1266.6 vs. gender incongruent = 1238.7; $p < 0.21$). Accuracy and RTs among conditions are reported in **Table 2**.

Discussion

Results of Experiment 2 showed that, on the one hand, even broadening the focus of attention to include the lateral stimuli did not trigger a bottom-up interference of non-emotional feature conflict (i.e., gender incongruence) on the main task. It is worth noting that, in Experiment 1, gender recognition was overall easier than emotion recognition, leaving open the possibility that the asymmetry in the relative influence of gender and emotional conflict could arise from different task difficulty. Conversely, in Experiment 2, the need to compare the central and the flanker stimuli increased the task demands for both the gender and the emotion features, and no difference between the two tasks emerged. This clears out the even remote scenario that greater interference effects are exerted by more difficult to encode features. On the other hand, results showed a differential influence of emotional features on an implicit level between faces and bodies. Indeed, while participants were less efficient in discriminating the gender in emotional incongruent than congruent trials when processing bodies, this difference was not found in response to faces (**Figure 3**). Alternative explanations may account for this result. First of all, it cannot be excluded that hampered gender comparison in emotional incongruent trials was driven by "intrinsically" postural, rather than emotional features. A static body is generally recognized as fearful in presence of specific anatomical characteristics, such as abdominal rotation, shoulders ad/abduction or backwards transfer of body weight (Coulson, 2004). As morphologic, sexually dimorphic cues as the waist-to-hip ratio or the shoulders' width have been demonstrated to be crucial for gender categorization (Lippa, 1983; Johnson and Tassinary, 2005), one can fairly assume that a bodily posture partially hiding or distorting these cues may hamper the discrimination of its gender. Therefore, the possibility that fearful, as opposed to happy, expressions may have interfered with gender comparison for the distortion of relevant sexually dimorphic cues rather than

TABLE 2 | Mean RTs and Accuracy for each experimental condition in Experiment 2.

	Emotion Comparison					**Gender Comparison**			
	Face		**Body**			**Face**		**Body**	
	RT (ms)	***Acc %***	***RT (ms)***	***Acc %***		***RT (ms)***	***Acc %***	***RT (ms)***	***Acc %***
Gender features					**Emotion features**				
Congruent	1033.05	0.87	1112.01	0.89	Congruent	1004.48	0.90	1070.40	0.80
Incongruent	1041.58	0.86	1109.65	0.90	Incongruent	1004.87	0.89	1094.69	0.78

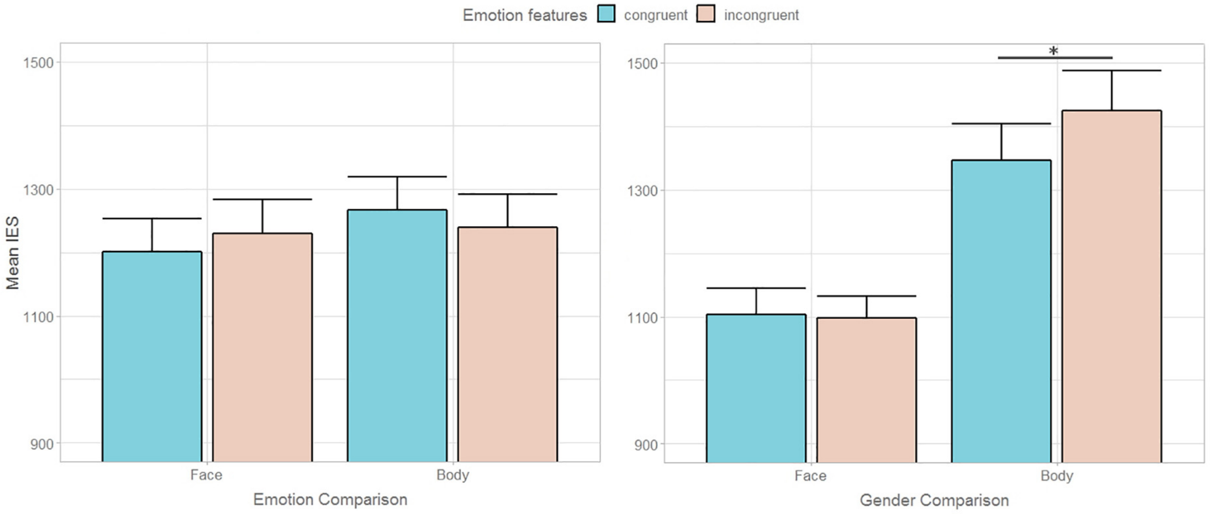


FIGURE 3 | Mean inverse efficiency scores (IES) during Emotion and Gender comparison according to task-relevant **(left graph)** and task-irrelevant **(right graph)** emotional congruency between central and lateral stimuli in Experiment 2. Error bars represent + 1 SEM.

for their emotional valence cannot be dismissed. Moreover, the different degree of variation in morphologic features between facial and bodily emotional expressions, more pronounced in the latter, may explain the differential effect between the two types of stimuli.

EXPERIMENT 3

A third control experiment was carried out specifically to examine the possibility that postural rather than emotional features may have affected gender comparison when task-irrelevant in Experiment 2. Here, participants were presented with images of bodies with an opening/up-ward or closing/down-ward posture with no emotional valence, selected with the purpose of resembling the approaching or avoidant movements associated with the expression of happiness or fear respectively. If the effect observed in Experiment 2 was likely to be driven by the emotional salience of the stimuli, then perceptual characteristics of body postures should not interfere with gender comparison processing. Conversely, a similar pattern of results as in Experiment 2 was expected if posture, rather than emotional valence, explained the task irrelevance interference in Experiment 2.

Participants

Twenty-four healthy volunteers (8 males, $M = 29$ $SD = 4.2$ years old) participated in this experiment. All participants had normal or corrected-to-normal vision. Prior to the beginning of the experiment, written informed consent was obtained from all participants. The experiment was approved by the local ethical committee and conducted in accordance with the declaration of Helsinki.

Stimuli, Task and Analysis

The same task structure, procedure and data handling of Experiment 1 and 2 were used in Experiment 3. However, in this control experiment only blocks with body images were used. Here, participants were asked to perform two same-or-different judgment tasks focusing either on the posture or on the gender of the stimuli. Two types of whole-body movements without emotional valence were selected: opening/up-ward and closing/down-ward body movements (**Figure 4**). Selected stimuli resembled the approaching or avoidant dynamics generally associated with the expression of happiness and fear respectively. Participants indicated whether the type of movement or the gender of the target stimulus matched with the stimuli at the side according to the task-relevant feature to be attended.

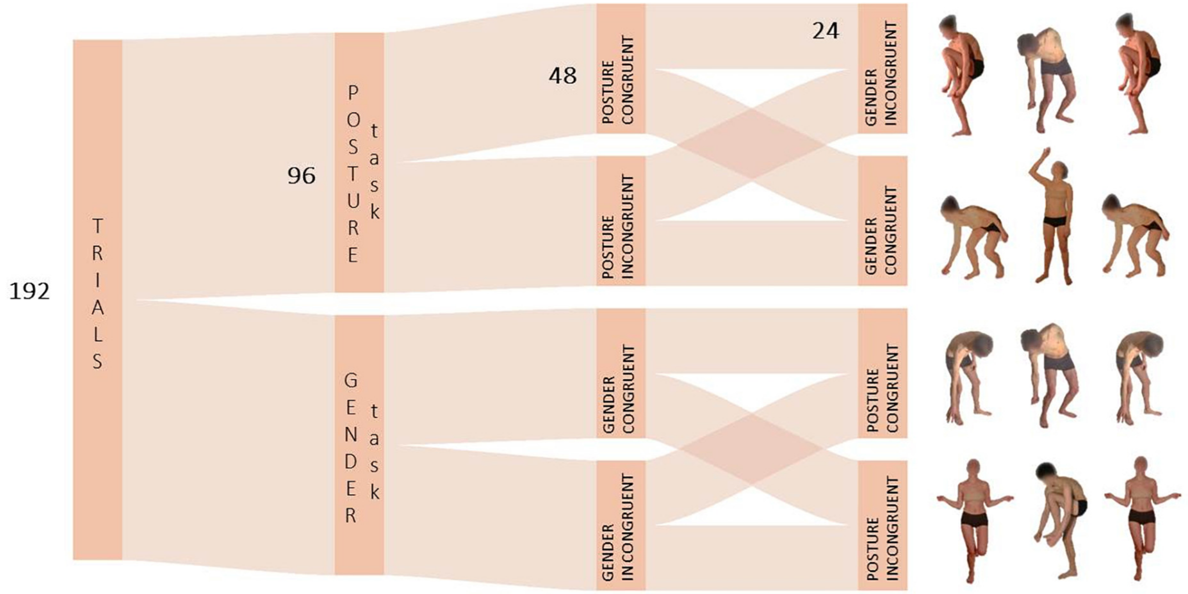


FIGURE 4 | Schematic depiction of the task structure in Experiment 3. In the Posture Task, the feature to be attended was the opening/up-ward vs. closing/down-ward movement of bodies with no emotional valence.

Results

A 2 (TASK, posture vs. gender comparison) × 2 (task-irrelevant CONGRUENCY) repeated measure ANOVA on IES yielded significant main effect of TASK [$F(1,23) = 51.91$, $p < 0.00001$, $\eta^2_p = 0.69$], indicating low efficiency in making gender as compared to postural features judgments. No other significant main nor interaction effect emerged. Accuracy and RTs among conditions are reported in **Table 3**.

Discussion

Results of Experiment 3 revealed that body posture did not influence gender comparison, suggesting that findings of Experiment 2 may indeed reflect an intrusion of information of emotional nature on the main task. In fact, even though opening/up-ward and closing/down-ward bodies were comparable to happy/fearful bodily expressions in terms of visibility of perceptual sexual cues, participants did not perform poorly in postural incongruent trials as compared to congruent trials (**Figure 5**). Simply put, emotional information and not postural features of body expressions could be held responsible for having interfered with gender comparison in Experiment 2. Conversely, task-irrelevant gender congruency did not affect posture comparison here, as it did not affect emotion comparison in Experiment 2.

TABLE 3 | Mean RTs and Accuracy for each experimental condition in Experiment 3.

	Gender Comparison			Posture Comparison	
	RT (ms)	*Acc %*		*RT (ms)*	*Acc %*
Posture features			**Gender features**		
Congruent	1153.63	0.77	Congruent	932.82	0.92
Incongruent	1158.49	0.78	Incongruent	910.62	0.91

GENERAL DISCUSSION

In the present study we tested the role of attentional modulation on the processing of bodily and facial emotion and gender features. We found that, while attention was focused on a central stimulus (flanker task; Experiment 1), the emotion features of two unattended lateral faces or bodies intruded in the main task, when either task relevant or task irrelevant. Conversely, the distracting gender features were successfully filtered out, in both the task relevant or task-irrelevant conditions, reducing the impact of ongoing mechanisms of conflict resolution. This suggests that the attentional control system can suppress distracting information that is void of any emotional valence via space- and/or feature-based filtering mechanisms. It can be speculated that filtering of flanker gender in Experiment 1 underpinned a space-based attention mechanism, confining attentional resources to a specific region of interest, selecting the stimuli displayed in that region and hindering this way the erroneous selection of distractors. Yet, our findings suggest that the emotional valence of a face or body stimulus can overcome this spatial attention filtering, in keeping with the evidence of reduced extinction of controlesional emotional vs. neutral faces (Vuilleumier et al., 2002) or bodies (Tamietto et al., 2007) in brain-lesion patients.

When the focus of attention was manipulated to include the processing of lateral stimuli (same-or-different task; Experiment 2 and 3), disabling what can be referred to as a proactive

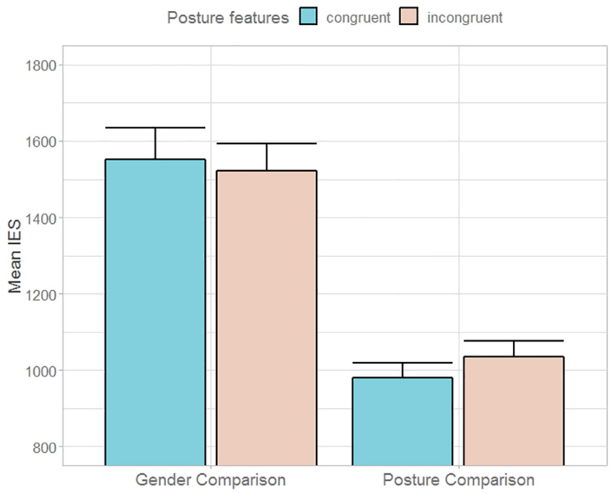


FIGURE 5 | Mean inverse efficiency scores (IES) during Gender and Posture comparison according to task-irrelevant postural congruency between central and lateral stimuli in Experiment 3. Error bars represent + 1 SEM.

control relying on space-based selection of input, only emotional expressions (Experiment 2) but not neutral postures (Experiment 3) of bodies affected the processing of their gender when task-irrelevant. However, task-irrelevant gender features of both faces and bodies and emotional features of faces were filtered out and did not intrude in the main task at hand. This may be due to the action of a feature-based filtering, operating a selection of the properties rather than (only) the location of the input. Both mechanisms were capable of an efficient filtering selection of non-emotional feature of both body and faces. However, the present findings suggest that the space-based mechanism, likely active in the Flanker task in Experiment 1, was pervious to emotional information conveyed by both faces and bodies, which both showed a comparable interference on the main task when the emotional dimension was either task relevant or irrelevant. Conversely, the feature-based mechanism was more effective in filtering out task-irrelevant emotional features of attended facial expressions, but it was pervious to task-irrelevant emotional information conveyed by body postures. This seems to point to a greater resilience of body than face emotional processing to feature-based filtering selection.

A question arises, thus, regarding the differential influence of implicit processing of emotional information between faces and bodies. At first, one may argue that bodily expressions conveyed emotion information more efficiently than faces. However, not only both types of stimuli were selected from validated databases among stimuli associated with the highest rates of recognition accuracy, but recent behavioral studies also demonstrated that the recognition rate of emotional expressions is similar for faces and bodies (de Gelder et al., 2014). Furthermore, the main effect of stimuli in both Experiments 1 and 2 may lead to consider the modulatory role of task difficulty to be at the basis of the discrepancy between faces and bodies. Indeed, participants were significantly less efficient when processing body as compared to face images. Nevertheless, behavioral indices of emotional interference have been reported at very different levels of cognitive engagement (Lim et al., 2008; Muller et al., 2008). Moreover, a meta-analytic investigation found the association between magnitude of the effect of emotional information intrusion on attention control and task difficulty to be non-significant (Carretié, 2014). Finally, while a differential performance in discriminating or matching faces and bodies was present in both Experiments 1 and 2, task-irrelevant emotional features of faces intruded in the main task in Experiment 1 but not in Experiment 2.

Alternatively, another explanation for this result calls into play the specific role of bodies in the communication of emotions. Understanding emotions expressed by other individuals is crucial to adapt our behavior in the physical and social world. Indeed, it has been suggested that emotional body images may boost the activation of brain regions involved both in emotional information processing and in the perception of our bodily experience (e.g., insula and somatosensory cortex) more than images of faces or affective pictures in general (de Gelder, 2006). This presumably occurs because, while emotional facial expression may trigger different responses, as they can signal the presence of a salient stimulus or, alternatively, invite other individuals to empathize, a body expression of emotional valence constitutes a more direct cue of how to cope with the environment. For example, compared to faces, fearful bodies provide not only information of the emotional state of the individual, but also additional information on how to cope with the threat that has been signaled through specifications of the postural schema.

Congruently, the processing of emotional cues conveyed by whole-body postures and movements implies the activation of a broad brain network, including not only regions typically involved in emotion perception, but also regions part of the motor system (Urgesi et al., 2006; de Gelder et al., 2014). Specifically, emotional bodily expressions have been demonstrated to automatically activate areas involved in action representation and preparation as compared to meaningful but emotionally neutral bodily postures (de Gelder et al., 2004). Furthermore, a motor response is elicited by bodily expressions even when presented in the blinded hemifield of patients with brain lesions (Tamietto et al., 2015) or in the unattended visual field of patients with hemispatial neglect (Tamietto et al., 2007, 2009). This piece of evidence supports that emotional body signals may be prioritized due to a processing bias, presumably because of their critical importance for survival.

Accordingly, evidence from brain stimulation studies showed altered corticospinal excitability, which might reflect action readiness to environmental input, during passive viewing of emotional images as compared to non-emotional images (Olivieri et al., 2003; Schutter et al., 2008; van Loon et al., 2010). A later brain stimulation study confirmed these findings using body images, strengthening the link between processing of emotional body expressions and action observation and preparation. Borgomaneri et al. (2012) observed a reduced cortico-spinal excitability in the left hemisphere during emotion categorization of both happy and fearful bodily expressions as compared to neutral body images. While this work did not find a differential effect of negative and positive emotion valences (but see Borgomaneri et al., 2015), it allowed controlling for the confounding effect of perceived motion in influencing early motor response, showing a differential effects for emotional and non-emotional postures despite comparable implied-motion effects. This is in keeping with our findings that emotional expressions, rather than postures, were driving the implicit, task-irrelevant interference on gender discrimination. Thus, as previously shown for affective pictures (Bannerman et al., 2009), emotional body language can capture attention automatically.

Previous research points to a prioritized processing of emotional stimuli over non-emotional ones. Nevertheless, there are studies that failed in finding a clear modulation of emotion processing on attention resources. This lack of modulation effect may be due to various factors, such as methodologic characteristics of paradigms and their ability to elicit distinct top–down regulation strategies, or type of stimuli. Accordingly, while a proactive top–down regulation was pervious to intrusion via bottom–up processing of emotional information conveyed by bodies and faces in the Flanker task used in Experiment 1, the link between emotional salience and action readiness may explain why, in Experiment 2, emotional information interfered with feature-based down-regulation mechanisms only when conveyed by bodies and not by faces.

The possible role of emotional valence and social relevance of the stimuli in attentional modulation may be relevant for exploring the mechanism of disordered attention. Indeed, compared to a consistent body of literature on the modulatory role of phobic and anxiety disorders on perceptual emotion processing (Sussman et al., 2016; Wante et al., 2016), limited evidence has been gathered so far on the influence of attention control and inhibition difficulties. A study on children and adolescents diagnosed with attention deficit and hyperactivity disorder (ADHD) showed that, during a digit categorization task, emotional distractors were associated with longer reaction time than neutral distractors in the clinical sample, while the very same difference was not observed in the control group (López-martín et al., 2013). Moreover, at an electrophysiological level this finding was accompanied with an increase in the amplitude of the event-related potential (ERP) component N200 for emotional as compared to neutral distractors, again only in the ADHD group. Likewise, other pathologies associated with attention problems may underpin an altered interaction between top–down inhibitory control and bottom-up affective processes. In a rehabilitation perspective, further research is still needed to shed a light on this interaction in abnormal neuropsychological profiles, in order to disentangle the reciprocal impact of attentional modulation on emotional processing.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Scientific Institute IRCCS E. Medea, Bosisio Parini, Italy. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

VO and CU conceived and designed the experiments, and analyzed and interpreted the data. AB and GP contributed to the design of the experiments. VO performed the experiments and wrote the original draft of the manuscript. VO, AB, GP, and CU reviewed and edited the manuscript. All authors contributed to the final approval of the manuscript.

FUNDING

This study was supported by the “5 per mille” funds 2018/2019 for biomedical research to AB and by the Italian Ministry of Health (Ricerca Corrente 2019 – Scientific Institute, IRCCS E. Medea).

REFERENCES

Abend, R., De Voogd, L., Salemink, E., Wiers, R. W., Pérez-edgar, K., Fitzgerald, A., et al. (2018). Association between attention bias to threat and anxiety symptoms in children and adolescents. *Depress. Anxiety* 35, 229–238. doi: 10.1002/da.22706.Association

Astudillo, C., Muñoz, K., and Maldonado, P. E. (2018). Emotional content modulates attentional visual orientation during free viewing of natural images. *Front. Hum. Neurosci.* 12:459. doi: 10.3389/fnhum.2018.00459

Atkinson, A. P., Tipples, J., Burt, D. M., and Young, A. (2005). Asymmetric interference between sex and emotion in face perception. *Percept. Psychophys.* 67, 1199–1213. doi: 10.3758/bf03193553

Bannerman, R. L., Milders, M., de Gelder, B., and Sahraie, A. (2009). Orienting to threat: faster localization of fearful facial expressions and body postures revealed by saccadic eye movements. *Proc. R. Soc. B* 276, 1635–1641. doi: 10.1098/rspb.2008.1744

Barratt, D., and Bundesen, C. (2012). Attentional capture by emotional faces is contingent on attentional control settings. *Cogn. Emot.* 26, 1223–1237. doi: 10.1080/02699931.2011.645279

Becker, D. V. (2017). Facial gender interferes with decisions about facial expressions of anger and happiness. *J. Exp. Psychol. Gen.* 146, 457–463. doi: 10.1037/xge0000279

Bishop, S. (2008). Neural mechanisms underlying selective attention to threat. *Ann. N. Y. Acad. Sci.* 152, 141–152. doi: 10.1196/annals.1417.016

Borgomaneri, S., Gazzola, V., and Avenanti, A. (2012). Motor mapping of implied actions during perception of emotional body language. *Brain Stimul.* 5, 70–76. doi: 10.1016/j.brs.2012.03.011

Borgomaneri, S., Vitale, F., and Gazzola, V. (2015). Seeing fearful body language rapidly freezes the observer ’ s motor cortex. *Cortex* 65, 232–245. doi: 10.1016/j.cortex.2015.01.014

Bruce, V., Burtonu, A. M., Hanna, E., Healey, P., and Mason, O. (1993). Sex discrimination: how do we tell the difference between male and female faces? *Perception* 22, 131–152. doi: 10.1068/p220131

Bruyer, R., and Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: is the inverse efficiency score (ies) a better dependent variable than the mean reaction time (rt) and the percentage of errors (pe)? *Psychol. Belg.* 51, 5–13.

Burra, N., and Kerzel, D. (2019). Task demands modulate effects of threatening faces on early perceptual encoding. *Front. Psychol.* 10:02400. doi: 10.3389/fpsyg.2019.02400

Carretié, L. (2014). Exogenous (automatic) attention to emotional stimuli: a review. *Cogn. Affect. Behav. Neurosci.* 14, 1228–1258. doi: 10.3758/s13415-014-0270-2

Carretié, L., Albert, J., López-martín, S., and Tapia, M. (2009). Negative brain: an integrative review on the neural processes activated by unpleasant stimuli. *Int. J. Psychophysiol.* 71, 57–63. doi: 10.1016/j.ijpsycho.2008.07.006

Cooper, R. M., Law, A. S., and Langton, S. R. H. (2013). Looking back at the stare-in-the-crowd effect: staring eyes do not capture attention in visual search. *J. Vis.* 13, 1–22. doi: 10.1167/13.6.10

Coulson, M. (2004). Attributing emotion to static body postures: recognition accuracy, confusions, and viewpoint dependence. *J. Nonverval Behav.* 28, 117–139. doi: 10.1023/b:jonb.0000023655.25550.be

de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nat. Rev. Neurosci.* 7, 242–249. doi: 10.1038/nrn1872

de Gelder, B., de Borst, A. W., and Watson, R. (2014). The perception of emotion in body expressions. *WIREs Cogn. Sci.* 6, 149–158. doi: 10.1002/wcs.1335

de Gelder, B., Snyder, J., Greve, D., Gerard, G., and Hadjikhani, N. (2004). Fear fosters flight: a mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proc. Natl. Acad. Sci. U.S.A.* 101, 16701–16706. doi: 10.1073/pnas.0407042101

Dolan, R. J., and Vuilleumier, P. (2003). Automaticity and the amygdala: nonconscious responses to emotional faces. *Ann. N. Y. Acad. Sci.* 1, 348–355. doi: 10.1111/j.1749-6632.2003.tb07093.x

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

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149. doi: 10.3758/BF03203267

Gandolfo, M., and Downing, P. E. (2019). Perceiving emotion and sex from the body: evidence from the Garner task for independent processes. *Cogn. Emot.* 33, 1–11. doi: 10.1080/02699931.2019.1634003

Gilboa-Schechtman, E., Ben-Artzi, E., Jeczemien, P., Marom, S., and Hermesh, H. (2004). Depression impairs the ability to ignore the emotional aspects of facial expressions: evidence from the Garner task. *Cogn. Emot.* 18, 209–231. doi: 10.1080/02699930341000176

Haas, S. A., Amso, D., and Fox, N. A. (2016). The effects of emotion priming on visual search in socially anxious adults. *Cogn. Emot.* 31, 1041–1054. doi: 10.1080/02699931.2016.1180281

Harris, D. A., Hayes-skelton, S. A., and Ciaramitaro, V. M. (2016). What ' s in a Face? How face gender and current affect influence perceived emotion. *Front. Psychol.* 7:1468. doi: 10.3389/fpsyg.2016.01468

Holthausen, B. S., Regenbogen, C., Turetsky, B. I., and Habel, U. (2016). The influence of task-irrelevant flankers depends on the composition of emotion categories. *Front. Psychol.* 7:712. doi: 10.3389/fpsyg.2016.00712

Horstmann, G., and Bauland, A. (2006). Search asymmetries with real faces: testing the anger-superiority effect. *Emotion* 6, 193–207. doi: 10.1037/1528-3542.6.2.193

John, Y. J., Zikopoulos, B., Bullock, D., and Barbas, H. (2016). The emotional gatekeeper: a computational model of attentional selection and suppression through the pathway from the amygdala to the inhibitory thalamic reticular nucleus. *PLoS Comput. Biol.* 12:e1004722. doi: 10.1371/journal.pcbi.1004722

Johnson, K., and Tassinary, L. G. (2005). Perceiving sex directly and indirectly meaning in motion and morphology. *Psychol. Sci.* 16, 890–897. doi: 10.1111/j.1467-9280.2005.01633.x

Kendall, L. N., Raffaelli, Q., Kingstone, A., and Todd, R. M. (2016). Iconic faces are not real faces: enhanced emotion detection and altered neural processing as faces become more iconic. *Cogn. Res. Princ. Implic.* 1:19. doi: 10.1186/s41235-016-0021-8

Kim, H., Somerville, L. H., Johnstone, T., Polis, S., Alexander, A. L., Shin, L. M., et al. (2004). Contextual modulation of amygdala responsivity to surprised faces. *J. Cogn. Neurosci.* 16, 1730–1745. doi: 10.1162/0898929042947865

Klein, A. M., Van Niekerk, R., Rapee, R. M., Hudson, J. L., Bogels, S. M., Becker, E. S., et al. (2017). Biases in attention, interpretation, memory, and associations in children with varying levels of spider fear: inter-relations and prediction of behavior. *J. Behav. Ther. Exp. Psychiatry* 54, 285–291. doi: 10.1016/j.jbtep.2016.10.001

Lim, S., Padmala, S., and Pessoa, L. (2008). Affective learning modulates spatial competition during low-load attentional conditions. *Neuropsychologia* 46, 1267–1278.

Lippa, R. (1983). Sex typing and the perception of body outlines. *J. Pers.* 51, 667–682. doi: 10.1111/j.1467-6494.1983.tb00873.x

Lobue, V. (2009). More than just another face in the crowd: superior detection of threatening facial expressions in children and adults. *Dev. Sci.* 12, 305–313. doi: 10.1111/j.1467-7687.2008.00767.x

López-martín, S., Albert, J., Fernández-jaén, A., and Carretié, L. (2013). Emotional distraction in boys with ADHD: neural and behavioral correlates. *Brain Cogn.* 83, 10–20. doi: 10.1016/j.bandc.2013.06.004

Mackiewicz, K. L., Kaiser, R. H., Deprince, A. P., and Banich, M. T. (2017). General and emotion-specific alterations to cognitive control in women with a history of childhood abuse. *NeuroImage Clin.* 16, 151–164. doi: 10.1016/j.nicl.2017.06.030

Macnamara, A., and Proudfit, G. H. (2014). Cognitive load and emotional processing in generalized anxiety disorder: electrocortical evidence for increased distractibility. *J. Abnorm. Psychol.* 123, 557–565. doi: 10.1037/a0036997

Maunsell, J. H. R., and Treue, S. (2006). Feature-based attention in visual cortex. *Trends Cogn. Sci.* 29, 317–322. doi: 10.1016/j.tins.2006.04.001

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

Minnebusch, D. A., and Daum, I. (2009). Neuropsychological mechanisms of visual face and body perception. *Neurosci. Biobehav. Rev.* 33, 1133–1144. doi: 10.1016/j.neubiorev.2009.05.008

Moro, V., Pernigo, S., Avesani, R., Bulgarelli, C., Urgesi, C., Candidi, M., et al. (2012). Visual body recognition in a prosopagnosic patient. *Neuropsychologia* 50, 104–117. doi: 10.1016/j.neuropsychologia.2011.11.004

Muller, M., Søren, K., and Keil, A. (2008). Time course of competition for visual processing resources between emotional pictures and foreground task. *Cereb. Cortex* 18, 1892–1899. doi: 10.1093/cercor/bhm215

Müsch, K., Siegel, M., Engel, A. K., and Schneider, T. R. (2017). Gamma-band activity reflects attentional guidance by facial expression. *NeuroImage* 146, 1142–1148. doi: 10.1016/j.neuroimage.2016.09.025

Nordström, H., and Wiens, S. (2012). Emotional event-related potentials are larger to figures than scenes but are similarly reduced by inattention. *BMC Neurosci.* 13:49. doi: 10.1186/1471-2202-13-49

Ohman, A., Flykt, A., and Esteves, F. (2001). Emotion drives attention: detecting the snake in the grass. *J. Exp. Psychol.* 130, 466–478. doi: 10.1037/AXJ96-3445.130.3.466

Olivieri, M., Babiloni, C., Filippi, M., Caltagirone, C., Babiloni, F., Cicinelli, P., et al. (2003). Influence of the supplementary motor area on primary motor cortex excitability during movements triggered by neutral or emotionally unpleasant visual cues. *Exp. Brain Res.* 149, 214–221. doi: 10.1007/s00221-002-1346-8

Padmala, S., Sambuco, N., Codispoti, M., and Pessoa, L. (2018). Attentional capture by simultaneous pleasant and unpleasant emotional distractors. *Emotion* 18, 1189–1194. doi: 10.1037/emo0000401

Park, G., and Thayer, J. F. (2014). From the heart to the mind: cardiac vagal tone modulates top-down and bottom-up visual perception and attention to emotional stimuli. *Front. Psychol.* 5:278. doi: 10.3389/fpsyg.2014.00278

Peelen, M. V., and Downing, P. E. (2019). Selectivity for the human body in the fusiform gyrus. *J. Neurophysiol.* 93, 603–608. doi: 10.1152/jn.00513.2004

Pessoa, L., Japee, S., and Sturman, D. (2006). Target visibility and visual awareness modulate amygdala responses to fearful faces. *Cereb. Cortex* 16, 366–375. doi: 10.1093/cercor/bhi115

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

Reed, C. L., Stone, V. E., Grubb, J. D., and Mcgoldrick, J. E. (2006). Turning configural processing upside down: part and whole body postures. *J. Exp. Psychol. Hum. Percept. Perform.* 32, 73–87. doi: 10.1037/0096-1523.32.1.73

Righart, R., and de Gelder, B. (2007). Impaired face and body perception in developmental prosopagnosia. *Proc. Natl. Acad. Sci. U.S.A.* 104, 17234–17238. doi: 10.1073/pnas.0707753104

Schimmack, U. (2005). Attentional interference effects of emotional pictures: threat, negativity, or arousal? *Emotion* 5, 55–66. doi: 10.1037/1528-3542.5.1.55

Schutter, D. J. L. G., Hofman, D., and Van Honk, J. (2008). Fearful faces selectively increase corticospinal motor tract excitability: a transcranial magnetic stimulation study. *Psychophysiology* 45, 345–348. doi: 10.1111/j.1469-8986.2007.00635.x

Shechner, T., Jarcho, J. M., Wong, S., Leibenluft, E., Pine, D. S., and Nelson, E. E. (2017). Threats, rewards, and attention deployment in anxious youth and adults: an eye tracking study. *Biol. Psychol.* 122, 121–129. doi: 10.1016/j.biopsycho.2015.10.004

Sussman, T. J., Heller, W., Miller, G. A., and Mohanty, A. (2013). Emotional distractors can enhance attention. *Psychol. Sci.* 24, 2322–2328. doi: 10.1177/0956797613492774.Emotional

Sussman, T. J., Jin, J., and Mohanty, A. (2016). Top-down and bottom-up factors in threat-related perception and attention in anxiety. *Biol. Psychol.* 121, 160–172. doi: 10.1016/j.biopsycho.2016.08.006

Tamietto, M., Castelli, L., Vighetti, S., Perozzo, P., Geminiani, G., and Weiskrantz, L. (2009). Unseen facial and bodily expressions trigger fast. *Proc. Natl. Acad. Sci. U.S.A.* 106, 17661–17666. doi: 10.1073/pnas.0908994106

Tamietto, M., Cauda, F., Celeghin, A., Diano, M., Costa, T., Cossa, F. M., et al. (2015). Once you feel it, you see it: insula and sensory-motor contribution to visual awareness for fearful bodies in parietal neglect. *Cortex* 62, 56–72. doi: 10.1016/j.cortex.2014.10.009

Tamietto, M., Geminiani, G., Genero, R., and De Gelder, B. (2007). Seeing fearful body language overcomes attentional deficits in patients with neglect. *J. Cogn. Neurosci.* 19, 445–454. doi: 10.1162/jocn.2007.19.3.445

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.The

Urgesi, C., Moro, V., Candidi, M., and Aglioti, S. M. (2006). Mapping implied body actions in the human motor system. *J. Neurosci.* 26, 7942–7949. doi: 10.1523/JNEUROSCI.1289-06.2006

van de Riet, W. A., Grèzes, J., and de Gelder, B. (2009). Specific and common brain regions involved in the perception of faces and bodies and the representation of their emotional expressions. *Soc. Neurosci.* 4, 37–41. doi: 10.1080/17470910701865367

van Loon, A., van den Wildenberg, P., van Stengeren, A., Hajcak, G., and Ridderinkhof, K. (2010). Emotional stimuli modulate readiness for action: a transcranial magnetic stimulation study. *Cogn. Affect. Behav. Neurosci.* 10, 174–181. doi: 10.3758/CABN.10.2.174

Victeur, Q., Huguet, P., and Silvert, L. (2019). Attentional allocation to task-irrelevant fearful faces is not automatic: experimental evidence for the conditional hypothesis of emotional selection selection. *Cogn. Emot.* doi: 10.1080/02699931.2019.1622512 [Epub ahead of print].

Vuilleumier, P., Armony, J. L., Clarke, K., Husain, M., Driver, J., and Dolan, R. J. (2002). Neural response to emotional faces with and without awareness: event-related fMRI in a parietal patient with visual extinction and spatial neglect. *Neuropsychologia* 40, 2156–2166. doi: 10.1016/s0028-3932(02)00045-3

Wante, L., Mueller, S. C., Demeyer, I., de Raedt, R., and Braet, C. (2016). The role of interference and inhibition processes in dysphoric early adolescents. *J. Clin. Child Adoles. Psychol.* 47, 608–619. doi: 10.1080/15374416.2015.1102068

Wendt, M., Kähler, S. T., Luna-rodriguez, A., and Jacobsen, T. (2017). Adoption of task-specific sets of visual attention. *Front. Psychol.* 8:687. doi: 10.3389/fpsyg.2017.00687

Zhou, P., and Liu, X. (2013). Attentional modulation of emotional conflict processing with flanker tasks. *PLoS One* 8:e60548. doi: 10.1371/journal.pone.0060548

Zhou, P., Yang, G., Nan, W., and Liu, X. (2016). The time course of attentional modulation on emotional conflict processing. *Cogn. Emot.* 30, 621–637. doi: 10.1080/02699931.2015.1020051

Conflict of Interest: 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.

frontiers in Psychology

BRIEF RESEARCH REPORT
published: 31 January 2020
doi: 10.3389/fpsyg.2020.00097

How Do Induced Affective States Bias Emotional Contagion to Faces? A Three-Dimensional Model

Andrés Pinilla[1]†, Ricardo M. Tamayo[2]† and Jorge Neira[2]*

[1] Quality and Usability Lab, Institute of Software Engineering and Theoretical Computer Science, Faculty of Electrical Engineering and Computer Science, Technische Universität Berlin, Berlin, Germany, [2] Laboratorio de Cognición Implícita, Departamento de Psicología, Universidad Nacional de Colombia, Bogotá, Colombia

Affective states can propagate in a group of people and influence their ability to judge others' affective states. In the present paper, we present a simple mathematical model to describe this process in a three-dimensional affective space. We obtained data from 67 participants randomly assigned to two experimental groups. Participants watched either an upsetting or uplifting video previously calibrated for this goal. Immediately, participants reported their baseline subjective affect in three dimensions: (1) positivity, (2) negativity, and (3) arousal. In a second phase, participants rated the affect they subjectively judged from 10 target angry faces and ten target happy faces in the same three-dimensional scales. These judgments were used as an index of participant's affective state after observing the faces. Participants' affective responses were subsequently mapped onto a simple three-dimensional model of emotional contagion, in which the shortest distance between the baseline self-reported affect and the target judgment was calculated. The results display a double dissociation: negatively induced participants show more emotional contagion to angry than happy faces, while positively induced participants show more emotional contagion to happy than angry faces. In sum, emotional contagion exerted by the videos selectively affected judgments of the affective state of others' faces. We discuss the directionality of emotional contagion to faces, considering whether negative emotions are more easily propagated than positive ones. Additionally, we comment on the lack of significant correlations between our model and standardized tests of empathy and emotional contagion.

Keywords: emotional contagion, facial expressions, evaluative space model, affective states, faces, emotions

OPEN ACCESS

Edited by:
Noemi Mazzoni,
University of Trento, Italy

Reviewed by:
Guillaume Dezecache,
Institut Jean Nicod, France
Elaine Hatfield,
University of Hawaii, United States

***Correspondence:**
Andrés Pinilla
andres.pinilla@qu.tu-berlin.de

†ORCID:
Andrés Pinilla
orcid.org/0000-0002-0812-7896
Ricardo M. Tamayo
orcid.org/0000-0002-8678-0145

Specialty section:
This article was submitted to
Consciousness Research,
a section of the journal
Frontiers in Psychology

Received: *31 July 2019*
Accepted: *14 January 2020*
Published: *31 January 2020*

Citation:
Pinilla A, Tamayo RM and Neira J (2020) How Do Induced Affective States Bias Emotional Contagion to Faces? A Three-Dimensional Model. Front. Psychol. 11:97. doi: 10.3389/fpsyg.2020.00097

INTRODUCTION

Emotions frequently involve reactions to events and people, influencing in turn our judgments about the affective states of others. Various prominent theoretical frameworks (Hatfield et al., 1993, 1994; Forgas, 1995; Peters and Kashima, 2015) comprise at least one type of psychological mechanism by which humans and other mammalians automatically converge to the affective states of their peers when it is socially appropriate (de Waal, 2012; Palagi et al., 2015). This mechanism has been defined as "a tendency to automatically mimic and synchronize expressions, vocalizations, postures and movements with those of another person's, and, consequently, to converge emotionally" (Hatfield et al., 1994, p. 5).

Evidence of automatic affective convergence for facial mimicry (Dimberg, 1982; Dimberg et al., 2000; Dimberg and Thunberg, 2012; Varcin et al., 2019), vocalizations (Cappella and Planalp, 1981; Fujiwara and Daibo, 2016) and body postures (Condon and Ogston, 1966; Schmidt et al., 2011), support the existence of an automatic mechanism for emotional contagion. Of course, automatic transmission of affect among dyads (Bruder et al., 2012; Butler, 2015) and beyond (Dezecache et al., 2013; Kramer et al., 2014) is not the only psychological process that modulates people's affective responses. Clearly, additional mechanisms linked to individual differences (Doherty, 1997; Künecke et al., 2014), specific appraisals (Fontaine et al., 2007), pragmatic goals (Fischer and Hess, 2017) and social roles (Murray, 1933) mediate the propagation of emotions. In fact, previous work has significantly contributed to clarify basic concepts in the field. For instance, clear differences exist today between deeply related concepts such as mood, affect, and emotion (Forgas, 1995; Russell, 2003; Peters and Kashima, 2015) and between related explanatory frameworks such as affect diffusion, emotional contagion, and affect infusion (Dezecache et al., 2015; Peters and Kashima, 2015).

However, to our knowledge, no specific framework exists to quantify the amount of subjective emotional contagion among individuals. Usually, emotional contagion is assumed by observing whether participants' affective responses (i.e., subjective ratings) or implicit automatic reactions (i.e., facial electromyographic responses) converge to the valence of previously presented stimuli (e.g., faces, videos, sounds, scenes, actors, or peers). In our view, an ideal framework for quantifying emotional contagion should describe how much affect is transferred from the participant's immediate emotional context to subsequent emotional responses.

We believe that such a model could contribute to understand the interaction between explicit and implicit cognitive mechanisms involved in emotional contagion. The need for a specific mathematical model for quantifying subjective emotional contagion might be illustrated by contemporary research concerned with the influence of automatic facial mimicry responses – predominantly related to implicit mechanisms – on subjective and affective ratings – predominantly related to explicit mechanisms – (e.g., Fujimura et al., 2010; Dezecache et al., 2013; Sato et al., 2013; Künecke et al., 2014; Deng and Hu, 2018). For instance, Deng and Hu (2018), collected subjective ratings of valence in a single bipolar scale together with electromyography (EMG) recordings from zygomaticus mayor (ZM) and corrugator supercilii (CS) muscles, from participants exposed to positive (happy) and negative (angry) faces. In two experiments, the authors found a lack of correlation between ZM activity and subjective ratings in response to positive stimuli. Similarly, in a conceptually analogous manipulation of transitive affect, Dezecache et al. (2013) found that explicit identifications of positive faces were not above chance, while ZM activity increased for positive faces (Dezecache et al., 2013). Furthermore, results from two experiments conducted by Fridlund (1991) support the claim that implicit facial expressions may be erratically associated with explicit subjective reports of emotional experience – for a detailed review, see Barrett et al. (2019). We believe these results are puzzling because subjective measures either collected in a single bipolar scale (Deng and Hu, 2018) or as discrete identifications (Dezecache et al., 2013) do not provide enough information to determine the amount of subjective contagion experienced by participants. In our view, research could benefit from disentangling the amount of contagion triggered by the stimuli from the amount of contagion that bias further affective judgments, pondering that a single bipolar measure of valence might not fully describe the subjective and affective processes of emotional contagion experienced by individuals. For instance, as suggested in the evaluative space model (ESM) (Cacioppo et al., 1997), increases in negative affect are not necessarily proportional to decreases in positive affect, and baseline subjective responses to positive stimuli usually offset baseline activation for negative stimuli. Therefore, using a single bipolar measure of valence as an index of subjective affect could have obscured genuine correlations between automatic mimicry and subjective affect in previous studies.

In the present paper, we propose a simple mathematical model to quantify emotional contagion to faces by computing the distance between two emotional coordinates in a three-dimensional affective space. The first coordinate represents the affective state that participants subjectively self-reported after observing a video. The second coordinate represents subjective judgments about the affective state of unfamiliar faces. Given that people tend to ascribe aspects of their own affective state to others (Forgas, 1995; Van Boven and Loewenstein, 2003), we assume that the distance between the first and second coordinates, indexes the amount of emotional contagion to faces experienced by participants. This is, how much convergence is there between participant's affective states before (T1) and after seeing the faces (T2). The measure at T1 is the evaluation of the video and the measure at T2 is the evaluation of each face. We interpret participant's evaluation of faces as an index of their affective state after watching each face, based on previous evidence suggesting that evaluation of other's facial expressions correlates with self-reported affective states (Blairy et al., 1999). Additional evidence shows that activation of the amygdala during a face evaluation-task correlates with participant's subjective evaluations of angry faces (Nomura et al., 2004). Therefore, subjective judgments about the affective state of other's faces can be considered an index of participants' affective state. This evidence is in turn consistent with predictions from the Affect as Information mechanism proposed in the Affect Infusion Model (AIM) (Forgas, 1995), which suggests that people tend to use their own emotional state as heuristic to attribute emotions to others.

In short, we propose a model to quantify the distance between two different *subjective evaluations*. The first one, induced by a video, the second one, involved in the evaluation of unfamiliar faces. In other words, the model calculates the distance between participants' self-reported affective state (triggered by the video) and the influence exerted by this affective state on the emotional contagion to unfamiliar faces. Shorter distances indicate more emotional contagion to faces and larger distances indicate the opposite.

Computing the shortest distance between these two affective states presupposes an efficient psychological mechanism for emotional contagion based on *minimal effort* (Forgas, 1995) or *energetic efficiency* (Cacioppo et al., 1997). This mechanism is conceived as an organism's tendency to adopt the shortest path and least effortful strategy to yield a subjective affective response, if it satisfies minimal adaptive requirements. We preferred a three-dimensional affective space based on monopolar ratings similar to the one suggested by Cacioppo et al. (1997) for two reasons. Firstly, because contemporary evidence suggests that a three-dimensional model does not exclude a bi-dimensional one (Mattek et al., 2017) and that monopolar measures do not preclude additional affective dimensions (e.g., dominance in the first case or net predisposition in the second case). Second, a three-dimensional space maps well into commonly used psychophysiological measures such as EMG activity from ZM and CS muscles for positivity and negativity, respectively, and electrodermal activity (EDA) or heart rate variability (HRV) for arousal (Bernat et al., 2006).

To extend the measurement framework for subjective responses of emotional contagion, we designed an experiment in which participants were induced to either a positive or a negative affective state by watching an affectively laden video clip. Immediately, we assessed their baseline subjective emotional state by asking them to self-report how the video made them feel in three dimensions: positivity, negativity, and arousal. In a second phase, participants were asked to judge the affective states of unfamiliar faces using the same three dimensions. As mentioned above, given that people tend to use their own affective states as a heuristic to judge the affective state of others (Forgas, 1995; Van Boven and Loewenstein, 2003), we reasoned that participant's evaluations of faces could index the amount of emotional convergence between the affective state exerted by the video and the affective state associated to each face. If emotional contagion is automatic and transitive (Dezecache et al., 2013), then it should facilitate emotional convergence by spreading the affective states triggered by the video to the affective state involved in the participant's judgments of unfamiliar faces. Therefore, based on the principle of least effort, we hypothesized that participants induced to a negative affective state would show more emotional contagion to angry than happy faces, while participants induced to a positive affective state would show more contagion to happy than angry faces.

Finally, we wanted to explore the relationship between our suggested measure of subjective emotional contagion and key psychometric measures of emotional contagion and empathy. Consequently, participants responded to the Emotional Contagion Scale (ECS; Doherty, 1997; Gouveia et al., 2007), the Interpersonal Reactivity Index (IRI; Davis, 1980; Pérez-Albéniz et al., 2003), the Questionnaire of Prosocial Conduct (QPC; Martorell et al., 2011), and the Basic Empathy Scale (BES; Jolliffe and Farrington, 2006; Merino-Soto and Grimaldo-Muchotrigo, 2015).

MATERIALS AND METHODS

Participants

Sixty-five undergraduate students from the National University of Colombia participated in the study. Their average age was 21.6 years old (*SD* = 3.7); 50.7% were women. Thirty-two participants were randomly assigned to the negative induction group and 35 to the positive induction group. Boxplots revealed outliers for two participants, whose data were excluded from further analysis. All participants provided written informed consent prior to participating in the experiment, in accordance with the ethical principles of the Declaration of Helsinki.

Materials

Inquisit 4 software was used to record subjective ratings and present all instructions, surveys and prompts on a 23-inch (1920 × 1980) computer screen. The approximate physical distance between the participant and the screen was 20 inches. Two scenes from the FilmStim (Schaefer et al., 2010) were used for emotional induction: *Dead Poets Society*[1] and *American History X*. Twenty images of the Pictures Of Facial Affect (POFA; Ekman, 1993) were used (10 happy and 10 angry). Affective responses to these stimuli and a similar set of stimuli were calibrated for the present population (see repository with **Supplementary Material**[2]). Additionally, the ECS (Doherty, 1997; Gouveia et al., 2007) was used to estimate the correlation between the proposed measurement model with individual differences. Finally, three empathy questionnaires adapted to Spanish were used: IRI (Davis, 1980; Pérez-Albéniz et al., 2003), QPC (Martorell et al., 2011), and BES (Jolliffe and Farrington, 2006; Merino-Soto and Grimaldo-Muchotrigo, 2015)[3].

Procedure

The experiment had four phases. In the first phase, participants were induced to either negative or positive affective states by watching a positive or a negative video clip taken from the FilmStim database (Schaefer et al., 2010) according to the randomly assigned experimental group. Immediately after watching the video, the following question was presented to participants: "How did this scene make you feel?" There were three Likert scales bellow the question. The scales had one label at each side and ranged from 0 to 100. In the scale that measured the negativity dimension, the labels were "0 – not bad at all," and "100 – very bad"; in the scale that measured the positivity dimension, the labels were "0 – not good at all," and "100 – very good"; in the scale that measured the arousal dimension, the labels were "0 – not restless at all," and "100 – very restless." In the present article, all prompts and instructions were translated from Spanish. To maximize participants' attention to the questions, their order randomly changed for each trial.

---

[1] The FilmStim database offers two video clips taken from *Dead Poets Society*. We used the second one.

[2] Videos and faces were quantitatively analyzed according to the assumptions of ESM (Cacioppo et al., 1997) during previous experiments. Datasets are available at: https://osf.io/52uj8/.

[3] Questionnaire responses were analyzed and made available at: https://osf.io/52uj8/.

The labels were red for the negativity dimension, green for the positivity dimension, and blue for the arousal dimension. The colors of labels remained constant throughout the experiment and were used to help participants discriminate the dimension to evaluate in each prompt.

In the second phase, 20 images of faces taken from the POFA were shown in random order for each participant (10 positive and 10 negative). The randomization compensated possible carryover effects produced by the faces. Before each image, a gray screen with a white cross in the middle was shown for 1 s. Every face was shown for 3 s. After each face, three questions with the same statement were shown. The statement was: "How do you think this person feels?" Participants answered using three Likert monopolar scales, just like the ones presented after the video. The scales had two labels (one in each side). The labels were different for each question. In the question that measured the negativity dimension, the labels were "0 – not bad at all," and "100 – very bad"; in the question that measured the positivity dimension, the labels were "0 – not good at all," and "100 – very good"; and in the question that measured the arousal dimension, the labels were "0 – not restless at all," and "100 – very restless."

In the third phase, participants were asked if they had previously seen the video clip from FilmStim (Schaefer et al., 2010), and two surveys used for the creation of the FilmStim (Schaefer et al., 2010) were applied as a manipulation check. These were the Positive and Negative Affect Schedule (PANAS) (Watson et al., 1988) and a self-reported emotional arousal scale (Schaefer et al., 2010).

Finally, in the fourth phase, participants were asked to answer the questionnaires in the following order: IRI, QPC, BES, and ECS.

Emotional Contagion Equation

We hypothesized that exposition to positive and negative videos could bias emotional contagion to happy versus angry faces. Therefore, we estimated the shortest distance between two affective coordinates (video and face) in a three-dimensional Euclidean space. The emotional distance between the affective state induced by the video and the subjective evaluation of the faces can be expressed using the equation:

$$ec = 1 - \left(\frac{\sqrt{\left(n_v - n_f\right)^2 + \left(p_v - p_f\right)^2 + \left(a_v - a_f\right)^2}}{100 \times \sqrt{3}} \right) \qquad (1)$$

where n_v, p_v, and a_v represent participants' affective state induced by the video in the negativity, positivity, and arousal dimensions, respectively, and n_f, p_f, and a_f represent participants' evaluation of the faces in the same dimensions. The value inside the parenthesis is subtracted from one because (a) it represents emotional distance analogously to the ESM (Cacioppo et al., 1997) and (b) because we assume that emotional contagion to faces is the inverse of this distance. To facilitate data analysis, we rescaled values to a range from 0 to 1. Thus, we divided everything inside the square root by the maximum possible distance between the two affective coordinates.

RESULTS

Data were analyzed using a 2 × 2 mixed ANOVA, with negative and positive emotional induction as between-group factor, happy and angry faces as within-group factor. The dependent variable was the magnitude of emotional contagion to faces, as defined in Eq. 1. Outliers were detected in two participants whose data were excluded from further analysis. Data were normally distributed for all cells of the experimental design, as assessed by Shapiro–Wilk test ($p > 0.05$). The assumption of homogeneity of variances was met, according to Brown–Forsythe test ($p > 0.05$). A statistically significant two-way interaction was found, $F(1, 65) = 103.957$, $p < 0.001$, $\eta_p^2 = 0.615$. A main effect for type of face was found, $F(1, 132) = 4.766$, $p = 0.031$, $\eta_p^2 = 0.035$. A paired-samples t-test was conducted to compare the magnitude of emotional contagion to happy and angry faces in each group. A Bonferroni correction was applied. Emotional contagion was significantly greater for angry ($M = 0.76$, $SD = 0.08$) than happy faces ($M = 0.41$, $SD = 0.18$) in the negative group, $t(31) =$, $p < 0.001$. The opposite tendency was found in the positive group, where emotional contagion was significantly lower for angry ($M = 0.54$, $SD = 0.15$) than for happy faces ($M = 0.71$, $SD = 0.13$), $t(34) =$, $p < 0.001$. No main effect for type of emotional induction was found, $F(1, 65) = 3.353$, $p = 0.072$, $\eta_p^2 = 0.049$.

We were interested in analyzing the effect of gender on emotional contagion to faces. This required a hierarchical model because our sample size differed between each group (Huta, 2014, p. 21). Therefore, a linear mixed effects model was run to analyze whether gender had an effect on emotional contagion to faces. The fixed effects of the model where type of emotional induction and gender. The random effects were the intercepts for participants and type of face, as well as by-participant and by-type-of-face random slopes for the effect of emotional contagion. This model was compared to another model that had the same parameters but did not have a fixed effect for gender. Results of both models were compared using a Chi-squared test. Given that no significant differences were found, we infer that participant's gender had no significant effect on emotional contagion to faces ($\chi^2(1) = 1.68$, $p = 0.195$).

Responses about familiarity with the videos were recorded for only 42 participants, due to a technical problem. Therefore, we excluded this question from the analysis. Otherwise, the manipulation check was successful. Overall, participants' evaluations (see **Table 1**) were similar to the scores reported in the FilmStim (Schaefer et al., 2010).

Results from the questionnaires were compared with the emotional contagion ratings. We predicted that participants with higher emotional contagion ratings would be more empathetic, as assessed with the empathy questionnaires, or more prone to emotional contagion, as assessed with the ECS (Doherty, 1997). However, we did not find this result (see **Table 2**). A Pearson correlation was performed between emotional contagion to faces and each questionnaire separately. No Bonferroni correction or similar was applied. There was a significant negative correlation between emotional contagion toward happy faces in the negative induction group and the IRI scores, $r(30) = -0.40$, $p = 0.02$ and

TABLE 1 | Summary of responses in the manipulation check and reported scores in the FilmStim (Schaefer et al., 2010) for the two videos used in the experiment.

	Experiment score	FilmStim score
American History X		
PA (positive affect)	2.55	2.04
NA (negative affect)	2.45	2.73
Arousal	5.44	5.84
Dead Poets Society		
PA (positive affect)	3.01	2.82
NA (negative affect)	1.56	1.21
Arousal	4.8	5.66

Positive and negative affect are measured using a five-point scale, taken from the PANAS (Watson et al., 1988), while arousal is measured using a seven-point scale, proposed by the authors of the FilmStim (Schaefer et al., 2010).

ECS scores (Doherty, 1997), $r(30) = -0.43$, $p = 0.01$. In all the other cases correlations were not significant.

DISCUSSION

Our findings suggest that participants minimized the emotional distance between the affective state triggered by the video and the affective state triggered by unfamiliar faces. When induced to a negative affective state, participants judged both angry and happy faces closer to a negative affective state. Conversely, when induced to a positive affective state, participants judged both happy and angry faces closer to a positive affective state. Plainly, participants' induced affective state biased their emotional contagion to unfamiliar faces. Additionally, the results also show a double dissociation, happy and angry faces were differently evaluated by participants previously exposed to either an uplifting or an upsetting video.

In our view, these results show key differences in the way emotional contagion operates, when it is triggered by negative versus positive emotions. Negative affective states generate more bias than positive states. This can be observed in **Figure 1**, where the highest emotional contagion is observed for the negative induction group toward angry faces. This finding is consistent with previous research suggesting that negative affective states trigger more behavioral changes than positive ones (Cacioppo et al., 1997).

In this experiment, we used two types of measures to assess participants' affective states. A direct measure for the video and an indirect measure for the faces. The first was *direct* in the sense that participants were asked explicitly about their feelings after watching the video ("How did this scene make you feel?"). The second was *indirect* in the sense that participants were asked to rate the feelings they perceived from the faces ("How do you think this person feels?"). We used this later measure to estimate the amount of convergence between the affective states before and after the presentation of faces. As mentioned above, we decided to use an indirect (implicit) measure because (1) people tend to use their own affective state as a heuristic to judge others' affective states (Forgas, 1995; Van Boven and Loewenstein, 2003) and (2) an indirect measure potentially reduces the noise caused by individual explicit preferences. For instance, if the second question would have been a direct measure, i.e., "how does this person make you feel?" the corresponding answers might have been biased by personal liking, and metacognitive judgments about the self. Similarly, if the question were "how do you feel?" the corresponding answer would imply a permanent higher-order judgment about the self and not a temporary automatic influence exerted by video on the judgment of each face. This methodological approach was inspired by previous research suggesting that people tend to misattribute the sources of their affective states, and that implicit measures tap earlier emotional processing-stages than direct measures do (e.g., Payne et al., 2010; Payne and Lundberg, 2014). Therefore, if emotional contagion involves basic and early emotional processing-stages as originally suggested by Hatfield et al. (1993, 1994), then there are more chances to measure it by indirectly estimating the influence it exerts on implicit (indirect) measures.

Our results provide independent support to the claim that emotional contagion is not always a linear process (Dezecache et al., 2015). Negative emotions seem to propagate more than positive emotions, and other people's affective states are perceived more negatively when the observer is also in a negative affective state. This suggests that stronger emotional contagion occurs when a person, in a negative affective state, observes someone in a similar affective state.

This is interesting because previous research (Kramer et al., 2014) suggests that Facebook users tend to display more positive emotions when exposed to positive emotional content and more negative emotions when exposed to negative emotional content. Our results are consistent with these findings as well. Therefore, increasing exposition to negative emotional content might intensify negative emotions in the observer, which in turn increases sensitivity to emotional contagion to negative stimuli. Conversely, increasing exposition to positive emotional content

TABLE 2 | Summary of Pearson correlations between questionnaires (*p*-values in parenthesis) and emotional contagion assessed with Eq. 1.

Questionnaire	Negative induction group		Positive induction group	
	Happy faces	Angry faces	Happy faces	Angry faces
ECS	−0.43 (0.015)*	−0.28 (0.127)	−0.28 (0.101)	−0.02 (0.929)
IRI	−0.40 (0.024)*	−0.18 (0.330)	−0.21 (0.219)	−0.04 (0.809)
BES	−0.30 (0.096)	−0.30 (0.099)	−0.10 (0.549)	−0.10 (0.559)
QPC	−0.19 (0.307)	−0.26 (0.148)	−0.23 (0.180)	−0.08 (0.658)

$^*p < 0.05$.

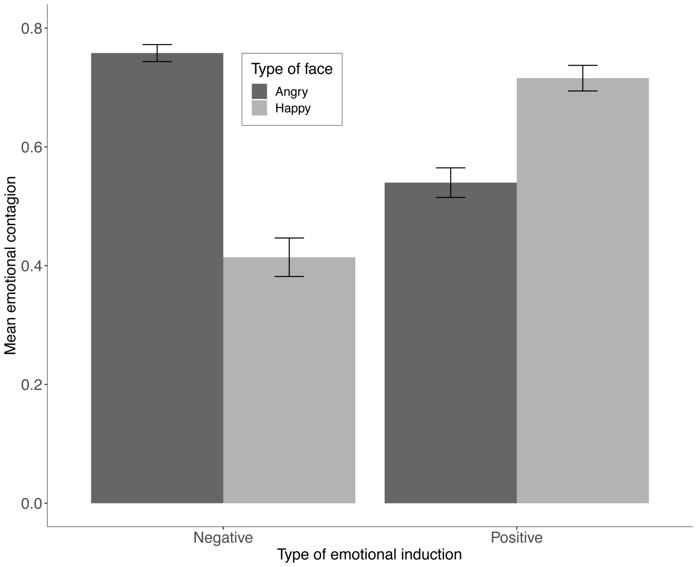


FIGURE 1 | Mean emotional contagion toward angry and happy faces for the groups exposed to a negative and a positive video clips, as assessed with Eq. 1. Error bars depict 95% CI.

might intensify positive emotions in the observer, magnifying emotional contagion to positive stimuli.

On the other hand, we did not find systematic correlations between psychometric questionnaires and our emotional contagion measure. In most cases, significant correlations were absent. In two cases where significant correlations were present, they were negative, suggesting that people displaying higher levels of empathy rated lower in our emotional contagion measure. However, the fact that these correlations were significant only for happy faces in the negative group (see **Table 2**) implies that those results cannot be generalized to all emotional contagion processes. In our view, these results are not completely surprising because questionnaires tap into a very different psychological process related to emotional contagion. Questionnaires are conceived as a personality-trait measure influenced by stable metacognitive judgments about the self in social situations. They are mainly based on explicit propositional knowledge (i.e., "I sometimes feel helpless when I am in the middle of a very emotional situation"; Davis, 1980). Instead, our measuring framework directly taps into specific momentarily states of emotional contagion hardly assessed by rationalized judgments about the self.

Based on our study, future experiments might investigate whether participants' affective states influence automatic imitation of facial expressions. Do people imitate negative faces faster or with greater accuracy when induced to a negative affective state? We believe that the answer would be affirmative. In those experiments, automatic emotional responses could be assessed using electrophysiological signals. Previous research points out that activation of CS and ZM muscles are associated with negative and positive emotions, respectively (Dimberg, 1982). Additional research has found that activation of the sympathetic system is associated with HRV (Appelhans and Luecken, 2006). Thus, activity of the CS-muscle, ZM-muscle, and HRV would be equivalent to the negativity, positivity, and arousal dimensions used in our study, which are similar to the dimensions of the ESM (Cacioppo et al., 1997). However, it is important to consider that facial expressions might not always correlate with self-reported subjective affective states (Fridlund, 1991). Yet, the magnitude of the dissociations between subjective (i.e., participant's ratings of positivity, negativity and arousal) and objective measures (e.g., EMG and HRV) could be assessed using a mathematical approach similar to Eq. 1. This approach would consist on calculating the emotional distance between subjective and objective emotional responses. This model would analyze individual differences in the magnitude of these dissociations, which in turn could help to quantify the relative contribution from each type of measure to the final affective state.

In short, our study provides evidence suggesting that people induced to a positive affective state show higher levels of emotional contagion to positive faces, while people induced to a negative affective state show higher levels of emotional contagion to negative faces. Furthermore, we provide evidence suggesting that subjective biases induced by current affective states are easily estimated by a simple mathematical model mapped onto a three-dimensional affective space. However, in scenarios where these two affective coordinates are similar due to factors not directly related to emotional stimulation (e.g., both states are neutral), the output of the model would indicate high emotional contagion, regardless of how much emotional change has been actually produced. This is a boundary condition of our model, which is exclusively useful for conditions where emotional contagion processes are reasonably assumed to be in operation.

Finally, the main contributions of this study are (1) to provide a measurement framework to analyze how affective states influence the directionality of emotional contagion and (2) to propose a methodological approach to analyze emotional contagion, not only as a binary outcome, but as a continuous quantitative variable.

AUTHOR'S NOTE

The dataset used in this article was previously used in the Master's Thesis of AP (Pinilla, 2017). This work was supervised by RT. The thesis is available at: http://bdigital.unal.edu.co/59467/1/1020759173.2017.pdf.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available at https://osf.io/52uj8/.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

RT and AP proposed the experimental design. JN worked on acquisition of the data. All authors analyzed and interpreted the results.

FUNDING

We acknowledge support by the Universidad Nacional de Colombia, Hermes Project No. 37456, the German Research Foundation, and the Open Access Publication Fund of TU Berlin.

ACKNOWLEDGMENTS

We appreciate help of Diego Useche formulating the mathematical model and acknowledge the effort of Deisy Valcárcel, Nicolás Buitrago, and Daniel Avilán gathering the data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.00097/full#supplementary-material

REFERENCES

Appelhans, B. M., and Luecken, L. J. (2006). Heart rate variability as an index of regulated emotional responding. *Rev. Gene. Psychol.* 10, 229–240. doi: 10.1037/1089-2680.10.3.229

Barrett, L. F., Adolphs, R. Marsella, S. Martinez, A. M. and Pollak, S. D. (2019). Emotional expressions reconsidered: challenges to inferring emotion from human facial movements. *Psychol. Sci. Public Interest* 20, 1–68. doi: 10.1177/1529100619832930

Bernat, E. Patrick, C. J., Benning, S. D., and Tellegen, A. (2006). Effects of picture content and intensity on affective physiological response. *Psychophysiology* 43, 93–103. doi: 10.1111/j.1469-8986.2006.00380.x

Blairy, S., Herrera, P., and Hess, U. (1999). Mimicry and the judgment of emotional facial expressions. *J. Nonverbal Behav.* 23, 5–41. doi: 10.1023/A:1021370825283

Bruder, M., Dosmukhambetova, D., Nerb, J., and Manstead, A. S. R. (2012). Emotional signals in nonverbal interaction: dyadic facilitation and convergence in expressions, appraisals, and feelings. *Cogn. Emot.* 26, 480–502. doi: 10.1080/02699931.2011.645280

Butler, E. A. (2015). Interpersonal affect dynamics: it takes two (and time) to tango. *Emot. Rev.* 7, 336–341. doi: 10.1177/1754073915590622

Cacioppo, J. T., Gardner, W. L., and Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: the case of attitudes and evaluative space. *Personal. Soc. Psychol. Rev.* 1, 3–25. doi: 10.1207/s15327957pspr0101_2

Cappella, J. N., and Planalp, S. (1981). Talk and silence sequences in informal conversations: III. Interspeaker influence. *Hum. Commun. Res.* 7, 117–132. doi: 10.1111/j.1468-2958.1981.tb00564.x

Condon, W. S., and Ogston, M. B. (1966). Sound film analysis of normal and pathological behavior patterns. *J. Nervous Ment. Dis.* 143, 338–347.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *Catalog Sel. Doc. Psychol.* 10, 1–17.

de Waal, F. B. M. (2012). The antiquity of empathy. *Science* 336, 874–876. doi: 10.1126/science.1220999

Deng, H., and Hu, P. (2018). Matching your face or appraising the situation: two paths to emotional contagion. *Front. Psychol.* 8:2278 doi: 10.3389/fpsyg.2017.02278

Dezecache, G., Conty, L., Chadwick, M., Philip, L., Soussignan, R., Sperber, D., et al. (2013). Evidence for unintentional emotional contagion beyond dyads. *PLoS One* 8:e67371. doi: 10.1371/journal.pone.0067371

Dezecache, G., Jacob, P., and Grèzes, J. (2015). Emotional contagion: its scope and limits. *Trends Cogn. Sci.* 19, 297–299. doi: 10.1016/j.tics.2015.03.011

Dimberg, U. (1982). Facial reactions to facial expressions. *Psychophysiology* 19, 643–647. doi: 10.1111/j.1469-8986.1982.tb02516.x

Dimberg, U., and Thunberg, M. (2012). Empathy, emotional contagion, and rapid facial reactions to angry and happy facial expressions: empathy and rapid facial reactions. *PsyCh J.* 1, 118–127. doi: 10.1002/pchj.4

Dimberg, U., Thunberg, M., and Elmehed, K. (2000). unconscious facial reactions to emotional facial expressions. *Psychol. Sci.* 11, 86–89. doi: 10.1111/1467-9280.00221

Doherty, R. W. (1997). The emotional contagion scale: a measure of individual differences. *J. Nonverbal Behav.* 21, 131–154. doi: 10.1023/A:1024956003661

Ekman, P.. (1993). *Pictures of Facial Affect (POFA)*. Washington, DC: Paul Ekman Group, Llc.

Fischer, A., and Hess, U. (2017). Mimicking emotions. *Curr. Opin. Psychol.* 17, 151–155. doi: 10.1016/j.copsyc.2017.07.008

Fontaine, J. R. J. Scherer, K. R., Roesch, E. B., and Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychol. Sci.* 18, 1050–1057. doi: 10.1111/j.1467-9280.2007.02024.x

Forgas, J. P. (1995). Mood and judgment: the affect infusion model (AIM). *Psychol. Bull.* 117, 39–66. doi: 10.1037/0033-2909.117.1.39

Fridlund, A. J. (1991). Sociality of solitary smiling: potentiation by an implicit audience. *J. Personal. Soc. Psychol.* 60, 229–240. doi: 10.1037/0022-3514.60.2.229

Fujimura, T., Sato, W., and Suzuki, N. (2010). Facial expression arousal level modulates facial mimicry. *Int. J. Psychophysiol.* 76, 88–92. doi: 10.1016/j.ijpsycho.2010.02.008

Fujiwara, K., and Daibo, I. (2016). Evaluating interpersonal synchrony: wavelet transform toward an unstructured conversation. *Front. Psychol.* 7:516. doi: 10.3389/fpsyg.2016.00516

Gouveia, V. V., Gouveia, R. S. V., Guerra, V. M., Santos, W. S., and de Medeiros, E.-D. (2007). Midiendo contagio emocional: adaptación de la escala de Doherty. *Rev. Psicol. Soc.* 22, 99–111. doi: 10.1174/021347407780705401

Hatfield, E., Cacioppo, J. T., and Rapson, R. L. (1993). Emotional contagion. *Curr. Dir. Psychol. Sci.* 2, 96–100. doi: 10.1111/1467-8721.ep10770953

Hatfield, E., Cacioppo, J. T., and Rapson, R. L. (1994). *Emotional Contagion*. Cambridge: Cambridge University Press.

Huta, V. (2014). When to use hierarchical linear modeling. *Q. Methods Psychol.* 10, 13–28. doi: 10.20982/tqmp.10.1.p013

Jolliffe, D., and Farrington, D. P. (2006). Development and validation of the basic empathy scale. *J. Adolesc.* 29, 589–611. doi: 10.1016/j.adolescence.2005.08.010

Kramer, A. D. I., Guillory, J. E.., and Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proc. Natl. Acad. Sci.* 111, 8788–8790. doi: 10.1073/pnas.1320040111

Künecke, J., Hildebrandt, A., Recio, G., Sommer, W., and Wilhelm, O. (2014). Facial EMG responses to emotional expressions are related to emotion perception ability. *PLoS One* 9:e84053. doi: 10.1371/journal.pone.0084053

Martorell, C., González, R., Ordóñez, A., and Gómez, O. (2011). Estudio confirmatorio del Cuestionario de Conducta Prosocial (CCP) y su relación con variables de personalidad y socialización. *Rev. Iberoam. Diagnóstico Eval. Psicol.* 2, 35–52.

Mattek, A. M., Wolford, G. L., and Whalen, P. J. (2017). A mathematical model captures the structure of subjective affect. *Perspect. Psychol. Sci.* 12, 508–526. doi: 10.1177/1745691616685863

Merino-Soto, C., and Grimaldo-Muchotrigo, M. (2015). Validación Estructural de la Escala Básica de Empatía (Basic Empathy Scale) modificada en adolescentes: un estudio preliminar. *Rev. Colomb. Psicol.* 24, 261–270. doi: 10.15446/rcp.v24n2.42514

Murray, H. A. (1933). The effect of fear upon estimates of the maliciousness of other personalities. *J. Soc. Psychol.* 4, 310–329. doi: 10.1080/00224545.1933.9919325

Nomura, M., Ohira, H., Haneda, K., Iidaka, T., Sadato, N., Okada, T., et al. (2004). Functional association of the amygdala and ventral prefrontal cortex during cognitive evaluation of facial expressions primed by masked angry faces: an event-related fMRI study. *NeuroImage*, 21, 352–363. doi: 10.1016/j.neuroimage.2003.09.021

Palagi, E., Nicotra, V., and Cordoni, G. (2015). Rapid mimicry and emotional contagion in domestic dogs. *Royal Soc. Open Sci.* 2:150505. doi: 10.1098/rsos.150505

Payne, B. K., Hall, D. L., Cameron, C. D., and Bishara, A. J. (2010). A process model of affect misattribution. *Personal. Soc. Psychol. Bull.* 36, 1397–1408. doi: 10.1177/0146167210383440

Payne, K., and Lundberg, K. (2014). The affect misattribution procedure: ten years of evidence on reliability, validity, and mechanisms: affect misattribution procedure. *Soc. Personal. Psychol. Compass* 8, 672–686. doi: 10.1111/spc3.12148

Pérez-Albéniz, A., de Paúl, J., Etxeberría, J., Montes Paz, M., and Torres, E. (2003). Adaptación de Interpersonal reactivity index (IRI) al español. *Psicothema* 15, 267–272.

Peters, K., and Kashima, Y. (2015). A multimodal theory of affect diffusion. *Psychol. Bull.* 141, 966–992. doi: 10.1037/bul0000020

Pinilla, A. (2017). *Contagio Emocional: Incidencia de los Estados Emocionales Sobre la Imitación de expresiones Faciales*. Master's Thesis, Universidad Nacional de Colombia, Bogotá.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychol. Rev* 110, 145–172. doi: 10.1037/0033-295X.110.1.145

Sato, W., Fujimura, T., Kochiyama, T., and Suzuki, N. (2013). Relationships among facial mimicry, emotional experience, and emotion recognition. *PLoS One* 8:e57889. doi: 10.1371/journal.pone.0057889

Schaefer, A., Nils, F., Sanchez, X., and Philippot, P. (2010). Assessing the effectiveness of a large database of emotion-eliciting films: a new tool for emotion researchers. *Cogn. Emot.* 24, 1153–1172. doi: 10.1080/02699930903274322

Schmidt, R. C., Fitzpatrick, P., Caron, R., and Mergeche, J. (2011). Understanding social motor coordination. *Hum. Mov. Sci.* 30, 834–845. doi: 10.1016/j.humov.2010.05.014

Van Boven, L., and Loewenstein, G. (2003). Social projection of transient drive states. *Personal. Soc. Psychol. Bull.* 29, 1159–1168. doi: 10.1177/0146167203254597

Varcin, K. J., Grainger, S. A., Richmond, J. L., Bailey, P. E., and Henry, J. D. (2019). A role for affectivity in rapid facial mimicry: an electromyographic study. *Soc. Neurosci.* 14 1–10. doi: 10.1080/17470919.2018.1564694

Watson, D., Clark, L. A., and Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *J. Personal. Soc. Psychol.* 54, 1063–1070. doi: 10.1037/0022-3514.54.6.1063

Conflict of Interest: 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.

frontiers
in Psychology

OPINION
published: 31 January 2020
doi: 10.3389/fpsyg.2020.00096

Implicit vs. Explicit Emotion Processing in Autism Spectrum Disorders: An Opinion on the Role of the Cerebellum

Libera Siciliano[1]* *and Silvia Clausi*[2,3]

[1] PhD Program in Behavioral Neuroscience, "Sapienza" University of Rome, Rome, Italy, [2] Ataxia Laboratory, IRCCS Fondazione Santa Lucia, Rome, Italy, [3] Department of Psychology, Sapienza University of Rome, Rome, Italy

Keywords: cerebellum, emotions, autism, implicit processing, explicit processing

OPEN ACCESS

Edited by:
Noemi Mazzoni,
University of Trento, Italy

Reviewed by:
Alessandro Crippa,
Eugenio Medea (IRCCS), Italy

****Correspondence:***
Libera Siciliano
libera.siciliano@uniroma1.it

Specialty section:
This article was submitted to
Consciousness Research,
a section of the journal
Frontiers in Psychology

Received: *05 August 2019*
Accepted: *14 January 2020*
Published: *31 January 2020*

Citation:
Siciliano L and Clausi S (2020) Implicit vs. Explicit Emotion Processing in Autism Spectrum Disorders: An Opinion on the Role of the Cerebellum. Front. Psychol. 11:96.
doi: 10.3389/fpsyg.2020.00096

INTRODUCTION

Human beings are continuously exposed to external and internal emotional stimuli which are processed to create adaptive social relations. Based on the nature of emotional phenomena and on the effort required for their elaboration, emotional processing could be viewed along a continuum, ranging from an implicit/unconscious level to an explicit/conscious level (Lane, 2000). The implicit processing of emotions is meant to be an automatic, procedural, and non-conceptual process that does not require conscious access to be executed. Instead, explicit processing requires declarative evaluation and involves higher cognitive resources to define conscious emotional states (Damasio, 1994; Lane, 2000).

Within the extended neural network involved in the emotional field, the implicit processing of emotions has been linked to the amygdala and the anterior cingulated cortex (Critchley et al., 2005; Webb et al., 2010), while the temporoparietal junction and the medial prefrontal cortex have been reported to be mainly involved when the conscious attribution of emotions and intentions is required (Saxe and Kanwisher, 2003; Saxe and Powell, 2006).

However, the neural networks underpinning these processes are still far from being well-defined (Schaller and Rauh, 2017).

Over the years, many scientific studies have recognized the cerebellum as being part of these brain networks (Stoodley and Schmahmann, 2010; Baumann and Mattingley, 2012; Leggio and Olivito, 2018; Clausi et al., 2019a; Van Overwalle et al., 2019a) and have evidenced its role in affective and emotional functioning (Adamaszek et al., 2015; Clausi et al., 2015, 2019b; Lupo et al., 2015, 2018). Accordingly, functional and anatomical connections have been found between the cerebellum and the cortical and subcortical structures involved in both implicit and explicit emotional processing (Critchley et al., 2000; Singer et al., 2004; Schutter et al., 2009; Stoodley and Schmahmann, 2010; Schraa-Tam et al., 2012).

The contribution of the cerebellum to implicit and explicit mechanisms underlying emotion processing has been recently acknowledged (Clausi et al., 2017).

Indeed, the cerebellum is involved in implicit aspects such as the modulation of autonomic reactions, the automatic component of emotional learning associated, e.g., with fear conditioning (Critchley et al., 2000; Sacchetti et al., 2005; Timmann et al., 2010), and in the implicit processing of emotional facial expression (Schutter et al., 2009; Clausi et al., 2017). The cerebellar vermis might play a role in this stage of emotion elaboration through its connections with the brainstem (catecholamine neurons), the hypothalamus (Snider and Maiti, 1976), and the limbic areas, such as the amygdala and hippocampus (Sacchetti et al., 2005).

Instead, regarding the explicit aspects, the cerebellum is involved in the self-perception of negative emotions and in the integration of internal state information with external environmental stimuli to consciously and adaptively elaborate emotions (Timmann et al., 2010; Clausi et al., 2017, 2019a).

The posterior portions of the cerebellum might contribute to these aspects by means of their connections with cortical areas involved in more complex features of emotional evaluation (i.e., the medial prefrontal cortex and temporoparietal junction) (Rudebeck et al., 2008; Buckner et al., 2011).

Evidence about the "emotional cerebellum" has also been conveyed by clinical studies. Indeed, a pattern of emotional and affective disorders has been found in patients with cerebellar damage as part of the well-known cerebellar cognitive affective syndrome (Schmahmann and Sherman, 1998; Tavano et al., 2007).

Intriguingly, converging clinical and neuroimaging evidence points to a cerebellar involvement in the emotional disturbances and social cognition impairments described in people with autism spectrum disorders (ASD) (Fatemi et al., 2012; D'Mello and Stoodley, 2015).

ASD are clinically complex and heterogeneous neurodevelopmental conditions characterized by core impairments in social interaction, repetitive behaviors, and restricted interests (American Psychiatric Association, 2013; Goldson, 2016). Emotional processing is a challenge for people with ASD, whose impairments have been reported for both implicit and explicit components (Senju, 2013; Lozier et al., 2014; Kana et al., 2016).

Nevertheless, these aspects are often separately investigated, leading to controversial conclusions (Ben-Shalom et al., 2006). Indeed, while most of the literature on people with ASD agrees on the existence of implicit emotional processing impairment, i.e., when the task requires the elaboration of emotional facial expression (Baron-Cohen et al., 1997, 2001), controversial results emerge when the task requires the explicit elaboration of emotions. In the latest case, the heterogeneous nature of the ASD condition and camouflaging phenomena due to learned compensatory strategies might lead to the fallacious assumption that people with high-functioning ASD (hf-ASD) do not show such deficits (Senju, 2012; Schuwerk et al., 2014; Schaller and Rauh, 2017). For example, this could happen in tasks in which context information can help the person explicitly attribute emotions to others (Frith and Frith, 2012; Senju, 2012; Schaller and Rauh, 2017). However, it must be considered that when the amount of context information increases, the environment becomes more confusing, and the compensation ability may no longer guarantee the processing of emotional information (Frith and Frith, 2012). Indeed, when the complexity increases, the person has to integrate emotional processing with mentalizing ability to correctly infer emotional states (Mier et al., 2010). This complex social elaboration is impaired in people with hf-ASD when dynamic video-based stimuli close to everyday life are used or when various aspects of social situations are analyzed (Dziobek et al., 2006; Schaller and Rauh, 2011, 2017). Thus, the heterogeneity of capacities shown in a complex condition such as ASD, together with the variety of tasks used in different studies, may lead to controversial results.

The complex behavioral outcome of people with ASD has been linked to the functional alteration of complex neural circuits encompassing several brain areas, such as the parietal, temporal, and frontal regions (Abell et al., 1999; Carper et al., 2002; Hazlett et al., 2006; Minshew and Williams, 2007), as well as subcortical structures (Sparks et al., 2002; Amaral et al., 2008; Cauda et al., 2011). Among these, the cerebellum has been consistently recognized as part of the distributed neural networks affected in people with ASD (Wang et al., 2014; D'Mello and Stoodley, 2015; Olivito et al., 2017; Stoodley et al., 2017). Indeed, evidence of structural and functional alterations in specific cerebellar regions and in cerebello-cortical networks underlying emotional processing has been reported in this population (Courchesne et al., 1988; Khan et al., 2015; Stoodley et al., 2017; Arnold Anteraper et al., 2019). In particular, a decrease in the Purkinje cell number in the cerebellar vermis and a gray matter reduction in the posterior cerebellum have been found in people with ASD (Ritvo et al., 1986; Bauman and Kemper, 2005; Fatemi et al., 2012), together with altered functional connectivity between the posterior cerebellum and the frontal and temporal areas involved in mentalizing abilities (Olivito et al., 2017).

In this framework, considering the suggested connections between the cerebellum and the cortical and subcortical structures involved in implicit and explicit emotional processing (Schutter et al., 2009; Stoodley and Schmahmann, 2010; Clausi et al., 2017) and the association between the cerebellum and the social-emotional impairments in ASD (Wang et al., 2014; D'Mello and Stoodley, 2015), in the present opinion, we will provide some novel insights into the possible nature of implicit and explicit emotion processing deficits in ASD related to cerebellar-specific involvement. To this aim, we will take into account the most primitive ability to process facial emotions to analyze the implicit emotional process and the capacity to integrate intricate environmental information with the theory of mind abilities to explicitly infer complex emotional states.

CEREBELLAR IMPLICIT EMOTIONAL PROCESSING IN AUTISM

One of the most primitive and implicitly learned processes crucial for social interactions is the ability to elaborate emotions conveyed by facial expressions (Schaller and Rauh, 2017). This ability and that of spontaneous emotional mimicry are part of the implicit processes that contribute to emotional contagion (McIntosh et al., 2006; Senju, 2013). An adaptive phenomenon linked to the implicit processing and learning of emotions is the habituation to recurrently presented faces, as measured by the time spent looking at them (Webb et al., 2010). This ability allows individuals to face predictable social stimuli in an automatic way, since facial expressions convey emotional cues that human beings typically process first when they are embedded in social situations (Shyman, 2017). In typically developing children, this phenomenon is associated with decreased neural responsiveness in the amygdala when repeated facial stimuli are presented (Webb et al., 2010).

Interplay between the cerebellum and the amygdala has been found to contribute to the process of implicit emotional learning (Snider and Maiti, 1976; Zhu et al., 2011). Animal studies have shown that a bidirectional interaction between the basolateral amygdala and the cerebellum allows learning-related plasticity in fearful conditions by increasing the firing in vermal Purkinje cells, thus inducing learning-related long-term potentiation in the cerebellum (Snider and Maiti, 1976; Zhu et al., 2011). In this way, the cerebellum integrates sensory and emotional information, enabling appropriate reactions to new fearful situations and maintaining them across time and contexts (Zhu et al., 2011). Accordingly, in fear conditioning paradigms, animal and human studies have shown an involvement of the cerebellar vermis in the associative mechanisms that contribute to the creation of memory traces with emotional valence (Sacchetti et al., 2002; Labrenz et al., 2015). Because of its extensive connections with limbic areas, the cerebellar vermis has been defined as the “limbic cerebellum” (Schmahmann, 1991).

Within this framework, it is important to emphasize that the automatization and implicit processing of emotional mechanisms require the repetitive and predictable patterns of stimuli.

Interestingly, the correct recognition of spatial and temporal relations among relevant actions has been associated with cerebellar predictive computing in motor and non-motor domains (Leggio et al., 2008; Molinari et al., 2008). Furthermore, cortical plastic changes mediated by cerebellar-driven facilitation have been described when a predictable pattern is conveyed by the current stimuli (Molinari et al., 2002; Ito, 2005).

Supporting cerebellar involvement in the emotional domain, clinical studies have revealed that damage in the cerebellar vermis is associated with autism-like behaviors in cerebellar patients, giving rise to difficulties in the automatic attribution of relevant emotional states regardless of the context, as in emotion recognition from facial expressions (Schmahmann and Sherman, 1998; Riva and Giorgi, 2000).

The elaboration of facial emotional expression and automatic emotional mimicry are reported to be the earliest social impairments in people with ASD (Baron-Cohen et al., 2000; McIntosh et al., 2006; Senju, 2013). These alterations are often related to dysfunction of the mirror neurons system (MNS), which is a set of brain regions active in both action execution and the observation of actions performed by others (Press et al., 2010).

Remarkably, it has been found that children with ASD show a slower habituation to recurrently presented faces compared to typically developing children (Webb et al., 2010). This diminished habituation has been correlated with social symptoms (Swartz et al., 2013). Consistently, in people with ASD, repeated exposure to faces, conveying both positive and negative emotions, has not been associated with decreased activity in the amygdala (Swartz et al., 2013; Tam et al., 2017).

Moreover, in these people, structural imaging studies have shown gray matter abnormalities in the vermis and hypoplasia of this cerebellar portion due to cellular defects in the Purkinje cells in early life (Courchesne et al., 2011; Fatemi et al., 2012; Wang et al., 2014).

We posit that, in people with ASD, cerebellar dysfunction and an altered interaction between the cerebellum and the amygdala might affect the implicit process of emotion elaboration, thus impeding habituation. Our assumption is that the alteration of the cerebellar vermis might prevent the detection of operative internal models for spatially and temporally organized emotional phenomena, thus impeding automatic responses for incoming emotional stimuli and their prediction and automated implementation when facing analogous events in the future (Leggio and Molinari, 2015).

CEREBELLAR EXPLICIT EMOTIONAL PROCESSING IN AUTISM

The explicit processing of emotions requires subjects to interpret several idiosyncratic and environmental stimuli and needs a detailed appraisal based on the interplay between present and past states (Schaller and Rauh, 2017).

When emotional stimuli (i.e., facial expressions) are embedded in complex social situations, the concomitant processing of mental states may be required (Schaller and Rauh, 2017).

In this case, the capacity to infer others’ emotional states is closely related to the attribution of their intentions and requires a declarative and a higher cognitive evaluation (Brothers and Ring, 1990).

The brain network supporting explicit emotional processing involves the cortico-pulvinar-cortical pathway (i.e., the medial frontal cortex, the superior temporal gyrus, and the cingulate cortex) and the supra-modal association cortices belonging to the default mode network (DMN) (Grimm et al., 2008; Sreenivas et al., 2012; Shobe, 2014).

It is well-known that the posterior cerebellar hemispheres, named the neocerebellum, play a role in high associative functions in synchrony with co-evoluted regions of the cerebral cortex recruited, for example, when more cognitive aspects of emotional processing are in demand (Adamaszek et al., 2017; Leggio and Olivito, 2018). Consistently, these cerebellar portions have been found to be functionally connected to cerebral association areas belonging to the DMN. In healthy subjects, fMRI studies have shown that these cerebello-cortical networks are involved in social-emotional tasks that require high-order cognitive reasoning (Buckner et al., 2011; Van Overwalle et al., 2015).

Furthermore, reduced functional connectivity (FC) between cerebellar posterior lobules, such as the right Crus-II, and cortical regions involved in complex social-emotional reasoning is present in patients with cerebellar degenerative atrophy, whose behavioral profile is characterized by mentalizing impairment and the inability to deduce emotional states (Clausi et al., 2019a).

As described above, studies on the ability to explicitly process emotions in people with ASD have yielded controversial conclusions (Senju, 2012; Schaller and Rauh, 2017). Indeed, some studies have shown that people with hf-ASD succeed in simple explicit tests in which it is required to attribute emotions to protagonists of social stories and where additional context information helps to deduce emotional states (Schaller and Rauh, 2017). However, when the context complexity is higher and people with hf-ASD are required to integrate facial emotion expressions, intentions and emotional states, their performance becomes worse (Schaller and Rauh, 2011, 2017).

The preserved processing and recognition of basic emotions in people with hf-ASD has usually been associated with compensatory strategies gained by declarative mechanisms, thanks to spared cognitive and executive functions and to "environmental scaffolding" built on external resources (Frith and Frith, 2012). Indeed, people with hf-ASD show typical outcomes in explicit tests when overtly instructed to attribute emotions to protagonists of social stories (Frith and Frith, 2012). Therefore, the recruitment of declarative/explicit strategies might support compensation in neurodevelopmental disorders and could explain the variability in symptom severity described in hf-ASD (Ullman and Pullman, 2015; Livingston and Happé, 2017). However, it has been speculated that the environment may either facilitate or inhibit compensation (Livingston and Happé, 2017), since growing environmental demands may exceed the compensatory ability (Schaller and Rauh, 2017).

Relatedly, when the complexity and novelty of a social context increases, the compensation of people with hf-ASD less efficiently overcomes their difficulties in emotion processing (Frith and Frith, 2012).

It has been proposed that underconnectivity between brain areas, including those of the DMN, could be the neural basis of the socioemotional impairment in ASD, as asserted by the theory of the underconnected brain in ASD (Just et al., 2007; Müller, 2007; Müller et al., 2011).

Moreover, among the various theories, the hypothesis of developmental diaschisis assumes that cerebellar dysfunction during critical periods could disrupt the development of far neocortical networks, including the aforementioned supra-modal associative cortices (Lai et al., 2014; Wang et al., 2014). The disruption of the fronto-cerebellar network across developmental stages may give rise to motor, emotional and social symptoms (Igelström et al., 2017), mainly associated with a general impairment of the predictive process (Sinha et al., 2014). Sinha et al. (2014) argued that prevented access to predictive processes impedes people with ASD from using previously learned relationships between events and stimuli, thus being constrained to interpret behaviors based only on basic environmental sensory signals.

Interestingly, fMRI studies on people with ASD have shown abnormal FC between regions of the DMN comprising the temporoparietal junction and the medial prefrontal cortex and posterior lobules of the cerebellum, such as Crus-I/II, involved in the mentalizing process (D'Mello and Stoodley, 2015; Olivito et al., 2018). Thus, we hypothesize that atypical interplay between these lobules and supra-modal cortices might impede explicit emotional processing when context complexity increases and that further social effort is required for the conscious attribution of emotions and intentions (Saxe and Powell, 2006).

THE CEREBELLAR PREDICTION IN EMOTIONAL PROCESSING

The cerebellum is known to act by generating operative internal models of spatially and temporally organized events comparing individuals' external events with their internal state (Ito, 2008; Molinari et al., 2008). These internal models allow us to predict incoming events and to modulate responses implicitly (Leggio and Molinari, 2015). Indeed, in the presence of cerebellar damage, the required rapid and continuous exchange of information between the internal model and the external stimuli might not be operative, thus interfering with the automatic processes (Clausi et al., 2019a).

According to the cerebellar sequence detection model, the detection and simulation of a sequence of events can either occur implicitly through experience or deliberately through overt effort (Leggio and Molinari, 2015). Indeed, analogous to information processing in the sensorimotor domain, the cerebellum might modulate high-order cortical activity (Middleton and Strick, 2001) by detecting predictable sequences of emotionally salient events based on internal models previously encoded and by allowing optimized feedforward control (Heleven et al., 2019; Van Overwalle et al., 2019b). When social novelty and environmental demands increase, explicit emotional processing may be supported by such a cerebellar operational mode; i.e., the cerebellum may exert a continuous checking on the accordance between the anticipated event based on social and emotional information and the ongoing behavior by means of projections from the cerebellar posterior regions to the areas of the cerebral cortex involved in high-order social behavior (Ito, 2008; Van Overwalle et al., 2019b). In this way, the cerebellum supports more sophisticated forms of prediction and guarantees fluid control of social-emotional processing and interactions (Clausi et al., 2019a; Van Overwalle et al., 2019b). Altogether, this mechanism allows for the regulation and adjustment of future emotional expectations and guarantees adaptive social behaviors.

As outlined in the previous section, Sinha et al. (2014) hypothesized the impairment of the predictive process as the central thread of core behavioral problems in ASD and suggested a role of the cerebellum among the key brain areas implicated in the prediction.

In this framework, we posit that in people with ASD, a dysfunction of the phylogenetically older portion of the cerebellum (i.e., the vermis), combined with its abnormal interaction with the amygdala, may impede the creation and detection of internal models for predictable and elementary social-affective signs, which are needed to process emotional stimuli automatically (Lozier et al., 2014; Wang et al., 2014). This dysfunction would prevent the correct attribution of emotional valence to upcoming social stimuli in a non-conceptual and advantageous way, thus not providing people with ASD with permanent and successful social-emotional implicit mechanisms (e.g., ability to elaborate emotions conveyed by facial expressions).

In addition, we hypothesize that when people with hf-ASD are required to explicitly process basic and "primitive" emotions, they would still be able to use their compensatory strategies. However, when social novelty and environmental demands increase, people with hf-ASD cannot benefit from cerebellar predictive control, which is crucial to adapting the previously learned responses across the context. Indeed, since the posterior cerebellum is implicated in the generation of internal models of social-emotional interactions, altered functional connectivity within the Crus-II–DMN network may affect the ability of people with hf-ASD to predict others' emotional states when the social demands require extending the encoded internal models to ambiguous or new contexts in a flexible way. This might result in difficulties engaging in adaptive social relations, typically entailing complex dynamics requiring the integration of multimodal emotional and mentalizing information.

CONCLUSION AND FUTURE INVESTIGATIONS

Neuroimaging data on ASD agree on the presence of both structural and functional alterations involving the cerebellar vermis and the more posterior part of the cerebellum (Crus I/II). These regions are known to be active when an implicit processing of emotional stimuli is required, thus contributing to emotional memory formation, and when the explicit interpretation of emotional states requires complex inferences and concomitant mentalizing processes.

We suggest that the impaired ability to implicitly process emotions in people with ASD could be due to a dysfunctional circuitry involving the vermis and the limbic system. On the other hand, a dysfunctional reorganization of the network comprising the CrusI/II and cortical regions involved in complex social reasoning may account for the inability to compensate for the difficulties in explicit tasks when the environmental complexity increases the emotional processing demands. See **Figure 1** for a schematic view of the cerebral-cerebellar networks proposed.

Conclusively, the underlying mechanisms of both implicit and explicit emotional impairments may be due to a dysfunctional cerebellar modulation on specific cerebral areas, as supported by the theory of the cerebellar operational mode described above. It is worth noting that since the cerebellum is anatomically and functionally connected to many of the cortical areas taking part in the MNS and the mentalizing network (Strick et al., 2009; Buckner et al., 2011; Van Overwalle et al., 2019a,b), it is conceivable that the hypothesis we developed for implicit and explicit emotional processing may be suitable for explaining the impairment in the mirroring system and in higher mentalizing abilities in people with ASD.

Overall, our theoretical hypothesis constitutes a framework for a new perspective on the role of the cerebellum in emotional processing dysfunctions in people with ASD.

Further experimental research is needed to better define the nature of cerebellar involvement in the socioemotional domain because it could shed further light on the pathogenesis of ASD and provide innovative insights into novel therapeutic interventions and neuromodulation targets. To this end, recognizing that the increasing complexity and novelty of contexts and stimuli could impact performance in experimental designs and everyday life, we suggest a stronger distinction between the investigation methods for implicit and explicit emotion processing. Finally,

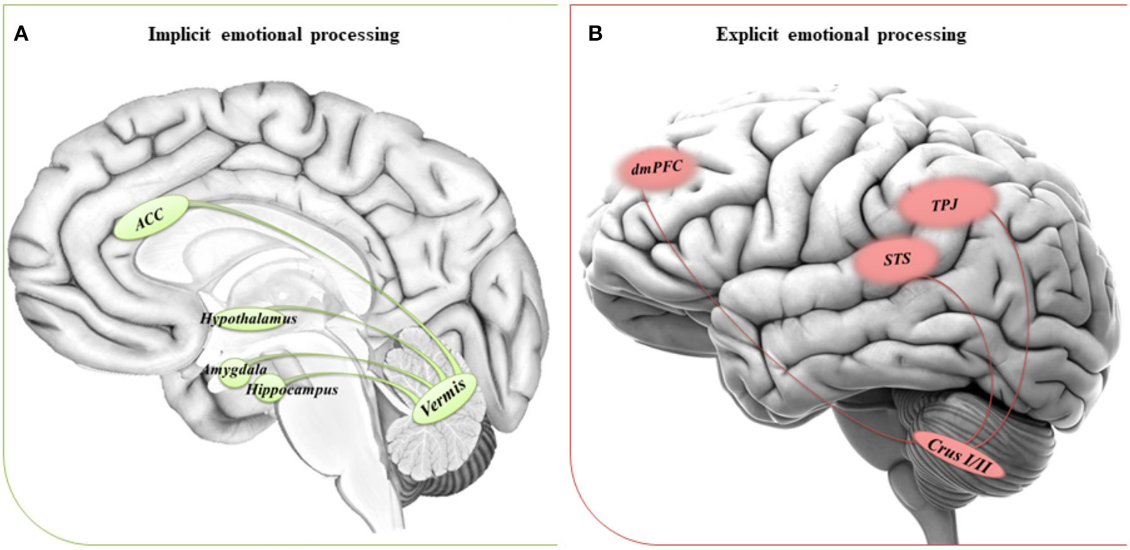


FIGURE 1 | Suggested networks for implicit and explicit emotional processing. **(A)** The circuitry involved in implicit processing of emotions comprises the cerebellar vermis and the anterior cingulate cortex (ACC), hypothalamus and areas of the limbic system, such as amygdala, and hippocampus. **(B)** The circuitry involved in explicit processing of emotions comprises the cerebellar posterior lobules Crus I/II and cerebral areas of the DMN, such as the dorsomedial prefrontal cortex (dmPFC), the temporoparietal junction (TPJ), and the superior temporal sulcus (STS).

comparisons between behavioral and neuroimaging data across different pathologies with both cerebellar and emotional dysfunctions will be crucial to confirm or reject our assumption.

AUTHOR CONTRIBUTIONS

LS: writing and preparation of manuscript. SC: supervising as expert on the topic and critical review of manuscript. All authors listed have contributed to the concept, design, revision, and approval of manuscript.

FUNDING

The present study was supported by grant from the Italian Ministry of Health to SC (Grant Number GR-2013-02354888) and by the Ph.D. Program in Behavioral Neuroscience from Sapienza University of Rome to LS.

ACKNOWLEDGMENTS

The editing support of American Journal Experts is acknowledged.

REFERENCES

Abell, F., Krams, M., Ashburner, J., Passingham, R., Friston, K., Frackowiak, R., et al. (1999). The neuroanatomy of autism: a voxel-based whole brain analysis of structural scans. *Neuroreport* 10, 1647–1651. doi: 10.1097/00001756-199906030-00005

Adamaszek, M., D'Agata, F., Ferrucci, R., Habas, C., Keulen, S., Kirkby, K. C., et al. (2017). Consensus paper: cerebellum and emotion. *Cerebellum*. 16, 552–576. doi: 10.1007/s12311-016-0815-8.

Adamaszek, M., Kirkby, K. C., D'Agata, F., Olbrich, S., Langner, S., Steele, C. J., et al. (2015). Neural correlates of disturbed emotional face recognition in cerebellar lesions. *Brain Res.* 1613, 1–12. doi: 10.1016/j.brainres.2015.01.027

Amaral, D. G., Schumann Mills, C., and Wu Nordahl, C. (2008). Neuroanatomy of autism. *Trends Neurosci.* 31, 137–145. doi: 10.1016/j.tins.2007.12.005

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5), 5th Edn*. Washington, DC: American Psychiatric Association. doi: 10.1176/appi.books.9780890425596

Arnold Anteraper, S., Guell, X., D'Mello, A., Joshi, N., Whitfield-Gabrieli, S., and Joshi, G. (2019). Disrupted cerebrocerebellar intrinsic functional connectivity in young adults with high-functioning autism spectrum disorder: a data-driven, whole-brain, high-temporal resolution functional magnetic resonance imaging study. *Brain Connect.* 9, 48–59. doi: 10.1089/brain.2018.0581

Baron-Cohen, S., Jolliffe, T., Mortimore, C., and Robertson, M. (1997). Another advanced test of theory of mind: evidence from very high-functioning adults with autism or Asperger Syndrome. *J. Child Psychol. Psychiatry* 38, 813–822. doi: 10.1111/j.1469-7610.1997.tb01599.x

Baron-Cohen, S., Ring, H. A., Bullmore, E. T., Wheelwright, S., Ashwin, C., and Williams, S. C. (2000). The amygdala theory of autism. *Neurosci. Biobehav. Rev.* 24, 355–364. doi: 10.1016/S0149-7634(00)00011-7

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

Bauman, M. L., and Kemper, T. L. (2005). *The Neurobiology of Autism*. Baltimore, MD: Johns Hopkins University Press.

Baumann, O., and Mattingley, J. B. (2012). Functional topography of primary emotion processing in the human cerebellum. *Neuroimage* 61, 805–811. doi: 10.1016/j.neuroimage.2012.03.044

Ben-Shalom, D., Mostofsky, S. H., Hazlett, R. L., Goldberg, M. C., Landa, R. J., Faran, Y., et al. (2006). Normal physiological emotions but differences in expression of conscious feelings in children with high-functioning autism. *J. Autism. Dev. Disord*. 36, 395–400. doi: 10.1007/s10803-006-0077-2

Brothers, L., and Ring, B. (1990). A neuroethological framework for the representation of mind. *J. Cogn. Neurosci*. 4, 107–118. doi: 10.1162/jocn.1992.4.2.107

Buckner, R. L., Krienen, F. M., Castellanos, A., Diaz, J. C., and Yeo, B. T. (2011). The organization of the human cerebellum estimated by intrinsic functional connectivity. *J. Neurophysiol*. 106, 2322–2345. doi: 10.1152/jn.00339.2011

Carper, R. A., Moses, P., Tigue, Z. D., and Courchesne, E. (2002). Cerebral lobes in autism: early hyperplasia and abnormal age effects. *Neuroimage* 16, 1038–1051. doi: 10.1006/nimg.2002.1099

Cauda, F., Geda, E., Sacco, K., D'Agata, F., Duca, S., Geminiani, G., et al. (2011). Grey matter abnormalities in autism spectrum disorder: an activation likelihood estimation meta-analysis study. *J. Neurol. Neurosur. Ps.* 82, 1304–1313. doi: 10.1136/jnnp.2010.239111

Clausi, S., Coricelli, G., Pisotta, I., Pavone, E. F., Lauriola, M., Molinari, M., et al. (2015). Cerebellar damage impairs the self-rating of regret feeling in a gambling task. *Front. Behav. Neurosci*. 9:113. doi: 10.3389/fnbeh.2015.00113

Clausi, S., Iacobacci, C., Lupo, M., Olivito, G., Molinari, M., and Leggio, M. (2017). The role of the cerebellum in unconscious and conscious processing of emotions: a review. *Appl. Sci*. 7:521, doi: 10.3390/app7050521

Clausi, S., Lupo, M., Olivito, G., Siciliano, L., Contento, M. P., Aloise, F., et al. (2019b). Depression disorder in patients with cerebellar damage: awareness of the mood state. *J. Affect. Disord*. 245, 386–393. doi: 10.1016/j.jad.2018.11.029

Clausi, S., Olivito, G., Lupo, M., Siciliano, L., Bozzali, M., and Leggio, M. (2019a). The cerebellar predictions for social interactions: theory of mind abilities in patients with degenerative cerebellar atrophy. *Front. Cell. Neurosci*. 12:510. doi: 10.3389/fncel.2018.00510

Courchesne, E., Webb, S. J., and Schumann, C. M. (2011). "From toddlers to adults: the changing landscape of the brain in autism," in *Autism Spectrum Disorders*, eds D. Amaral, D. Geschwind, and G. Dawson (Oxford: Oxford University Press), 611–631. doi: 10.1093/med/9780195371826.003.0040

Courchesne, E., Yeung-Courchesne, R., Press, G. A., Hesselink, J. R., and Jernigan, T. L. (1988). Hypoplasia of cerebellar vermal lobules VI and VII in autism. *N. Engl. J. Med.* 318, 1349–1354. doi: 10.1056/NEJM198805263182102

Critchley, H. D., Corfield, D. R., Chandler, M. P., Mathias, C. J., and Dolan, R. J. (2000). Cerebral correlates of autonomic cardiovascular arousal: a functional neuroimaging investigation in humans. *J. Physiol.* 523, 259–270. doi: 10.1111/j.1469-7793.2000.t01-1-00259.x

Critchley, H. D., Rotshtein, P., Nagai, Y., O'Doherty, J., Mathias, C. J., and Dolan, R. J. (2005). Activity in the human brain predicting differential heart rate responses to emotional facial expressions. *Neuroimage* 24, 751–762. doi: 10.1016/j.neuroimage.2004.10.013

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York, NY: G.P. Putnam's Press.

D'Mello, A. M., and Stoodley, C. J. (2015). Cerebro-cerebellar circuits in autism spectrum disorder. *Front. Neurosci*. 9:408. doi: 10.3389/fnins.2015.00408

Dziobek, I., Fleck, S., Kalbe, E., Brand, M., Kessler, J., Woike, J. K., et al. (2006). Introducing MASC: a movie for the assessment of social cognition. *J. Autism Dev. Disord*. 36, 623–636. doi: 10.1007/s10803-006-0107-0

Fatemi, S. H., Aldinger, K. A., Ashwood, P., Bauman, M. L., Blaha, C. D., Blatt, G. J., et al. (2012). Consensus paper: pathological role of the cerebellum in autism. *Cerebellum* 11, 777–807. doi: 10.1007/s12311-012-0355-9

Frith, C. D., and Frith, U. (2012). Mechanisms of social cognition. *Annu. Rev. Psychol*. 63, 287–313. doi: 10.1146/annurev-psych-120710-100449

Goldson, E. (2016). Advances in autism-2016. *Adv. Pediatr*. 63, 333–355. doi: 10.1016/j.yapd.2016.04.014

Grimm, S., Boesiger, P., Beck, J., Schuepbach, D., Bermpohl, F., Walter, M., et al. (2008). Altered negative BOLD responses in the default-mode network during emotion processing in depressed subjects. *Neuropsychopharmacology* 34, 843–932. doi: 10.1038/npp.2008.81

Hazlett, H. C., Poe, M. D., Gerig, G., Smith, R. G., and Piven, J. (2006). Cortical gray and white brain tissue volume in adolescents and adults with autism. *Biol. Psychiat.* 59, 1–6. doi: 10.1016/j.biopsych.2005.06.015

Heleven, E., van Dun, K., and Van Overwalle, F. (2019). The posterior cerebellum is involved in constructing social action sequences: an fMRI study. *Sci Rep*. 31:11110. doi: 10.1038/s41598-019-46962-7

Igelström, K. M., Webb, T. W., and Graziano, M. S. A. (2017). Functional connectivity between the temporoparietal cortex and cerebellum in autism spectrum disorder. *Cereb. Cortex*. 27, 2617–2627. doi: 10.1093/cercor/bhw079

Ito, M. (2005). Bases and implications of learning in the cerebellum–adaptive control and internal model mechanism. *Prog. Brain. Res.* 148, 95–109. doi: 10.1016/S0079-6123(04)48009-1

Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nat. Rev. Neurosci.* 9, 304–313. doi: 10.1038/nrn2332

Just, M. A., Cherkassky, V. L., Keller, T. A., Kana, R. K., and Minshew, N. J., et al. (2007). Functional and anatomical cortical underconnectivity in autism: evidence from an FMRI study of an executive function task and corpus callosum morphometry. *Cereb. Cortex* 17, 951–961. doi: 10.1093/cercor/bhl006

Kana, R. K., Patriquin, M. A., Black, B. S., Channell, M. M., and Wicker, B. (2016). Altered medial frontal and superior temporal response to implicit processing of emotions in autism. *Autism Res.* 9, 55–66. doi: 10.1002/aur.1496

Khan, A. J., Nair, A., Keown, C. L., Dakto, M. C., Lincoln, A. J., and Muller, R. A. (2015). Cerebro-cerebellar resting state functional connectivity in children and adolescents with autism spectrum disorder. *Biol. Psychiat.* 78, 625–634. doi: 10.1016/j.biopsych.2015.03.024

Labrenz, F., Icenhour, A., Thürling, M., Schlamann, M., Forsting, M., Timmann, D., et al. (2015). Sex differences in cerebellar mechanisms involved in pain-related safety learning. *Neurobiol. Learn. Mem*. 123, 92–99. doi: 10.1016/j.nlm.2015.05.006

Lai, M. C., Lombardo, M. V., and Baron-Cohen, S. (2014). Autism. *Lancet* 8, 896–910. doi: 10.1016/S0140-6736(13)61539-1

Lane, R. D. (2000). "Neural correlates of conscious emotional experience," in *Cognitive Neuroscience of Emotion,* eds. R. D. Lane and L. Nadel (New York, NY: Oxford University Press), 345–370.

Leggio, M., and Molinari, M. (2015). Cerebellar sequencing: a trick for predicting the future. *Cerebellum* 14, 35–38. doi: 10.1007/s12311-014-0616-x

Leggio, M., and Olivito, G. (2018). Topography of the cerebellum in relation to social brain regions and emotions. *Handb. Clin. Neurol*. 154, 71–84. doi: 10.1016/B978-0-444-63956-1.00005-9

Leggio, M. G., Tedesco, A. M., Chiricozzi, F. R., Clausi, S., Orsini, A., and Molinari, M. (2008). Cognitive sequencing impairment in patients with focal or atrophic cerebellar damage. *Brain* 131, 1332–1343. doi: 10.1093/brain/awn040

Livingston, L. A., and Happé, F. (2017). Conceptualising compensation in neurodevelopmental disorders: reflections from autism spectrum disorder. *Neurosci. Biobehav. Rev*. 80, 729–742. doi: 10.1016/j.neubiorev.2017.06.005

Lozier, L. M., Vanmeter, J. W., and Marsh, A. (2014). Impairments in facial affect recognition associated with autism spectrum disorders: a meta-analysis. *Dev. Psychopathol.* 26, 1–13. doi: 10.1017/S0954579414000479

Lupo, M., Olivito, G., Siciliano, L., Masciullo, M., Molinari, M., Cercignani, M., et al. (2018). Evidence of cerebellar involvement in the onset of a manic state. *Front. Neurol.* 9:774. doi: 10.3389/fneur.2018.00774

Lupo, M., Troisi, E., Chiricozzi, F. R., Clausi, S., Molinari, M., and Leggio, M. (2015). Inability to process negative emotions in cerebellar damage: a functional transcranial doppler sonographic study. *Cerebellum* 14, 663–669. doi: 10.1007/s12311-015-0662-z

McIntosh, D. N., Reichmann-Decker, A., Winkielman, P., and Wilbarger, J. L. (2006). When the social mirror breaks: deficits in automatic, but not voluntary, mimicry of emotional facial expressions in autism. *Dev. Sci.* 9, 295–302. doi: 10.1111/j.1467-7687.2006.00492.x

Middleton, F. A., and Strick, P. L. (2001). Cerebellar projections to the prefrontal cortex of the primate. *J. Neurosci.* 21, 700–712. doi: 10.1523/JNEUROSCI.21-02-00700.2001

Mier, D., Lis, S., Neuthe, K., Sauer, C., Esslinger, C., Gallhofer, B., et al. (2010). The involvement of emotion recognition in affective theory of mind. *Psychophysiology* 47, 1028–1039. doi: 10.1111/j.1469-8986.2010.01031.x

Minshew, N. J., and Williams, D. L. (2007). The new neurobiology of autism: cortex, connectivity and neuronal organization. *Arch. Neurol.* 6, 945–950. doi: 10.1001/archneur.64.7.945

Molinari, M., Chiricozzi, F., Clausi, S., Tedesco, A., De Lisa, M., and Leggio, M. (2008). Cerebellum and detection of sequences, from perception to cognition. *Cerebellum* 7, 611–615. doi: 10.1007/s12311-008-0060-x

Molinari, M., Filippini, V., and Leggio, M. G. (2002). Neuronal plasticity of interrelated cerebellar and cortical networks. *Neuroscience* 111, 863–870. doi: 10.1016/S0306-4522(02)00024-6

Müller, R. A. (2007). The study of autism as a distributed disorder. *Ment. Retard. Dev. Disabil. Res. Rev.* 13, 85–95. doi: 10.1002/mrdd.20141

Müller, R. A., Shih, P., Keehn, B., Deyoe, J. R., Leyden, K. M., and Shukla, D. K. (2011). Underconnected, but how? A survey of functional connectivity MRIstudies in autism spectrum disorders. *Cereb. Cortex* 21, 2233–2243. doi: 10.1093/cercor/bhq296

Olivito, G., Clausi, S., Laghi, F., Tedesco, A. M., Baiocco, R., Mastropasqua, C., et al. (2017). Resting-state functional connectivity changes between dentate nucleus and cortical social brain regions in autism spectrum disorders. *Cerebellum* 16, 283–292. doi: 10.1007/s12311-016-0795-8

Olivito, G., Lupo, M., Laghi, F., Clausi, S., Baiocco, R., Cercignani, M., et al. (2018). Lobular patterns of cerebellar resting-state connectivity in adults with Autism Spectrum Disorder. *Eur. J. Neurosci.* 47, 729–735. doi: 10.1111/ejn.13752

Press, C., Heyes, C., and Kilner, J. M. (2010). Learning to understand others' actions. *Biol. Lett*. 23, 457–460. doi: 10.1098/rsbl.2010.0850

Ritvo, E. R., Freeman, B. J., Scheibel, A. B., Duong, T., Robinson, H., Guthrie, D., et al. (1986). Lower Purkinje cell counts in the cerebella of four autistic subjects: initial findings of the UCLA–NSAC autopsy research project. *Am. J. Psychiatry* 143, 862–866. doi: 10.1176/ajp.143.7.862

Riva, D., and Giorgi, C. (2000). The cerebellum contributes to higher function during development: evidence from a series of children surgically treated for posterior fossa tumors. *Brain* 123, 1051–1061. doi: 10.1093/brain/123.5.1051

Rudebeck, P. H., Bannerman, D. M., and Rushworth, M. F. (2008). The contribution of distinct subregions of the ventromedial frontal cortex to emotion, social behavior, and decision making. *Cogn. Affect. Behav. Neurosci*. 8, 485–497. doi: 10.3758/CABN.8.4.485

Sacchetti, B., Baldi, E., Lorenzini, C. A., and Bucherelli, C. (2002). Cerebellar role in fear-conditioning consolidation. *Proc. Natl. Acad. Sci. U.S.A*. 99, 8406–8411. doi: 10.1073/pnas.112660399

Sacchetti, B., Scelfo, B., and Strata, P. (2005). The cerebellum: Synaptic changes and fear conditioning. *Neuroscientist* 11, 217–227. doi: 10.1177/1073858405276428

Saxe, R., and Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". *Neuroimage* 19, 1835–1842. doi: 10.1016/S1053-8119(03)00230-1

Saxe, R., and Powell, L. J. (2006). It's the thought that counts: specific brain regions for one component of theory of mind. *Psychol. Sci.* 17, 692–699. doi: 10.1111/j.1467-9280.2006.01768.x

Schaller, M. U., and Rauh, R. (2017). What difference does it make? Implicit, explicit and complex social cognition in autism spectrum disorders. *J. Autism. Dev. Disord*. 47, 961–979. doi: 10.1007/s10803-016-3008-x

Schaller, U. M., and Rauh, R. (2011). *Animated-Theory-of-Mind (AToM): A Test for Measuring Implicit and Explicit Theory of Mind Taking Account of Dynamic and Contextual Aspects (Technical Report ALU-KJPP-2011-001)*. Freiburg: Research Group for Developmental Cognitive Psychiatry, Department of Child and Adolescent Psychiatry and Psychotherapy, University of Freiburg.

Schmahmann, J. D. (1991). An emerging concept. The cerebellar contribution to higher function. *Arch. Neurol.* 48, 1178–1187. doi: 10.1001/archneur.1991.00530230086029

Schmahmann, J. D., and Sherman, J. (1998). The cerebellar cognitive affective syndrome. *Brain* 121, 561–579. doi: 10.1093/brain/121.4.561

Schraa-Tam, C. K. L., Rietdijk, W. J. R., Verbeke, W. J. M. I., Dietvorst, R. C., Van Den Berg, W. E., Bagozzi, R. P., et al. (2012). fMRI activities in the emotional cerebellum: a preference for negative stimuli and goal directed behavior. *Cerebellum*. 11, 233–245. doi: 10.1007/s12311-011-0301-2

Schutter, D. J., Enter, D., and Hoppenbrouwers, S. S. (2009). High-frequency repetitive transcranial magnetic stimulation to the cerebellum and implicit processing of happy facial expressions. *J. Psychiatry Neurosci*. 34, 60–65.

Schuwerk, T., Vuori, M., and Sodian, B. (2014). Implicit and explicit Theory of Mind reasoning in autism spectrum disorders: the impact of experience. *Autism* 19, 459–468. doi: 10.1177/1362361314526004

Senju, A. (2012). Spontaneous theory of mind and its absence in autism spectrum disorders. *Neuroscientist* 18, 108–113. doi: 10.1177/1073858410397208

Senju, A. (2013). Atypical development of spontaneous social cognition in autism spectrum disorders. *Brain. Dev.* 35, 96–101. doi: 10.1016/j.braindev.2012.08.002

Shobe, E. R. (2014). Independent and collaborative contributions of the cerebral hemispheres to emotional processing. *Front. Hum. Neurosci*. 8:230. doi: 10.3389/fnhum.2014.00230

Shyman, E. (2017). Please wait, processing: a selective literature review of the neurological understanding of emotional processing in ASD and its potential contribution to neuroeducation. *Brain Sci*. 7:153. doi: 10.3390/brainsci7110153

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., and Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*. 303, 1157–1162. doi: 10.1126/science.1093535

Sinha, P., Kjelgaard, M. M., Gandhi, T. K., Tsourides, K., Cardinaux, A. L., Pantazis, D., et al. (2014). Autism as a disorder of prediction. *Proc. Natl. Acad. Sci. U.S.A.* 111, 15220–15225. doi: 10.1073/pnas.1416797111

Snider, R. S., and Maiti, A. (1976). Cerebellar contributions to the Papez circuit. *J. Neurosci. Res*. 2, 133–146. doi: 10.1002/jnr.490020204

Sparks, B. F., Friedman, S. D., Shaw, D. W., Aylward, E. H., and Echelard, D., Artru, A. A., et al. (2002). Brain structural abnormalities in young children with autism spectrum disorder. *Neurology* 59, 184–192. doi: 10.1212/WNL.59.2.184

Sreenivas, S., Boehm, S. G., and Linden, D. E. J. (2012). Emotional faces and the default mode network. *Neurosci. Lett*. 506, 229–234. doi: 10.1016/j.neulet.2011.11.012

Stoodley, C. J., D'Mello, A. M., Ellegood, J., Jakkamsetti, V., Liu, P., Nebel, M. B., et al. (2017). Altered cerebellar connectivity in autism and cerebellar-mediated rescue of autism-related behaviors in mice. *Nat. Neurosci.* 20, 1744–1751. doi: 10.1038/s41593-017-0004-1

Stoodley, C. J., and Schmahmann, J. D. (2010). Evidence for topographic organization in the cerebellum of motor control versus cognitive and affective processing. *Cortex* 46, 831–844. doi: 10.1016/j.cortex.2009.11.008

Strick, P. L., Dum, R. P., and Fiez, J. A. (2009). Cerebellum and nonmotor function. *Annu. Rev. Neurosci* 32, 413–434. doi: 10.1146/annurev.neuro.31.060407.125606

Swartz, J. R., Wiggins, J. L., Carrasco, M., Lord, C., and Monk, C. S. (2013). Amygdala habituation and prefrontal functional connectivity in youth with autism spectrum disorders. *J. Am. Acad. Child Adolesc. Psychiatr*. 52, 84–93. doi: 10.1016/j.jaac.2012.10.012

Tam, F. I., King, J. A., Geisler, D., Korb, F. M., Sareng, J., Ritschel, F., et al. (2017). Altered behavioral and amygdala habituation in high-functioning adults with autism spectrum disorder: a fMRI Study. *Sci. Rep.* 7:13611. doi: 10.1038/s41598-017-14097-2

Tavano, A., Grasso, R., Gagliardi, C., Triulzi, F., Bresolin, N., Fabbro, F., et al. (2007). Disorders of cognitive and affective development in cerebellar malformations. *Brain* 130, 2646–2660. doi: 10.1093/brain/awm201

Timmann, D.,Drepper, J., Frings, M., Maschke, M., Richter, S., Gerwig, M., et al. (2010). The human cerebellum contributes to motor, emotional and cognitive associative learning. A review. *Cortex* 46, 845–857. doi: 10.1016/j.cortex.2009.06.009

Ullman, M. T., and Pullman, M. Y. (2015). A compensatory role for declarative memory in neurodevelopmental disorders. *Neurosci. Biobehav. Rev*. 51, 205–222. doi: 10.1016/j.neubiorev.2015.01.008

Van Overwalle, F., D'aes, T., and Mariën, P. (2015). Social cognition and the cerebellum: a meta analytic connectivity analysis. *Hum. Brain. Mapp*. 36, 5137–5154. doi: 10.1002/hbm.23002

Van Overwalle, F., Manto, M., Leggio, M., and Delgado-García, J. M. (2019a). The sequencing process generated by the cerebellum crucially contributes to social interactions. *Med. Hypotheses* 128, 33–42. doi: 10.1016/j.mehy.2019.05.014

Van Overwalle, F., Van de Steen, F., and Mariën, P. (2019b). Dynamic causal modeling of the effective connectivity between the cerebrum and cerebellum in social mentalizing across five studies. *Cogn. Affect. Behav. Neurosci*. 19, 211–223. doi: 10.3758/s13415-018-00659-y

Wang, S. S. H., Kloth, A. D., and Badura, A. (2014). The cerebellum, sensitive periods, and autism. *Neuron* 83, 518–532. doi: 10.1016/j.neuron.2014.07.016

Webb, S. J., Jones, E. J., Merkle, K., Namkung, J., Toth, K., Greenson, J., et al. (2010). Toddlers with elevated autism symptoms show slowed habituation to faces. *Child Neuropsychol*. 16, 255–278. doi: 10.1080/09297041003601454

Zhu, L., Sacco, T., Strata, P., and Sacchetti, B. (2011). Basolateral amygdala inactivation impairs learning-induced long-term potentiation in the cerebellar cortex. *PLoS ONE* 6:e16673. doi: 10.1371/journal.pone.0016673

Conflict of Interest: 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.

frontiers in Psychology

BRIEF RESEARCH REPORT
published: 17 March 2020
doi: 10.3389/fpsyg.2020.00332

Alexithymia, Not Autism Spectrum Disorder, Predicts Perceived Attachment to Parents in School-Age Children

Michele Giannotti, Simona de Falco and Paola Venuti*

Department of Psychology and Cognitive Sciences, University of Trento, Trento, Italy

Alexithymia is defined as a limited ability in the cognitive processing of emotions. Literature suggested its negative influence on interpersonal relationship, documenting elevated alexithymia in individuals with Autism Spectrum Disorder (ASD) compared to control groups. However, the study of alexithymia in school-age children with ASD remains largely unexplored as well as its effect on specific child socioemotional outcomes such as quality of attachment relationships. This study examines alexithymia and perceived attachment to parents in twenty-four children with ASD (without intellectual disability) and 24 typically developing (TD) children (mean age 10 years) using the self-reported Alexithymia Questionnaire for Children (AQC) and the Inventory of Parent and Peer Attachment (IPPA). Measures of family SES as well as child intelligence were collected. Data revealed that ASD children showed higher levels of Alexithymia compared to TD group. In addition, 21% of participants with ASD exceed alexithymia categorical cut-off. By contrast, no difference emerged in the perception of attachment to parents. Moreover, alexithymia, but not ASD status, was found to predictive of child perception of attachment to parents. We observed no significant effect of child age and verbal IQ. Our findings showed that alexithymia was more common in children with ASD, whereas attachment was similar between groups. Difficulties in identifying and describing one's own feelings may hinder the construction of a positive representation of parent-child attachment relationship regardless of child clinical status. Thus, alexithymia seems to play a key role on the way school-age children with and without ASD perceive their relationship with their parents.

Keywords: autism spectrum disorder (ASD), attachment, alexithimia, emotion processing, school-age, parent-child relationship and ASD

Edited by:
Giulia Mattavelli,
University Institute of Higher Studies in Pavia, Italy

Reviewed by:
Michela Di Trani,
Sapienza University of Rome, Italy
Hannah Madaleine Hobson,
University of Greenwich,
United Kingdom

****Correspondence:***
Michele Giannotti
michele.giannotti@unitn.it

Specialty section:
This article was submitted to Consciousness Research, a section of the journal Frontiers in Psychology

Received: *11 October 2019*
Accepted: *12 February 2020*
Published: *17 March 2020*

Citation:
Giannotti M, de Falco S and Venuti P (2020) Alexithymia, Not Autism Spectrum Disorder, Predicts Perceived Attachment to Parents in School-Age Children. Front. Psychol. 11:332. doi: 10.3389/fpsyg.2020.00332

INTRODUCTION

Autism spectrum disorder (ASD) is a neurodevelopmental condition characterized by sociocommunicational impairments and restricted and repetitive patterns of behaviors and interests (American Psychiatric Association [APA], 2013). Empirical studies have widely documented the presence of significant difficulties in the domain of emotion processing and regulation in ASD (Silani et al., 2008). In particular, during the last two decades, the construct of alexithymia has received greater attention in the arena of emotion processing in ASD (Kinnaird et al., 2019). Namely, the concept of alexithymia refers to individual difficulties in identifying, describing, and distinguishing one's own feelings, which are often accompanied by an externally oriented thinking instead of a focus on internal experience (Sifneos, 1973). Previous studies have consistently shown elevated alexithymia in individuals with ASD compared to control groups, regardless of the level of intellectual abilities (Hill et al., 2004; Lombardo et al., 2007; Griffin et al., 2016), reporting higher rates of adults with ASD above the clinical level (Bird and Cook, 2013). Nevertheless, only a few studies investigated alexithymia in children and adolescents showing similar findings (Griffin et al., 2016; Milosavljevic et al., 2016).

Interestingly, a series of studies found that alexithymia and not ASD status underpinned the specific emotion processing difficulties reported in this clinical population (Bird et al., 2010; Heaton et al., 2012). This area of research did not attribute emotion-processing and interoception difficulties to ASD core symptoms, but suggested instead a predictive effect of alexithymia (Bird and Cook, 2013; Shah et al., 2016). In addition, several studies underlined a considerable overlap between ASD and alexithymia neurobiological and anatomical correlates (van der Velde et al., 2013), including altered activation of specific brain areas such as amygdala and anterior cingulate cortex (Bernhardt et al., 2013; Caria and de Falco, 2015). Moreover, alexithymia as well as ASD are associated with impaired mentalizing abilities (Moriguchi et al., 2006), possibly in light of the reduced integration between physiological states interoception and emotional consciousness (Gaigg et al., 2018). For these reasons, scholars suggest considering alexithymia as a significant predictor of developmental outcomes of individuals with ASD and a key concept to identifying cognitive profiles of specific subgroups within the ASD heterogeneity (Lai et al., 2013; Fietz et al., 2018). Despite its clinical relevance, prior research on ASD focused almost exclusively on adulthood – hence, the investigation of alexithymia in childhood remains largely unexplored. To date, only one study (Griffin et al., 2016) has been conducted on young children, confirming a higher level of alexithymia for the ASD group using both self- and parent-reports.

Attachment and Alexithymia in ASD

Among the several factors associated with alexithymia, quality of attachment to parental figures constitutes a crucial predictor of children's healthy psychological development (Carlson and Sroufe, 1995), showing a significant contribution to several indicators of socioemotional adjustment and adaptation. In fact, developmental research has documented a robust relationship between insecure attachment to caregiver and emotion processing difficulties (Laible, 2007; Thompson and Meyer, 2007; Brumariu et al., 2012). However, the association between attachment and alexithymia has been mostly investigated in adulthood (Picardi et al., 2005). In fact, studies on children and adolescents focused on this link are still scarce, particularly on clinical samples (Oskis et al., 2013; Koelen et al., 2015). Therefore, there is a substantial lack of research concerning the association between alexithymia and attachment in children with ASD. A recent study (Costa et al., 2019) found that alexithymia predicts reduced parent–child interaction more than ASD status, suggesting the need to consider its impact on quality of relationship in future research. In fact, the study did not include a specific measure of attachment relationship to parents, which represents a core feature of child socioemotional development.

In this regard, literature on attachment showed no differences between ASD and their typical counterparts in the perception of quality of attachment to parents (Bauminger et al., 2010: Chandler and Dissanayake, 2014; Sivaratnam et al., 2018). Nonetheless, the study of attachment in ASD during middle childhood remains poorly investigated and further replications are needed to clarify the absence of significant differences with typical controls. Additionally, understanding the mechanism underlying the perception of attachment security to parents in children with ASD may elucidate the association among key aspects of socioemotional development in this clinical population, explaining which child features may contribute to positive explicit representations of trustworthy, sensitive, and available parents. Given that neither ASD status nor symptoms severity showed a significant influence on the perception of attachment security in ASD, it could be interesting to consider the role of alexithymia in predicting this socioaffective domain. According to a bidirectional perspective, child characteristics may influence the quality of parent – child interaction (Costa et al., 2019), altering parental attitudes and caregiving behaviors. Therefore, specific subclinical phenomenon such as impairments in identifying and describing one's own feelings as well as difficulty in distinguishing emotion from bodily sensations (Silani et al., 2008) can affect the way in which children perceive their attachment relationship with their parents.

The present study aimed at investigating alexithymia in school-age children with ASD (without intellectual disability) by examining its predictive role on child perception of attachment security to parents. Firstly, we explored potential differences between groups for both attachment and alexithymia. To this purpose, we also estimated the percentage of children above the normative cut-off of alexithymia. Next, the predictive effect of child features, including alexithymia, were tested with the aim to identify which mechanisms intervene in shaping the perception of attachment security to parents in ASD. We primarily hypothesized to find higher levels and rates of alexithymia in children with ASD compared to the controls, according to previous studies on this topic (Kinnaird et al., 2019). No differences were expected with respect to the perception of attachment security to parents as highlighted by prior research (Teague et al., 2017). In regard to our second aim, we expect to find a significant link between alexithymia and attachment according to literature on neurotypical population.

METHOD

Procedure

Participants with ASD diagnosis were recruited through two different clinical centers for children with neurodevelopmental disorders. Clinicians informed parents of children with a certified

ASD diagnosis who meet inclusion criteria about the possibility to be involved in this study. We used snowball sampling and specific advertisements in the University area to recruit children with typical development. In response to the expression of interest of the parents, we invited the families to the clinical centers to participate in this study. After reading a detailed informative of the study, parents signed the informed consent, including the form concerning protection of personal data. This procedure has been accomplished according to the EU General Data Protection Regulation (GDPR) no. 2016/679. This study was given ethical approval by the Ethic Committee on Experiments involving human beings of the University of Trento. The administration of the questionnaires was conducted by an experimenter in a quiet room of the centers involved in the study. Mothers of children with TD were asked to complete the Social Responsiveness Scale 2 (Constantino and Gruber, 2005) in order to screen participants of the control group for sociocommunicational impairment. Similarly, the Autism Diagnostic Observation Schedule (ADOS, Module 3; Lord et al., 2015) was used to confirm diagnosis in children of the ASD group. We also collected measures of verbal and non-verbal intelligence and family socioeconomic status in both groups. Moreover, children completed two self-report questionnaires for the assessment of alexithymia and quality of attachment to parents.

Participants

Out of the 52 contacted families, 4 refused to participate (2 of ASD and 2 of the TD group) – hence, our final sample included 24 children with ASD (without intellectual disability) and 24 children with typical development (TD). Children mean age is 126.4 months (*SD* = 16.45) for the clinical group and 115.88 months (*SD* = 25.14) for the normative group. The majority of the participants involved in this study were males (75% of the total sample and 62.5% of TD), particularly in the ASD group (*n* = 19; 87.5%). Family Socioeconomic Status (SES; Hollingshead, 1975) is similar for both groups ranging from medium to high. All the children of the clinical group had a certified clinical diagnosis of ASD without intellectual disability (IQ greater than 70), based on clinical judgment according to the *Diagnostic Statistical Manual of Mental Disorders 5th Edition* (DSM V; American Psychiatric Association [APA], 2013). Children with intellectual disability, severe impairment of cognitive functioning, co-occurring psychiatric disorders, and deficits in expressive and receptive language were excluded from the study. For the TD group, we did not include children with a history of psychiatric disorder.

Measures

Verbal IQ

Wechsler Scale of Intelligence for Children (WISC-IV; Wechsler, 2003) is the most widely used standardized tool in the field of developmental psychological assessment. According to the methodology and aims of this research, we used two WISC core subscales (Similarities and Vocabulary) to generate an index of child verbal intelligence. Similarities provide an estimation of child verbal abstract reasoning. This core subtest also involves language development, lexical knowledge, auditory comprehension, memory, and the ability to discriminate between essential and non-essential features. Vocabulary (VOC) offers a measure of child lexical knowledge and formation of verbal concepts.

Non-verbal IQ

Raven Colored Progressive Matrices (CPM; Raven et al., 1962) is a widely recognized individual non-verbal assessment of children intelligence based on figural materials. Specifically, this tool evaluates non-verbal perceptual and inductive reasoning in children from 3 to 11 years old independently from the culture or cognitive impairment. It consists of three series of 12 items developed to measure the main characteristic processes of this age group. A general weighted score was calculated by adding the child's correct answers.

Attachment to Parents

Inventory of Parent and Peer Attachment (IPPA; Armsden and Greenberg, 1987) is a self-reported measure aimed to assess how children and adolescents perceive their parents and close friends as a source of psychological security. In this study, we used the version related to the relationship with parents only. It comprises 28 items rated on a five-point Likert-type scale from 1 = "almost never or never true" to "almost always or always true," which generates three subscales (Trust, Communication, and Alienation) and a total score. Higher scores indicate a positive perception of the attachment relationship. This questionnaire showed adequate psychometric properties (Jewell et al., 2019), and it has been used to investigate the perception of attachment security in children with ASD (without intellectual disability) during middle childhood (Teague et al., 2017).

Alexithymia

Alexithymia Questionnaire for Children (AQC; Rieffe et al., 2006) is a self-reported measure to assess alexithymia in young children. It is adapted from the well-validated measure used for the assessment in adulthood (TAS-20; Bagby et al., 1994) to be developmentally appropriate preserving similar structure and content. The AQC is composed of 20 items rated on a three-point Likert-type scale (ranging from 0 = not true to 2 = often true) representing three core factors: (a) Difficulty Identifying Feelings (DIF; seven items): (b) Difficulty Describing Feelings (DDF; five items); and (c) Externally Oriented Thinking (EOT; eight items). An example of an AQC item is "I can easily say how I feel inside." Five items of the scale were formulated positively, for example, "It is important to understand how you feel inside" and thus the scoring was reversed.

In this study, we used the Italian version of the questionnaire (Di Trani et al., 2009, 2018). Higher scores correspond to the elevated presence of this factor. The validity and reliability of this measure were confirmed by empirical data (Rieffe et al., 2006; Di Trani et al., 2009, 2018), except for EOT, which showed low reliability. In our sample, AQC internal consistency was good for the total score (α = 0.678), DIF (α = 0.718), and DDF (α = 623) and poor for EOT (α = 0.035). This measure has already been used in children with ASD and other neurodevelopmental disorders (Donfrancesco et al., 2013; Griffin et al., 2016).

Data Analysis

The statistical analysis of the data was carried out using the statistical package SPSS (22.0 for Windows). As a preliminary analysis, we checked for the normality of the distribution, outliers, and linearity and we tested bivariate correlations among alexithymia and attachment scores. A one-way multivariate analysis of variance (MANOVA) was used to test group differences on the control variable in order to detect potential covariates. Similarly, with chi-squared test, we assessed potential differences between groups regarding gender distribution. In addition, we transformed IPPA and AQC total scores into z scores in order to standardize both variables of interest. With respect to the first aim of the study, we performed a Mann–Whitney test to compare the two study groups on attachment total scores. For group differences on alexithymia, we performed a Student's t-test for the total score and a one-way MANOVA for the three subscales of the questionnaire. One outlier (1 TD) was removed from the statistical analysis. Using a categorical cut-off score based on the normative values of AQC Italian validation study (Di Trani et al., 2018), we calculated rates of children at risk of alexithymia. Specifically, we determined the threshold parameters using the normative means + 1 SD clustered by two age groups (8–10 and 11–14 years old) and child gender. Differences between groups on this categorical variable were tested using the Fisher exact test. Next, we used a hierarchical linear regression to test which child variables contributed significantly in predicting child attachment to parent scores. Specifically, the first step of the regression model included child age, verbal IQ, and ASD status (presence/absence). In the second step of analysis, we added child self-reported alexithymia score to test its independent predictive effect on the overall IPPA model. A third step was included to test whether the effect of the alexithymia on attachment is moderated by the ASD status.

RESULTS

The AQC scores were normally distributed in both ASD and control group as opposed to IPPA total scores. Children with ASD did not differ from the TD group with respect to the control variables such as child age, verbal and non-verbal ability, and family SES (see **Table 1**). Similarly, no difference was found in the gender distribution between groups (Fisher exact test, $p = 0.09$). Correlational analysis is shown in **Table 2**. By addressing the first study aim, there was a significant difference between the two groups on AQC total score [$t_{(45)} = -2.36$, $p = 0.022$, Partial eta squared = 0.111] with children of the ASD group showing a higher level of alexithymia. The MANOVA was significant for the group effect on the AQC subscales [Wilk's Lambda = 0.821, $F(3,43) = 3.12$, $p = 0.035$]. Among the univariate tests, only the DDF subscale was significant, $F(1,45) = 5.65$, $p = 0.022$, whereas a prominent trend that approached significance was found for the EOT subscale, $F(1,45) = 3.76$, $p = 0.059$.

TABLE 1 | Descriptive statistics of the study variable for the ASD and TD group.

	ASD		TD		Group differences	
	M	***SD***	***M***	***SD***	***p***	***Eta***
Age (months)	126.43	16.45	117.52	24.32	0.185	0.045
Verbal IQ	22.45	4.00	22.91	4.94	0.722	0.003
Perceptual reasoning^{a}	0.52	80	0.67	0.78	0.590	0.008
Family SES	43.52	13.79	45.39	15.23	0.607	0.007
Alexithymia total score	39.50	5.75	36.09	3.90	0.022*	0.111
Difficulty identifying feelings	13.12	3.55	12.13	2.41	0.022*	0.112
Difficulty describing feelings	10.16	2.38	8.65	1.94	0.270	0.027
Externally oriented thinking	16.20	2.62	14.91	1.88	0.059	0.077
Attachment to parents	61.19	8.442	66.43	13.75	0.090	0.064

ASD, autism spectrum disorder; TD, typical development; IQ, intelligence quotient; SES, socioeconomic status. ^{a}*z scores;* **p* < *0.05.*

TABLE 2 | Spearman correlations among perceived attachment to parents and alexithymia total score and subscales.

Variable	1	2	3	4	5
1. Alexithymia total score	–				
2. Difficulty identifying feelings	0.810***	–			
3. Difficulty describing feelings	0.745***	−0.501***	–		
4. Externally oriented thinking	0.483**	0.139	0.124	–	
5. Attachment to parents	−0.552***	−0.402**	−0.398**	−0.321*	–

**p* < *0.05;* ***p* < *0.01;* ****p* < *0.001.*

In addition, we found higher rates of alexithymia in the ASD group; specifically 20.8% of children with ASD were above the cut-off compared to 8.3% of the typical development group. However, only a marginal trend toward significance emerged from the analysis (Fisher exact test, $p = 0.091$). The same analysis was conducted without removing the outlier showing similar results. In this case, though no statistical differences emerged on AQC total score, we found a substantial trend toward significance [$t_{(46)} = -1.90$, $p = 0.063$]. No significant differences were found between children with ASD and the control group on IPPA total scores ($Z = 1.69$, $p = 0.090$, Partial eta squared = 0.064). To ascertain the effect of child features and the contribution of alexithymia, a hierarchical linear regression including two separate steps was carried out (see **Table 3**). The first step that included child age, verbal IQ, and ASD status was not statistically significant [$F_{(3,42)} = 2.44$, $p = 0.078$]. Nevertheless, child age was positively associated with attachment to parents ($\beta = 0.349$), whereas the contribution of verbal IQ and ASD status was not statistically significant. By entering AQC total score as an independent predictor in the second step of the linear regression, the overall model was significant [$F_{(4,42)} = 3.54$, $p = 0.015$], explaining 19.5% of the variance. The p-value associated with adjusted R squared change for the second step is also statistically significant [$F_{(1,38)} = 5.93$, $p = 0.020$]. Specifically, the data revealed that alexithymia was the significant independent negative regressor ($\beta = -0.361$), whereas no effect was found for child age, verbal IQ, and ASD status. Similarly,

TABLE 3 | Hierarchical regression analysis of perceived attachment to parents by child age, verbal IQ, ASD status, and self-reported alexithymia.

Predictors	**Attachment to parents**				
	β	***SE***	***p***	ΔR^2	***p ch***
Step 1			0.078	0.158	
Age	0.349	0.084	0.030*		
Verbal IQ	0.122	0.398	0.424		
ASD status	−0.285	3.51	0.066		
Step 2			0.015*	0.114	0.020*
Age	0.255	0.082	0.099		
Verbal IQ	0.110	0.375	0.446		
ASD status	−0.172	3.48	0.257		
Alexithymia	−0.361	0.350	0.020*		
Step 3			0.13*	0.041	0.144
Age	0.250	0.080	0.100		
Verbal IQ	0.071	0.375	0.620		
ASD status	−1.85	26.58	0.111		
Alexithymia	−1.10	1.23	0.040*		
Alexithymia ∗ ASD status	2.02	0.713	0.144		

ASD, autism spectrum disorder; TD, typical development, IQ, intelligence quotient; SES, socioeconomic status; $^*p < 0.05$.

alexithymia was still significant in the third step of the regression ($\beta = -1.10$; $p = 0.040$), whereas no interaction effect with ASD status was observed.

DISCUSSION

The current study examined alexithymia in school-age children with ASD (without intellectual disability), exploring its influence on the perception of attachment security to parents. To this aim, we tested the hypothesis that alexithymia may contribute to a negative view of the quality of the relationship with parents from the child's perspective. To date, there are no studies that have investigated this specific link in children with atypical neurodevelopmental condition, including ASD.

Firstly, we found that children with ASD and TD showed no significant differences in the perception of attachment security to parents. Although a limited number of studies have been conducted on this topic, our results are consistent with previous findings. According to earlier meta-analytic findings, recent empirical evidence revealed that school-age children with ASD (without intellectual disability) reported similar levels of security in the relationship with their parents to those found in typically developing children (Teague et al., 2017). However, considering the negative impact of ASD sociocommunicational difficulties and emotional reactivity on quality of attachment bond and parent – child interaction, these findings raised questions about which mechanism may explain child-positive perception of the attachment relationship.

Secondly, in line with the only study available (Griffin et al., 2016), our results showed that school-age children with ASD reported higher levels of alexithymia compared to their typical counterpart. Specifically, children with ASD reported more difficulties in describing their feelings and inner states.

Moreover, we found that alexithymia is more common in ASD, also in school-age, with approximately one in five reporting scores of alexithymia above the cut-off. Our finding confirms the difficulties of children with ASD in cognitive processing of their own emotions, documented by previous research on adolescent and adulthood (Bird and Cook, 2013; Milosavljevic et al., 2016). As expected, we did not find a large effect on this group difference given that the children's self-report may be less reliable compared to other ratings provided by child informants. In fact, the limited abilities of individuals with ASD with respect to self-referential cognition (Lombardo et al., 2007) may undermine the accuracy of self-reported measurement. It is essential to consider other aspects associated to ASD phenotype that are strictly interrelated to alexithymia (Fitzgerald and Bellgrove, 2006) such as impairment in mentalizing and self-reflection, less coherent representations of emotional experience (Losh and Capps, 2006), absence of emotional vocalization (Heaton et al., 2012), behavioral rigidity, and impaired inhibitory control (Mosconi et al., 2009; D'Cruz et al., 2013). Thus, even if we checked for linguistic abilities, elevated levels and rates of alexithymia in children with ASD may be explained at least to some extent by these specific disorders of cognitive, emotional, and behavioral functioning.

According to the second aim of this study, we found a significant link between alexithymia and perception of attachment security in children with and without ASD. In particular, among child characteristics, alexithymia level and not ASD diagnosis predicts the extent to which children perceive their relationship with their parents as a source of security in middle childhood. It is conceivable that a specific deficit in identifying and describing one's own feelings may hinder the construction of a positive representation of parent – child attachment relationship regardless of child clinical status. With respect to ASD, despite a growing area of research linking alexithymia and children psychological outcomes (Brewer et al., 2015; Morie et al., 2019), this is the first investigation documenting the significant impact of alexithymia on the perception of attachment security to parents. An ongoing debate in ASD is whether the occurrence of alexithymia affects social motivation, influencing attitudes and behaviors at an interpersonal level (Pastore et al., 2019). Studies reported that alexithymia in children with ASD was associated with less expressive coherence (Costa et al., 2017), empathy, and perspective taking (Lartseva et al., 2015) as well as lower enjoyment of prosocial interactions (Gebauer et al., 2014). Moreover, as highlighted by Costa et al. (2019), alexithymia in children (more than ASD status) can negatively affect parent–child relationships, explaining the reduced amount of dyadic exchanges. The mismatch between arousal activation and subjective experience of feelings (Gaigg et al., 2018) may also contribute to the formation of less coherent child representation influencing how information is encoded and processed. Therefore, the possibility to develop unbalanced representations of the attachment figures may significantly increase given the potential negative consequences of alexithymia on different levels of emotion processing and relational exchanges. Furthermore, giving a coherent meaning to their own interpersonal experiences with parental figures may be more complex for the limited personal resources in emotional self-understanding, regulation, and expression. In fact, children with ASD and alexithymia may also exhibit difficulties in interpreting and responding to emotion in others (Poquérusse et al., 2018) as in the case of child caregivers. Impaired mentalizing associated with low self-memory in ASD (Lombardo et al., 2007) and difficulties with episodic autobiographical memories (Lind, 2010) constitute additional risk factors for the construction of coherent explicit representations based on past relational experiences. Following this direction, our findings extend literature on the influence of alexithymia on socioemotional development by considering a child's self-perspective on attachment to parents during school age.

In sum, our findings showed higher levels of alexithymia compared to the control group, whereas the perception of attachment security was similar between groups. Notably, alexithymia, not autism, was found to be the only significant predictor of child attachment to parents. Thus, alexithymia seems to play a key role on the way children with ASD perceive their relationship with their mothers and fathers. In this regard, several aspects associated with alexithymia such as impaired emotion processing, neurophysiological atypical processes, reduced mentalizing, and low self-memory may hamper parent – child relationship and consequently child-explicit representations.

Lastly, some limitations of this study need to be acknowledged. We do not include a measure of alexithymia rated by child informants as suggested by prior research (Griffin et al., 2016). Thus, a comparison between two different sources of information on alexithymia scores was not possible.

Additionally, a small sample size, a cross-sectional design, and the lack of a continuous measure of ASD symptoms severity represent other specific drawbacks of this study. Future research can expand these findings including a measure of child-implicit attachment representations. In general, our results confirmed that alexithymia could be useful in subgrouping and identifying specific cognitive profiles within the autism spectrum condition. We also suggest alexithymia as a potential covariate in the comparative study on ASD. Another possible indication is to assess alexithymia in parents to examine whether there are direct associations with child outcomes. In terms of clinical implications, we recommend assessing alexithymia adequately in school-age children with ASD in order to acquire specific information on their emotional functioning. Children with ASD and co-occurring higher level of alexithymia may benefit from interventions that combine evidence-based treatment for sociocommunicational difficulties and specific strategies aimed at enhancing the cognitive processing of their own emotions. In conclusion, our findings confirmed the importance to target children alexithymia in ASD, considering its clinical significance not only on emotion processing but also on other significant domains of socioemotional development, as is the case of attachment to parental figures. Interventions involving mothers and fathers aimed at boosting the quality of relational exchanges and child emotional capacities should evaluate and address alexithymia given its prominent contribution to child adjustment in ASD during middle childhood.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Human Research Ethics Committee of the University of Trento. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

MG, SF, and PV contributed to the design and implementation of the research. MG collected and analyzed the data. All authors discussed the results and commented on the manuscript.

REFERENCES

American Psychiatric Association [APA] (2013). *Diagnostic and Statistical Manual of Mental Disorders*, 5th Edn, Arlington, VA: American Psychiatric Association [APA].

Armsden, G., and Greenberg, M. (1987). The inventory of parent and peer attachment: individual differences and their relationship to psychological well-being in adolescence. *J. Youth Adolesc.* 16, 427–454. doi: 10.1007/BF02202939

Bagby, R. M., Parker, J. D., and Taylor, G. J. (1994). 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

Bauminger, N., Solomon, M., and Rogers, S. J. (2010). Externalizing and internalizing behaviors in ASD. *Autism Res.* 3, 101–112. doi: 10.1002/aur.131

Bernhardt, B. C., Valk, S. L., Silani, G., Bird, G., Frith, U., and Singer, T. (2013). Selective disruption of sociocognitive structural brain networks in autism and alexithymia. *Cereb. Cortex* 24, 3258–3267. doi: 10.1093/cercor/bht182

Bird, G., and Cook, R. (2013). Mixed emotions: the contribution of alexithymia to the emotional symptoms of autism. *Transl. Psychiatr.* 3:e285. doi: 10.1038/tp.2013.61

Bird, G., Silani, G., Brindley, R., White, S., Frith, U., and Singer, T. (2010). Empathic brain responses in insula are modulated by levels of alexithymia but not autism. *Brain* 133, 1515–1525. doi: 10.1093/brain/awq060

Brewer, R., Marsh, A. A., Catmur, C., Cardinale, E. M., Stoycos, S., Cook, R., et al. (2015). The impact of autism spectrum disorder and alexithymia on judgments of moral acceptability. *J. Abnorm. Psychol.* 124:589. doi: 10.1037/abn0000076

Brumariu, L. E., Kerns, K. A., and Seibert, A. (2012). Mother–child attachment, emotion regulation, and anxiety symptoms in middle childhood. *Pers. Relat.* 19, 569–585. doi: 10.1016/j.bbr.2016.07.035

Caria, A., and de Falco, S. (2015). Anterior insular cortex regulation in autism spectrum disorders. *Front. Behav. Neurosci.* 9:38. doi: 10.3389/fnbeh.2015.00038

Carlson, E. A., and Sroufe, L. A. (1995). "Contribution of attachment theory to developmental psychopathology," in *Developmental Psychopathology: Theory and Method*, Vol. 1, eds D. Cicchetti, and D. J. Cohen (Oxford: Wiley), 581–616.

Chandler, F., and Dissanayake, C. (2014). An investigation of the security of caregiver attachment during middle childhood in children with high-functioning autistic disorder. *Autism* 18, 485–492. doi: 10.1177/1362361313486205

Constantino, J. N., and Gruber, C. P. (2005). *Social Responsiveness Scale*. Los Angeles: Western Psychological Services.

Costa, A. P., Steffgen, G., and Samson, A. C. (2017). Expressive incoherence and alexithymia in autism spectrum disorder. *J. Autism Dev. Disord.* 47, 1659–1672. doi: 10.1007/s10803-017-3073-9

Costa, A. P., Steffgen, G., and Vögele, C. (2019). The role of alexithymia in parent–child interaction and in the emotional ability of children with autism spectrum disorder. *Autism Res.* 12, 458–468. doi: 10.1002/aur.2061

D'Cruz, A. M., Ragozzino, M. E., Mosconi, M. W., Shrestha, S., Cook, E. H., and Sweeney, J. A. (2013). Reduced behavioral flexibility in autism spectrum disorders. *Neuropsychology* 27:152. doi: 10.1037/a0031721

Di Trani, M., Presaghi, F., Renzi, A., Greenmanc, P. S., and Solanoa, L. (2018). The Italian version of the alexithymia questionnaire for children (AQC): factor structure and reliability. La versione italiana dell'Alexithymia Questionnaire for Children (AQC): struttura fattoriale e attendibilità. *Rassegna Psicol.* 35, 47–60.

Di Trani, M., Tomassetti, N., Bonadies, M., Capozzi, F., De Gennaro, L., Presaghi, F., et al. (2009). Un Questionario Italiano per l'Alessitimia in Età Evolutiva: struttura fattoriale e attendibilità. *Psicologia Della Salute* 2, 131–143. doi: 10.3280/PDS2009-002009

Donfrancesco, R., Di Trani, M., Gregori, P., Auguanno, G., Melegari, M. G., Zaninotto, S., et al. (2013). Attention-deficit/hyperactivity disorder and alexithymia: a pilot study. *ADHD Attent. Def. Hyperact. Disord.* 5, 361–367.

Fietz, J., Valencia, N., and Silani, G. (2018). Alexithymia and autistic traits as possible predictors for traits related to depression, anxiety, and stress: a multivariate statistical approach. *J. Eval. Clin. Pract.* 24, 901–908. doi: 10.1111/jep.12961

Fitzgerald, M., and Bellgrove, M. A. (2006). The overlap between alexithymia and Asperger's syndrome. *J. Autism Dev. Disord.* 36:573. doi: 10.1007/s10803-006-0096-z

Gaigg, S. B., Cornell, A. S., and Bird, G. (2018). The psychophysiological mechanisms of alexithymia in autism spectrum disorder. *Autism* 22, 227–231. doi: 10.1177/1362361316667062

Gebauer, L., Skewes, J., Westphael, G., Heaton, P., and Vuust, P. (2014). Intact brain processing of musical emotions in autism spectrum disorder, but more cognitive load and arousal in happy vs. sad music. *Front. Neurosci.* 8:192. doi: 10.3389/fnins.2014.00192

Griffin, C., Lombardo, M. V., and Auyeung, B. (2016). Alexithymia in children with and without autism spectrum disorders. *Autism Res.* 9, 773–780. doi: 10.1002/aur.1569

Heaton, P., Reichenbacher, L., Sauter, D., Allen, R., Scott, S., and Hill, E. (2012). Measuring the effects of alexithymia on perception of emotional vocalizations in autistic spectrum disorder and typical development. *Psychol. Med.* 42, 2453–2459. doi: 10.1017/S0033291712000621

Hill, E., Berthoz, S., and Frith, U. (2004). Brief report: cognitive processing of own emotions in individuals with autistic spectrum disorder and in their relatives. *J. Autism Dev. Disord.* 34, 229–235. doi: 10.1023/b:jadd.0000022613.41399.14

Hollingshead, A. A. (1975). *Four-Factor Index Of Social Status*. New Haven, CT: Yale University.

Jewell, T., Gardner, T., Susi, K., Watchorn, K., Coopey, E., Simic, M., et al. (2019). Attachment measures in middle childhood and adolescence: a systematic review of measurement properties. *Clin. Psychol. Rev*. 68, 71–82. doi: 10.1016/j.cpr.2018.12.004

Kinnaird, E., Stewart, C., and Tchanturia, K. (2019). Investigating alexithymia in autism: a systematic review and meta-analysis. *Eur. Psychiatr.* 55, 80–89. doi: 10.1016/j.eurpsy.2018.09.004

Koelen, J. A., Eurelings-Bontekoe, E. H., Stuke, F., and Luyten, P. (2015). Insecure attachment strategies are associated with cognitive alexithymia in patients with severe somatoform disorder. *Int. J. Psychiatr. Med.* 49, 264–278. doi: 10.1177/0091217415589303

Lai, M. C., Lombardo, M. V., Chakrabarti, B., and Baron-Cohen, S. (2013). Subgrouping the autism "spectrum": reflections on DSM-5. *PLoS Biol.* 11:e1001544. doi: 10.1371/journal.pbio.1001544

Laible, D. (2007). Attachment with parents and peers in late adolescence: links with emotional competence and social behavior. *Pers. Individ. Differ.* 43, 1185–1197. doi: 10.1016/j.paid.2007.03.010

Lartseva, A., Dijkstra, T., and Buitelaar, J. K. (2015). Emotional language processing in autism spectrum disorders: a systematic review. *Front. Hum. Neurosci.* 8:991. doi: 10.3389/fnhum.2014.00991

Lind, S. E. (2010). Memory and the self in autism: a review and theoretical framework. *Autism* 14, 430–456. doi: 10.1177/1362361309358700

Lombardo, M. V., Barnes, J. L., Wheelwright, S. J., and Baron-Cohen, S. (2007). Self-referential cognition and empathy in autism. *PLoS One* 2:e883. doi: 10.1371/journal.pone.0000883

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., Bishop, S. L., et al. (2015). *Autism Diagnostic Observation Schedule ADOS-2. Manual (Part I): Modules*. Torrance, CA: Western Psychological Services.

Losh, M., and Capps, L. (2006). Understanding of emotional experience in autism: insights from the personal accounts of high-functioning children with autism. *Dev. Psychol.* 42:809. doi: 10.1037/0012-1649.42.5.809

Milosavljevic, B., Leno, V. C., Simonoff, E., Baird, G., Pickles, A., Jones, C. R., et al. (2016). Alexithymia in adolescents with autism spectrum disorder: its relationship to internalising difficulties, sensory modulation and social cognition. *J. Autism Dev. Disord.* 46, 1354–1367. doi: 10.1007/s10803-015-2670-8

Morie, K. P., Jackson, S., Zhai, Z. W., Potenza, M. N., and Dritschel, B. (2019). Mood disorders in high-functioning autism: the importance of alexithymia and emotional regulation. *J. Autism Dev. Disord.* 49, 2935–2945. doi: 10.1007/s10803-019-04020-1

Moriguchi, Y., Ohnishi, T., Lane, R. D., Maeda, M., Mori, T., Nemoto, K., et al. (2006). Impaired self-awareness and theory of mind: an fMRI study of mentalizing in alexithymia. *Neuroimage* 32, 1472–1482. doi: 10.1016/j.neuroimage.2006.04.186

Mosconi, M. W., Kay, M., D'cruz, A. M., Seidenfeld, A., Guter, S., Stanford, L. D., et al. (2009). Impaired inhibitory control is associated with higher-order repetitive behaviors in autism spectrum disorders. *Psychol. Med.* 39, 1559–1566. doi: 10.1017/S0033291708004984

Oskis, A., Clow, A., Hucklebridge, F., Bifulco, A., Jacobs, C., and Loveday, C. (2013). Understanding alexithymia in female adolescents: the role of attachment style. *Pers. Individ. Differ.* 54, 97–102. doi: 10.1177/0004867414546386

Pastore, L., Dellantonio, S., Mulatti, C., and Esposito, G. (2019). To be or not to be emotionally aware and socially motivated : how alexithymia impacts autism spectrum disorders. *Behav. Brain Sci.* 42:e106. doi: 10.1017/S0140525X18002315

Picardi, A., Toni, A., and Caroppo, E. (2005). Stability of alexithymia and its relationships with the 'big five'factors, temperament, character, and attachment style. *Psychother. Psychosom.* 74, 371–378. doi: 10.1159/000087785

Poquérusse, J., Pastore, L., Dellantonio, S., and Esposito, G. (2018). Alexithymia and autism spectrum disorder: a complex relationship. *Front. Psychol.* 9:1196. doi: 10.3389/fpsyg.2018.01196

Raven, J. C., Raven, J., and Court, J. (1962). *Coloured Progressive Matrices*. Oxford: Oxford Psychologists Press.

Rieffe, C., Oosterveld, P., and Terwogt, M. M. (2006). An alexithymia questionnaire for children: factorial and concurrent validation results. *Pers. Individ. Differ.* 40, 123–133. doi: 10.1016/j.paid.2005.05.013

Shah, P., Hall, R., Catmur, C., and Bird, G. (2016). Alexithymia, not autism, is associated with impaired interoception. *Cortex* 81, 215–220. doi: 10.1016/j.cortex.2016.03.021

Sifneos, P. E. (1973). The prevalence of 'alexithymic'characteristics in psychosomatic patients. *Psychother. Psychosom.* 22, 255–262. doi: 10.1159/000286529

Silani, G., Bird, G., Brindley, R., Singer, T., Frith, C., and Frith, U. (2008). Levels of emotional awareness and autism: an fMRI study. *Soc. Neurosci.* 3, 97–112. doi: 10.1080/17470910701577020

Sivaratnam, C., Newman, L., and Rinehart, N. (2018). Emotion-recognition and theory of mind in high-functioning children with ASD: relationships with attachment security and executive functioning. *Res. Autism Spectr. Disord.* 53, 31–40. doi: 10.1016/j.rasd.2018.05.005

Teague, S. J., Gray, K. M., Tonge, B. J., and Newman, L. K. (2017). Attachment in children with autism spectrum disorder: a systematic review. *Res. Autism Spectr. Disord.* 35, 35–50. doi: 10.1016/j.ridd.2012.06.015

Thompson, R. A., and Meyer, S. (2007). "Socialization of emotion regulation in the family," in *Handbook of Emotion Regulation*, ed. J. J. Gross (New York, NY: Guilford), 249–268.

van der Velde, J., Servaas, M. N., Goerlich, K. S., Bruggeman, R., Horton, P., Costafreda, S. G., et al. (2013). Neural correlates of alexithymia: a meta-analysis of emotion processing studies. *Neurosci. Biobehav. Rev.* 37, 1774–1785. doi: 10.1016/j.neubiorev.2013.07.008

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children Fourth Edition*. San Antonio, TX: Pearson.

Conflict of Interest: 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.

frontiers
in Psychology

ORIGINAL RESEARCH
published: 25 March 2020
doi: 10.3389/fpsyg.2020.00478

"Motion or Emotion? Recognition of Emotional Bodily Expressions in Children With Autism Spectrum Disorder With and Without Intellectual Disability"

Noemi Mazzoni[1]*, Isotta Landi[1,2], Paola Ricciardelli[3,4], Rossana Actis-Grosso[3,4] and Paola Venuti[1]

[1] ODFLab – Department of Psychology and Cognitive Sciences, University of Trento, Rovereto, Italy, [2] MPBA, Fondazione Bruno Kessler, Trento, Italy, [3] Department of Psychology, University of Milano – Bicocca, Milan, Italy, [4] Milan Centre for Neuroscience, Milan, Italy

The recognition of emotional body movement (BM) is impaired in individuals with Autistic Spectrum Disorder ASD, yet it is not clear whether the difficulty is related to the encoding of body motion, emotions, or both. Besides, BM recognition has been traditionally studied using point-light displays stimuli (PLDs) and is still underexplored in individuals with ASD and intellectual disability (ID). In the present study, we investigated the recognition of happy, fearful, and neutral BM in children with ASD with and without ID. In a non-verbal recognition task, participants were asked to recognize pure-body-motion and visible-body-form stimuli (by means of point-light displays-PLDs and full-light displays-FLDs, respectively). We found that the children with ASD were less accurate than TD children in recognizing both the emotional and neutral BM, either when presented as FLDs or PLDs. These results suggest that the difficulty in understanding the observed BM may rely on atypical processing of BM information rather than emotion. Moreover, we found that the accuracy improved with age and IQ only in children with ASD without ID, suggesting that high level of cognitive resources can mediate the acquisition of compensatory mechanisms which develop with age.

Keywords: autism spectrum disorder, biological motion, emotional bodily expressions, emotion recognition, intellectual disability

OPEN ACCESS

Edited by:
Katsumi Watanabe,
Waseda University, Japan

Reviewed by:
Masahiro Hirai,
Jichi Medical University, Japan
Frank Pollick,
University of Glasgow,
United Kingdom

****Correspondence:***
Noemi Mazzoni
noemi.mazzoni@unitn.it;
noemimax@gmail.com

Specialty section:
This article was submitted to Consciousness Research, a section of the journal Frontiers in Psychology

Received: *03 October 2019*
Accepted: *02 March 2020*
Published: *25 March 2020*

Citation:
Mazzoni N, Landi I, Ricciardelli P, Actis-Grosso R and Venuti P (2020) "Motion or Emotion? Recognition of Emotional Bodily Expressions in Children With Autism Spectrum Disorder With and Without Intellectual Disability". Front. Psychol. 11:478. doi: 10.3389/fpsyg.2020.00478

INTRODUCTION

As social individuals, we are constantly surrounded by other people and daily involved in social interactions. Hence, the ability to comprehend correctly the social signals is fundamental to our species. Among the social signals, the body movement (BM) plays a key role in conveying other people's feeling, intentions, and emotions and its importance is comparable to that of facial

---

Abbreviations: ASD, autism spectrum disorder; BM, body movement; FLDs, full-light display stimuli; HFA, high-functioning autism spectrum disorder; ID, intellectual disability; LFA, low-functioning autism spectrum disorder; PLDs, point-light display stimuli; TD, typical developing.

expressions (de Gelder et al., 2010; de Gelder and de Borst, 2015), or even more (Van den Stock et al., 2007; Aviezer et al., 2012).

A solely kinematic description of BM is conveyed by biological motion, traditionally represented by means of point light displays (PLDs, Johansson, 1973). In typical developing individuals (TD), such a description is sufficient for inferring a great amount of social information, such as identity (Kozlowski and Cutting, 1977; Troje et al., 2005); gender (Kozlowski and Cutting, 1977; Pollick et al., 2002; Johnson et al., 2011); the nature of the actions (Johansson, 1973; Dittrich, 1993; Alaerts et al., 2011); intentions (Roché et al., 2013); and – more importantly to the present study – emotions (Dittrich et al., 1996; Atkinson et al., 2004; Clarke et al., 2005; Roether et al., 2009).

In contrast, the ability to perceive and recognize the BM seems to be impaired in autisms spectrum disorder (ASD). ASD is a neurodevelopmental disorder characterized by persistent deficits in social communication and social interaction, in social-emotional reciprocity, and in non-verbal communicative behavior (American Psychiatric Association, 2013). In the last two decades, an increasing amount of studies have suggested that social difficulties in ASD may be associated with difficulties in processing the BM (Blake et al., 2003; Dakin and Frith, 2005; Annaz et al., 2010) and in understanding the information that it conveys (Grèzes et al., 2009; Kaiser et al., 2010; Mcpartland et al., 2012; Pavlova, 2012). For instance, children with ASD showed early difficulties in orienting toward BM (Klin and Jones, 2008; Klin et al., 2009) and no preference for biological over scrambled motion (Blake et al., 2003; Annaz et al., 2012). Moreover, difficulties in recognizing *emotional* BM has been consistently reported in children (Moore et al., 1997; Fridenson-hayo et al., 2016), adolescents (Hubert et al., 2007; Parron et al., 2008) and adults with ASD (Atkinson, 2009; Philip et al., 2010; Nackaerts et al., 2012; Alaerts et al., 2014) or with high autistic traits (Actis-Grosso et al., 2015). However, other studies failed to find differences between neurotypical individuals and people with ASD when emotionally *neutral* movements were presented (Moore et al., 1997; Hubert et al., 2007; Parron et al., 2008; Murphy et al., 2009; Saygin et al., 2010). Therefore, it is still not clear whether the social difficulty in ASD is related to difficulties in processing the BM, the emotional content, or both.

Notably, most of the previous studies have focused on PLDs recognition. However, in daily life the body form is fully visible and this should be taken into account when social difficulties are investigated. Studies that have used stimuli with visible body form (full-light display stimuli – FLDs) together with PLDs, are scarce. Research in TD children (Ross et al., 2012) and adults (Atkinson et al., 2004) found an advantage in recognizing emotional FLDs compared to PLDs, suggesting a better recognition of emotional BM when more natural – although more complex – stimuli were used. In ASD, the investigation of differences in recognizing emotional PLDs and FLDs was limited to adults and showed an impaired recognition of both PLDs and FLDs (Atkinson, 2009). Poorer accuracy in recognizing emotional FLDs of BM have been shown in children with ASD (Fridenson-hayo et al., 2016). Yet, differences in recognizing emotional FLDs and PLDs have not been investigated.

Furthermore, individuals with ASD and intellectual disabilities (ID; namely with an IQ < 70) and/or with language impairment have rarely been included in past research on emotional BM recognition. Indeed, most of the previous studies have investigated the recognition of emotional BM in people without ID (namely with an IQ ≥ 70), using verbal tasks and verbal abilities as groups matching criteria. Critically, the cognitive profile of people with ASD is commonly uneven with non-verbal abilities higher than verbal ones (e.g., Happé, 1994; Joseph et al., 2002; Black et al., 2009). Therefore, it is possible that the groups of TD and ASD participants were not well matched in past research, since similar verbal skills might correspond to higher non-verbal abilities in people with ASD. Moreover, the social skills of people with ASD vary considerably according to IQ as good cognitive abilities may mediate the acquisition of compensatory strategies – that are declarative and learned explicitly through experience and/or training – for dealing with social and emotional signals (McKay et al., 2012; Rutherford and Troje, 2012). Altogether, these issues could explain why some studies failed to find differences between people with ASD and controls (Herrington et al., 2007; Freitag et al., 2008; Murphy et al., 2009; Saygin et al., 2010). An examination of emotional BM recognition in children with ASD using non-verbal task and non-verbal IQ as group-matching criteria is thus desirable. In the absence of verbal components, such procedure could also be used to compare children with ASD with and without ID or verbal disorders.

On these premises, the aim of the present study was to investigate which are the critical aspects of emotional BM that are impaired in children with ASD (i.e., difficulty in the processing of the emotional content, the BM, or both). Building on previous studies, we presented both stimuli conveying motion-kinematic-only (PLDs) and stimuli where the body form was visible (FLDs); we included children with ASD with and without intellectual disabilities (ID); we adopted the non-verbal IQ as matching criteria; and we asked the participants to perform a non-verbal simplified emotion recognition task. To assess the role of body motion information processing and emotion recognition, we presented emotional (happy–positive; fearful-negative) and neutral BM. If the difficulty in recognizing the meaning of BM in ASD were associated with difficulties in the encoding and/or processing of body motion, we would expect an impaired recognition of neutral as well as emotional BMs. Conversely, if the difficulty in recognizing the meaning of BM were related to the processing and recognition of the emotional content, children with ASD would show a specific impairment in the recognition of happy and/or fearful, but not neutral, movements. Likewise, by comparing PLDs and FLDs, we aimed at testing whether the difficulty in children with ASD was related to the visual characteristics (i.e., pure kinematic or visible body form) of the stimuli or instead to difficulties in processing the BM *per se*. If that were the case, we would expect an impaired performance in both display types (PLDs and FLDs). By contrast, if the impairment were specifically related to a difficulties in perceiving a human being depicted as PLDs (due to the tendency of individuals with ASD to process parts of the stimuli and their difficulty to integrate these part into a gestalt, e.g., Happé and Frith, 2006), an impaired performance in PLDs, but not in FLDs, should be found. Finally, regarding the age- and the non-verbal IQ- factors, we predicted that the ability to recognize emotional BM would improve with age in TD children (Ross et al., 2012) but to a lesser extent in children with ASD (Blake et al., 2003; Annaz et al., 2010). In addition, we expected that higher cognitive resources would lead to a more accurate and faster BM recognition in TD children, and, in turn, would allow ASD children to compensate and obviate, at least in part, their impairment in BM comprehension.

MATERIALS AND METHODS

Stimuli

Participants were presented with short video-clips displaying PLDs and FLDs of whole-body expressions of happiness, fear, and neutral actions. The stimuli were adapted from a larger set of stimuli (Atkinson et al., 2004, 2007). The FLDs consisted of 3-s digital movies depicting a gray-scale actor moving against a black background. To ensure that the BM was the only source of social signals available to the observer, the actors' face was covered. The PLDs lasted 2 s and consisted of 13 lighting dots placed over the main joints of the actor, moving against a black background, and were created by converting the FLDs stimuli to PLDs (Atkinson et al., 2012). Examples of the stimuli can be viewed at https://atkinsonap.github.io/stimuli/. Each display type included 10 different bodily expressions of each emotion, which vary in intensity and type of movements. This variability made the stimuli representative of a broad range of bodily emotions, but has the potential problem of a different ease of recognition for the stimuli with different intensity, given that a direct correlation between emotions' intensity and recognizability has been demonstrated in a similar set of stimuli (Dittrich et al., 1996; Atkinson et al., 2004). Thus, in a separate pilot study, we asked to 21 TD children between 5 and 11 years old (10 females and 11 males; M_{age} = 9.29; SD_{age} = 1.45) to rate the intensity of the expressed emotion by using a 9-point Likert scale. Stimuli were presented according to the intensity index obtained from this pilot study (see section "Procedure"). Specifically, the order of the video was based on the recognition ease and ranged from the easiest to the most difficult video, according to the rate established in the pilot study.

Participants

A total of 27 TD children, 25 children with ASD without ID (HFA) and 17 children with ASD and ID (LFA) participated in this study, for a total of 69 children (**Table 1**). Children with ASD were recruited at the Laboratory of Observation, Diagnosis, and Education (ODFLab) – University of Trento, the Autism Parents Association in Trento (AGSAT), and the Istituto Dosso Verde in Milan. All the children with ASD met the established criteria for ASD as specified in DSM-IV (American Psychiatric Association, 2000) or DSM-5 (American Psychiatric Association, 2013). The diagnosis of ASD was made by experienced clinicians, on the basis also of the administration of the Autism Diagnostic Observation Schedule (Lord et al., 2000), or based on the Autism Diagnostic Interview (ADI/ADI-R) (Lord et al., 1994). The non-verbal IQ score was measured by Raven's progressive matrices (Raven, 1941), the Colored progressive matrices (John and Raven, 2003), or the subscales of Leiter-R brief IQ (Roid and Miller, 1997) according to the participants' age and cognitive difficulties. TD children were recruited from the general population via the personal network of the researchers and master students involved. Specific data on socioeconomic status were not recorded. The study was approved by the Ethical Committee of the University of Milano-Bicocca. Before the experiment, the participants and their parents received a detailed explanation of the procedure and provided a signed informed consent, in accordance with the Declaration of Helsinki.

Procedure

The participants were tested individually in a quiet room. They sat in front of a PC, at a distance of 60 centimeters from the computer monitor. The experimenter sat next to the participant for the entire duration of the study. To overcome the verbal impairment and the difficulties in focusing and in maintaining the attention often described in children with ASD (Goldstein et al., 2001; Happé et al., 2006), we asked the participants to perform a non-verbal recognition task. Moreover, to limit the task attentional and working memory load, the number and the duration of the experimental blocks were minimized, as well as the response options. Thanks to this simplified procedure, the children with LFA and language difficulties were also able to perform the task. The task consisted of a two-forced choice categorization of the emotional content of BMs, made by pressing a designated key on the computer keyboard. Sticky emoticons reproducing the facial expression corresponding to the type of emotions presented through BM was stuck on the response keys (**Figure 1**). At the beginning of each block, the participants performed a brief practice session. Each movie was preceded by a fixation cross, only when the child paid attention to the monitor the experimenter started the video. Each video was presented once. At the end of each video, the video disappeared and the question "Which emotion was expressed?" was presented until the participant made the response (**Figure 1**). The participant was asked to categorize the observed video by pressing the key with the corresponding facial emoticon, as accurately and fast as possible. Accuracy and response times (RTs) were recorded. If the participant did not recognize the emotion expressed in the video and was not able to choose between the two response options, the experimenter pressed a defined key (namely, the key corresponding to the letter "g" in the block "Fear-Happiness"; "a" in the block "Neutral-Happiness"; and "l" in the block "Fear-Neutral") and a new video was presented. This key was different between the blocks to prevent the children from learning which was the key that would have allowed them to skip the response. Stimulus presentation was controlled and behavioral responses were recorded with E-Prime 2.0 software® (Psychology Software Tools, Inc). The task was divided into three blocks, each block included videos of only two types of emotion and only two response options were possible (Fear-Happiness; Fear-Neutral; Happiness-Neutral). The order of the blocks was counterbalanced among the participants. Each block included 10

TABLE 1 | Descriptive statistics: numerosity (N), female:male ratio (F:M), means and standard deviations (SD) of the chronological age, mental age, and IQ are reported in the three groups of children (HFA, LFA, and TD).

Group of functioning	**N**	**Chronological age (years)**	**Mental age (years)**	**IQ**
	(F:M)	***Mean (SD)***	***Mean (SD)***	***Mean (SD)***
ASD	42 (3:39)			
Without ID (HFA)	25 (1:24)	9.88 (2.96)	9.77 (3.26)	100.16 (19.48)
With ID (LFA)	17 (2:15)	12.33 (2.19)	5.429 (1.76)	44.87 (14.87)
TD	27 (13:14)	8.81 (1.84)	9.69 (2.33)	110.00 (14.74)

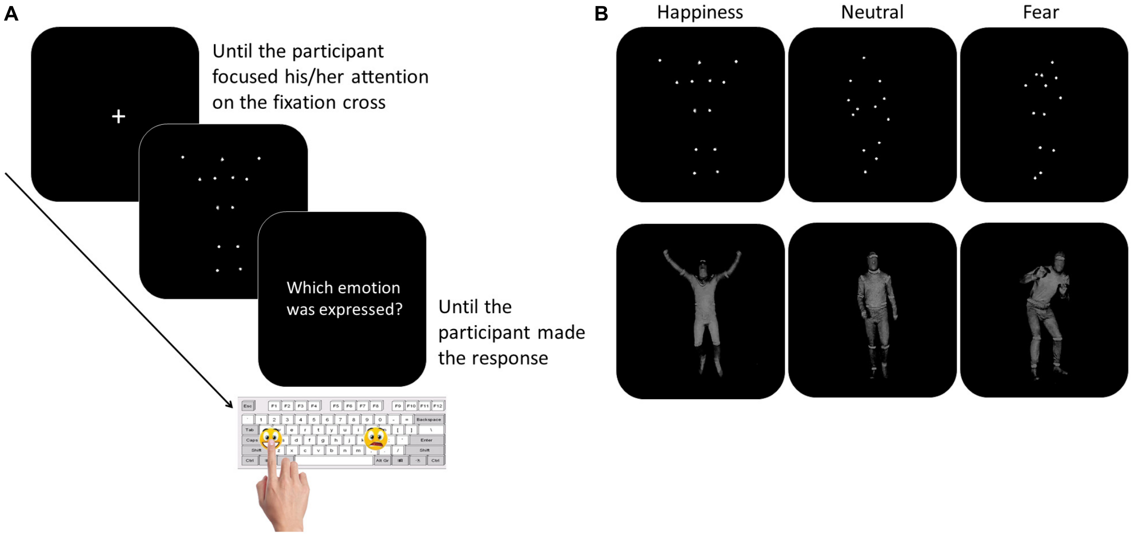


FIGURE 1 | Participants performed a two forced-choice non-verbal task, during which they were asked to recognize the emotion content of whole- BMs by pressing the corresponding emoticon on the keyboard (Panel **A**). The stimuli included happy, neutral, or fearful body movements depicted as PLDs (Panel **B**, top row) and FLDs (Panel **B**, bottom row).

PLDs and 10 FLDs for each type of emotion, for a total of 40 videos per block (20 FLDs and 20 PLDs). Within each block, PLDs and FLDs were presented separately, half of the participants saw PLDs first and then FLDs, the other half vice versa. Within each block, for both PLDs and FLDs and for each emotion type, the order of the video was based on the recognition ease and ranged from the easiest to the most difficult video, according to the rate established in the pilot study (see section "Stimuli"). By adopting this pseudo-randomized order of presentation, we aimed at meeting the ASDs' attentional limits, thus optimizing their performance. In fact, a complete randomized order of stimulus presentation might have resulted in two – or even more – difficult videos to be consecutively presented within the first trials of the block, inducing participants with ASD to prematurely disengage their attention from the task.

Statistical Analysis

All the analyses were performed with the software R, package 3.3.1 (R Core Team, 2016). Wilcoxon-rank-sum test with continuity correction showed that TD children were matched with the group of children with HFA for chronological age ($W = 297.5$, $p = 0.246$), mental age ($W = 389.5$, $p = 0.671$), and IQ ($W = 471$, $p = 0.067$). The group of children with LFA was chronologically older but mentally younger than the group of TD children (chronological age: $W = 367$, $p < 0.001$; mental age: $W = 28$, $p < 0.001$) and children with HFA (chronological age: $W = 117$, $p = 0.025$; mental age: $W = 365$, $p < 0.001$). The IQ level in children with LFA was lower than the IQ in children with HFA ($W = 405$, $p < 0.001$) and TD children ($W = 0$, $p < 0.001$) (**Table 1**). The study sample size was empirically set at ∼25 subjects per group, namely TD, LFA, and HFA, for a total of 75 subjects. In particular, this numerosity was determined taking into account the sample size of previous studies in the field, the complexity of the task that was presented, and its feasibility for individuals with LFA. Despite being influenced by group composition, recruitment areas, and available means, the sample size selected was consistent with previous literature on the subject (Moore et al., 1997; Atkinson et al., 2004; Hubert et al., 2007; Parron et al., 2008; Atkinson, 2009; Annaz et al., 2010, 2012; Fridenson-hayo et al., 2016). Eventually, we successfully included in the study 69 subjects (27 TD; 25 HFA; 17 LFA) that completed the task. We performed a *post hoc* power analysis to check whether this sample size was acceptable for the 3 × 2 × 3 mixed ANCOVA model to detect an effect. We ran a *post hoc* *F*-test power analysis (R package *pwr)* with medium effect size of 0.15 and sample size equal to 69 (see estimated effect size from

Federici et al., 2019). We obtained that the power of detecting such an effect at 0.05 level is equal to 0.90.

As a measure of *accuracy*, the percentages of the videos correctly categorized were calculated for each participant for each emotion type both for FLDs and PLDs. Since the mean percentages of correct responses was not normally distributed, the arcsine square root transformation of the proportions of correct responses was performed and all the following analyses were made on the transformed data.

Response times (RTs) were defined as the total amount of time (in milliseconds) from the offset of the video to the participant's response. Only the RTs relative to correct responses were included in the analyses. For each participant, the RTs considered as outliers according to the Tukey's method were discharged (Tukey, 1977; Hoaglin et al., 1986.; Rousseeuw and Leroy, 1987; Ratcliff, 1993; Hoaglin, 2003; Hodge et al., 2004), for a total of 22.82% of the trials (specifically 12.26% in the TD group, 25.91% in the HFA and 36.76% in the LFA). Since the mean RTs of correct responses was not normally distributed, the natural logarithms transformation of the averaged RTs was performed and all the following analyses were made on the transformed data.

Two participants with LFA were excluded from the final analysis because they did not terminate the task. Since the female:male ratio was higher in TD group than in the ASD groups, differences between TD females and males were investigated using *t*-tests. No differences were found both in Accuracy ($t = -0.033$, df = 24.79, *p*-value = 0.974) and in RTs ($t = 1.378$, df = 24.07, *p*-value = 0.181), therefore the variable "sex" was not considered in the further analysis. Analyses of both Accuracy and RTs were performed using 3 × 2 × 3 mixed ANCOVAs with Group as the between-factor (TD, HFA, and LFA); Display (FLDs, PLDs) and Emotion (Fear, Happiness, Neutral) as within-factors; non-verbal IQ and Age (chronological age) as covariates. The generalized eta squared values (η^2_G) (Olejnik and Algina, 2003) were also reported as an additional metric of effect size for all significant or marginally significant effects or interaction. Homoscedasticity of variances was tested with Levene-Test. Standard residuals were examined with Shapiro-Wilk normality test and by means of a quantile-quantile plot (QQ-plots) to check that data were normally distributed before parametric statistics were applied. *Post hoc* comparisons were performed using pairwise *t*-tests and the significance of alpha level was adjusted with Bonferroni's correction (adjusted p-values are reported). Finally, to simultaneously investigate the trade-off between the speed and the accuracy of a response and better understand how the strategies vary among functioning groups we conducted a speed-accuracy trade-off (SATO) analysis.

RESULTS

Analysis of Accuracy

Overall Analysis

Results of the mixed ANCOVA showed that overall the accuracy significantly increased with Age and IQ [$F_{(1,58)} = 12.576$, $p < 0.001$, $\eta^2\text{G} = 0.091$; $F_{(1,58)} = 12.077$, $p < 0.001$, $\eta^2\text{G} = 0.088$, respectively]. There was a significant effect of Group [$F_{(2,58)} = 21.051$; $p < 0.001$; $\eta^2\text{G} = 0.251$], with TD children outperforming both the HFA ($p < 0.001$) and the LFA ones ($p < 0.001$), and the children with HFA outperforming the LFA ones ($p < 0.001$). Furthermore, a significant effect of Display [$F_{(1,62)} = 11.211$; $p = 0.001$; $\eta^2\text{G} = 0.013$], with higher accuracy for FLDs compared to PLD; and a significant effect of Emotion [$F_{(2,124)=}6.228$; $p = 0.003$; $\eta^2\text{G} = 0.036$], with lower accuracy for happy stimuli than fearful ($p = 0.001$) and neutral ($p = 0.001$) stimuli, were found. The interaction between Display and Emotion was also significant [$F_{(2,128)} = 3.216$; $p = 0.043$; $\eta^2\text{G} = 0.005$], with higher accuracy in FLDs than in PLDs only for happy stimuli ($p = 0.004$). Interestingly, also the interaction between Group and Emotion closely approached significance [$F_{(4,124)} = 2.36$; $p = 0.057$; $\eta^2_G = 0.027$]. TD were more accurate than ASDs not only in recognizing emotional (happy and fearful) but also neutral stimuli, while differences between the two groups with ASD were found only in neutral stimuli (**Table 2** and **Figure 2A**).

Analysis Within Single Groups

The high variability of accuracy in the ASD groups' (see **Figure 2** and **Table 3** columns SD and SE) and the reduced sample size, especially in the LFA group, could have masked some effect, prevented some interactions to reach significance, and reduced the effect size. Therefore, to better understand the effect of Emotion and Display in the three groups of children, we performed three separated ANCOVAs with Emotion and Display as within-factors and IQ and Age as covariates (**Figures 2B,C**). The results showed a significant effect of Display only in TD children, with FLDs recognized more accurately than PLDs [$F_{(1,24)} = 7.23$, $p = 0.013$, $\eta^2\text{G} = 0.02$]. The effect of Emotion resulted significant in TD children [$F_{(2,48)} = 5.78$, $p = 0.006$, $\eta^2\text{G} = 0.098$], with happy stimuli recognized less accurately than fearful ($p = 0.003$) and neutral ($p = 0.001$) ones, while no difference between fearful and neutral stimuli emerged. The effect of Emotion tended toward significance also in children with ASD (HFA: $F_{(2,44)} = 2.97$, $p = 0.062$, $\eta^2\text{G} = 0.05$; LFA: $F_{(2,24)} = 3.14$; $p = 0.061$, $\eta^2\text{G} = 0.04$), therefore exploratory *post hoc* comparisons were performed also in these groups. Similarly to the TD group, in children with HFA the happy stimuli were recognized less accurately than neutral ($p = 0.012$) and fearful ones ($p = 0.048$), while fearful did not differ from neutral stimuli. Differently, in children with LFA the neutral stimuli were recognized with lower accuracy than fearful ($p = 0.018$) and marginally than happy ones ($p = 0.056$), while happy did not differ from fearful stimuli. The interaction between

TABLE 2 | The table shows the results of Bonferroni adjusted pairwise *t*-test in the three groups of children (HFA, LFA, and TD) in every emotional category.

	Post hoc comparison Emotion*Group		
	Happiness	**Fear**	**Neutral**
TD vs. HFA	$p = 0.014$	$p = 0.044$	$p = 0.019$
TD vs. LFA	$p = 0.003$	$p < 0.001$	$p < 0.001$
HFA vs. LFA	n.s	n.s.	$p < 0.001$

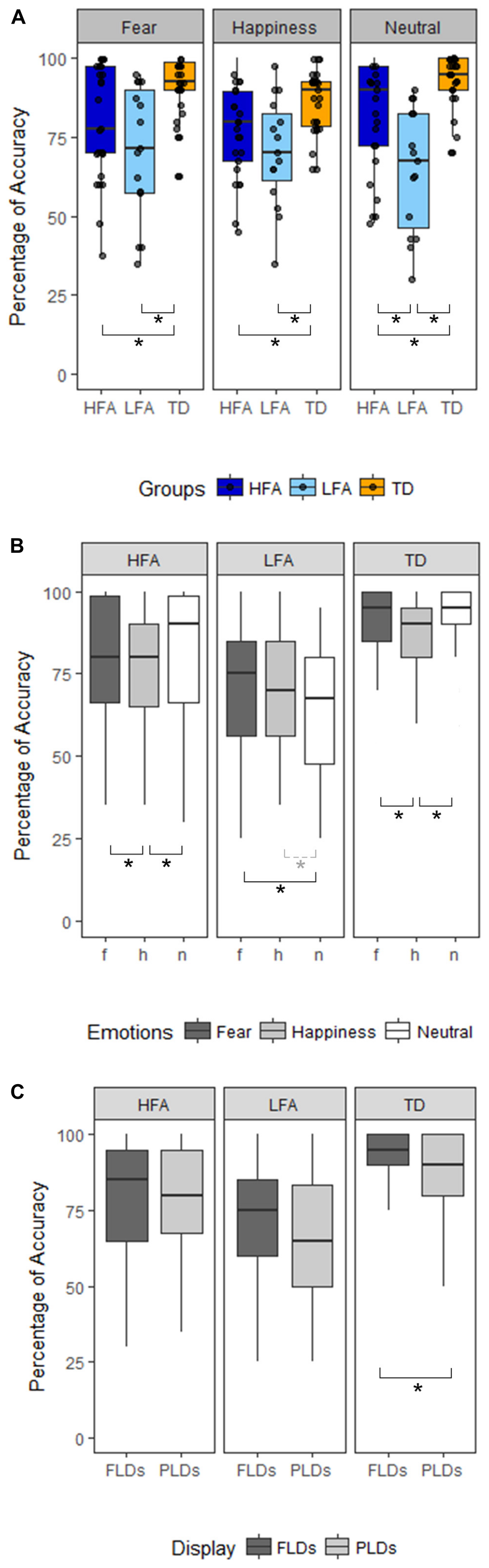


FIGURE 2 | Continued

FIGURE 2 | Panel **(A)** shows the differences between the three groups of children (HFA: ASD without ID, LFA: ASD with ID; TD: typical developing) in the three emotional categories (Happiness, Fear, and Neutral). The black dots represent the percentage of Accuracy of every participant in the relative emotional category. Panel **(B)** shows within-group differences between the emotional categories (dark-gray for Fear, light-gray for Happiness, white for Neutral stimuli). Panel **(C)** shows within-group differences between the Display type (FLDs in dark-gray, PLDs in light-gray). In all the Panels, the boxes represent the interquartile range, the black horizontal lines represent the median, the black vertical bars represent the standard error. The black stars represent the significant differences ($p < 0.05$), the dashed line and gray star represent the marginally significant differences ($0.05 < p < 0.01$).

Display and Emotion was not significant in any of the group. Finally, the effects of Age [$F_{(1,22)} = 8.66$, $p = 0.007$, $\eta^2G = 0.156$] and IQ [$F_{(1,22)} = 7.18$, $p = 0.014$, $\eta^2G = 0.133$] resulted significant only in children with HFA.

Analysis of Response Times (RTs)

The results of the mixed ANCOVA showed that overall the RTs significantly decreased with Age and IQ [$F_{(1,58)} = 4.83$, $p = 0.032$, $\eta^2{}_G = 0.07$; $F_{(1,58)} = 8.803$, $p = 0.004$, $\eta^2{}_G = 0.11$, respectively]. The effect of Emotion was significant [$F_{(2,124)} = 6.954$; $p = 0.001$; $\eta^2{}_G = 0.008$], with fearful stimuli recognized faster than happy ($p < 0.001$) and neutral ones ($p = 0.003$). The effect of Display closely approached significance [$F_{(1,62)} = 3.806$; $p = 0.056$; $\eta^2{}_G = 0.003$], with faster RTs for FLDs than for PLDs. The effect of Group [$F_{(2,58)} = 1.585$; $p = 0.214$; $\eta^2{}_G = 0.043$] and all the interactions were not significant (all $ps > 0.05$). Coherently with the analysis in Accuracy, we performed also three within-group ANCOVAs separately in the three groups, with Emotion and Display as within-factors and IQ and Age as covariates. Results showed a main effect of Emotion in TD children [$F_{(2,48)} = 4.01$, $p = 0.024$; $\eta^2{}_G = 0.03$], with fearful stimuli recognized faster than happy ones ($p = 0.002$). The effect of Emotion was significant also in children with LFA [$F_{(2,24)} = 3.78$; $p = 0.037$; $\eta^2{}_G = 0.003$], with fearful stimuli recognized faster than happy ($p = 0.028$) and neutral ones ($p = 0.033$). The effect of Display and the interaction between Display and Emotion were not significant in any of the group. Finally, the effects of Age and IQ were significant only in children with LFA [$F_{(1,12)} = 6.64$, $p = 0.024$, $\eta^2{}_G = 0.264$; $F_{(1,12)} = 18.17$, $p = 0.001$, $\eta^2{}_G = 0.495$, respectively].

Speed-Accuracy Trade-Off (SATO)

To investigate the SATO, we plotted the RTs in ms against the percentage of Accuracy for the three groups in recognizing the emotions with different Displays (**Figure 3**). Furthermore, we performed an ANOVA with logRT as dependent variable, Group as between-group factor (TD, HFA, and LFA), and Type of response (correct vs. incorrect) as within-group factor. The results showed a significant main effect of Group [$F(2,62) = 5.80$, $p = 0.005$ $\eta^2 = 0.115$], with a greater logRT in HFA and LFA compared to TD [$p = 0.001$ and $p = 0.005$, respectively], but no differences between HFA and LFA. The main effect of Type of response was also significant [$F(1,61) = 72.133$, $p < 0.001$, $\eta^2 = 0.134$], with greater logRT for incorrect vs. correct responses. Notably, the interaction between Group and Type of response

TABLE 3 | The table shows the mean, standard deviations (SD), and standard errors (SE) of Accuracy and RTs in each group (HFA, LFA, and TD) averaged by emotions (Fear, Happiness, and Neutral) and display types (FLDs and PLDs).

Group	**Emotion**	**Display**	**Accuracy (%)**			**RTs (msec)**		
			Mean	***SD***	***SE***	***Mean***	***SD***	***SE***
HFA	Fear	*FLDs*	81.18	19.31	4.55	1867.44	1438.07	338.96
	Fear	*PLDs*	77.80	21.02	4.95	1846.25	1474.54	347.55
	Happiness	*FLDs*	78.16	15.16	3.57	1956.79	1889.27	445.31
	Happiness	*PLDs*	75.20	16.49	3.89	2018.46	1655.59	390.23
	Neutral	*FLDs*	81.00	20.77	4.89	2058.36	2080.01	490.26
	Neutral	*PLDs*	83.20	16.70	3.94	1996.76	1508.31	355.51
LFA	Fear	*FLDs*	72.33	21.78	5.13	1709.65	1258.32	296.59
	Fear	*PLDs*	68.89	23.31	5.50	1920.32	1198.63	282.52
	Happiness	*FLDs*	75.00	20.35	4.80	1796.61	1144.20	269.69
	Happiness	*PLDs*	66.00	16.60	3.91	2018.21	1171.10	276.03
	Neutral	*FLDs*	66.00	20.63	4.86	1906.29	1183.49	278.95
	Neutral	*PLDs*	63.33	20.15	4.75	2028.01	1062.12	250.34
TD	Fear	*FLDs*	93.70	9.86	2.32	1199.03	315.38	74.34
	Fear	*PLDs*	89.26	10.80	2.55	1206.61	323.09	76.15
	Happiness	*FLDs*	89.26	8.63	2.03	1287.90	243.78	57.46
	Happiness	*PLDs*	83.89	13.33	3.14	1269.94	295.19	69.58
	Neutral	*FLDs*	94.07	7.08	1.67	1266.62	227.39	53.60
	Neutral	*PLDs*	92.78	10.32	2.43	1230.26	281.97	66.46

was not significant, showing that in all the tree groups the incorrect answers corresponded to greater RTs. These results suggest that the greater number of errors in children with ASD are not due to a tendency to respond faster compared to TD ones.

DISCUSSION

We initially predicted that if the impairment in recognizing the emotional BM was due to a specific difficulty in emotion processing, the children with ASD should recognize the *emotional* stimuli with poorer accuracy and higher RTs, while they should not differ from TD children in recognizing the *neutral* stimuli. Our results showed that children with ASD, independently of their IQ level, were less accurate than TD children in recognizing both the emotional and the neutral BM, either when displayed as FLDs or PLDs. When the emotional content was correctly recognized, the RTs did not differ between the three groups. Additional analysis showed that the vision of the body form improved the recognition of BM only in TD children, while in children with ASD the accuracy did not differ between the two display types. According to our initial hypothesis, these results suggest that the impairment in recognizing the emotional content of BM in ASD is associated to the processing of BM information rather than being specific for the processing of the emotional content or the visual characteristics of the stimuli.

Differences in Recognizing Emotional and Neutral BM

Previous studies about recognition of emotional BM in ASD found a specific difficulty in naming the emotional but not the neutral BM either in children and adults (Moore et al., 1997; Hubert et al., 2007; Parron et al., 2008). The discrepancies between our and existing results may be explained by methodological differences. For instance, in previous studies the participants were asked to verbally describe the observed PLDs using emotional words (Moore et al., 1997; Hubert et al., 2007; Parron et al., 2008). However, labeling the emotions may present more than a challenge to participants with ASD (Lartseva et al., 2014). Hence, it is possible that some language or memory difficulties (e.g., in knowing/recall/select the specific emotion label) have constrained the performance of individuals with ASD on the verbal emotion recognition tasks. For this reason, in the present study we tried to overcome the limitations of spontaneous verbal labeling of emotions by asking participants to perform a non-verbal two-forced-choice task. The forced-choice task reduces the cognitive load and helps to obviate the potential difficulties related to the recall/selection of the emotional label, which may have led to poorer performance in previous studies. This simplified non-verbal procedure, together with the non-verbal IQ as matching criteria, was crucial to allow also the participants with LFA to perform the emotion recognition task and to highlight differences and commonalities between children with ASD with and without ID.

Using the forced-choice task, we found that TD children were more accurate than children with ASD in recognizing both the emotional and the neutral stimuli. This shows that the attribution of an affective significance to BM is difficult for children with ASD, even when no verbal label is required, either when they have to recognize the presence (fearful and happy BM) or the absence (neutral BM) of emotions. Our results parallel the findings of two other studies that adopted a forced-choice task to investigate the emotional BM recognition in adults with ASD (Atkinson, 2009; Nackaerts et al., 2012) and supports the hypothesis that

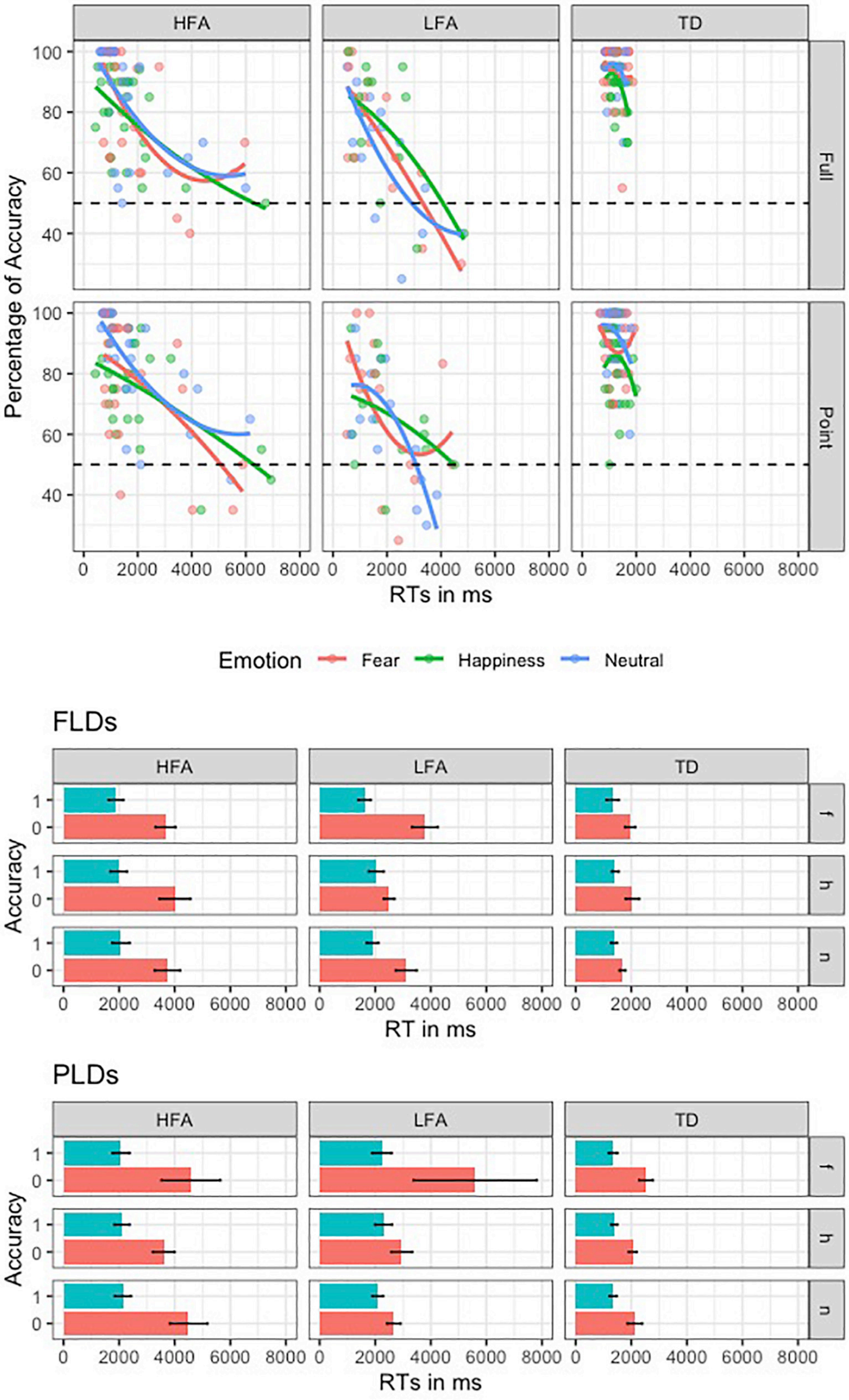


FIGURE 3 | This figure shows the speed-accuracy trade-off in the three groups of children. In the first plot, the second order polynomial distribution of the RTs against Accuracy rate is represented in the three groups (columns) for the two display types (rows); lines and dots are colored according to the three emotions. The second and the third plots represent the mean RT of correct (blue) and incorrect (red) responses, relative to the three emotions (rows, "f" is for fear, "h" is for happiness, and "n" is for neutral) in the three groups (columns).

the difficulty in recognizing the affective meaning of BM is related to the processing of body motion – rather than emotion – information. This is not surprising, as in bodily expressions the emotional content is conveyed by the BMs, thus the correct recognition of the bodily-expressed emotion requires the correct encoding of the observed motion information (Chouchourelou et al., 2006; Roether et al., 2009). Nevertheless, this is consistent with the idea that abnormalities in processing high-order motor information can affect the social functioning in ASD and lead to difficulties in social domains (Oberman and Ramachandran, 2007; Rizzolatti and Fabbri-Destro, 2010; Casartelli et al., 2016). Evidence from neuroimaging studies seems to corroborate this hypothesis. At the brain level, the vision of bodily expressions activates a network of areas responsible for action understanding that comprises the superior temporal sulcus (pSTS) and the putative mirror-neuron system (MNS), together with visual, limbic and subcortical structures (Heberlein et al., 2004; Grèzes et al., 2007; de Gelder et al., 2010; Van den Stock et al., 2011). In particular, the MNS contains high-level motor representations that are active both during the execution and the observation of an action (for an extensive review, see Rizzolatti et al., 2014). Interestingly, transcranial magnetic stimulation (TMS) studies provided evidence that the perturbation of the BM-related areas alters the recognition of bodily expressions (Candidi et al., 2011; Engelen et al., 2015; Mazzoni et al., 2017), proving that these areas are causally involved in recognizing the affective meaning of BM. Consistently, fMRI findings showed abnormal activation of the neural network responsible for action understanding in individuals with ASD when presented with emotional BM (Freitag et al., 2008; Grèzes et al., 2009; McKay et al., 2012). Furthermore, results of an electromyographic study suggest that deficits related to the parietal node of MNS may underlie the difficulties in action comprehension in children with ASD (Cattaneo et al., 2007). This body of evidence suggests that structural, functional, and connectivity abnormalities in the action-related areas in ASD may prevent the correct encoding of the observed BM. In turn, this may produce difficulties in recognizing the observed BM, both emotional and neutral, that is indeed what we found in the present study.

Differences Between Emotions

We found that the three emotional types were recognized with different accuracy and different RTs in the three groups: in children with HFA, the effect of Emotions was significant in accuracy, with TD-like poorer accuracy in happy stimuli compared to fearful and neutral ones, but it was not significant in RTs; in children with LFA, the recognition of the neutral BM was accomplished with the poorest accuracy, beside they showed TD-like faster RTs for fearful stimuli. Notwithstanding the nature of the additional within-group analysis was exploratory and should be interpreted with caution, these results highlight the importance of including both the children with HFA and LFA in research aimed at investigating the recognition of emotional BM in ASD and to consider them as separate groups.

Notably, in all the three groups there was an advantage for the fearful stimuli. This is in line with other behavioral studies on recognition of emotional BM (Atkinson et al., 2004; Bannerman et al., 2009; Philip et al., 2010). When a body expression is perceived, the priority in our brain is to create a representation of the observed emotional movements in order to react promptly and adaptively (Grèzes et al., 2007). In particular, fear communicates the presence of potential threats in the environment and it is associated with increased vigilance and attention (Phelps et al., 2006; Tamietto et al., 2007; Kret et al., 2013), enhanced visual processing (van Heijnsbergen et al., 2007; Borhani et al., 2015), and modulation of the motor system (Borgomaneri et al., 2012, 2015). Our results are in line with these studies and showed that fearful BM was recognized more accurately in TD children and in children with HFA, and its recognition was faster in TD children and in children with LFA, compared to the recognition of happy or neutral stimuli.

Another interesting result is that the children with LFA recognized with poorest accuracy the neutral stimuli. Our hypothesis it that this could be due to the fact that the recognition of emotional states is reinforced in social interaction, whilst the recognition of "neutral" state is not. Indeed, since very early in development, the emotions are highlighted and reinforced in natural social exchange. For instance, the parents categorize the emotions and the internal states of the child while she/he is experiencing them, using the emotional label corresponding to the facial or bodily expressions (e.g., "You play with the ball! Oh look at you, you are smiling, you are happy! You like it!"). This reinforces the coupling of emotional non-verbal stimuli (such as facial and bodily expressions) with the concurrent action and internal state, and serves the acquisition of the ability to recognize the emotional expressions. Differently, it is rare that this reinforcement is done with the neutral expressions, neither during daily social interaction, nor during training or intervention with children with ASD. Indeed, when the attention is focused on the neutral actions, it is more probable reinforced the type of action and its functional significance, rather than its non-emotional content. As a consequence, children with ASD can create internal representations of the emotional expressions, but not of the neutral ones. Nevertheless, it is possible that the children with HFA, thanks to the higher IQ, can learn to associate the neutral expressions with the absence of any emotions. However, if the IQ level is poor, as it is in individuals with LFA, this explicit strategies will be learned with difficulty or not learned at all. This might be the reason why the recognition of the neutral stimuli was particularly difficult in our group of children with LFA.

Differences Between PLDs and FLDs

Most of the previous studies have investigated the BM recognition in ASD using only PLDs. However, in the real life the body form is fully visible and the use of realistic stimuli should hence be considered – if not encouraged – in research aimed at explaining the origin of daily social difficulty in clinical populations. Indeed, to ascribe the social impairment in ASD to deficit in processing the BM information, the difficulty should be consistent independently of the visual representation of BM (e.g., pure kinematic or fully visible body form). The present work contributes to the small number of studies that have investigated the emotional BM recognition using both FLDs and PLDs. So far, this evidence in children with ASD has been missed. Our results showed a difficulty in children with ASD, compared to TD children, in recognizing the emotional BM when presented as PLDs (i.e., simplified stimuli). This is consistent with previous studies with PLDs (Moore et al., 1997; Hubert et al., 2007; Parron et al., 2008; Nackaerts et al., 2012) and confirm a difficulty in children with ASD in understanding the significance of BM represented as PLDs. In addition, we strikingly found that the accuracy in children with ASD was poorer than that of TD children also in recognizing the FLDs. According to our initial predictions, this result supports the hypothesis that the impaired recognition of BM in ASD is related to higher-order motion processing, which underlie the encoding of BM independently of its visual representation.

Furthermore, within group analysis showed that TD's accuracy was higher when the body FLDs, while in children with ASD no difference between PLDs and FLDs emerged. According to the models for BM processing (e.g., Giese and Poggio, 2003), the BM identification requires the integration of form and motion information, likely associated with the activation and involvement of the superior temporal sulcus (pSTS). Specifically, the authors proposed a hierarchical model for BM processing with two pathways, one for form and the other one for motion information. In the model, the form pathway goes from visual cortex to inferior temporal cortex, the fusiform cortex, and STS. The motion pathway goes from visual cortex to MT and subsequently to STS and fusiform cortex. The two pathways merge into STS that integrate information from both streams. From STS, the BM information are then projected to the frontal and parietal areas. The posterior part of STS is indeed part of the action observation network and provide the main visual input to the fronto-parietal regions of the putative mirror neuron system (namely, the inferior frontal gyrus and the inferior parietal lobule) (for an exernsive review, see Rizzolatti et al., 2014). Our results suggest that this integration of form and motion information occurs, in TD children and facilitates the recognition of the emotional content of BM, while it seems to be impaired in children with ASD. Consistently, neuroimaging findings in ASD have shown structural abnormalities in STS (McAlonan et al., 2005; Zilbovicius et al., 2006; Courchesne et al., 2007); atypical activation in STS during BM perception (Freitag et al., 2008; Kaiser et al., 2010; McKay et al., 2012); and significant reduced connectivity between STS and the fronto-parietal areas processing higher-order motion information (McKay et al., 2012; Alaerts et al., 2014). Hence, we hypothesize that these abnormalities in STS and in its connections with the motor areas of the MNS may result in a lack of integration of form and motion cues, which in turn may lead to the impaired comprehension of BM and, thus, may explain the different recognition of FLDs and PLDs that we found between the children with and without ASD.

Relation Between IQ, Age, and the Recognition of Body Movements

As predicted, results of the overall analysis showed that the accuracy and speed in recognizing the emotional BM improved with age and IQ. This finding matches another study on basic and complex emotion recognition (Fridenson-hayo et al., 2016). In this study, the FLDs were employed in 5–9 years old children with TD and HFA, a significant relation with this ability and the age was found.

In children with ASD, we found that the effect of IQ and Age were significant only in children with HFA for what concerns the accuracy, and only in children with LFA for what concerns the RTs. These results suggested that higher cognitive resources in children with HFA might subtend the acquisition of alternative strategies useful for BM recognition (Rutherford and Troje, 2012), which proficiency seems to improve with age. It is plausible, in fact, that when the mechanism for BM perception is altered, as it is in children with ASD (Kaiser et al., 2010; Kaiser and Pelphrey, 2012; Mcpartland et al., 2012), the recognition of emotional BMs it is likely mediated by alternative strategies (McKay et al., 2012). For instance, using a point-light direction discrimination paradigm, McKay et al. (2012) found a comparable behavioral performance, but different brain activation, in adults with and without ASD. This suggests that individuals with ASD may accomplish the same task using different brain regions compared to TD participants, yet reaching similar proficiency levels. In particular, in TD group a unitary network of areas, which included temporal and parietal regions, was found to be active. On the contrary, this activation was missing in the ASD group and two other distinct networks were activated, involving brain areas selective for motion and form encoding, respectively. In other words, in people with ASD there seems to be a lack of implicit and automatic process of emotional BM. To compensate this lack, individuals with high level of IQ may develop explicit strategies that serve the emotional BM recognition by recruiting cognitive processes that are specialized for the encoding of non-social stimuli. Presumably, these strategies are declarative and learned explicitly through experience and/or training, they are sophisticated and require considerable cognitive efforts. Therefore, they can be developed only in presence of high cognitive resources. Our results are consistent with these findings and, in addition, we found that this alternative and compensatory strategies improved with age through childhood and, potentially, would reach TD-like performance later in life. Future studies that investigate the differences in performance between individuals with HFA at different ages (e.g., childhood, adolescence, and adulthood) and age matched control groups would help better understanding this developmental trend.

With regard to the children with LFA, we found that the accuracy was not predicted neither by the age nor by the IQ, but these effects resulted significant in RTs analysis. This suggests that, even though the difficulties of children with LFA in understanding the emotional meaning of BM remain stable during the development, when they are successful in comprehending the BM they do it more rapidly with age, and the rapidity is increased by higher non-verbal IQ.

In TD children we found that this ability was not modulated by age and IQ, nor for the accuracy neither for the RTs. This result is partially contrary to our initial hypothesis and to previous findings in children with TD (Ross et al., 2012). However, this incongruence may result from the narrower range of age and the limited sample size of our study, which could have masked some effects. Indeed, whereas our study involved 27 TD children between the age of 5 and 11 years, the study of Ross and collaborators included a larger sample of 107 children aged between 4 and 17 years old. Another possible explanation could be ascribed to the type of task: our task had a reduced load of working memory with respect to that of Ross and collaborators, as it was aimed at testing children with LFA, therefore it could have been less sensible to detect age-related changes in TD.

A few limitations of the study should be noted: the reduced sample size and the high variability of responses, especially in children with LFA, could have prevented the interaction between group and displays and between group and emotion to reach significance in the overall analysis of accuracy and RTs. Whereas some effects clearly emerged from the additional within-group analysis, the nature of these analysis is exploratory and, therefore, the data should be interpreted with caution. For these reasons, further studies with larger sample size are needed before firm conclusions can be drawn. Despite these limitations, our findings contribute to the very small number of studies on emotional BM recognition that have involved people with LFA and could be the starting point for future research aimed to better understand this ability in subgroups of people with ASD with different characteristics. To better understand differences and commonalities between individuals LFA and HFA, future investigation of emotional BM recognition in ASD should also be extended to other emotional contents (e.g., including anger and sadness), comparisons between different sources of social and emotional stimuli (e.g., faces, bodies, and voices). Finally, different range of age and longitudinal studies are desirable to examine the developmental trajectories of BM recognition in ASD.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committee of the University of Milano-Bicocca. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

NM, PR, RA-G, and PV discussed the original idea of the project and designed the study. NM collected the data, performed the statistical analysis, and drafted the manuscript. IL performed the statistical analysis and edited the manuscript. PV, PR, and RA-G edited the manuscript. All authors revised the article.

ACKNOWLEDGMENTS

We are indebted to all participants and their families, and to ODFLab in Rovereto, AGSAT in Trento, and Istituto Dosso Verde in Milan, without whose collaboration this study would never have been realized. We are much obliged to Anthony P. Atkinson (Durham University) for sharing with us the FLDs and PLDs stimuli. We would like to thank Elena Ricca, Elisa Covella, Iuri Mastrandrea, Barbara Caola, Samuel Baruzzi, Elena Miorandi, and Lorenza Antonini for their help with the data collection. Finally, we are grateful to Michele Filosi for the statistical support. The content of the manuscript has been previously published as a Ph.D. Thesis (Mazzoni, 2017). Partial results of this study have been presented during the 14th European Congress of Psychology 2015. Implications of Research on the Neuroscience of Affect Attachment and Social Cognition Conference 2015 and IMFAR 2016.

REFERENCES

Actis-Grosso, R., Bossi, F., and Ricciardelli, P. (2015). Emotion recognition through static faces and moving bodies: a comparison between typically developed adults and individuals with high level of autistic traits. *Front. Psychol.* 6:1570. doi: 10.3389/fpsyg.2015.01570

Alaerts, K., Nackaerts, E., Meyns, P., Swinnen, S. P., and Wenderoth, N. (2011). Action and emotion recognition from point light displays: an investigation of gender differences. *PLoS One* 6:e020989. doi: 10.1371/journal.pone.0020989

Alaerts, K., Woolley, D. G., Steyaert, J., Di Martino, A., Swinnen, S. P., and Wenderoth, N. (2014). Underconnectivity of the superior temporal sulcus predicts emotion recognition deficits in autism. *Soc. Cogn. Affect. Neurosci.* 9, 1589–1600. doi: 10.1093/scan/nst156

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders Fourth Edition Text Revision (DSM-IV-TR)*. Washington, DC: American Psychiatric Association.

American Psychiatric Association (2013). *The Diagnostic and Statistical Manual of Mental Disorders: DSM 5*. Arlington, VA: APA.

Annaz, D., Campbell, R., Coleman, M., Milne, E., and Swettenham, J. (2012). Young children with autism spectrum disorder do not preferentially attend to biological motion. *J. Autism Dev. Disord.* 42, 401–408. doi: 10.1007/s10803-011-1256-3

Annaz, D., Remington, A., Milne, E., Coleman, M., Campbell, R., Thomas, M. S. C., et al. (2010). Development of motion processing in children with autism. *Dev. Sci.* 13, 826–838. doi: 10.1111/j.1467-7687.2009.00939.x

Atkinson, A. P. (2009). Impaired recognition of emotions from body movements is associated with elevated motion coherence thresholds in autism spectrum disorders. *Neuropsychologia* 47, 3023–3029. doi: 10.1016/j.neuropsychologia.2009.05.019

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., and Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception* 33, 717–746. doi: 10.1068/p5096

Atkinson, A. P., Tunstall, M. L., and Dittrich, W. H. (2007). Evidence for distinct contributions of form and motion information to the recognition of emotions from body gestures. *Cognition* 104, 59–72. doi: 10.1016/j.cognition.2006.05.005

Atkinson, A. P., Vuong, Q. C., and Smithson, H. E. (2012). Modulation of the face- and body-selective visual regions by the motion and emotion of point-light face and body stimuli. *Neuroimage* 59, 1700–1712. doi: 10.1016/j.neuroimage.2011.08.073

Aviezer, H., Trope, Y., and Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science* 338, 1225–1229. doi: 10.1126/science.1224313

Bannerman, R. L., Milders, M., de Gelder, B., and Sahraie, A. (2009). Orienting to threat: faster localization of fearful facial expressions and body postures revealed by saccadic eye movements. *Proc. Biol. Sci.* 276, 1635–1641. doi: 10.1098/rspb.2008.1744

Black, D. O., Wallace, G. L., Sokoloff, J. L., and Kenworthy, L. (2009). Brief report: IQ split predicts social symptoms and communication abilities in high-functioning children with autism spectrum disorders. *J. Autism Dev. Disord.* 39, 1613–1619. doi: 10.1007/s10803-009-0795-3

Blake, R., Turner, L. M., Smoski, M. J., Pozdol, S. L., and Stone, W. L. (2003). Visual recognition of biological motion is impaired in children with autism. *Psychol. Sci.* 14, 151–157. doi: 10.1111/1467-9280.01434

Borgomaneri, S., Gazzola, V., and Avenanti, A. (2012). Motor mapping of implied actions during perception of emotional body language. *Brain Stimulat.* 5, 70–76. doi: 10.1016/j.brs.2012.03.011

Borgomaneri, S., Vitale, F., and Avenanti, A. (2015). Early changes in corticospinal excitability when seeing fearful body expressions. *Sci. Rep.* 5:14122. doi: 10.1038/srep14122

Borhani, K., Làdavas, E., Maier, M. E., Avenanti, A., and Bertini, C. (2015). Emotional and movement-related body postures modulate visual processing. *Soc. Cogn. Affect. Neurosci.* 10, 1092–1101. doi: 10.1093/scan/nsu167

Candidi, M., Stienen, B. M. C., Aglioti, S. M., and de Gelder, B. (2011). Event-related repetitive transcranial magnetic stimulation of posterior superior temporal sulcus improves the detection of threatening postural changes in human bodies. *J. Neurosci.* 31, 17547–17554. doi: 10.1523/JNEUROSCI.0697-11.2011

Casartelli, L., Molteni, M., and Ronconi, L. (2016). So close yet so far: motor anomalies impacting on social functioning in autism spectrum disorder. *Neurosci. Biobehav. Rev.* 63, 98–105. doi: 10.1016/j.neubiorev.2016.02.001

Cattaneo, L., Fabbri-Destro, M., Boria, S., Pieraccini, C., Monti, A., Cossu, G., et al. (2007). Impairment of actions chains in autism and its possible role in intention understanding. *Proc. Natl. Acad. Sci. U.S.A.* 104, 17825–17830. doi: 10.1073/pnas.0706273104

Chouchourelou, A., Matsuka, T., Harber, K., and Shiffrar, M. (2006). The visual analysis of emotional actions. *Soc. Neurosci.* 1, 63–74. doi: 10.1080/17470910600630599

Clarke, T. J., Bradshaw, M. F., Field, D. T., Hampson, S. E., and Rose, D. (2005). The perception of emotion from body movement in point-light displays of interpersonal dialogue. *Perception* 34, 1171–1180. doi: 10.1068/p5203

Courchesne, E., Pierce, K., Schumann, C. M., Redcay, E., Buckwalter, J. A., Kennedy, D. P., et al. (2007). Mapping early brain development in autism. *Neuron* 56, 399–413. doi: 10.1016/j.neuron.2007.10.016

Dakin, S., and Frith, U. (2005). Vagaries of visual perception in autism. *Neuron* 48, 497–507. doi: 10.1016/j.neuron.2005.10.018

de Gelder, B., and de Borst, A. W. (2015). "Body perception," in *Brain Mapping: An Encyclopedic Reference*, ed. A. W. Toga (London: Elsevier Academic Press), 107–114. doi: 10.1016/B978-0-12-397025-1.00162-7

de Gelder, B., Van den Stock, J., Meeren, H. K. M., Sinke, C. B. A., Kret, M. E., and Tamietto, M. (2010). Standing up for the body. Recent progress in uncovering the networks involved in the perception of bodies and bodily expressions. *Neurosci. Biobehav. Rev.* 34, 513–527. doi: 10.1016/j.neubiorev.2009.10.008

Dittrich, W. H. (1993). Action categories and the perception of biological motion. *Perception* 22, 15–22. doi: 10.1068/p220015

Dittrich, W. H., Troscianko, T., Lea, S. E., and Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception* 25, 727–738. doi: 10.1068/P250727

Engelen, T., de Graaf, T. A., Sack, A. T., and de Gelder, B. (2015). A causal role for inferior parietal lobule in emotion body perception. *Cortex* 73, 195–202. doi: 10.1016/j.cortex.2015.08.013

Federici, A., Parma, V., Vicovaro, M., Radassao, L., Casartelli, L., and Ronconi, L. (2019). Anomalous perception of biological motion in autism: a conceptual review and meta-analysis. *bioRxiv* [Preprint]. Available online at: https://www.biorxiv.org/content/10.1101/530030v2

Freitag, C. M., Konrad, C., Häberlen, M., Kleser, C., von Gontard, A., Reith, W., et al. (2008). Perception of biological motion in autism spectrum disorders. *Neuropsychologia* 46, 1480–1494. doi: 10.1016/j.neuropsychologia.2007.12.025

Fridenson-hayo, S., Berggren, S., Lassalle, A., Tal, S., Pigat, D., Bölte, S., et al. (2016). Basic and complex emotion recognition in children with autism?: cross cultural findings. *Mol. Autism* 7:52. doi: 10.1186/s13229-016-0113-9

Giese, M. A., and Poggio, T. (2003). Cognitive neuroscience: neural mechanisms for the recognition of biological movements. *Nat. Rev. Neurosci.* 4, 179–192. doi: 10.1038/nrn1057

Goldstein, G., Johnson, C. R., and Minshew, N. J. (2001). Attentional processes in autism. *J. Autism Dev. Disord.* 31, 433–440. doi: 10.1023/A:1010620820786

Grèzes, J., Pichon, S., and de Gelder, B. (2007). Perceiving fear in dynamic body expressions. *Neuroimage* 35, 959–967. doi: 10.1016/j.neuroimage.2006.11.030

Grèzes, J., Wicker, B., Berthoz, S., and de Gelder, B. (2009). A failure to grasp the affective meaning of actions in autism spectrum disorder subjects. *Neuropsychologia* 47, 1816–1825. doi: 10.1016/J.NEUROPSYCHOLOGIA.2009.02.021

Happé, F., Booth, R., Charlton, R., and Hughes, C. (2006). Executive function deficits in autism spectrum disorders and attention-deficit/hyperactivity disorder: examining profiles across domains and ages. *Brain Cogn.* 61, 25–39. doi: 10.1016/j.bandc.2006.03.004

Happé, F., and Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *J. Autism Dev. Disord.* 36, 5–25. doi: 10.1007/s10803-005-0039-0

Happé, F. G. E. (1994). Wechsler IQ profile and theory of mind in autism: a research note. *J. Child Psychol. Psychiatr.* 35, 1461–1471. doi: 10.1111/j.1469-7610.1994.tb01287.x

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

Herrington, J. D., Baron-Cohen, S., Wheelwright, S. J., Singh, K. D., Bullmore, E. T., Brammer, M., et al. (2007). The role of MT+/V5 during biological motion perception in Asperger Syndrome: an fMRI study. *Res. Autism Spectr. Disord.* 1, 14–27. doi: 10.1016/j.rasd.2006.07.002

Hoaglin, D. C. (2003). John W. Tukey and data analysis. *Statist. Sci.* 18, 311–318.

Hoaglin, D. C., Iglewicz, B., and Tukey, J. W. (1986). Performance of some resistant rules for outlier labeling. *J. Am. Statist. Assoc.* 81:991. doi: 10.1080/01621459.1986.10478363

Hodge, V. J., Hodge, V. J., and Austin, J. (2004). A survey of outlier detection methodologies. *Artif. Intellig. Rev.* 22, 85–126. doi: 10.1023/b:aire.0000045502.10941.a9

Hubert, B., Wicker, B., Moore, D. G., Monfardini, E., Duverger, H., Da Fonseca, D., et al. (2007). Brief report: recognition of emotional and non-emotional biological motion in individuals with autistic spectrum disorders. *J. Autism Dev. Disord.* 37, 1386–1392. doi: 10.1007/s10803-006-0275-y

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Percept. Psychophys.* 14, 201–211. doi: 10.3758/BF03212378

John, and Raven, J. (2003). "Raven progressive matrices," in *Handbook of Nonverbal Assessment*, ed. R. S. McCallum (Boston, MA: Springer), 223–237. doi: 10.1007/978-1-4615-0153-4_11

Johnson, K. L., McKay, L. S., and Pollick, F. E. (2011). He throws like a girl (but only when he's sad): emotion affects sex-decoding of biological motion displays. *Cognition* 119, 265–280. doi: 10.1016/j.cognition.2011.01.016

Joseph, R. M., Tager-Flusberg, H., and Lord, C. (2002). Cognitive profiles and social-communicative functioning in children with autism spectrum disorder. *J. Child Psychol. Psychiatry Allied Discipl.* 43, 807–821. doi: 10.1111/1469-7610.00092

Kaiser, M. D., Hudac, C. M., Shultz, S., Lee, S. M., Cheung, C., Berken, A. M., et al. (2010). Neural signatures of autism. *Proc. Natl. Acad. Sci. U.S.A.* 107, 21223–21228. doi: 10.1073/pnas.1010412107

Kaiser, M. D., and Pelphrey, K. A. (2012). Disrupted action perception in autism: behavioral evidence, neuroendophenotypes, and diagnostic utility. *Dev. Cogn. Neurosci.* 2, 25–35. doi: 10.1016/j.dcn.2011.05.005

Klin, A., and Jones, W. (2008). Altered face scanning and impaired recognition of biological motion in a 15-month-old infant with autism. *Dev. Sci.* 11, 40–46. doi: 10.1111/j.1467-7687.2007.00608.x

Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G., and Jones, W. (2009). Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature* 459, 257–261. doi: 10.1038/nature07868

Kozlowski, L. T., and Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Percept. Psychophys.* 21, 575–580. doi: 10.3758/BF03198740

Kret, M. E., Roelofs, K., Stekelenburg, J. J., and de Gelder, B. (2013). Emotional signals from faces, bodies and scenes influence observers' face expressions, fixations and pupil-size. *Front. Hum. Neurosci.* 7:810. doi: 10.3389/fnhum.2013.00810

Lartseva, A., Dijkstra, T., and Buitelaar, J. K. (2014). Emotional language processing in autism spectrum disorders: a systematic review. *Front. Hum. Neurosci.* 8:991. doi: 10.3389/FNHUM.2014.00991

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., et al. (2000). The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J. Autism Dev. Disord.* 30, 205–223.

Lord, C., Rutter, M., and Le Couteur, A. (1994). Autism diagnostic interview-revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J. Autism Dev. Disord.* 24, 659–685. doi: 10.1007/bf02172145

Mazzoni, N. (2017). *The Recognition of Emotional Biological Movement in Individuals with Typical Development and Autism Spectrum Disorder (ASD)*. Ph. D thesis, University of Trento, Trento.

Mazzoni, N., Jacobs, C., Venuti, P., Silvanto, J., and Cattaneo, L. (2017). State-dependent TMS reveals representation of affective body movements in the anterior intraparietal cortex. *J. Neurosci.* 37, 7231–7239. doi: 10.1523/JNEUROSCI.0913-17.2017

McAlonan, G. M., Cheung, V., Cheung, C., Suckling, J., Lam, G. Y., Tai, K. S., et al. (2005). Mapping the brain in autism. A voxel-based MRI study of volumetric differences and intercorrelations in autism. *Brain* 128, 268–276. doi: 10.1093/brain/awh332

McKay, L. S., Simmons, D. R., McAleer, P., Marjoram, D., Piggot, J., and Pollick, F. E. (2012). Do distinct atypical cortical networks process biological motion information in adults with Autism Spectrum Disorders? *Neuroimage* 59, 1524–1533. doi: 10.1016/j.neuroimage.2011.08.033

Mcpartland, J. C., Coffman, M., and Pelphrey, K. A. (2012). Recent advances in understanding the neural bases of autism spectrum disorders. *Curr. Opin. Peadiatr.* 23, 628–632. doi: 10.1097/MOP.0b013e32834cb9c9.Recent

Moore, D. G., Hobson, R. P., and Lee, A. (1997). Components of person perception: an investigation with autistic, non-autistic retarded and typically developing children and adolescents. *Br. J. Dev. Psychol.* 15, 401–423. doi: 10.1111/j.2044-835X.1997.tb00738.x

Murphy, P., Brady, N., Fitzgerald, M., and Troje, N. F. (2009). No evidence for impaired perception of biological motion in adults with autistic spectrum disorders. *Neuropsychologia* 47, 3225–3235. doi: 10.1016/j.neuropsychologia.2009.07.026

Nackaerts, E., Wagemans, J., Helsen, W., Swinnen, S. P., Wenderoth, N., and Alaerts, K. (2012). Recognizing biological motion and emotions from point-light displays in autism spectrum disorders. *PLoS One* 7:e044473. doi: 10.1371/journal.pone.0044473

Oberman, L. M., and Ramachandran, V. S. (2007). The simulating social mind: the role of the mirror neuron system and simulation in the social and communicative deficits of autism spectrum disorders. *Psychol. Bull.* 133, 310–327. doi: 10.1037/0033-2909.133.2.310

Olejnik, S., and Algina, J. (2003). Generalized Eta and omega squared statistics: measures of effect size for some common research designs. *Psychol. Methods* 8, 434–447. doi: 10.1037/1082-989x.8.4.434

Parron, C., Da Fonseca, D., Santos, A., Moore, D. G., Monfardini, E., and Deruelle, C. (2008). Recognition of biological motion in children with autistic spectrum disorders. *Autism* 12, 261–274. doi: 10.1177/1362361307089520

Pavlova, M. A. (2012). Biological motion processing as a hallmark of social cognition. *Cereb. Cortex* 22, 981–995. doi: 10.1093/cercor/bhr156

Phelps, E. A., Ling, S., and Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychol. Sci.* 17:292. doi: 10.1111/J.1467-9280.2006.01701.X

Philip, R. C. M., Whalley, H. C., Stanfield, A. C., Sprengelmeyer, R., Santos, I. M., Young, A. W., et al. (2010). Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. *Psychol. Med.* 40, 1919–1929. doi: 10.1017/S0033291709992364

Pollick, F. E., Lestou, V., Ryu, J., and Cho, S. B. (2002). Estimating the efficiency of recognizing gender and affect from biological motion. *Vis. Res.* 42, 2345–2355. doi: 10.1016/S0042-6989(02)00196-7

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychol. Bull.* 114, 510–532. doi: 10.1037/0033-2909.114.3.510

Raven, J. C. (1941). Standardization of progressive matrices, 1938. *Br. J. Med. Psychol.* 19, 137–150. doi: 10.1111/j.2044-8341.1941.tb00316.x

Rizzolatti, G., Cattaneo, L., Fabbri-Destro, M., and Rozzi, S. (2014). Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. *Physiol. Rev.* 94, 655–706. doi: 10.1152/physrev.00009.2013

Rizzolatti, G., and Fabbri-Destro, M. (2010). Mirror neurons: from discovery to autism. *Exp. Brain Res.* 200, 223–237. doi: 10.1007/s00221-009-2002-3

Roché, L., Hernandez, N., Blanc, R., Bonnet-Brilhault, F., Centelles, L., Schmitz, C., et al. (2013). Discrimination between biological motion with and without social intention: a pilot study using visual scanning in healthy adults. *Intern. J. Psychophysiol.* 88, 47–54. doi: 10.1016/j.ijpsycho.2013.01.009

Roether, C. L., Omlor, L., Christensen, A., and Giese, M. A. (2009). Critical features for the perception of emotion from gait. *J. Vis.* 9, 15–32. doi: 10.1167/9.6.15

Roid, G. H., and Miller, L. J. (1997). *The Leiter International Performance Scale-revised Edition*. Lutz, FL: Psychological Assessment Resources.

Ross, P. D., Polson, L., and Grosbras, M. H. (2012). Developmental changes in emotion recognition from full-light and point-light displays of body movement. *PLoS One* 7:e044815. doi: 10.1371/journal.pone.0044815

Rousseeuw, P. J., and Leroy, A. (1987). *Robust Regression And Outlier Detection*. Hoboken, NJ: Wiley.

Rutherford, M. D., and Troje, N. F. (2012). IQ predicts biological motion perception in autism spectrum disorders. *J. Autism Dev. Disord.* 42, 557–565. doi: 10.1007/s10803-011-1267-0

Saygin, A. P., Cook, J., and Blakemore, S.-J. (2010). Unaffected perceptual thresholds for biological and non-biological form-from-motion perception in autism spectrum conditions. *PLoS One* 5:e13491. doi: 10.1371/journal.pone.0013491

Tamietto, M., Geminiani, G., Genero, R., and de Gelder, B. (2007). Seeing fearful body language overcomes attentional deficits in patients with neglect. *J. Cogn. Neurosci.* 19, 445–454. doi: 10.1162/jocn.2007.19.3.445

Troje, N. F., Westhoff, C., and Lavrov, M. (2005). Person identification from biological motion: effects of structural and kinematic cues. *Percept. Psychophys.* 67, 667–675. doi: 10.3758/BF03193523

Tukey, J. W. (1977). *Exploratory Data Analysis*. Boston, MA: Addison-Wesely.

Van den Stock, J., Righart, R., and de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion* 7, 487–494. doi: 10.1037/1528-3542.7.3.487

Van den Stock, J., Tamietto, M., Sorger, B., Pichon, S., Grézes, J., and de Gelder, B. (2011). Cortico-subcortical visual, somatosensory, and motor activations for perceiving dynamic whole-body emotional expressions with and without striate cortex (V1). *Proc. Natl. Acad. Sci. U.S.A.* 108, 16188–16193. doi: 10.1073/pnas.1107214108

van Heijnsbergen, C. C. R. J., Meeren, H. K. M., Grèzes, J., and de Gelder, B. (2007). Rapid detection of fear in body expressions, an ERP study. *Brain Res.* 1186, 233–241. doi: 10.1016/j.brainres.2007.09.093

Zilbovicius, M., Meresse, I., Chabane, N., Brunelle, F., Samson, Y., and Boddaert, N. (2006). Autism, the superior temporal sulcus and social perception. *Trends Neurosci.* 29, 359–366. doi: 10.1016/j.tins.2006.06.004

Conflict of Interest: 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.

frontiers in Psychology

ORIGINAL RESEARCH
published: 31 March 2020
doi: 10.3389/fpsyg.2020.00433

Explicit and Implicit Responses of Seeing Own vs. Others' Emotions: An Electromyographic Study on the Neurophysiological and Cognitive Basis of the Self-Mirroring Technique

Alessandra Vergallito[1†]***, Giulia Mattavelli***[2†]***, Emanuele Lo Gerfo***[3]***, Stefano Anzani***[4]***, Viola Rovagnati***[1]***, Maurizio Speciale***[5]***, Piergiuseppe Vinai***[5,6,7]***, Paolo Vinai***[5]***, Luisa Vinai***[8] ***and Leonor J. Romero Lauro***[1,6]*****

[1] Dipartimento di Psicologia, Università degli Studi di Milano-Bicocca, Milan, Italy, [2] NETS, Scuola Universitaria Superiore IUSS, Pavia, Italy, [3] Clinical Psychology Service of Mediterranean Institute for Transplantation and Advanced Specialized Therapies (IRCSS IsMeTT), Palermo, Italy, [4] Department of Neuroscience, Imaging and Clinical Sciences, "G. d'Annunzio" University of Chieti–Pescara, Chieti, Italy, [5] "GNOSIS" Research and Psychotherapy Group, Mondovì, Italy, [6] Studi Cognitivi, Cognitive Psychotherapy School and Research Center, Milan, Italy, [7] Psicologia Scientifica – Centro di Ricerca e Promozione Sociale, Milan, Italy, [8] Pôle de Psychiatrie et Psychothérapie du Centre Hospitalier du Valais Romand, Monthey, Switzerland

OPEN ACCESS

Edited by:
Antonino Vallesi,
University of Padova, Italy

Reviewed by:
Michel Thiebaut De Schotten,
Brain Connectivity and Behaviour Group (BCBG), France
Paola Sessa,
University of Padova, Italy

****Correspondence:***
Leonor J. Romero Lauro
leonor.romero1@unimib.it

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Consciousness Research, a section of the journal Frontiers in Psychology

Received: *11 October 2019*
Accepted: *25 February 2020*
Published: *31 March 2020*

Citation:
Vergallito A, Mattavelli G, Lo Gerfo E, Anzani S, Rovagnati V, Speciale M, Vinai P, Vinai P, Vinai L and Romero Lauro LJ (2020) Explicit and Implicit Responses of Seeing Own vs. Others' Emotions: An Electromyographic Study on the Neurophysiological and Cognitive Basis of the Self-Mirroring Technique. Front. Psychol. 11:433. doi: 10.3389/fpsyg.2020.00433

Facial mimicry is described by embodied cognition theories as a human mirror system-based neural mechanism underpinning emotion recognition. This could play a critical role in the Self-Mirroring Technique (SMT), a method used in psychotherapy to foster patients' emotion recognition by showing them a video of their own face recorded during an emotionally salient moment. However, dissociation in facial mimicry during the perception of own and others' emotions has not been investigated so far. In the present study, we measured electromyographic (EMG) activity from three facial muscles, namely, the zygomaticus major (ZM), the corrugator supercilii (CS), and the levator labii superioris (LLS) while participants were presented with video clips depicting their own face or other unknown faces expressing anger, happiness, sadness, disgust, fear, or a neutral emotion. The results showed that processing self vs. other expressions differently modulated emotion perception at the explicit and implicit muscular levels. Participants were significantly less accurate in recognizing their own vs. others' neutral expressions and rated fearful, disgusted, and neutral expressions as more arousing in the self condition than in the other condition. Even facial EMG evidenced different activations for self vs. other facial expressions. Increased activation of the ZM muscle was found in the self condition compared to the other condition for anger and disgust. Activation of the CS muscle was lower for self than for others' expressions during processing a happy, sad, fearful, or neutral emotion. Finally, the LLS muscle showed increased activation in the self condition compared to the other condition for sad and fearful expressions but increased activation in the other condition compared to the self condition for happy and neutral expressions. Taken together, our complex pattern of results suggests a dissociation at both the explicit and implicit levels in emotional processing of self vs. other emotions that, in the light of the Emotion in Context view, suggests that STM effectiveness is primarily due to a contextual–interpretative process that occurs *before* that facial mimicry takes place.

Keywords: facial mimicry, Self-Mirroring Technique, facial expression, EMG, emotion recognition

INTRODUCTION

Difficulty in accessing and recognizing emotions is a primary problem in many psychiatric and psychological diseases (Keltner and Kring, 1998; Tse and Bond, 2004), such as major depression (see Leppänen, 2006; Bourke et al., 2010 for a review), anxiety (Zeitlin and McNally, 1993; Silvia et al., 2006; Demenescu et al., 2010; Karukivi et al., 2010), and eating (Schmidt et al., 1993; Corcos et al., 2000; Speranza et al., 2007; Harrison et al., 2010) and personality (Domes et al., 2009; Loas, 2012; Loas et al., 2012) disorders. Indeed, alexithymia, which is a clinical condition characterized by difficulty in identifying and describing emotions, is present in over 50% of patients seeking psychological help (Sifneos, 1973; Kojima, 2012). This condition not only contributes to the emergence of symptoms (Güleç et al., 2013) but also influences the psychotherapeutic process (Ogrodniczuk et al., 2004; Leweke et al., 2009) and patients' compliance to treatments (Speranza et al., 2011).

Therefore, improving patients' ability to recognize and elaborate on their own emotions is a central goal of psychotherapy, regardless of the specific orientation. In the Cognitive Behavioral Therapy (CBT), for instance, achieving emotional awareness is a primary and fundamental step, since the core of the therapeutic process relies on instructing patients to monitor their feelings and thoughts in different situations, unveiling the connection between thoughts, emotions, and actions. Therefore, patients with poor introspective and self-reflective abilities might encounter great difficulty in the proper detection, description, and naming of their emotional experience. Trying to provide a solution for this issue, Vinai et al. (2015) created a video-based methodology called the Self-Mirroring Technique (SMT). SMT has been developed in the clinical setting and can be used as a coadjutant in different psychotherapeutic approaches. It is based on the audio-visual recording of the therapeutic session, with the aim of showing to the patients the emotions conveyed by their facial expressions. The procedure consists of asking patients to recall an emotionally significant event in their lives while their face is video recorded. Immediately after recall, the psychotherapist shows the patients their own video on the screen and again videos their face. The clinician then shows the patients the effects of seeing their own emotions (for more details on the clinical protocol, see Vinai et al., 2015). The observation of and listening to the video recordings are useful for patients to observe their own thoughts and emotions from an external position and has the effect of increasing their ability to recognize own emotions (Vinai et al., 2015), thus enhancing metacognition (Lorenzini et al., 2006).

Previous studies have reported the effectiveness of SMT in the psychotherapeutic setting (Vinai et al., 2016; Frau et al., 2018). However, the cognitive and neurophysiological mechanisms underlying this success are not completely clear. On the one hand, authors have suggested that observing a video depicting their own emotional expressions may help patients to recognize their own emotions by employing the innate system, which is typically used to understand others' emotions (Frau et al., 2018), thus improving the patients' poor ability in the self-reflective and introspective functions. Another possible—and not mutually exclusive—hypothesis suggests that giving patients the opportunity to view their face while listening to their words might allow them to add new—supplementary—information to their previous knowledge, thus facilitating the transition from an unidentified emotion to emotional awareness. Moreover, viewing themselves while feeling an emotion might elicit another emotion, which could be helpful not only for understanding their own emotional state but also for accepting or managing it. For instance, the observation of their own face expressing sadness might induce self-compassion (e.g., Petrocchi et al., 2017).

The STM foundation can be explored more thoroughly by referring to the phenomenon of *facial* (or emotional) *mimicry*, which has been defined as "*the imitation of emotional (facial) expressions of another person*" (Hess and Fischer, 2013; Hess and Fischer, 2017). In the literature, facial mimicry has been investigated by presenting participants with emotional stimuli and recording the activity of specific facial muscles, typically through electromyography (EMG; e.g., Dimberg, 1982; Larsen et al., 2003), and less frequently using the Facial Action Coding System[1] (Ekman and Friesen, 1978; Murata et al., 2016).

In EMG studies, participants are typically presented with static pictures (e.g., Dimberg and Thunberg, 1998; Scarpazza et al., 2018) and more rarely with dynamic stimuli (Sato et al., 2008; Rymarczyk et al., 2011) and face-to-face interactions (Künecke et al., 2017). Crucially, most studies focused on two emotions, namely, anger and happiness, and rarely included other discrete emotions (see Hess and Fischer, 2013, for a review). The most robust pattern emerging across the studies is the double dissociation between *corrugator supercilii* (CS) and *zygomaticus major* (ZM) muscle activity, which is dependent on the expression valence. Indeed, the presentation of angry—negative valence—faces increased activity in the CS, namely, the muscle used to approximate the eyebrows when frowning, while happy—positive valence—expressions induced higher EMG activity in the ZM, which is the muscle used when smiling, combined with decreased activity in the CS (Dimberg and Thunberg, 1998; De Wied et al., 2006). Other emotions have been less systematically investigated, and weaker link between emotions and muscular activation have been established (see Hess and Fischer, 2013).

---

[1] Given the purpose of the present experiment, we will focus on EMG studies.

For example, increased activity in the CS has also been linked to sadness (e.g., Weyers et al., 2009—neutrally primed group; Hess and Blairy, 2001) and fear (e.g., Van Der Schalk et al., 2011), while these patterns did not emerge in other studies (e.g., Lundqvist, 1995; Oberman et al., 2009). Increased activity in the levator labii superioris (LLS) has sometimes been reported in response to a disgust expression (Lundqvist and Dimberg, 1995), but not in a consistent way (Hess and Blairy, 2001).

Noteworthily, over the years, different proposals have addressed the interpretation of facial mimicry, generally focusing on different aspects of the phenomenon and—respective to our aim of deepening STM functioning—leading to different predictions.

According to the *embodied hypothesis*, viewing an emotional expression triggers activity in the same brain regions and peripheral efferent involved in the execution of similar expressions, thus eliciting—through a feedback process—the corresponding emotional state in the mimicker, a process known as *sensorimotor simulation* (see for a recent review Wood et al., 2016). Sensorimotor simulation can lead to a motor output—though overt mimicry is not a necessary component (Goldman and Sripada, 2005)—thus facilitating emotion recognition (Stel and Van Knippenberg, 2008; Neal and Chartrand, 2011; but see Hess and Fischer, 2017 for a critical review) and understanding (Niedenthal, 2007; Oberman and Ramachandran, 2007; Bastiaansen et al., 2009; Niedenthal et al., 2010; Gallese and Sinigaglia, 2011).

This interpretation is in line with the classical view, named by Hess and Fischer (2013) the *Matched Motor Hypothesis*. According to this view, facial mimicry is an automatic motor response and is independent of the intentions of both the expresser and the observer (Chartrand and Bargh, 1999; Preston and De Waal, 2002; for a review, see Hess and Fischer, 2013).

However, previous evidence suggested that mimicry can be influenced by different contextual cues, such as the type of emotion and the expresser's and observer's identity, relationship, and emotional state. These points are considered by an alternative account of facial mimicry known as the *Emotion Mimicry in Context view* (Hess and Fischer, 2013, 2017). Its authors suggested that facial expressions are *intrinsically meaningful*; indeed, they convey information about the feelings, thoughts, and intentions of others. The authors suggested that facial mimicry cannot be based merely on a perception-behavior link; rather, it requires the interpretation of the intention of a specific emotional stimulus in each context.

Following the evidence described so far, we created an experimental setting to investigate the cognitive and neurophysiological mechanisms underlying STM functioning by exploring how healthy participants process others' and their own emotional facial expressions both at an implicit and explicit level. To do so, we created a two-step study. In the preliminary experiment, we presented film excerpts (Schaefer et al., 2010) to induce specific discrete emotions in participants. Participants were video recorded during film viewing in order to create ecologic and dynamic facial videos, which were then validated by a group of 15 judges and were used as stimuli in the "others" condition of the main experiment. Following the same procedure, in the main experiment, we first recorded participants' faces while they watched the same movie excerpts, thus creating the dynamic stimuli belonging to the "self" condition, and then measured EMG facial activity during the observation of video depicting their own vs. others' facial expressions. Emotion recognition accuracy, valence, and arousal ratings were collected during the experiment as explicit measures.

Following the Motor Match Hypothesis prediction, we hypothesize that the EMG activity elicited by the different emotions will be congruent with the muscles involved in expressing the same emotion, confirming the specificity of sensorimotor simulation. Indeed, according to the theory, facial mimicry is an automatic match motor response that is not influenced by contextual or interpretative information. Within this framework, the beneficial effect of STM in the therapeutic process might be primarily caused by a more primitive and implicit sensorimotor/embodied re-experience of the seen expressions, which promotes the transition to a more mature and explicit self-reflective and interpretative ability, leading to the possibility of becoming aware of own emotions.

Conversely, finding a dissociation between the emotion seen and the corresponding EMG activity and/or the emotion expressed by own vs. others' facial expressions would support the Emotion Mimicry in Context view, thus corroborating the idea that a contextual–interpretative process occurs *before* facial mimicry takes place (Hess and Fischer, 2014) and is a prerequisite for STM effectiveness. In this case, the clinical efficacy of SMT could be primarily due to the supplementary information provided by observing their own emotions and hearing their own words, which helps patients to integrate their emotional experience at a richer and multisensory level.

Of course, the two mechanisms should be considered as not mutually exclusive and can coexist in the same patients during specific events, emotions, or moments of their lives.

MATERIALS AND METHODS

Preliminary Experiment: Stimuli Preparation and Validation

Phase 1: Stimuli Preparation

In the first phase, we recorded the faces of 15 volunteers while they were viewing 12 emotion-eliciting film excepts. The aim of this phase was twofold. First, we wanted to validate the efficacy of the selected movie clips in eliciting specific emotions. Second, we needed to create stimuli of dynamic and ecological facial expressions for the main experiment.

Participants

Fifteen volunteers (six males, M age = 23.1, range 20–30 years old) took part in the experiment in exchange for course credits. Participants were Caucasian, without beard or mustache, with normal or contact lens-corrected eyesight. Participants were naïve to the purpose of the study.

The entire study was approved by the local ethical committee, and participants were treated in accordance with the Declaration of Helsinki. All participants provided written informed consent to be recorded during the experiment and granted their authorization for the use of recordings for scientific purposes.

Procedure and analysis

Participants took part in the experiment individually. They sat in an artificially illuminated room at 60 cm from a 15.7" laptop monitor on which 12 film excerpts were presented. Ten of the clips were selected from a validated sample of emotion-eliciting film excerpts (Schaefer et al., 2010), and the other two were added in order to obtain more specific and time-locked emotional reactions (for a complete list of selected videos, see the Supplementary Material—Section A).

The duration of each clip ranged from 1 to 4 min, and, according to Schaefer et al. (2010), each of them elicited to a greater extent the following specific emotion: disgust, happiness, anger, sadness, and fear.

Participants were instructed to watch the film excerpts and then answer 11 order-randomized questions, asking to what extent they felt happiness, anxiety, anger, calm, disgust, joy, embarrassment, fear, engagement, surprise, and sadness during the clip. Each emotion intensity was rated using an analogical visual scale ranging from 0 (=not at all) to 100 (=to a very great extent).

The movie and question presentation was controlled by the software E-Prime 2.0 (Psychology Software Tools Inc., Pittsburgh, PA, United States) with the order of video clips randomized across participants. Participants' faces were recorded during movie observation using an HD Pro Webcam c 920 full HD 1080p (Logitech, Newark, CA, United States), which was fixed at the top of the laptop. The webcam was placed in front of the participant, so to have their face in the center of the recording area.

A white panel was placed behind the participant in order to create a uniform background. Participants were asked to pay attention to the movie, fixating the screen without covering their faces with their hands. In this way, we recorded 180 video clips of participants' faces (12 for each one) during movie observation. From these stimuli, we then selected the videos for the main experiment belonging to the condition "other." A manipulation check was run, confirming that each clip elicited the intended emotion to a greater extent as compared to others, thus replicating previous findings (Schaefer et al., 2010; Vergallito et al., 2018) (see **Supplementary Table S1**). Each "other" video was offline analyzed using FaceReader 6 software (Noldus, 2014), which automatically recognizes and codes the six basic universal facial expressions (plus a neutral state expression) with an accuracy of 89% (Lewinski et al., 2014). The aim of this procedure was to select the temporal window in which the specific emotion was maximally expressed. Indeed, the software allows videos (or pictures) to be analyzed and coded frame-by-frame, producing a summary table in which emotions are expressed on a scale from 0 to 1, where 0 corresponds to the absence of emotion indexes in the facial expression and 1 indicates maximum intensity (for a description of the algorithm used by FaceReader, see van Kuilenburg et al., 2005). This procedure allowed the exact time at which participants maximally expressed a specific emotion to be selected in order to cut 3-s time-window clips from each video (for this procedure, we used video editing software, namely, Windows Movie Maker). Then, we selected six videos for each participant, each one representing a discrete emotion (disgust, happiness, anger, sadness, fear, or a neutral condition), for a total of 90 clips.

Phase 2: Stimuli Validation

Participants and procedure

Fifteen Caucasian volunteers (five males; M age = 22.7, range 21–25), naïve to the purpose of the study, took part in the experiment in exchange for course credits.

Participants sat in an artificially illuminated room at 60 cm from a 15.7" laptop monitor, where the 90 video clips previously created were presented using the E-Prime 2 software. After each video, participants were asked: (i) to rate the valence of the actor's facial expression on an analogical visual scale ranging from 0 (negative) to 100 (positive); (ii) to select the specific emotion conveyed by the actor's facial expression among seven alternatives (disgust, happiness, anger, sadness, fear, neutral, none of the previous options); (iii) to rate how much they felt confident about their choice (0 = not confident at all, 100 = confident to a very great extent); (iv) how intense the emotion expressed by the facial expression was (0 = not intense at all, 100 = very intense); (v) how aroused they felt aroused during the video presentation (0 = not aroused at all, 100 = very aroused). We selected the three participants whose emotions, conveyed by facial expressions, were most accurately identified by the fifteen judges. In this way, we obtained six clips for each participant, each one representative of a specific emotion, for a total of 18 clips, which were used as videos belonging to the *other-expression condition* in the main experiment (see **Supplementary Table S2** for the judges' ratings).

Main Experiment: EMG Recordings

Participants and Procedure

Eighteen healthy volunteers (15 females and three males, M age = 22.4, range = 19–26) participated in the study. Participants were Caucasian, right-handed, and had normal vision or contact lens-corrected vision; males were shaved and without mustache. All participants were naïve to the purpose of the study. Participants took part in a two-session experiment. The first session aimed at collecting video of the participants' depicting their own facial expressions (*self-expression condition*). The second session represented the core of the present research project, consisting of facial EMG recording during the view of *self* and *other* videos.

First Session Procedure: Self-Stimuli Creation

The first session procedure was the same as phase one of the preliminary experiment (see section "Phase 1: Stimuli Preparation"), summarized as follows: (i) participants watched film excerpts, during which we recorded their faces; (ii) the videos were analyzed using FaceReader Software in order to select the time-window at which a given emotion was maximally expressed; (iii) for each of the six discrete emotions, a 3-s video was cut from the entire registration, thus obtaining six videos for each participant.

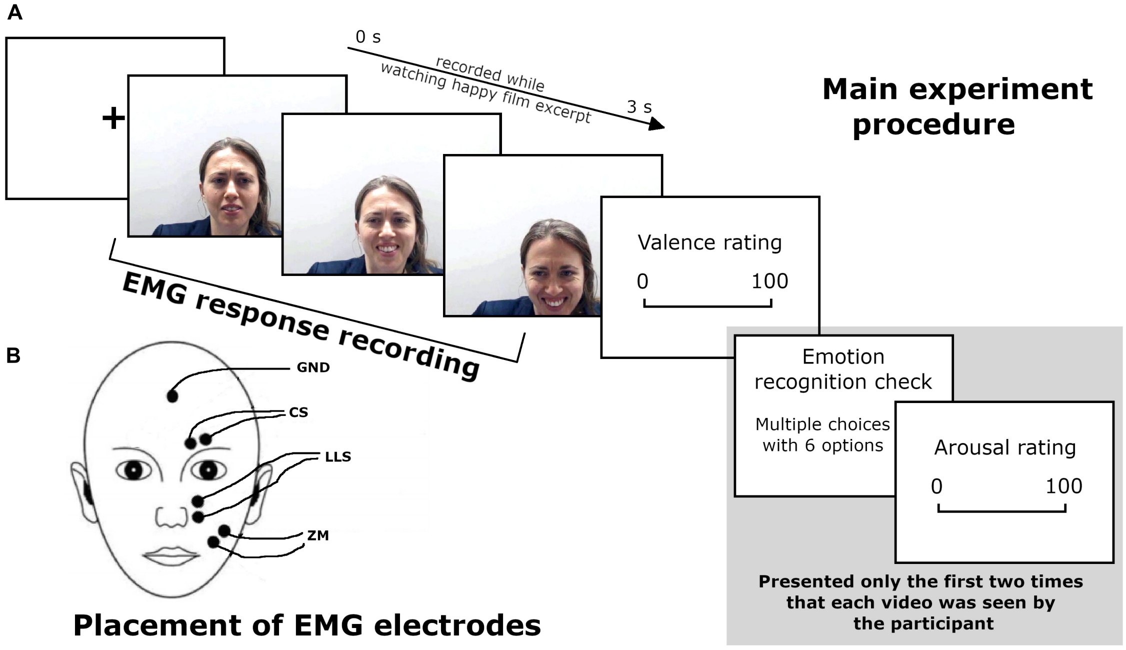


FIGURE 1 | (A) Main experiment procedure during the EMG recording. **(B)** Schematic representation of EMG electrode placement. *Written informed consent was obtained from the participant for the publication of this image.*

Second Session Procedure: Facial EMG Registration

The second session took place 1 week after the first one. Participants sat in a comfortable chair, in an artificially illuminated room at 60 cm from a 15.7" laptop monitor. Videos depicting their own facial expressions (*self-expression condition*) and videos of emotional expressions of three actors (*other-expression condition*) were presented during EMG recording (see **Figure 1** for a scheme of the procedure). Stimuli were randomly presented using E-Prime 2 software. Each video was presented 20 times, for a total of 480 trials divided into four blocks of 120 trials each. The total number of stimuli comprised 120 trials from each of the four expressers, namely, the participant—who changed for each of the 18 subjects—and the three actors chosen in phase 1.

Second Session Procedure: Explicit Measure

Following each video, participants were asked to rate the valence of stimuli on an analogical visual scale from 0 (negative) to 100 (positive). Moreover, we assessed participants' accuracy in recognizing the specific emotion and the arousal induced by each video. In particular, the first two times that a given video was presented (for a total of 48 trials), participants were asked to indicate which emotion was conveyed in a multiple-choice question with six options: (disgust, happiness, anger, sadness, fear, and neutral) and to what extent they were aroused by the video (analogical visual scale from 0—not at all to100—to a very great extent).

Electromyographic Recordings and Pre-processing

Facial surface EMG was recorded from three pairs of 4-mm diameter surface Ag/AgCl active electrodes corresponding to three distinct bipolar montages using a Digitimer D360 amplifier (Digitimer Ltd., Welwyn Garden City, Hertfordshire, United Kingdom). Electrodes were filled with Ten20 conductive paste (Weaver and Co., Aurora, CO, United States) and attached over the left[2] ZM, CS, and LLS, in accordance with guidelines from Fridlund and Cacioppo (1986, see also Cattaneo and Pavesi, 2014 for an overview of facial muscle anatomy). The ground electrode was placed at the midline, at the border of the hairline (Van Boxtel, 2010). The EMG signal was recorded by a computer using SIGNAL software (Cambridge Electronic Devices, Cambridge, United Kingdom) with online filters set at 50 Hz and 2 kHz, and a sampling rate of 5 kHz. Offline, signals were then digitally band-pass filtered between 20 and 400 Hz (van Boxtel, 2001) with SIGNAL software and were segmented into 15 time-bins of 200 ms, in addition to a 500-ms pre-stimulus baseline. The magnitude of the EMG signal was computed by calculating the root-mean-square (RMS) over 200-ms time-bins after the onset of each video. Trials in which the mean RMS was greater than three standard deviations from the mean value of that specific muscle were rejected. The RMS of each trial and bin was then divided by the baseline RMS. Finally, trials were averaged for each muscle based on emotion and self/other condition.

Statistical Analysis

Data analysis was carried out in the statistical programming environment R (R Development Core Team, 2013), using a linear mixed-effects model as the statistical procedure (Baayen et al., 2008).

---

[2] Muscular activity was recorded only from the left side of the face, which is more involved in emotional expressions (e.g., Rinn, 1984; Dimberg and Petterson, 2000).

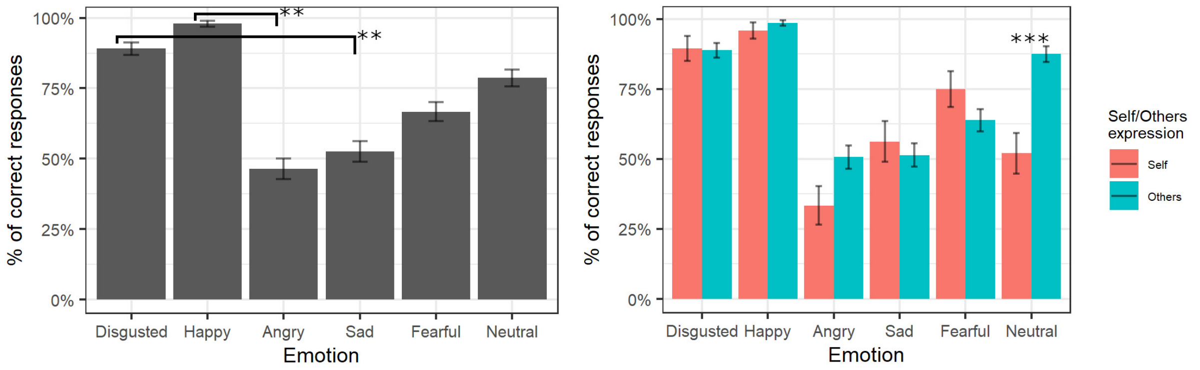


FIGURE 2 | Percentage of correctly identified emotions in the main experiment. Significant differences between emotions are highlighted with asterisks (left panel, ** <0.01). Percentage of correct identifications for self/other expression are depicted on the right panel, with asterisks highlighting self vs. other differences (*** <0.001).

For behavioral data, valence and arousal ratings were submitted to a series of linear mixed-effects regressions using the LMER procedure, whereas accuracy was submitted to a series of binomial logistic regression using the GLME procedure in the lme4 R package (version 1.1-17; Bates et al., 2015).

As fixed effects, emotion (factorial, six levels), self/other expression (factorial, two levels: self, other), bin (factorial, 15 levels[3]), and their interactions were tested with a series of likelihood ratio tests (LRT) to assess the inclusion of the effects, which significantly increased the model's goodness of fit (Gelman and Hill, 2006). See **Supplementary Tables S3–S5** for the model selection. Concerning the random effect structure, by-subject and a by-video[4] random intercepts were included; moreover, random by-subject and by-video random slopes for emotion and self/other expression random effects were included only when the model's goodness of fit increased.

For EMG data, pre-processed ZM, CS, and LLS activity was submitted to a series of linear mixed-effects regressions using the LMER procedure, testing in a forward stepwise LRT procedure the fixed effects of emotion (factorial, six levels), self/other expression (factorial, two levels) and bin (factorial, 15 levels), and their interactions. The random-effects structure included by-subject, by-video, and by-trial intercepts. After having fitted the full model for each muscle, influential outliers were removed via model-criticism (2.5 SD of standardized residuals).

The results of the LRT procedures for model selection and the parameters of the final best-fitting models are reported in **Supplementary Tables S6–S8**. *Post hoc* procedures on the final best-fitting model, applying Bonferroni correction for multiple comparisons, were carried out for direct pairwise contrasts on significant main effects and interactions using the "phia" R package (version 0.2-1, De Rosario-Martinez, 2015). See Supplementary Material—Section C for tables summarizing average and standard error mean values for explicit and implicit measures.

RESULTS

Explicit Measure Results

Accuracy Ratings

The final model of logistic regression on accuracy included the fixed effects of emotion [$\chi^2(5) = 228.88$, $p < 0.001$], self/other expression [$\chi^2(1) = 2.38$, $p = 0.12$], and their interaction [$\chi^2(6) = 30.08$, $p < 0.001$] (see **Supplementary Table S3** for the model selection). *Post hoc* tests on the main effect of emotion showed that videos showing happy expressions were better recognized than videos showing angry expressions ($p = 0.007$), and videos representing disgust expressions were better-recognized than videos showing sad expressions ($p = 0.006$). The significant interaction showed that self/other expressions differently affected accuracy in specific emotions; in particular, *post hoc* comparisons revealed higher accuracy for other than self faces only for the neutral expressions ($p < 0.001$; see **Figure 2**).

Valence Ratings

The final model for valence ratings included the fixed effects of emotion [$\chi^2(5) = 10872$, $p < 0.001$] and self/other expression [$\chi^2(1) = 0.001$, $p = 0.97$] as well as their interaction [$\chi^2(6) = 162.01$, $p < 0.001$] (see **Supplementary Table S4** for the model selection). *Post hoc* tests on the main effect of emotion highlighted that videos showing happy expressions were rated as more positive than videos showing all other expressions (all *ps* < 0.001) and neutral videos as more positive than those showing angry, sad, fear, and disgust facial expressions (all *ps* < 0.001). Moreover, videos showing sad expressions were rated as more positive than videos displaying anger ($p < 0.001$) and fear ($p = 0.004$), and videos showing a fear expression were rated as more negative than videos representing disgust ($p = 0.02$). The significant interaction showed that self/other

---

[3] Due to the ecologic and dynamic nature of our face recording, the maximal point of emotion expressions was not time locked across emotions and actors. Therefore, the notion of bin was reasoned to consider potential effects due to the different time points at which different emotion expressions might occur at the maximum intensity.

[4] Here we referred to video to indicate each single stimulus presented to participants.

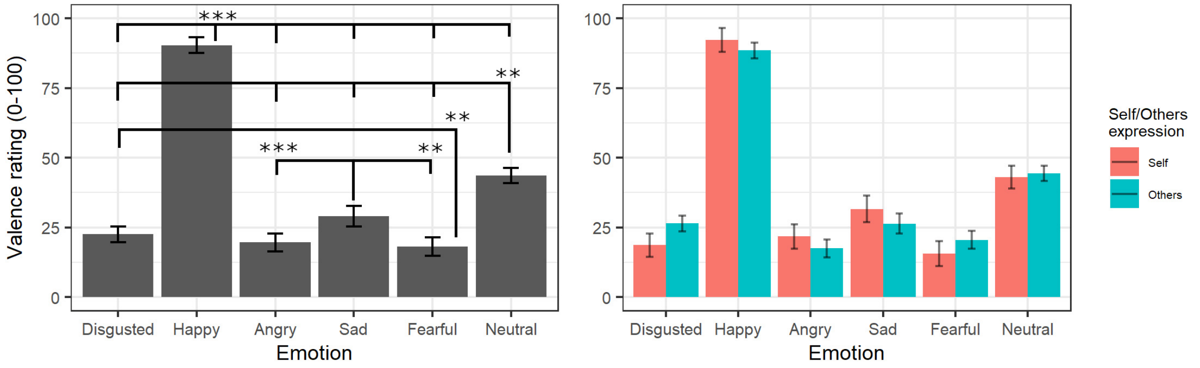


FIGURE 3 | Estimated marginal means for Valence ratings of videos in the main experiment. Vertical bars represent standard error. Significant differences between emotions are highlighted with asterisks (left panel, $^{**}<0.01$, $^{***}<0.001$). Valence ratings for self/other expression are depicted on the right panel.

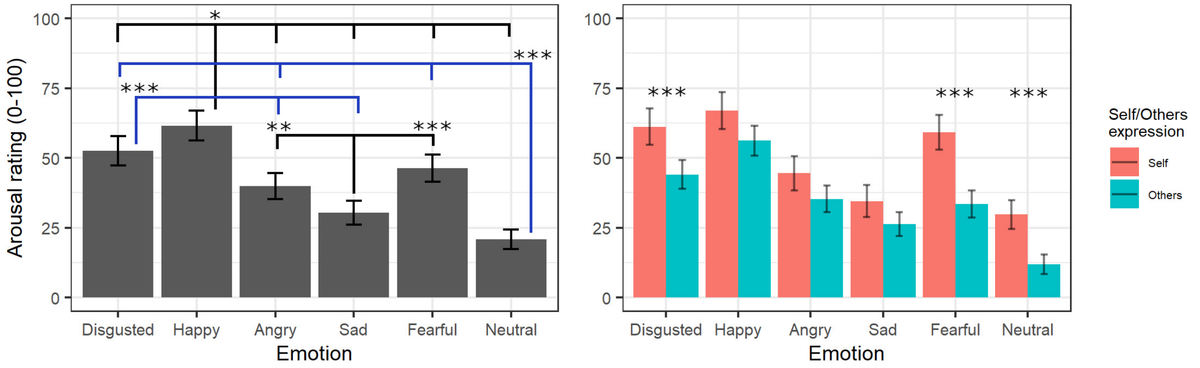


FIGURE 4 | Estimated marginal means for Arousal ratings of videos in the main experiment. Vertical bars represent standard error. Significant differences between emotions are highlighted with asterisks (left panel, $^{*}<0.05$, $^{**}<0.01$, $^{***}<0.001$). Arousal ratings for self/other expression are depicted on the right panel, with asterisks highlighting self vs. other differences ($^{***}<0.001$).

expressions differently affected emotion valence ratings; however, *post hoc* tests showed no significant differences between self and other videos in any of the specific emotions ($p > 0.05$; see **Figure 3**).

Arousal Ratings

The final model on arousal ratings included the fixed effects of emotion [$\chi^2(5) = 508.73$, $p < 0.001$] and self/other expression [$\chi^2(1) = 6.04$, $p = 0.014$] as well as their interaction [$\chi^2(5) = 21.27$, $p < 0.001$] (see **Supplementary Table S5** for the model selection). *Post hoc* tests on the main effect of emotion evidenced greater arousal scores for videos showing happy expressions compared to those showing all other emotions ($ps < 0.02$), whereas neutral videos were rated with lower arousal scores compared to those showing all other emotions ($ps < 0.001$) except for sad ($p = 0.053$). Moreover, videos showing expressions of disgust were given greater arousal scores than videos with angry and sad ($p < 0.001$) expressions, and sad videos were rated as less arousing than videos with angry ($p = 0.001$) and fearful ($p < 0.001$) expressions. Parameters for the interaction effects showed that the self/other conditions differently affected arousal scores in specific emotions. Indeed, *post hoc* tests showed significantly higher arousal scores for self compared to others' facial expressions showing fear, disgust, and neutral emotions ($ps < 0.001$) but not with angry, happy, or sad ones ($ps > 0.2$; see **Figure 4**).

EMG Results

Zygomatic Major Activity

The final model for ZM activity included the main effects of emotion [$\chi^2(5) = 105.43$, $p < 0.001$], self/other expression [$\chi^2(1) = 0.52$, $p = 0.470$], bin [$\chi^2(14) = 28.31$, $p = 0.013$], and the interaction between emotion and self/other expression [$\chi^2(5) = 39.42$, $p < 0.001$] (see **Supplementary Table S6** for the model selection). *Post hoc* tests on the main effect of emotion revealed significant greater ZM activity for videos with happy, anger, and disgust expressions compared to videos with sad, fear, and neutral expressions (all $ps < 0.001$). *Post hoc* analyses on bin main effect showed no significant difference. Direct *post hoc* contrasts on the effect of self/other expression in each emotion showed significantly greater ZM activity for self compared to other videos when angry ($p = 0.001$) and disgust ($p = 0.013$) emotions were presented (see **Figure 5**).

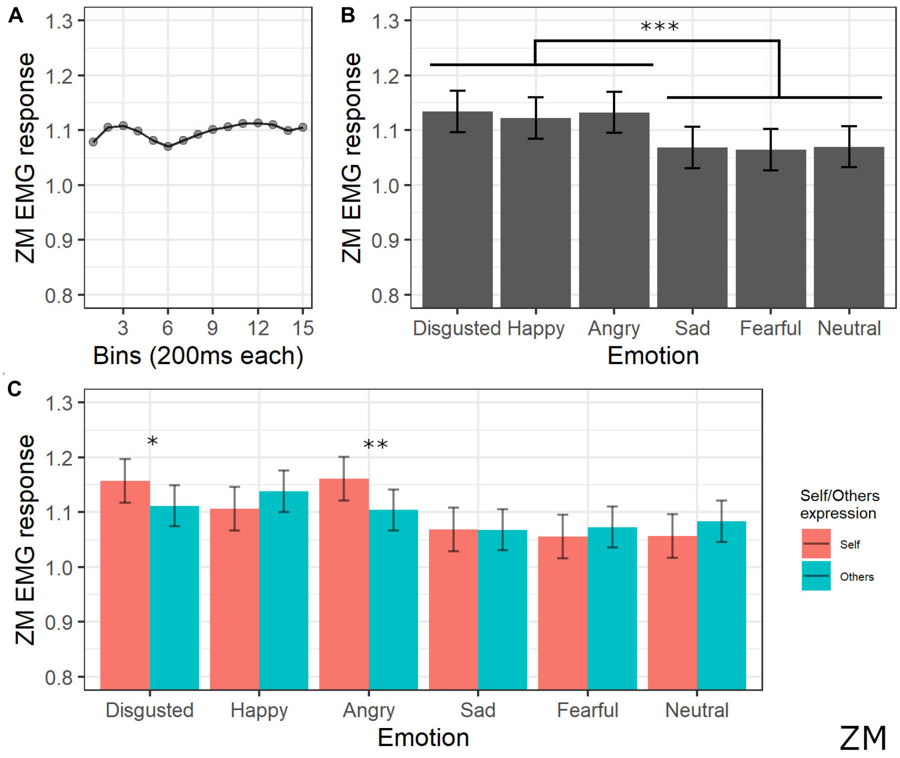


FIGURE 5 | Results for ZM activity. **(A)** Main effect of bin. Each dot represents the estimated marginal mean of EMG response within the timeframe of each bin. **(B)** Main effect of emotion. Estimated marginal means of EMG response for each emotion. Vertical bars represent standard error. Significant differences between emotions are highlighted with asterisks (***<0.001). **(C)** Interaction between emotions and self/other condition. Significant differences between self vs. other are highlighted with asterisks (*<0.05, **<0.01).

Corrugator Supercilii Activity

The final model for CS activity included the main effects of emotion [$\chi^2(5) = 205.64$, $p < 0.001$], self/other expression [$\chi^2(1) = 63.36$, $p < 0.001$], and bin [$\chi^2(14) = 19.59$, $p = 0.144$], as well as emotion by self/other expression [$\chi^2(5) = 86.19$, $p < 0.001$] and emotion by bin [$\chi^2(70) = 91.83$, $p = 0.041$] interactions (see **Supplementary Table S7** for the model selection). *Post hoc* tests performed on the main effect of self/other expression showed that CS activity was lower for self compared to other videos ($p < 0.001$). *Post hoc* testing performed on emotion main effect showed significantly lower activity of CS for videos with a happy expression compared to all emotions (*ps* < 0.001) and for disgust compared to angry, fear, and neutral expressions (*ps* < 0.001). Moreover, CS activity was significantly higher for angry compared to sad ($p < 0.001$), neutral ($p < 0.001$), and fearful ($p = 0.029$) expressions and for fear compared to sad videos ($p < 0.001$). Finally, activity was lower for sad compared to neutral videos ($p = 0.029$). Self/other expression differently affected CS activity depending on the specific emotion: *post hoc* tests revealed significantly lower activity for self compared to others' faces when happy ($p < 0.001$), sad ($p < 0.001$), fear ($p = 0.001$), and neutral ($p < 0.001$) expressions were presented. Finally, *post hoc* on the emotion by bin interaction revealed a difference between CS activity for happy videos compared to angry ones starting from bin 6 (*ps* < 0.05; see **Figure 6**).

Levator Labii Superioris Activity

The final model for LLS activity included the main effects of emotion [$\chi^2(5) = 183.1259$, $p < 0.001$], self/other expression [$\chi^2(1) = 1.6476, p = 0.19$], and bin [$\chi^2(14) = 117.8613, p < 0.001$], as well as emotion by self/other expression [$\chi^2(5) = 53.7085$, $p < 0.001$] and emotion by bin [$\chi^2(70) = 154.7846$, $p < 0.001$] interactions (see **Supplementary Table S8** for the model selection). *Post hoc* tests on the main effect of emotion showed significant higher LLS activity for happy expressions compared to all emotions (*ps* < 0.001, except for disgust with $p = 0.002$), for disgust compared to sad ($p < 0.001$) and neutral ($p < 0.001$) expressions, and for fear compared to sad ($p = 0.005$) and neutral ($p = 0.001$) emotions. *Post hoc* analyses of interaction between the emotion and self/other expressions showed that LLS was significantly more activated when seeing others' compared to self faces expressing happiness ($p < 0.001$) and neutral ($p = 0.014$) emotions; higher LLS activity for self compared to others emerged for sad ($p = 0.014$) and fear ($p < 0.01$) expressions. Finally, *post hoc* testing on bin main effect revealed higher activity in later bins of the videos (from 2000 to 3000 ms) compared to the earlier bins (from 0 to 1200 ms) (*ps* < 0.05); in particular, analysis on the emotion by bin interaction revealed that LLS activity increased for happy videos compared to all presented emotions starting from bin 10 (*ps* < 0.05; see **Figure 7**).

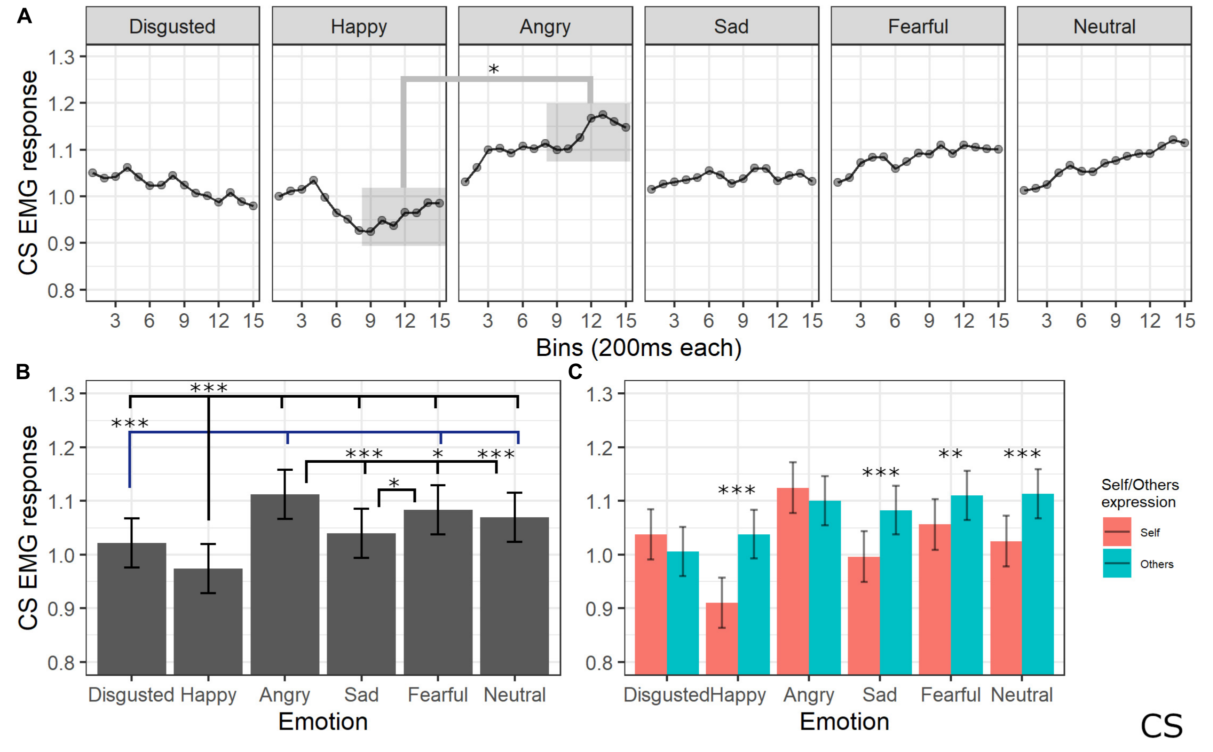


FIGURE 6 | Results for CS activity. **(A)** Main effect of the self/other condition. Estimated marginal means of EMG response for video of self vs. others. Significant difference highlighted with asterisks (*<0.05). Vertical bars represent standard error. **(B)** Main effect of emotion. Estimated marginal means of EMG response for each emotion. Significant differences between emotions are highlighted with asterisks (*<0.05, ***<0.001). **(C)** Interaction between emotions and self/other condition. Significant differences between self vs. other are highlighted with asterisks (**<0.01, ***<0.001).

DISCUSSION

The present study aimed at exploring, for the first time, the neurophysiological and cognitive underpinnings of SMT by assessing explicit and implicit responses to the view of own vs. others' faces dynamically expressing discrete emotions. Ecological dynamic stimuli representing the category "other" were created in a preliminary experiment by video-recording participants while watching film excerpts eliciting sad, angry, fearful, happy, disgusted, or neutral emotions. The same procedure was used in the main experiment to create "own" stimuli. Then, in the main experiment, explicit and implicit measures were collected. Explicit measures concerned emotion identification, valence, and arousal ratings, whereas facial EMG activity (implicit measure) was recorded from the ZM, CS, and LLS as measurements of facial mimicry induced by perceiving own and others' emotional expressions.

Results on explicit measures confirmed the validity of our stimuli, which were recognized and rated for valence and arousal according to the previous literature on emotion processing. Indeed, happiness was the emotion with the highest accuracy score, confirming a longstanding finding on the advantage of happy-face recognition over all negative facial expressions (Feyereisen et al., 1986; Kirita and Endo, 1995; Leppänen and Hietanen, 2003; Palermo and Coltheart, 2004; Juth et al., 2005; Calvo and Lundqvist, 2008; Tottenham et al., 2009; Calvo et al., 2010), likely due to its highly salient and distinctive facial features. Interestingly, an interaction between emotion and the self/other condition was found, consisting of better accuracy in recognizing other compared to own neutral emotion, which is considered the most ambiguous facial expression. Crucially, 34.7% of participants categorized their own neutral face as representing an emotion, which was negative in 88% of the incorrect categorizations (sadness in 73%, fear in 9%, and angry in 18% of trials), while it was rated as positive (happiness) in 12% of trials. Only in 15.3% of trials did participants evaluate others' neutral faces as depicting an emotion. The tendency to attribute the status of emotion even to a minimum sign of muscular activation in own faces can be interpreted in the light of the embodied simulation theory and is a well-known and frequent effect in the clinical application of SMT. Clinical experience with SMT, indeed, suggests that the patient recognizes an emotion on his/her face even when highly expert therapists are not able to detect it (Vinai et al., 2016). When requested to pick the frame, patients are incredibly competent at indicating on the screen the exact moment in which they see the minimal movement of the lips or of the eyes indicating the emotional state.

Concerning valence ratings, as expected, happiness was the most positively rated expression, followed by neutral, sadness, disgust, anger, and fear, which received lower ratings. From

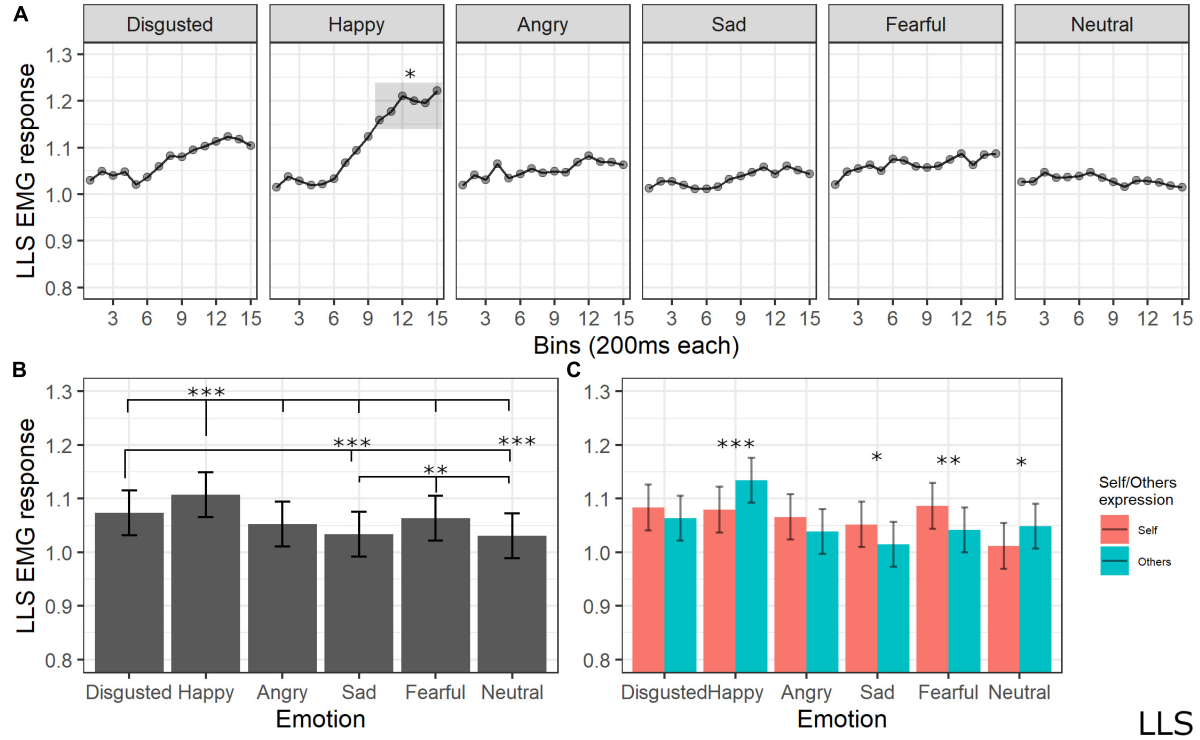


FIGURE 7 | (A) Interaction between emotions and bin. Each dot represents the estimated marginal mean of EMG response within the timeframe of each bin. Highlighted bins are significantly higher than the same bins from other emotions (*<0.05). **(B)** Main effect of emotion. Estimated marginal means of EMG response for each emotion. Significant differences between emotions are highlighted with asterisks (***<0.001, **<0.01). **(C)** Interaction between emotions and self/other condition. Significant differences between self vs. other are highlighted with asterisks (***<0.001, **<0.01, *<0.05).

a clinical perspective, it is interesting that sadness was not considered the most negative expression. This result is in line with previous studies: indeed, the expression of sadness—as happiness—signals an affiliative intention of the expresser, inducing an increase of liking, prosocial behavior, and other positive actions (Jakobs et al., 2001; Hess and Fischer, 2013, 2017).

Concerning arousal ratings, happiness in the *own condition* was rated as the most arousing condition, whereas neutral *other* was rated as the lowest. Interestingly, own expressions were rated as more arousing overall, with fear, disgust, and neutral expressions reaching a statistically significant difference as compared to the other condition. It is of note that the observation of own sadness was less arousing as compared to the other negative expressions (and less negatively rated), in line with clinical consideration. Indeed, patients' observation of their own face while recalling sad life events typically induces a positive sense of self-compassion (Vinai et al., 2015; Petrocchi et al., 2017). In the light of the attachment theory (Ainsworth and Bowlby, 1991), we can speculate that the sight of ourselves in trouble induces a positive desire of caregiving (more than desperation), which would be useful from an evolutionary point of view.

Regarding implicit measures, the EMG pattern that emerged in our study is in line with previous research concerning the muscular-specific activation induced by emotion valence, thus confirming the validity of our experimental apparatus. Indeed, CS activity was maximally increased by angry faces and reduced during the observation of happy expressions (e.g., Dimberg and Thunberg, 1998; Dimberg and Petterson, 2000; see Hess and Fischer, 2013 for a review). This pattern is the most consistently reported when analyzing facial EMG, suggesting that the CS tenses during negative emotion processing and relaxes during processing positive emotions (Ekman, 2007). Moreover, our results on explicit measures suggested that the expression of anger (together with fear) was rated as more negative compared to the other emotions. ZM activity was higher for stimuli depicting happiness, in line with previous studies suggesting the involvement of this muscle during the processing of positive valence stimuli (e.g., Larsen et al., 2003; Tan et al., 2012; Hess and Fischer, 2013). LLS activity was maximally increased by the expression of happiness as compared to all other emotions and of disgust as compared to sadness and neutral emotions. In line with our findings, LLS activity has been reported to be specifically involved in the expression and facial mimicry of disgust (Lundqvist and Dimberg, 1995; Van Boxtel, 2010), but some previous studies also found an increased activity of this muscle during the processing of positive valence emotions (Lundqvist, 1995; Vergallito et al., 2019).

Crucially, the emotion expressed in the videos interacted with the self-other condition in influencing EMG activation. The CS was more relaxed during the observation of own expressions of happiness, sadness, fear, and the neutral condition as compared to the same emotions expressed by others. As previously mentioned, the typical pattern of CS activation consists of reducing its activity during positive emotions; indeed, it seems reasonable that processing one's own emotion of happiness induces a stronger effect, namely, a lower activation. Less clear is which mechanisms reduce CS even for other emotions that are considered as negative. A possible speculation, based on the clinical evidence previously discussed (Vinai et al., 2015), is that the reduced activity of the CS might reflect a positive feeling of self-compassion induced by observation of one's own emotion. For the ZM, the interaction between emotion expression and the self vs. other condition was significant during the observation of angry and disgusted faces, with higher activity for the own condition as compared to the other condition. To the best of our knowledge, no previous studies indexed the mimicry of these emotions to the ZM; thus, further studies are needed to account for this dissociation. Finally, in the LLS, we found higher activation for the other condition as compared to the self condition for happy and neutral expressions, while increased activity in the self condition emerged for sadness and fear expressions.

These results contrast with the hypothesis that facial mimicry is fully automatic, as predicted by the Motor Match Hypothesis since, if the process were simply based on a perception–behavior link, it should be independent of who is the expresser and who is the perceiver. Our results suggest instead that EMG recorded activity also reflected a post-interpretative stage, corresponding to the emotion experienced as a consequence of the one observed on the expresser's face.

Despite the specific pattern of ratings and muscular activations found in the present work, the main result that clearly emerged is a dissociation at both explicit and implicit levels in emotional processing of self vs. other emotions. Few studies have systematically investigated this issue. What we know from past research is that self-related stimuli are more relevant to us than those related to others (e.g., Ross and Sicoly, 1979; Brédart et al., 2006) and that this sense of self is intrinsically linked to one's own face (e.g., Porciello et al., 2014). Crucially, even though we can perceive our face only by looking at a mirror, the view of our own face is more effective in activating mirror neurons than is perceiving the face of another person (Uddin et al., 2005, 2007). Differences in processing self vs. other expressions have also been detected at the multisensory integration level: indeed, viewing their own face being touched modulates participants' tactile experience more strongly than viewing other participants being touched (Serino et al., 2008).

To the best of our knowledge, no previous studies directly investigated facial mimicry occurring while viewing own as compared to others' facial expressions. Considering our data and previous research, however, it is possible that at least two different mechanisms form the basis of STM efficacy in clinical settings. On one side, we have an automatic and more primitive facial mimicry process, which elicited in the observers the activation of the mirror neuron system and thus the previously described embodied simulation phenomenon, fostering own emotion recognition via this sensorimotor simulation. This process could be useful from an adaptive point of view: indeed, simulation helps one to understand what the other person is thinking and prepare an appropriate behavioral response. At the same time, we know from previous studies that facial mimicry is not only simulation (i.e., smiling when observing another person smiling) but can also be reactive (i.e., expressing fear when observing an angry face; Dimberg, 1982; McIntosh et al., 2006). In line with this evidence, viewing ourselves suffering might induce a different type of emotion, such as a feeling of self-compassion (Petrocchi et al., 2017). Indeed, the fact that observing own negative emotions causes greater arousal and modulation of electrophysiological responses might be the basis for a deeper comprehension and self-compassion, leading patients to recognize their suffering through the emotions depicted on their faces and becoming more aware of their emotional experience, thus providing hints on the neurophysiological mechanism at the root of SMT success.

It must be said that our findings are more in line with this second interpretation, but we ran this exploratory study on healthy participants, in which self-reflective and introspective abilities are expected to be less compromised as compared to patients.

Finally, it is possible that in our experimental participants considered their own videos differently from those showing others. Indeed, they had contextual information that made own expressions different from those of others; namely, they remembered how and when their expressions were produced. However, we do not think this possibility undermines our results: indeed, it would reinforce the idea that facial mimicry occurs only secondarily to an automatic-interpretative stage. Moreover, this would make our paradigm even more similar to the clinical setting: indeed, patients know how and when the recalled experience took place.

Taken together, our results open new avenues for future research on the contrast between explicit and implicit reactions to facial expressions and on the difference between processing own vs. others' emotions.

Study Limitations

The main limitation of the study is the unbalanced number of males and females (3 vs. 15) in our sample. Indeed, it is known that females are more facially expressive than males in emotion-evoking situations (e.g., Buck et al., 1972; Buck, 1984). This difference translates to larger facial muscular activity in females (e.g., Schwartz et al., 1980; Dimberg and Lundquist, 1990), which is present from childhood (Cattaneo et al., 2018) and according to previous researchers may be valence-dependent, with females being more reactive to positive emotional stimuli and males to negative ones (McDuff et al., 2017). Moreover, gender differences have also been suggested in explicit face processing, with females evaluating human faces as more positive and arousing as compared to men (e.g., Proverbio, 2017).

DATA AVAILABILITY STATEMENT

The dataset generated from the current study is available as **Supplementary Material**.

ETHICS STATEMENT

The study was approved by the Ethical Committee of the Department of Psychology of the University of Milano-Bicocca, and participants' ethical treatment was in accordance with the principles stated in the Declaration of Helsinki. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

MS, PiV, LR, AV, and EG conceived and designed the study. VR, SA, and EL run the experiment. GM performed statistical analysis. GM, LR, and AV interpreted results. AV, GM, and LR drafted the paper, which was critically revised by all authors.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.00433/full#supplementary-material

REFERENCES

Ainsworth, M. D. S., and Bowlby, J. (1991). An ethological approach to personality development. *Am. Psychol.* 46, 331–341.

Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* 59, 390–412. doi: 10.1016/j.jml.2007.12.005

Bastiaansen, J. A., Thioux, M., and Keysers, C. (2009). Evidence for mirror systems in emotions. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 2391–2404. doi: 10.1098/rstb.2009.0058

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., et al. (2015). "Linear mixed-effects models using 'Eigen' and S4," in *Package 'lme4'. Convergence* 12:2 (accessed January 24, 2020).

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

Brédart, S., Delchambre, M., and Laureys, S. (2006). Short article: one's own face is hard to ignore. *Q. J. Exp. Psychol.* 59, 46–52. doi: 10.1080/17470210500343678

Buck, R. (1984). *The Communication of Emotion*. New York, NY: Guilford Press.

Buck, R. W., Savin, V. J., Miller, R. E., and Caul, W. F. (1972). Communication of affect through facial expressions in humans. *J. Pers. Soc. Psychol.* 23, 362–371. doi: 10.1037/h0033171

Calvo, M. G., and Lundqvist, D. (2008). Facial expressions of emotion (KDEF): Identification under different display-duration conditions. *Behav. Res. Methods* 40, 109–115. doi: 10.3758/brm.40.1.109

Calvo, M. G., Nummenmaa, L., and Avero, P. (2010). Recognition advantage of happy faces in extrafoveal vision: featural and affective processing. *Vis. Cogn.* 8, 1274–1297. doi: 10.1080/13506285.2010.481867

Cattaneo, L., and Pavesi, G. (2014). The facial motor system. *Neurosci. Biobehav. Rev.* 38, 135–159. doi: 10.1016/j.neubiorev.2013.11.002

Cattaneo, L., Veroni, V., Boria, S., Tassinari, G., and Turella, L. (2018). Sex differences in affective facial reactions are present in childhood. *Front. Integr. Neurosci.* 12:19. doi: 10.3389/fnint.2018.00019

Chartrand, T. L., and Bargh, J. A. (1999). The chameleon effect: the perception–behavior link and social interaction. *J. Pers. Soc. Psychol.* 76, 893–910. doi: 10.1037/0022-3514.76.6.893

Corcos, M., Guilbaud, O., Speranza, M., Paterniti, S., Loas, G., Stephan, P., et al. (2000). Alexithymia and depression in eating disorders. *Psychiatry Res.* 93, 263–266. doi: 10.1016/s0165-1781(00)00109-8

De Rosario-Martinez, H. (2015). *Phia: Post-hoc Interaction Analysis. R Package Version 0.2-1* (accessed January 24, 2020).

De Wied, M., van Boxtel, A., Zaalberg, R., Goudena, P. P., and Matthys, W. (2006). Facial EMG responses to dynamic emotional facial expressions in boys with disruptive behavior disorders. *J. Psychiatr. Res.* 40, 112–121. doi: 10.1016/j.jpsychires.2005.08.003

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

Dimberg, U. (1982). Facial reactions to facial expressions. *Psychophysiology* 19, 643–647. doi: 10.1111/j.1469-8986.1982.tb02516.x

Dimberg, U., and Lundquist, L. O. (1990). Gender differences in facial reactions to facial expressions. *Biol. Psychol.* 30, 151–159. doi: 10.1016/0301-0511(90)90024-q

Dimberg, U., and Petterson, 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

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

Ekman, P. (2007). *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*. New York, NY: St. Martin's Griffin.

Ekman, P., and Friesen, W. V. (1978). *The Facial Action Coding System (FACS)*. Palo Alto, Ca: Consulting Psychologists Press.

Feyereisen, P., Malet, C., and Martin, Y. (1986). "Is the faster processing of expressions of happiness modality-specific?," in *Aspects of Face Processing*, eds H. D. Ellis, M. A. Jeeves, and F. Newcombe (Dordrecht: Springer), 349–355. doi: 10.1007/978-94-009-4420-6_37

Frau, C., Cocco, P., Navarra, R., Caruso, C., and Ross, C. (2018). The use of self-mirroring therapy in phase 1 treatment of a patient with other specified dissociative disorder. *Eur. J. Trauma Dissoc.* 2, 55–61. doi: 10.1016/j.ejtd.2017.10.004

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

Gallese, V., and Sinigaglia, C. (2011). What is so special about embodied simulation? *Trends Cogn. Sci.* 15, 512–519. doi: 10.1016/j.tics.2011.09.003

Gelman, A., and Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge university 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

Güleç, M. Y., Altintaş, M., Ýnanç, L., Bezgin, Ç. H., Koca, E. K., and Güleç, H. (2013). Effects of childhood trauma on somatization in major depressive disorder: The role of alexithymia. *J. Affect. Disord.* 146, 137–141. doi: 10.1016/j.jad.2012.06.033

Harrison, A., Sullivan, S., Tchanturia, K., and Treasure, J. (2010). Emotional functioning in eating disorders: attentional bias, emotion recognition and emotion regulation. *Psychol. Med.* 40, 1887–1897. doi: 10.1017/s0033291710000036

Hess, U., and Blairy, S. (2001). Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy. *Int. J. Psychophysiol.* 40, 129–141. doi: 10.1016/s0167-8760(00)00161-6

Hess, U., and Fischer, A. (2013). Emotional mimicry as social regulation. *Pers. Soc. Psychol. Rev.* 17, 142–157. doi: 10.1177/1088868312472607

Hess, U., and Fischer, A. (2014). Emotional mimicry: why and when we mimic emotions. *Soc. Pers. Psychol. Comp.* 8, 45–57. doi: 10.1111/spc3.12083

Hess, U., and Fischer, A. (2017). "The role of emotional mimicry in intergroup relations," in *Oxford Research Encyclopedia of Communication*, eds H. Giles and J. Harwood (New York, NY: Oxford University Press).

Jakobs, E., Manstead, A. S., and Fischer, A. H. (2001). Social context effects on facial activity in a negative emotional setting. *Emotion* 1, 51–69. doi: 10.1037/1528-3542.1.1.51

Juth, P., Lundqvist, D., Karlsson, A., and Öhman, A. (2005). Looking for foes and friends: perceptual and emotional factors when finding a face in the crowd. *Emotion* 5, 379–395. doi: 10.1037/1528-3542.5.4.379

Karukivi, M., Hautala, L., Kaleva, O., Haapasalo-Pesu, K. M., Liuksila, P. R., Joukamaa, M., et al. (2010). Alexithymia is associated with anxiety among adolescents. *J. Affect. Disord.* 125, 383–387. doi: 10.1016/j.jad.2010.02.126

Keltner, D., and Kring, A. M. (1998). Emotion, social function, and psychopathology. *Rev. Gen. Psychol.* 2, 320–342. doi: 10.1037/1089-2680.2.3.320

Kirita, T., and Endo, M. (1995). Happy face advantage in recognizing facial expressions. *Acta Psychol.* 89, 149–163. doi: 10.1016/0001-6918(94)00021-8

Kojima, M. (2012). Epidemiologic studies of psychosocial factors associated with quality of life among patients with chronic diseases in Japan. *J. Epidemiol.* 22, 7–11. doi: 10.2188/jea.je20110114

Künecke, J., Wilhelm, O., and Sommer, W. (2017). Emotion recognition in nonverbal face-to-face communication. *J. Nonverb. Behav.* 41, 221–238. doi: 10.1007/s10919-017-0255-2

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

Leppänen, J. M. (2006). Emotional information processing in mood disorders: a review of behavioral and neuroimaging findings. *Curr. Opin. Psychiatry* 19, 34–39. doi: 10.1097/01.yco.0000191500.46411.00

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

Leweke, F., Bausch, S., Leichsenring, F., Walter, B., and Stingl, M. (2009). Alexithymia as a predictor of outcome of psychodynamically oriented inpatient treatment. *Psychother. Res.* 19, 323–331. doi: 10.1080/10503300902870554

Lewinski, P., den Uyl, T. M., and Butler, C. (2014). Automated facial coding: validation of basic emotions and FACS AUs in FaceReader. *J. Neurosci. Psychol. Econ.* 7, 227–236. doi: 10.1037/npe0000028

Loas, G. (2012). Alexithymia and dependent personality disorder. *Psychiatry Res.* 196, 325–326. doi: 10.1016/j.psychres.2011.12.030

Loas, G., Speranza, M., Pham-Scottez, A., Perez-Diaz, F., and Corcos, M. (2012). Alexithymia in adolescents with borderline personality disorder. *J. Psychosom. Res.* 72, 147–152.

Lorenzini, R., Sassaroli, S., and Ruggiero, G. M. (eds) (2006). *Psicoterapia Cognitiva Dell'ansia:(rimuginio), Controllo ed Evitamento*. Cortina: Raffaello Cortina editore.

Lundqvist, L. O. (1995). Facial EMG reactions to facial expressions: a case of facial emotional contagion? *Scand. J. Psychol.* 36, 130–141. doi: 10.1111/j.1467-9450.1995.tb00974.x

Lundqvist, L. O., and Dimberg, U. (1995). Facial expressions are contagious. *J. Psychophysiol.* 9, 203–211.

McDuff, D., Kodra, E., el Kaliouby, R., and LaFrance, M. (2017). A large-scale analysis of sex differences in facial expressions. *PLoS One* 12:e0173942. doi: 10.1371/journal.pone.0173942

McIntosh, D. N., Reichmann-Decker, A., Winkielman, P., and Wilbarger, J. L. (2006). When the social mirror breaks: deficits in automatic, but not voluntary, mimicry of emotional facial expressions in autism. *Dev. Sci.* 9, 295–302. doi: 10.1111/j.1467-7687.2006.00492.x

Murata, A., Saito, H., Schug, J., Ogawa, K., and Kameda, T. (2016). Spontaneous facial mimicry is enhanced by the goal of inferring emotional states: evidence for moderation of "automatic" mimicry by higher cognitive processes. *PLoS One* 11:e0153128. doi: 10.1371/journal.pone.0153128

Neal, D. T., and Chartrand, T. L. (2011). Embodied emotion perception: amplifying and dampening facial feedback modulates emotion perception accuracy. *Soc. Psychol. Pers. Sci.* 2, 673–678. doi: 10.1177/1948550611406138

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

Noldus (2014). *FaceReader: Tool for Automatic Analysis of Facial Expression: Version 6.0*. Wageningen: Noldus Information Technology.

Oberman, L. M., and Ramachandran, V. S. (2007). The simulating social mind: the role of the mirror neuron system and simulation in the social and communicative deficits of autism spectrum disorders. *Psychol. Bull.* 133, 310–327. doi: 10.1037/0033-2909.133.2.310

Oberman, L. M., Winkielman, P., and Ramachandran, V. S. (2009). Slow echo: facial EMG evidence for the delay of spontaneous, but not voluntary, emotional mimicry in children with autism spectrum disorders. *Dev. Sci.* 12, 510–520. doi: 10.1111/j.1467-7687.2008.00796.x

Ogrodniczuk, J. S., Piper, W. E., and Joyce, A. S. (2004). Alexithymia as a predictor of residual symptoms in depressed patients who respond to short-term psychotherapy. *Am. J. Psychother.* 58, 150–161. doi: 10.1176/appi.psychotherapy.2004.58.2.150

Palermo, R., and Coltheart, M. (2004). Photographs of facial expression: accuracy, response times, and ratings of intensity. *Behav. Res. Methods Inst. Comput.* 36, 634–638. doi: 10.3758/bf03206544

Petrocchi, N., Ottaviani, C., and Couyoumdjian, A. (2017). Compassion at the mirror: exposure to a mirror increases the efficacy of a self-compassion manipulation in enhancing soothing positive affect and heart rate variability. *J. Posit. Psychol.* 12, 525–536. doi: 10.1080/17439760.2016.1209544

Porciello, G., Holmes, B. S., Liuzza, M. T., Crostella, F., Aglioti, S. M., and Bufalari, I. (2014). Interpersonal Multisensory Stimulation reduces the overwhelming distracting power of self-gaze: psychophysical evidence for 'engazement'. *Sci. Rep.* 4:6669.

Preston, S. D., and De Waal, F. B. (2002). Empathy: Its ultimate and proximate bases. *Behav. Brain Sci.* 25, 1–20. doi: 10.1017/s0140525x02000018

Proverbio, A. M. (2017). Sex differences in social cognition: the case of face processing. *J. Neurosci. Res.* 95, 222–234. doi: 10.1002/jnr.23817

R Development Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rinn, W. E. (1984). The neuropsychology of facial expression: a review of the neurological and psychological mechanisms for producing facial expressions. *Psychol. Bull.* 95, 52–77.

Ross, M., and Sicoly, F. (1979). Egocentric biases in availability and attribution. *J. Pers. Soc. Psychol.* 37, 322–336. doi: 10.1037/0022-3514.37.3.322

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

Sato, W., Fujimura, T., and Suzuki, N. (2008). Enhanced facial EMG activity in response to dynamic facial expressions. *Int. J. Psychophysiol.* 70, 70–74. doi: 10.1016/j.ijpsycho.2008.06.001

Scarpazza, C., Làdavas, E., and Cattaneo, L. (2018). Invisible side of emotions: somato-motor responses to affective facial displays in alexithymia. *Exp. Brain Res.* 236, 195–206. doi: 10.1007/s00221-017-5118-x

Schaefer, A., Nils, F., Sanchez, X., and Philippot, P. (2010). Assessing the effectiveness of a large database of emotion-eliciting films: a new tool for emotion researchers. *Cogn. Emot.* 24, 1153–1172. doi: 10.1080/02699930903274322

Schmidt, U., Tiller, J., and Treasure, J. (1993). Psychosocial factors in the origins of bulimia nervosa. *Int. Rev. Psychiatry* 5, 51–60. doi: 10.3109/09540269309028294

Schwartz, G. E., Brown, S. L., and Ahern, G. L. (1980). Facial muscle patterning and subjective experience during affective imagery: sex differences. *Psychophysiology* 17, 75–82. doi: 10.1111/j.1469-8986.1980.tb02463.x

Serino, A., Pizzoferrato, F., and Ladavas, E. (2008). Viewing a face (especially one's own face) being touched enhances tactile perception on the face. *Psychol. Sci.* 19, 434–438. doi: 10.1111/j.1467-9280.2008.02105.x

Sifneos, P. E. (1973). The prevalence of 'alexithymic'characteristics in psychosomatic patients. *Psychother. Psychosom.* 22, 255–262. doi: 10.1159/000286529

Silvia, P. J., Allan, W. D., Beauchamp, D. L., Maschauer, E. L., and Workman, J. O. (2006). Biased recognition of happy facial expressions in social anxiety. *J. Soc. Clin. Psychol.* 25, 585–602. doi: 10.1521/jscp.2006.25.6.585

Speranza, M., Loas, G., Guilbaud, O., and Corcos, M. (2011). Are treatment options related to alexithymia in eating disorders? Results from a three-year naturalistic longitudinal study. *Biomed. Pharmacother.* 65, 585–589. doi: 10.1016/j.biopha.2010.01.009

Speranza, M., Loas, G., Wallier, J., and Corcos, M. (2007). Predictive value of alexithymia in patients with eating disorders: a 3-year prospective study. *J. Psychosom. Res.* 63, 365–371. doi: 10.1016/j.jpsychores.2007.03.008

Stel, M., and Van Knippenberg, A. (2008). The role of facial mimicry in the recognition of affect. *Psychol. Sci.* 19, 984–985. doi: 10.1111/j.1467-9280.2008.02188.x

Tan, J. W., Walter, S., Scheck, A., Hrabal, D., Hoffmann, H., Kessler, H., et al. (2012). Repeatability of facial electromyography (EMG) activity over corrugator supercilii and zygomaticus major on differentiating various emotions. *J. Ambient Intell. Humaniz. Comput.* 3, 3–10. doi: 10.1007/s12652-011-0084-9

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

Tse, W. S., and Bond, A. J. (2004). The impact of depression on social skills: a review. *J. Nerv. Ment. Dis.* 192, 260–268. doi: 10.1097/01.nmd.0000120884.60002.2b

Uddin, L. Q., Iacoboni, M., Lange, C., and Keenan, J. P. (2007). The self and social cognition: the role of cortical midline structures and mirror neurons. *Trends Cogn. Sci.* 11, 153–157. doi: 10.1016/j.tics.2007.01.001

Uddin, L. Q., Kaplan, J. T., Molnar-Szakacs, I., Zaidel, E., and Iacoboni, M. (2005). Self-face recognition activates a frontoparietal "mirror" network in the right hemisphere: an event-related fMRI study. *Neuroimage* 25, 926–935. doi: 10.1016/j.neuroimage.2004.12.018

van Boxtel, A. (2001). Optimal signal bandwidth for the recording of surface EMG activity of facial, jaw, oral, and neck muscles. *Psychophysiology* 38, 22–34. doi: 10.1111/1469-8986.3810022

Van Boxtel, A. (2010). "Facial EMG as a tool for inferring affective states," in *Proceedings of the Measuring Behavior*, (Wageningen: Noldus Information Technology), 104–108.

Van Der Schalk, J., Fischer, A., Doosje, B., Wigboldus, D., Hawk, S., Rotteveel, M., et al. (2011). Convergent and divergent responses to emotional displays of ingroup and outgroup. *Emotion* 11, 286–298. doi: 10.1037/a0022582

van Kuilenburg, H., Wiering, M., and Den Uyl, M. (2005). "A model based method for automatic facial expression recognition," in *Proceedings of the European Conference on Machine Learning*, (Berlin: Springer), 194–205. doi: 10.1007/11564096_22

Vergallito, A., Petilli, M. A., Cattaneo, L., and Marelli, M. (2019). Somatic and visceral effects of word valence, arousal and concreteness in a continuum lexical space. *Sci. Rep.* 9:20254.

Vergallito, A., Riva, P., Pisoni, A., and Lauro, L. J. R. (2018). Modulation of negative emotions through anodal tDCS over the right ventrolateral prefrontal cortex. *Neuropsychologia* 119, 128–135. doi: 10.1016/j.neuropsychologia.2018.07.037

Vinai, P., Speciale, M., and Alibrandi, M. (2016). L'integrazione della Self Mirroring Therapy nel protocollo CBT per il trattamento del disturbo ossessivo compulsivo: un caso clinico. *Quad. Psicoter. Cogn.* 38, 107–116. doi: 10.3280/qpc2016-038008

Vinai, P., Speciale, M., Vinai, L., Vinai, P., Bruno, C., Ambrosecchia, M., et al. (2015). The clinical implications and neurophysiological background of useing self-mirroring technique to enhance the identification of emotional experiences: an example with rational emotive behavior therapy. *J. Rational Emot. Cogn. Behav. Ther.* 33, 115–133. doi: 10.1007/s10942-015-0205-z

Weyers, P., Mühlberger, A., Kund, A., Hess, U., and Pauli, P. (2009). Modulation of facial reactions to avatar emotional faces by nonconscious competition priming. *Psychophysiology* 46, 328–335. doi: 10.1111/j.1469-8986.2008.00771.x

Wood, A., Rychlowska, M., Korb, S., and Niedenthal, P. (2016). Fashioning the face: sensorimotor simulation contributes to facial expression recognition. *Trends Cogn. Sci.* 20, 227–240. doi: 10.1016/j.tics.2015.12.010

Zeitlin, S. B., and McNally, R. J. (1993). Alexithymia and anxiety sensitivity in panic disorder and obsessive-compulsive disorder. *Am. J. Psychiatry* 150, 658–660. doi: 10.1176/ajp.150.4.658

Conflict of Interest: 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.

frontiers in Psychology

ORIGINAL RESEARCH
published: 03 April 2020
doi: 10.3389/fpsyg.2020.00442

Auditory Contagious Yawning Is Highest Between Friends and Family Members: Support to the Emotional Bias Hypothesis

Ivan Norscia[1]*, Anna Zanoli[1], Marco Gamba[1] and Elisabetta Palagi[2]

[1] Department of Life Sciences and Systems Biology, School of Nature Sciences, University of Turin, Turin, Italy, [2] Unit of Ethology, Department of Biology, University of Pisa, Pisa, Italy

Contagious yawning differs from spontaneous yawning because it occurs when an individual yawns in response to someone else's yawn. In *Homo sapiens* and some non-human primates contagious yawning is higher between strongly than weakly bonded individuals. Up to date, it is still unclear whether this social asymmetry underlies emotional contagion (a basic form of empathy preferentially involving familiar individuals) as predicted by the Emotional Bias Hypothesis (EBH) or is linked to a top-down, selective visual attention bias (with selective attention being preferentially directed toward familiar faces) as predicted by the Attentional Bias Hypothesis (ABH). To verify whether the visual attentional bias explained the yawn contagion bias or not, in this study, we considered only yawns that could be heard but not seen by potential responders (auditory yawns). Around 294 of auditory yawning occurrences were extrapolated from over 2000 yawning bouts collected in free ranging humans for over nine years. Via GLMM, we tested the effect of intrinsic features (i.e., gender and age) and social bond (from strangers to family members) on yawn. The individual identity of the subjects (trigger and potential responder) was included as random factor. The social bond significantly predicted the occurrence of auditory yawn contagion, which was highest between friends and family members. A gender bias was also observed, with women responding most frequently to others' yawns and men being responded to most frequently by others. These results confirm that social bond is per se one of the main drivers of the differences in yawn contagion rates between individuals in support of the EBH of yawn contagion.

Keywords: emotional contagion, bottom-up attention, selective attention, top-down attention, yawn contagion, mimicry

Edited by:
Noemi Mazzoni,
University of Trento, Italy

Reviewed by:
Barbara Schmidt,
Friedrich Schiller University Jena, Germany
Teresa Romero,
University of Lincoln, United Kingdom

****Correspondence:***
Ivan Norscia
ivan.norscia@unito.it

Specialty section:
This article was submitted to Consciousness Research, a section of the journal Frontiers in Psychology

Received: *10 November 2019*
Accepted: *25 February 2020*
Published: *03 April 2020*

Citation:
Norscia I, Zanoli A, Gamba M and Palagi E (2020) Auditory Contagious Yawning Is Highest Between Friends and Family Members: Support to the Emotional Bias Hypothesis. Front. Psychol. 11:442. doi: 10.3389/fpsyg.2020.00442

INTRODUCTION

Yawning is an involuntary sequence of mouth opening, deep inspiration, brief apnea, and more or less slow expiration (Baenninger, 1997; Walusinski and Deputte, 2004; Guggisberg et al., 2010; Krestel et al., 2018). When elicited, a yawn cannot be totally suppressed. Therefore, it has been defined as a stereotyped or reflex-like pattern (Lehmann, 1979; Provine, 1986). In *Homo sapiens*, several hypotheses have been put forth with variable support to explain mechanisms and functions of spontaneous yawning, such as oxygenation (respiratory function caused by hypoxia), stress-related behavior (caused by arousal), or thermoregulation (caused by hyperthermia; Guggisberg et al., 2010; Massen et al., 2014; Gallup and Gallup, 2019). Being a physiological response, yawning can be affected by internal and external factors such as the time of the day (Giganti and Zilli, 2011) or intracranial/brain temperature (Gallup and Eldakar, 2013).

Yawning can be self-directed and/or displayed to others (Moyaho et al., 2017; Palagi et al., 2019). In human and non-human primates, depending on the species, when yawning is shown to others, it can communicate threat (Troisi et al., 1990; Deputte, 1994) and/or physiological and behavioral changes (Provine et al., 1987; Leone et al., 2015; Zannella et al., 2015). In humans, yawning is a socially modulated response because it can be inhibited by actual—and not virtual—social presence (Gallup et al., 2019) and because a yawn can be triggered by someone else's yawn, as a result of a phenomenon known as contagious yawning (Provine, 1989, 2005). Yawn contagion can be elicited even if the yawn is heard but not seen (Arnott et al., 2009; Massen et al., 2015).

In humans, their phylogenetically closest ape species (chimpanzees: *Pan paniscus*; bonobos: *Pan troglodytes*) and the African monkey *Theropithecus gelada*, contagious yawning is not only present (Provine, 1986; Palagi et al., 2009; Tan et al., 2017; but see: Amici et al., 2014) but also socially modulated because the yawning response is highest when certain categories of individuals are involved (e.g., kin, group-members, dominants; Palagi et al., 2009; Campbell and de Waal, 2011, 2014; Norscia and Palagi, 2011; Demuru and Palagi, 2012; Massen et al., 2012). Two main arguments have been presented to explain this social asymmetry in contagious yawning, which have been grouped into two main hypotheses: the *Emotional Bias Hypothesis* (EBH), linking contagious yawning to emotional transfer, and the *Attentional Bias Hypothesis* (ABH), which considers contagious yawning as a motor response that is subject to differences in top-down attentional processes (Palagi et al., 2020).

The EBH predicts that the social asymmetry observed in yawn contagion rates reflects differences in the different social bonding, a proxy of emotional bonding, between individuals. This hypothesis is supported by evidence that yawn contagion rates follow an empathic trend (*sensu* Preston and de Waal, 2002), being highest between individuals sharing a strongest emotional bond. Specifically, Norscia and Palagi (2011) found that in humans yawn contagion rates are greatest in response to kin and friends than in response to acquaintances and strangers. In adult chimpanzees yawn contagion is higher between in-group compared to out-group members (Campbell and de Waal, 2011) and in bonobos yawn contagion rates are greatest between individuals that affiliate more with one another (Demuru and Palagi, 2012). In a comparative investigation including both humans and bonobos, Palagi et al. (2014) found that the yawn contagion rates were affected by the relationship quality between individuals more than by the species the subjects belonged to. Additionally, in humans yawn contagion increases with age when the ability to identify others' emotions increases and declines with old age when such ability declines (Wiggers and van Lieshout, 1985; Anderson and Meno, 2003; Saxe et al., 2004; Singer, 2006; Millen and Anderson, 2011; Bartholomew and Cirulli, 2014). Yawn contagion rates increase from infancy to adulthood also in chimpanzees (Madsen and Persson, 2013).

The ABH predicts that the social asymmetry observed in yawn contagion can be due to differences in social, visual attention (Massen and Gallup, 2017). In particular, highest levels of contagious yawning would be due to the extra top-down, selective visual attention paid to individuals that are more relevant to the observer, such as familiar subjects, as it occurs in humans and geladas, or dominants, as it occurs in chimpanzees or bonobos (Yoon and Tennie, 2010; Massen et al., 2012; Massen and Gallup, 2017). According to Massen and Gallup (2017), ABH would be backed up by the existing evidence on the different visual detection and visual perceptive encoding of faces of familiar and/or in-group subjects compared to unfamiliar ones (e.g., Buttle and Raymond, 2003; Ganel and Goshen-Gottstein, 2004; Jackson and Raymond, 2006; Michel et al., 2006).

In this study, we analyzed data on yawning collected over 9 years on humans in their natural settings and we extrapolated the cases in which the yawn emitted by a subject could be heard but not seen by a potential responder (auditory yawn). By considering only the cases in which the visual cue of the yawning stimulus was not detectable, we verified whether the social asymmetry previously observed in yawn contagion rates persisted or not. In particular, we tested the following alternative predictions derived from the two hypotheses presented above (EBH and ABH).

Prediction 1a: according to the EBH, the rates of yawn contagion are influenced by the strength of the inter-individual social bond—a proxy of the emotional bond—*per se* and not by a different top-down, selective visual attention paid to certain individuals in particular. If this hypothesis is supported, we expect to observe the social bias also when the visual cue of the yawning stimulus is excluded and the rates of auditory contagious yawning to be higher between strongly bonded compared to weakly bonded individuals.

Prediction 1b: according to the ABH, the higher levels of yawn contagion between strongly bonded compared to weakly bonded individuals would be linked to the closest top-down, selective visual attention that individuals pay to individuals that are relevant to them, e.g., family and friends. If this hypothesis is supported, the social bias observed in the yawn contagion should disappear when only auditory yawns are considered because the visual cue cannot be attended by the potential responder.

MATERIALS AND METHODS

Data Collection and Operational Definitions

For this study, we considered the vocalized yawns emitted by a subject that could only be heard—but not seen—by a potential responder (hereafter: auditory yawns). The emitter and the potential responders had to be in a range of ≤ 5 m. Vocalized yawns involved the use of vocal folds and the yawns that only involved heavy inspiration/expiration were not considered as vocalized; 294 cases of auditory yawns were extrapolated from a dataset of a total of 2001 yawning bouts collected over 9 years—from 2010 to 2019—by using the all occurrences sampling method (Altmann, 1974). Specifically, auditory yawns were collected from November 2010 to May 2019, from 05.30 am to 02.30 am, on human Caucasian subjects, aged from 18 to 77, during their routinely activities, e.g., in work places, over meals, during social meetings, etc., with the subjects being unaware of being observed and in absence of any evident external source of anxiety. The auditory yawn database included 193 yawner-potential responder dyads. Depending on the situation, the information was recorded, unnoted, through alphanumerical codes and entered directly into calculation sheets, typed in mobile phones or written on paper, and then entered in calculation sheets for subsequent elaboration. Basic information such as age and the relationship between people was known to the authors. The potential responders were coded as in the non-sight condition when their head was rotated by 180° with respect to the trigger or when a physical, sight-blocking obstacle was present preventing the potential responder from seeing the trigger's face and body. Trigger and responder were never completely isolated (e.g., in two separate rooms with closed doors) from one another. The social closeness was collected on four levels: 0 = strangers, who had never met before; 1 = acquaintances, who exclusively shared an indirect relationship based on a third external element, that is work duty (colleagues) or friends in common (friends of friends); 2 = friends, non-related individuals sharing a direct relationship not exclusively related to a third external element; 3 = regular partners and kin ($r \geq 0.25$). Previous literature reports that yawn responses can be elicited within 5 min after watching someone else's yawn (the trigger's yawn) (Provine, 1986), with a maximum in the first minute (Provine, 2005; Palagi et al., 2014). Literature also reports that from the fourth minute there is a highest probability of autocorrelation (meaning that the presence of a yawn performed by a subject at t_0 increases the probability to have another yawn by the same subject at $t_{(0+X)}$ where X is the increasing unit of time; Kapitány and Nielsen, 2017). Hence, we considered the yawn responses occurring within a 3 min time window from the yawn emitted by the trigger. To further reduce the autocorrelation bias, in case of a chain of yawns emitted by the trigger (more yawns emitted in the 3-min time window) we registered as a response only the first yawn performed after the perception of the last yawn. We coded a yawn as "spontaneous" when no other subject had yawned in the 5 min preceding the yawning event.

Statistical Analyses

For the analyses, the following variables were considered: occurrence of contagion, coded as: 1 = presence, 0 = absence; the social bond was entered with the four levels defined above (0 = strangers; 1 = acquaintances; 2 = friends; 3 = kin); trigger's and observer's sex were labeled as: M = male, F = female; the age classes of the trigger and the responder were coded as follow: yo = youth (18–24 years old); ad = adult (25–64 years old); se = senior (above 65 years old) (Statistics Canada, 2009); the time slots were coded as follows: 1 = 05:30–09:00 am; 2 = 09:01 am–12:30 pm; 3 = 12:31–16:00 pm; 4 = 16:01–19:30 pm; 5 = 19:31–23:00 pm; 6 = 23:01–02:30 (Giganti and Zilli, 2011). The database (see **Supplementary Data Sheet**) included 84 males, 69 females, 16 youngsters (yo), 122 adults (ad), and 15 senior (se). To test whether the occurrence of yawn contagion was influenced by the factors bond (0 = strangers; 1 = acquaintances; 2 = friends; 3 = kin), sex of the trigger (Trigger_sex), sex of the responder (Responder_sex), the age class of the trigger (Trigger_ageclass), and the age class of the responder (Responder_ageclass), and time slot (from 1 to 6), we used a generalized linear mixed model (GLMM) that included these five predictors as fixed effects and triggers (Trigger) and responders (Responder) identities as random effects. We fitted the models in R (R Core Team, 2018; version 3.5.1) using the function lmer of the R-package lme4 (Bates et al., 2015). We established the significance of the full model by comparison to a null model comprising only the random effects (Forstmeier and Schielzeth, 2011). We used a likelihood ratio test (Dobson, 2002) to test this significance (ANOVA with argument "Chisq"). We calculated the *p*-values for the individual predictors based on likelihood ratio tests between the full and the null model by using the R-function "drop1" (Barr et al., 2013). As the response variable was binary, we used a binomial error distribution. We tested whether the interaction between the sexes or the age classes of the trigger and the responder were significant, but as they were not, we did not include them in the model. We used a multiple contrast package (multcomp) to perform all pairwise comparisons for each bonding levels with the Tukey test (Bretz et al., 2010). We reported the Bonferroni-adjusted *p*-values, estimate (Est), standard error (SE), and *z*-values.

RESULTS

We compared the model fitted versus a null model comprising only the random factors (likelihood ratio test: $\chi^2 = 149.995$, $df = 17$, $p < 0.001$). As we found at least one predictor was having a significant impact on the response, we moved on with a drop1 procedure. The GLMM indicated a significant effect of bond across four comparisons (Tukey test; 2 = friends versus 0 = strangers, Est = 5.4810; SE = 0.9807, $z = 5.589$, $p < 0.001$; 3 = kin versus 0 = strangers, Est = 6.6872, SE = 1.1602, $z = 5.764$, $p < 0.001$; 2 = friends versus 1 = acquaintances, Est = 3.7643, SE = 0.7201, $z = 5.227$, $p < 0.001$; kin = 3 versus 1 = acquaintances, Est = 4.9706, SE = 0.9381, $z = 5.299$, $p < 0.001$) (**Figure 1**). The pairwise comparisons of bonding levels revealed that yawn contagion was significantly higher in family and friends than in strangers and acquaintances, with no significant differences between strangers and acquaintances and between family and friends (1 = acquaintances versus 0 = strangers; Est = 1.7167, SE = 0.7062, $z = 2.431$, $p = 0.0664$; 3 = family versus 2 = friends; Est = 1.2063, SE = 0.8960, $z = 1.346$, $p = 0.5156$). The GLMM also indicated a significant effect of the sex of both triggers and responders, and of bonding (see **Table 1**): yawn contagion of female responders was higher compared to males (**Figure 3**), and males, as triggers, were responded to more frequently by

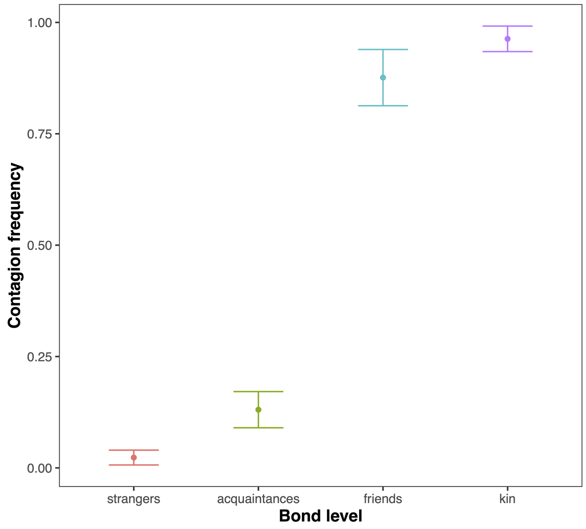


FIGURE 1 | Line plot of the effect of the social bond between trigger and responder (X-axis) on the mean occurrence of acoustic yawn contagion (Y-axis). Friends and kin show significantly higher yawn contagion frequencies than strangers and acquaintances (Tukey test: friends versus strangers $p < 0.001$; kin versus strangers $p < 0.001$; friends versus acquaintances $p < 0.001$; kin versus acquaintances $p < 0.001$; other combinations, ns).

TABLE 1 | Results of the GLMM, including the following fixed factors: bond (0 = strangers; 1 = acquaintances; 2 = friends; 3 = kin), trigger and responder sex (M = male; F = female), trigger and responder age class (yo = youth, 15–24 years old; ad = adult, 25–64 years old; se = senior, above 65 years old), and time slot (1 = 05:30–09:00 am; 2 = 09:01 am–12:30 pm; 3 = 12:31–16:00 pm; 4 = 16:01–19:30 pm; 5 = 19:31–23:00 pm; 6 = 23:01 pm–02:30 am).

	Estimate	SE	χ^2	*P*
(Intercept)[a]	−2.404	0.956	a	a
Bond (acquaintances)[b,c]	1.844	0.722	2.554	0.000
Trigger sex (male)[b,c]	0.919	0.451	2.036	0.037
Responder sex (male)[b,c]	−1.207	0.512	−2.358	0.012
Trigger age class (senior)[b,c]	0.255	0.949	0.268	0.930
Responder age class (senior)[b,c]	−0.934	1.045	−0.893	0.722
Time slot (09:01–12:30)[b,c]	−1.011	0.929	−1.089	0.326

The identity of triggers (Trigger) and responders (Responder) was included as random factors. Full versus null model: chisq = 149.995, df = 17, p < 0.001. [a]*Not shown as not having a meaningful interpretation.* [b]*Estimate ± SE refer to the difference of the response between the reported level of this categorical predictor and the reference category of the same predictor.* [c]*These predictors were dummy coded, with the "Bond (Strangers)," "Trigger sex (Female)," "Responder sex (Female)," "Trigger age class (Adult)," "Responder age class (Adult)," and "Time slot (05:30–09:00)" being the reference categories.*

others compared to females (**Figure 2**). In contrast, we found no significant main effects of the age of both triggers and responders, the time slot in which yawns were emitted, and interaction between the sex of the subjects (**Table 1**).

DISCUSSION

This study shows for the first time that yawn contagion is significantly affected by the social bond between individuals (**Table 1**) even when the triggering stimuli are auditory yawns, which we defined as vocalized yawns that could be heard but not seen (visual cue undetectable, auditory cue detectable). In particular, auditory contagious yawning is significantly more frequent between kin and friends than between strangers and acquaintances (**Figure 1**). This finding supports prediction 1a based on the EBH and not prediction 1b based on the ABH, leading to the conclusion that in humans top-down, selective visual attention cannot be the main driver of the social asymmetry observed in yawn contagion rates (Norscia and Palagi, 2011; Norscia et al., 2016a). Also the sex of the trigger and the sex of the receiver had a significant effect on yawn contagion rates, with men—as triggers—being responded to by others more frequently than women (**Figure 2**) and women responding more frequently to others' yawns than men (**Figure 3**).

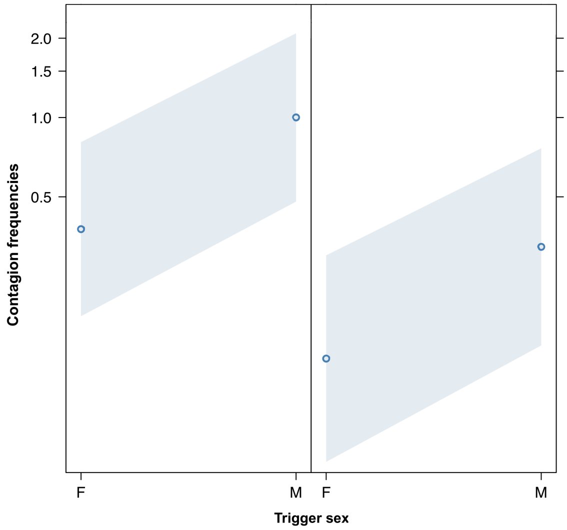


FIGURE 2 | Line plot of the effect of the sex of the trigger (X-axis) on acoustic yawn contagion mean occurrence (Y-axis). Effect of the trigger sex on acoustic yawn contagion when the responder is a female (**right**) and a male (**left**). Males' yawns elicit more yawns than females' ones regardless of the sex of the responder (result of the GLMM, $p = 0.022$). Points represent the effect that the response variable has on the independent variable "sex of the trigger," based on the value predicted by the model. Colored bands show the 95% confidence interval.

Contrary to Bartholomew and Cirulli (2014), we found no age effect on yawn contagion, most probably because our database on auditory yawns had a strong prevalence of adults (25–64 years old). The highest levels of auditory yawn contagion in women compared to men confirm the gender bias observed in naturalistic conditions on humans susceptible to yawn contagion by Norscia et al. (2016a,b) when considering a larger dataset that also included yawns that could be seen by the potential responder (with yawning sensory modality—vision, hearing, or both—not affecting the response). The gender bias is also in partial agreement with previous results obtained in controlled settings, including the visual cue (Chan and Tseng, 2017; but see Norscia and Palagi, 2011; Bartholomew and Cirulli, 2014). It has been hypothesized that the high degree of yawn contagion in

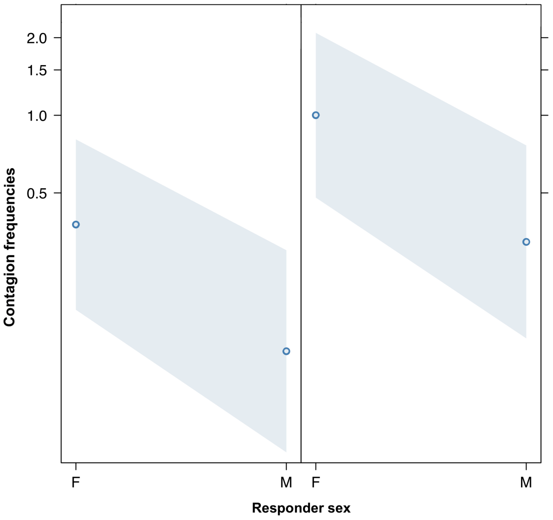


FIGURE 3 | Line plot of the effect of the sex of the responder (*X*-axis) on acoustic yawn contagion mean occurrence (*Y*-axis). Effect of the responder sex on acoustic yawn contagion when the trigger is a female (**right**) or a male (**left**). Females respond significantly more than males regardless of the sex of the trigger (result of the GLMM, $p = 0.021$). Points represent the effect that the response variable has on the independent variable "sex of the responder," based on the value predicted by the model. Colored bands show the 95% confidence interval.

women might inform emotional contagion (Norscia et al., 2016a), in the light of reportedly higher empathic abilities—related to maternity—of women compared to men (Christov-Moore et al., 2014). However, this issue is still under debate because cultural differences across human societies can mold social bonding dynamics in a different way. It is therefore complicate, at this stage of knowledge, to disentangle cultural factors, inter-personal relationship quality, and gender influence in the distribution of yawn contagion. Our results also show that men perform better than women as triggers and the most parsimonious hypothesis for this might be that men's vocalizations can be better heard in natural settings, often characterized by background noises. Indeed, the perception of voice gender primarily relies on the fundamental frequency that is on average lower by an octave in male than female voices, with lower frequency vocalizations traveling further than high frequency ones (Marten and Marler, 1977; Latinus and Taylor, 2012). However, to our knowledge, there is no specific study addressing the possible gender bias in yawn audibility and further investigation with experimental trials in controlled condition is therefore necessary to verify this speculation.

In this study, we also found that the differences in yawn contagion rates across categories (family and friends, strangers, and acquaintances) cannot be explained by differences in top-down, selective visual attention. This finding is in line with previous literature. Contagious yawning appears to involve brain areas that are more related to the orienting-bottom up network [temporoparietal junction (TPJ), brainstem nuclei, ventrolateral prefrontal cortex (vlPC)] than top-down related areas [frontal eye fields (FEFs), intraparietal sulcus (IPS), parietal areas; for a review: Palagi et al., 2020]. Moreover, yawn contagion is neither sensitive to the sensory cues present in the signal (auditory, visual, or audio-visual) (Arnott et al., 2009; Norscia and Palagi, 2011) nor affected by the visual perspective of the triggering stimulus (yawns in orientations of 90°, 180°, and 270° are able to trigger yawning responses as frontal, 0° yawns; Provine, 1989, 1996). Chan and Tseng (2017) found that the ability to detect a yawn as such (perceptual detection sensitivity) was related to the duration of gaze to the eyes of the stimulus releasing face, but eye-gaze patterns were not able modulate contagious yawning. In chimpanzees, contagious yawning frequencies were highest between same-group than different-group individuals, even if the responders looked longer at out-group chimpanzee videos (Campbell and de Waal, 2011). The argument that visual selective attention can bias yawn contagion rates in a specific direction (subjects responding more to family than strangers) is also undermined by the absence of any specific pattern of social attention in human and non-human primate. Via eye-tracking (applied to measure the viewing time) and by showing unknown faces to their experimental subjects, Méary et al. (2014) observed that humans were skewed toward own-race faces whereas rhesus macaques' attention was more attracted by new than by same species faces. Kawakami et al. (2014) observed that human subjects paid more attention to the eyes of ethnic in-group members and to nose and mouth of ethnic out-group members. The same study also revealed that visual attention did not depend on the target race. By measuring how long the experimental subjects gazed at the screen, Whitehouse et al. (2016) observed that Barbary macaques paid more attention to scratching videos of non-stranger than stranger individuals but also noted that within the non-strangers, macaques paid most attention to those individuals with which they shared a weak social bonding. By measuring glance rates, Schino and Sciarretta (2016) observed that mandrills looked more at their own kin than at non-kin but also more at dominant than at subordinate group mates. Therefore, these studies (used to support ABH) describe no single pattern of selective attention. One further important point to consider is the very definition of familiarity and group-membership adopted by most of the studies used to support ABH (Massen and Gallup, 2017). These studies showed better visual detection and visual perceptive encoding of faces of familiar/in-group subjects compared to unfamiliar ones but defined familiarity and group membership *not* on the basis of the personal relationships between individuals. Instead, familiarity or group-membership were defined on the ground of indirect knowledge (e.g., photo of famous people or of a subject already shown in pre-trial phases) or common race (e.g., Buttle and Raymond, 2003; Ganel and Goshen-Gottstein, 2004; Jackson and Raymond, 2006; Michel et al., 2006). This definition is fine for the purposes of these studies but it is not as much fine if the results are used to propose alternative explanations for the influence that real social bonding—based on real relationships—may have on a phenomenon, in this case yawn contagion. For example, Michel et al. (2006) observed that Caucasian and Asian subjects could better recognize same-race faces but this difference was not present in Asian subjects who had been living for about a year among Caucasians. Another point of discussion concerns the presence of yawn contagion in children with autism spectrum disorder (ASD), which frequently show alterations in visual attention (Richard and Lajiness-O'Neill, 2015). In ASD children, yawn contagion can be absent (Senju et al., 2007), impaired (Helt et al., 2010), or similar to typically developing children when the subjects are induced to redirect their attention the video stimulus during the experimental trials (Usui et al., 2013). In a recent study, Mariscal et al. (2019) found that yawn contagion in ASD children was positively related to the blood concentration of oxytocin, the hormone involved in parental and social attachment (Decety et al., 2016) and posited that yawn contagion in ASD children may be related to variable mean oxytocin concentrations across different study cohorts (Mariscal et al., 2019). This finding is in line with the EBH hypothesis that links yawn contagion rates to social bonding, which can reflect emotional bonding.

CONCLUSION

Our study adds to the discussion over the mechanisms underlying the social asymmetry in yawn contagion (for a critical reviews: see Adriaense et al., 2020; Palagi et al., 2020), by showing that yawn contagion is probably associated with bottom-up, rather than with top-down, selective attention. Bottom-up attention is primarily lead by the sensory perception of the eliciting stimulus whereas top-down, selective attention is a voluntary, sustained process in which a particular item is selected internally and focused upon or examined (Katsuki and Constantinidis, 2014). In this respect, the acoustic stimulus (auditory yawn) emitted by the trigger was heard and could elicit a yawning response in the receiver, even though the receiver was not paying any voluntary visual attention to the trigger. Moreover, the yawning response rates were socially modulated, with auditory yawn contagion being highest in individuals that were most strongly bonded to one another. Hence, top-down selective attention is not the main driver of the social asymmetry observed in yawn contagion, which appears to be a stimulus driven phenomenon-related to bottom-up attention processes. Further investigation is necessary to understand whether and in what way other forms of attention or pre-attentive stages are able to affect yawn contagion.

DATA AVAILABILITY STATEMENT

The dataset used for this study is provided in the **Supplementary Material**.

ETHICS STATEMENT

The study involving human participants were reviewed and approved by Comitato di Bioetica d'Ateneo (University Bioethical Committee)—University of Turin (ref. no. 451945). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

IN and EP carried out data collection and conceived and wrote the manuscript. AZ helped with data collection and manuscript revision, figures, and tables. MG carried out statistical analyses and wrote the related part of the manuscript.

ACKNOWLEDGMENTS

The authors wish to thank Prof. Roberto Barbuti and Prof. Cristina Giacoma for having supported the research at the University of Pisa and the University of Turin, respectively.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.00442/full#supplementary-material

REFERENCES

Adriaense, J. E. C., Koski, S. E., Huber, L., and Lamm, C. (2020). Challenges in the comparative study of empathy and related phenomena in animals. *Neurosci. Biobehav. Rev.* 112, 62–82. doi: 10.1016/j.neubiorev.2020.01.021

Altmann, J. (1974). Observational study of behavior: sampling methods. *Behaviour* 49, 227–266. doi: 10.1163/156853974x00534

Amici, F., Aureli, F., and Call, J. (2014). Response facilitation in the four great apes: is there a role for empathy? *Primates* 55, 113–118. doi: 10.1007/s10329-013-0375-1

Anderson, J. R., and Meno, P. (2003). Psychological influences on yawning in children. *Curr. Psychol. Lett.* 11, 1–7.

Arnott, S. R., Singhal, A., and Goodale, M. A. (2009). An investigation of auditory contagious yawning. *Cogn. Affect. Behav. Neurosci.* 9, 335–342. doi: 10.3758/CABN.9.3.335

Baenninger, R. (1997). On yawning and its functions. *Psychon. Bull. Rev.* 4, 198–207. doi: 10.3758/bf03209394

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang.* 68, 255–278. doi: 10.1016/j.jml.2012.11.001

Bartholomew, A. J., and Cirulli, E. T. (2014). Individual variation in contagious yawning susceptibility is highly stable and largely unexplained by empathy or other known factors. *PLoS One* 9:e91773. doi: 10.1371/journal.pone.0091773

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., et al. (2015). *Package lme4*. Available online at: http://lme4.r-forge.r-project.org/

Bretz, F., Hothorn, T., and Westfall, P. (2010). *Multiple Comparisons Using R*. Boca Raton, FL: Chapman & Hall.

Buttle, H., and Raymond, J. E. (2003). High familiarity enhances visual change detection for face stimuli. *Percept. Psychophys.* 65, 1296–1306. doi: 10.3758/BF03194853

Campbell, M. W., and de Waal, F. B. M. (2011). Ingroup-outgroup bias in contagious yawning by chimpanzees supports link to empathy. *PLoS One* 6:e18283. doi: 10.1371/journal.pone.0018283

Campbell, M. W., and de Waal, F. B. M. (2014). Chimpanzees empathize with group mates and humans, but not with baboons or unfamiliar chimpanzees. *Proc. R. Soc. B Biol. Sci.* 281:20140013. doi: 10.1098/rspb.2014.0013

Chan, M. H., and Tseng, C. H. (2017).). Yawning detection sensitivity and yawning contagion. *Iperception.* 8:2041669517726797. doi: 10.1177/2041669517726797

Christov-Moore, L., Simpson, E. A., Coudé, G., Grigaityte, K., Iacoboni, M., and Ferrari, P. F. (2014). Empathy: gender effects in brain and behavior. *Neurosci. Biobehav. Rev.* 46, 604–627. doi: 10.1016/j.neubiorev.2014.09.001

Decety, J., Bartal, I. B. A., Uzefovsky, F., and Knafo-Noam, A. (2016). Empathy as a driver of prosocial behaviour: highly conserved neurobehavioural mechanisms across species. *Philos. Trans. R. Soc. B Biol. Sci.* 371:20150077. doi: 10.1098/rstb.2015.0077

Demuru, E., and Palagi, E. (2012). In bonobos yawn contagion is higher among kin and friends. *PLoS One* 7:e49613. doi: 10.1371/journal.pone.0049613

Deputte, B. L. (1994). Ethological study of yawning in primates. I. Quantitative analysis and study of causation in two species of Old World monkeys (*Cercocebus albigena and Macaca fascicularis*). *Ethology* 98, 221–245. doi: 10.1111/j.1439-0310.1994.tb01073.x

Dobson, A. J. (2002). *An Introduction to Generalized Linear Models*, 2nd Edn. Boca Raton, FL: Chapman and Hall.

Forstmeier, W., and Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: overestimated effect sizes and the winner's curse. *Behav. Ecol. Sociobiol.* 65, 47–55. doi: 10.1007/s00265-010-1038-5

Gallup, A. C., and Eldakar, O. T. (2013). The thermoregulatory theory of yawning: what we know from over 5 years of research. *Front. Neurosci.* 6:188. doi: 10.3389/fnins.2012.00188

Gallup, A. C., and Gallup, G. G. Jr. (2019). Methylphenidate-induced yawning: a thermoregulatory connection. *Am. J. Therap.* [Epub ahead of print]. doi: 10.1097/MJT.0000000000001000

Gallup, A. C., Vasilyev, D., Anderson, N., and Kingstone, A. (2019). Contagious yawning in virtual reality is affected by actual, but not simulated, social presence. *Sci. Rep.* 9:294. doi: 10.1038/s41598-018-36570-2

Ganel, T., and Goshen-Gottstein, Y. (2004). Effects of familiarity on the perceptual integrality of the identity and expression of faces: the parallel-route hypothesis revisited. *J. Exp. Phsycol. Hum. Percept. Perform.* 30, 583–597. doi: 10.1037/0096-1523.30.3.583

Giganti, F., and Zilli, I. (2011). The daily time course of contagious and spontaneous yawning among humans. *J. Ethol.* 29, 215–219. doi: 10.1007/s10164-010-0242-0

Guggisberg, A. G., Mathis, J., Schnider, A., and Hess, C. W. (2010). Why do we yawn? *Neurosci. Biobehav. Rev.* 34, 1267–1276. doi: 10.1016/j.neubiorev.2010.03.008

Helt, M. S., Eigsti, I. M., Snyder, P. J., and Fein, D. A. (2010). Contagious yawning in autistic and typical development. *Child Dev.* 81, 1620–1631. doi: 10.1111/j.1467-8624.2010.01495.x

Jackson, M. C., and Raymond, J. E. (2006). The role of attention and familiarity in face identification. *Percept. Psychophys.* 68, 543–557. doi: 10.3758/BF03208757

Kapitány, R., and Nielsen, M. (2017). Are yawns really contagious? A critique and quantification of yawn contagion. *Adapt. Human. Behav. Physiol.* 3, 134–155. doi: 10.1007/s40750-017-0059-y

Katsuki, F., and Constantinidis, C. (2014). Bottom-up and top-down attention: different processes and overlapping neural systems. *Neuroscientist* 20, 509–521. doi: 10.1177/1073858413514136

Kawakami, K., Williams, A., Sidhu, D., Choma, B. L., Rodriguez-Bailón, R., Cañadas, E., et al. (2014). An eye for the I: preferential attention to the eyes of ingroup members. *J. Pers. Soc. Psychol.* 107, 1–20. doi: 10.1037/a0036838

Krestel, H., Bassetti, C. L., and Walusinski, O. (2018). Yawning - its anatomy, chemistry, role, and pathological considerations. *Prog. Neurobiol.* 161, 61–78. doi: 10.1016/j.pneurobio.2017.11.003

Latinus, M., and Taylor, M. J. (2012). Discriminating male and female voices: differentiating pitch and gender. *Brain Topogr.* 25, 194–204. doi: 10.1007/s10548-011-0207-9

Lehmann, H. E. (1979). Yawning: a homeostatic reflex and its psychological significance. *Bull. Meninger Clin.* 43, 123–136.

Leone, A., Ferrari, P., and Palagi, E. (2015). Different yawns, different functions? Testing social hypotheses on spontaneous yawning in *Theropithecus gelada. Sci. Rep.* 4:4010. doi: 10.1038/srep04010

Madsen, E. A., and Persson, T. (2013). Contagious yawning in domestic dog puppies (*Canis lupus* familiaris): the effect of ontogeny and emotional closeness on low-level imitation in dogs. *Anim. Cogn.* 16, 233–240. doi: 10.1007/s10071-012-0568-9

Mariscal, M. G., Oztan, O., Rose, S. M., Libove, R. A., Jackson, L. P., Sumiyoshi, R. D., et al. (2019). Blood oxytocin concentration positively predicts contagious yawning behavior in children with autism spectrum disorder. *Autism Res.* 12, 1156–1161. doi: 10.1002/aur.2135

Marten, K., and Marler, P. (1977). Sound transmission and its significance for animal vocalization. *Behav. Ecol. Sociobiol.* 2, 271–290. doi: 10.1007/bf00299740

Massen, J. J. M., Church, A. M., and Gallup, A. C. (2015). Auditory contagious yawning in humans: an investigation into affiliation and status effects. *Front. Psychol.* 6:1735. doi: 10.3389/fpsyg.2015.01735

Massen, J. J. M., Dusch, K., Eldakar, O. T., and Gallup, A. C. (2014). A thermal window for yawning in humans: yawning as a brain cooling mechanism. *Physiol. Behav.* 130, 145–148. doi: 10.1016/j.physbeh.2014.03.032

Massen, J. J. M., and Gallup, A. C. (2017). Why contagious yawning does not (yet) equate to empathy. *Neurosci. Bio. Rev.* 80, 573–585. doi: 10.1016/j.neubiorev.2017.07.006

Massen, J. J. M., Vermunt, D. A., and Sterck, E. H. M. (2012). Male yawning is more contagious than female yawning among chimpanzees (*Pan troglodytes*). *PLoS One* 7:e40697. doi: 10.1371/journal.pone.0040697

Méary, D., Li, Z., Li, W., Guo, K., and Pascalis, O. (2014). Seeing two faces together: preference formation in humans and rhesus macaques. *Anim. Cogn.* 17, 1107–1119. doi: 10.1007/s10071-014-0742-3

Michel, C., Caldara, R., and Rossion, B. (2006). Same-race faces are perceived more holistically than other-race faces. *Vis. Cogn.* 14, 55–73. doi: 10.1080/13506280500158761

Millen, A., and Anderson, J. R. (2011). Neither infants nor toddlers catch yawns from their mothers. *Biol. Lett.* 7, 440–442. doi: 10.1098/rsbl.2010.0966

Moyaho, A., Urbina, A. F., Guzmán, E. M., and Walusinski, O. (2017). Yawning: a cue and a signal. *Heliyon* 3:e00437. doi: 10.1016/j.heliyon.2017.e00437

Norscia, I., Demuru, E., and Palagi, E. (2016a). She more than he: gender bias supports the empathic nature of yawn contagion in *Homo sapiens*. *R. Soc. Open Sci.* 3:150459. doi: 10.1098/rsos.150459

Norscia, I., Demuru, E., and Palagi, E. (2016b). Difference in contagious yawning between susceptible men and women: why not? *R. Soc. Open Sci.* 3:160477. doi: 10.1098/rsos.160477

Norscia, I., and Palagi, E. (2011). Yawn contagion and empathy in *Homo sapiens*. *PLoS One* 6:e28472. doi: 10.1371/journal.pone.0028472

Palagi, E., Celeghin, A., Tamietto, M., Winkielman, P., and Norscia, I. (2020). The neuroethology of spontaneous mimicry and emotional contagion in human and non-human animals. *Neurosci. Biobehav. Rev.* 111, 149–165. doi: 10.1016/j.neubiorev.2020.01.020

Palagi, E., Leone, A., Mancini, G., and Ferrari, P. F. (2009). Contagious yawning in gelada baboons as a possible expression of empathy. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19262–19267. doi: 10.1073/pnas.0910891106

Palagi, E., Norscia, I., and Cordoni, G. (2019). Lowland Gorillas (*Gorilla gorilla* gorilla) failed to respond to others' yawn: experimental and naturalistic evidence. *J. Comp. Psychol.* 133, 406–416. doi: 10.1037/com0000175

Palagi, E., Norscia, I., and Demuru, E. (2014). Yawn contagion in humans and bonobos: emotional affinity matters more than species. *Peer J.* 2:e519. doi: 10.7717/peerj.519

Preston, S. D., and de Waal, F. B. (2002). Empathy: its ultimate and proximate bases. *Behav. Brain Sci.* 25, 1–20. doi: 10.1017/S0140525X02000018

Provine, R. R. (1986). Yawning as a stereotyped action pattern and releasing stimulus. *Ethology* 72, 448–455. doi: 10.1111/j.1439-0310.1986.tb00611.x

Provine, R. R. (1989). Faces as releasers of contagious yawning: an approach to face perception using normal human subjects. *Bull. Psychon. Soc.* 27, 211–214. doi: 10.3758/BF03334587

Provine, R. R. (1996). "Contagious yawning and laughter: significance for sensory feature detection," in *Social Learning in Animals: The Roots of Culture*, eds C. M. Heyes and B. G. Galef Jr. (Cambridge, MA: Academic Press, Inc), 179. doi: 10.1016/b978-012273965-1/50010-8

Provine, R. R. (2005). Yawning: the yawn is primal, unstoppable and contagious, revealing the evolutionary and neural basis of empathy and unconscious behavior. *Am. Sci.* 93, 532–539. doi: 10.1511/2005.56.980

Provine, R. R., Hamernik, H. B., and Curchack, B. C. (1987). Yawning: relation to sleeping and stretching in humans. *Ethology* 76, 152–160. doi: 10.1111/j.1439-0310.1987.tb00680.x

R Core Team (2018). *R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing*. Available online at: http://www.R-project.org/

Richard, A. E., and Lajiness-O'Neill, R. (2015). Visual attention shifting in autism spectrum disorders. *J. Clin. Exp. Neuropsychol.* 37, 671–687. doi: 10.1080/13803395.2015.1042838

Saxe, R., Carey, S., and Kanwisher, N. (2004). Understanding other minds: linking developmental psychology and functional neuroimaging. *Ann. Rev. Psychol.* 55, 87–124. doi: 10.1146/annurev.psych.55.090902.142044

Schino, G., and Sciarretta, M. (2016). Patterns of social attention in mandrills *Mandrillus Sphinx*. *Int. J. Primatol.* 37, 752–761. doi: 10.1007/s10764-016-9936-7

Senju, A., Maeda, M., Kikuchi, Y., Hasegawa, T., Tojo, Y., and Osanai, H. (2007). Absence of contagious yawning in children with autism spectrum disorder. *Biol. Lett.* 3, 706–708. doi: 10.1098/rsbl.2007.0337

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: review of literature and implications for future research. *Neurosci. Biobehav. Rev.* 30, 855–863. doi: 10.1016/j.neubiorev.2006.06.011

Statistics Canada (2009). *Age Categories, Life Cycle Groupings*. Ottawa: Statistics Canada.

Tan, J., Ariely, D., and Hare, B. (2017). Bonobos respond prosocially toward members of other groups. *Sci. Rep.* 7:14733. doi: 10.1038/s41598-017-15320-w

Troisi, A., Aureli, F., Schino, G., Rinaldi, F., and de Angelis, N. (1990). The influence of age, sex, and rank on yawning behavior in two species of macaques (*Macaca fascicularis* and *M. fuscata*). *Ethology* 86, 303–310. doi: 10.1111/j.1439-0310.1990.tb00437.x

Usui, S., Senju, A., Kikuchi, Y., Akechi, H., Tojo, Y., Osanai, H., et al. (2013). Presence of contagious yawning in children with autism spectrum disorder. *Autism. Res. Treat.* 2013:971686. doi: 10.1155/2013/971686

Walusinski, O., and Deputte, B. L. (2004). The phylogeny, ethology and nosology of yawning. *Rev. Neurol.* 160, 1011–1021. doi: 10.1016/S0035-3787(04)71138-8

Whitehouse, J., Micheletta, J., Kaminski, J., and Waller, B. M. (2016). Macaques attend to scratching in others. *Anim. Behav.* 122, 169–175. doi: 10.1016/j.anbehav.2016.10.020

Wiggers, M., and van Lieshout, F. M. (1985). Development of recognition of emotions: children's reliance on situational and facial expressive cues. *Dev. Psychol.* 21, 338–349. doi: 10.1037/0012-1649.21.2.338

Yoon, J. M., and Tennie, C. (2010). Contagious yawning: a reflection of empathy, mimicry, or contagion? *Anim. Behav.* 79, e1–e3. doi: 10.1016/j.anbehav.2010.02.011

Zannella, A., Norscia, I., Stanyon, R., and Palagi, E. (2015). Testing yawning hypotheses in wild populations of two strepsirrhine species: *Propithecus verreauxi* and *Lemur catta*. *Am. J. Primatol.* 77, 1207–1215. doi: 10.1002/ajp.22459

Conflict of Interest: 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.

frontiers
in Psychology

PERSPECTIVE
published: 10 July 2020
doi: 10.3389/fpsyg.2020.01630

Motivated Cue-Integration and Emotion Regulation: Awareness of the Association Between Interoceptive and Exteroceptive Embodied Cues and Personal Need Creates an Emotion Goal

*Idit Shalev**

Laboratory for Embodiment and Self-Regulation, Department of Psychology, Ariel University, Ariel, Israel

Research on emotion suggests that individuals regulate their emotions to attain hedonic or instrumental goals. However, little is known of emotion regulation under low emotional clarity. The theory of motivated cue integration (MCI) suggests that emotion regulation under low emotional clarity should be understood as dissociation between a high-level individual hierarchical system of goals and low level interoceptive and exteroceptive embodied cues. MCI conceptualizes low emotional clarity as the product of low access to signals of emotion that result in prediction error associated with mismatch between the current bodily state and the predicted state. This deficit in emotional processing could be understood as a problem of means substitution, suggesting that use of alternative multisensory data may facilitate situational evaluation. Based on this reasoning, a new perspective on emotion regulation under low emotional clarity is presented, according to which interchangeable attention to multisensory data associated with words, associations, and images may help in cue integration, enabling the creation of a link between concrete bodily cues, abstract mental representation, and a more accurate prediction. Based on the idea that emotional episodes are conceptualized as special types of goal-directed action episodes, this process will lead to the creation of broader integrative meaning, results in the development of emotion goal.

Keywords: motivated cue-integration, emotion regulation, emotional clarity, awareness of sensation techniques, self-regulation, multisensory, interoception

OPEN ACCESS

Edited by:
Alessia Celeghin,
University of Turin, Italy

Reviewed by:
Andrea Zaccaro,
University of Pisa, Italy
Giorgia Ponsi,
Sapienza University of Rome, Italy
Federica Vellante,
University of Studies G. d'Annunzio
Chieti and Pescara, Italy

****Correspondence:***
Idit Shalev
shalev.idit@gmail.com

Specialty section:
This article was submitted to
Consciousness Research,
a section of the journal
Frontiers in Psychology

Received: *24 September 2019*
Accepted: *16 June 2020*
Published: *10 July 2020*

Citation:
Shalev I (2020) Motivated Cue-Integration and Emotion Regulation: Awareness of the Association Between Interoceptive and Exteroceptive Embodied Cues and Personal Need Creates an Emotion Goal.
Front. Psychol. 11:1630.
doi: 10.3389/fpsyg.2020.01630

THE PROBLEM OF LOW EMOTIONAL CLARITY

Emotion regulation is an essential aspect of mental health and refers to processes that amplify, attenuate, or maintain the strength of emotional reactions (Gross, 1998; Davidson et al., 2000) such that individuals are capable of controlling their behavior (see Linehan, 1993; Melnick and Hinshaw, 2000) or accepting and valuing emotional responses (Zajonc, 1980; Schwarz, 1990; Linehan, 1993; Cole et al., 1994). To date, ample research on emotion regulation has mainly been focused on top-down strategies aimed to change current emotions into desired emotions (see Gross, 2015). For example, the instrumental approach of emotion regulation suggests that individual emotions are motivated by hedonic or instrumental benefits to promote goal pursuit (e.g., Tamir and Ford, 2009, 2012; Tamir and Bigman, 2018). Research sheds light on the differences between implicit and explicit regulation of emotion, indicating the extent of effort under explicit regulation of emotion (Gyurak et al., 2011; Koole and Rothermund, 2011). However, little is known about the transition mechanisms between low and high emotional clarity to establish an emotional goal and better regulatory control. There is evidence of the differentiation between six levels of regulatory control (Lane et al., 2000; Smith and Lane, 2015), indicating that the lowest level of regulatory control is the automatic level of somatic and visceral reflexes associated with body state- and behavioral-regulation (Longhurst, 2011; Strominger et al., 2012). In fact, only the sixth and highest level of regulation is associated with voluntary emotion regulation (Smith and Lane, 2015), with its activation depending on having an emotion goal and strategic goal-directed attempts to suppress or reappraise emotional responses. This level is associated with top-down voluntary control over cognition, attention, and behavior, which activates regions such as dorsolateral prefrontal cortex (DLPFC), the dorsomedial prefrontal cortex (DMPFC), the ventrolateral prefrontal cortex (VLPFC) and parts of the dorsal anterior cingulate cortex (dACC) and the posterior parietal cortex (Gross, 1998; Ochsner and Gross, 2005).

Emotional clarity is defined as the extent to which people unambiguously identify, label, and characterize their own emotions. A majority of theories consider awareness and clarity of emotions as the building blocks of emotion regulation (e.g., Salovey et al., 1995; Gratz and Roemer, 2004, 2008; Boden and Thompson, 2015). For example, the "process model" (Gross, 1998) suggests that appraisal of a situation involves assessment of cues associated with emotional response against top-down prevailing factors (e.g., goals, social, cultural and familial influences, personality, etc.,). Accordingly, there is evidence that individuals high in clarity of emotion have meta-emotional knowledge that facilitates efficient and effective selection and implementation of emotion regulation strategies which have the greatest likelihood of achieving emotion regulation goals (Barrett et al., 2001; Tamir et al., 2008; Gross and Jazaieri, 2014). By contrast, when clarity of emotion is low, people tend to not regulate emotions or to fail in effective selection and implementation of emotion regulation strategies (Boden et al., 2013). Research demonstrates that emotional clarity deficits are associated with symptoms of depression, social anxiety, borderline personality, binge eating, and alcohol use, indicating that emotional clarity deficits may be viewed as a transdiagnostic phenomenon with divergent pathways causing problems in controlling emotions ranging from condition to condition (Vine and Aldao, 2014). These deficits come into play especially in alexithymia (Murphy et al., 2018), viewed as a continuum which manifests as difficulty in identifying feelings and in distinguishing them from bodily sensations of emotional arousal, difficulty in describing one's own feelings, and an externally oriented cognitive style, i.e., a focusing of one's attention externally with little introspection or insight into one's own feelings (Nemiah et al., 1976). Following this view, there is evidence that lower vagally mediated heart rate variability (vmHRV) resting is linked to enhanced emotional regulatory difficulties. Particularly, deficient in emotional clarity and impulse regulation, indicating that emotional regulation and autonomic regulation share neuronal networks within the brain (Williams et al., 2015). The influence of poor emotional clarity was also identified in psoriasis patients with high psychological distress. They were lower in emotional clarity than control and displayed less performance discrepancy under low vs. high cognitive load conditions (Panasiti et al., 2019).

Motivated cue integration (MCI) theory (Shalev, 2015, 2018, 2019) provides a perspective on the potential transformation of low emotional clarity into the creation of emotion goals. In what follows, I will first present the theory of MCI. I will then apply the principles of the theory on the problem of low emotional clarity. Based on the principles of the theory, I will finally propose therapeutic strategies to increase emotional clarity.

THEORY OF MOTIVATED CUE INTEGRATION

The theory of MCI (Shalev, 2015, 2018, 2019) attests that individuals construe meaning by integrating high-level control processes and low-level cues, such that, on the one hand, active top-down goals influence attention to interoceptive and exteroceptive signals. On the other hand, the perceiver's likelihood of drawing a specific inference is based on her unique system of goals, suggesting that active goals increase the accessibility of specific interoceptive and exteroceptive cues (Shalev, 2015).

According to MCI, this process occurs both in the bottom-up and the top-down directions through selective attention to multisensory information and affective signals, which, in turn, are integrated into meaning and result in action generation. Awareness of the association between high-level processes and low-level processes has been demonstrated in various studies. Based on the rationale that many voluntary actions aimed at ensuring homeostatic properties, Ainley and Tsakiris (2013) have shown that participants with higher levels of interoceptive awareness have demonstrated a greater intentional binding that reflects a stronger sense of agency. In addition, lower levels of interoceptive awareness were associated with greater self-other blurring during the enfacement illusion (Tajadura-Jiménez et al., 2012; Tajadura-Jiménez and Tsakiris, 2014). Additionally, Barrett et al. (2001) argued that persons with highly differentiated emotional experience could better regulate their emotions, pointing to the fact that a greater sensitivity to ongoing bodily changes will facilitate the regulation of emotional responses. This view is consistent with recent findings on the association between interoceptive ability and emotion regulation (Zamariola et al., 2019).

LOW-LEVEL CUES AND THE CREATION OF MEANING

The distinction between high-level and low-level processes has been illustrated in earlier emotion theories (James and Dennis, 1884; Damasio, 1994; Craig, 2004). For example, psychological theory of James and Dennis (1884) associated visceral-afferent feedback and emotional experience. The assumption that higher level of processing is grounded in the organism's lower level sensory and motor experiences is also at the core of embodied cognition research (Barsalou, 1999, 2008; Meier et al., 2012; Winkielman et al., 2015), enabling new ways to conceptualize and understand emotional processes (Damasio, 1999; Bechara and Damasio, 2005; Niedenthal, 2007). There is evidence that embodied cognition is influenced by various sources of information including innate processes, personal history, and culture (Meier et al., 2012). The general idea is that exeroceptive contextual cues activate associated mental representations (e.g., Bargh and Morsella, 2010; Loersch and Payne, 2011), suggesting that activation automatically spreads from concepts driven by experiences in the physical world to their metaphorically-related social concepts (for reviews, see Williams et al., 2009; Meier et al., 2012). For example, research of the emotion-related associations between physical warmth/coldness and psychological warmth/coldness was heterogeneous, indicating that momentary motivational states can lead to different patterns of activation across different situations (Bargh and Shalev, 2012; Shalev, 2015).

This variability in findings and the challenge of replication (Shalev and Bargh, 2015) help explain why MCI emphasized the relevance of both internal and external cues. Research on interoception and psychopathology indicates various associations between psychopathology and over-sensitivity vs. under-sensitivity to interoceptive cues (Khalsa et al., 2018). Accordingly, recent developments in neuroscience research indicate the existence of two types of inputs: (1) exteroceptive inputs associated with the perception of the body from the outside, based on multisensory integration and (2) interoceptive inputs, defined as the sense of the internal physiological state that supports homeostatic regulation of the body, resulting in physiological integrity and associated affective states, drives, and emotions (Garfinkel et al., 2015).

HIGH-LEVEL GOALS, EMOTION, AND ACTION GENERATION

MCI suggests that the monitoring of the flow of interoceptive and exeroceptive cues, associated with words, mental images, accessible memories along with the individual's goals, will contribute to the formation of integrative meaning (Shalev, 2018). Goals are defined as hierarchically organized cognitive representations of desired endpoints that affect evaluations, emotions, and behaviors (Fishbach and Ferguson, 2007). Several alternative means to the same goal potentially exist and could substitute for each other (Kruglanski et al., 2002). The degree and length of situational activation of individual's goals are influenced by sources of motivational relevance regardless of representation content. For example, value relevance is the extent to which acting on a mental representation produces desired results and/or prevents unwanted results; control relevance is related to the efficacy with which the activated representation makes things happen; and truth relevance determines what is real (Eitam and Higgins, 2010).

Individual differences are conveyed by the unique associations that individuals make between goals and means of attainment (Shalev, 2015), the repeated coupling of sensory signals (Rescorla, 1985), and the strength of the association between particular physical sensations and psychological concepts such as the combination of homeostatic cues (e.g., temperature and dryness). In addition, situational demands, history, and psychiatric and neuropsychological conditions (e.g., cognitive flexibility) influence cue integration (Shalev, 2015). Recent findings indicate, for example, that avoidance behavior in social anxiety may be related to biased distance estimation (Givon-Benjio and Okon-Singer, 2020).

There is evidence that both emotional and non-emotional action tendencies are determined by high-level goal-directed processes, which differ only in the degree of control precedence that they have (Hommel et al., 2017). Accordingly, competition between different goal-directed processes results in action control loops in which the degree to which a given stimulus event is related to and discrete from a current goal is assessed (Hommel et al., 2017; Moors et al., 2017). This helps to explain that emotional episodes are conceptualized as special types of goal-directed action episodes (Hommel et al., 2017). Similarly, Solms (2014) suggested that emotion is extended to exteroception (i.e., contextualized: "I feel like that") and transformed into a goal which comes into play through voluntary action.

MCI, PREDICTIVE CODING, AND LOW EMOTIONAL CLARITY

According to the predictive coding model (e.g., Rao and Ballard, 1999; Clark, 2013; Spratling, 2017), emotional responses depend on the continuous updated process of prediction of internal signal causes. This process occurs because, to navigate the body in the world and minimize a free energy functional of internal states, the brain must discover information about the likely reasons for sensory cues (i.e., perception) without direct access to these causes (Friston, 2010). This process is also influenced by the need to minimize the cost of prediction error either by updating generative models or by performing actions to associate sensory states in line with predictions (Hutchinson and Barrett, 2019). For example, upon detecting physical or cognitive change, probabilistic inference on the reasons for sensory cues is computed according to Bayesian principles such that the brain may represent the concept of "anger" as having a 90% chance of accounting for that pattern, while representing "sadness" as having a 5% chance, happiness a 1% chance, and so forth (Barrett, 2017; Smith et al., 2017). From the perspective of interoceptive predictive coding models (Seth, 2013), interoceptive prediction error arises when the "actual" state of the body does not match the predicted state.

MCI suggests that dissociation between high-level personal goals and low-level multisensory data results in a prediction error. Under this condition, automatic emotion regulation of preferences does not contribute to relief of psychological distress. While automatic prediction may be efficient in different contexts, low emotional clarity requires increased awareness.

Based on this reasoning, according to MCI, low access to multisensory data could be viewed as a means substitution problem (Kruglanski et al., 2002), indicating that when goal progress *via* a prior means was unrecognized or thwarted, alternative means should be used. Following this view, alternative types of encoding (e.g., mental images and multisensory data) may increase access to the dissociated aspects of the psychological experience and clarify emotion. Supportive evidence exists for the possible substitutability of emotion and alternative types of encoding, indicating the association between multisensory data and mental images as alternative types of encoding (Holmes and Mathews, 2005). For example, research shows that mental images in different modalities are primarily sensory-perceptual representations. There is overlap between the brain areas activated during imagery and those involved in processing the equivalent sensory and perceptual events. For example, there is evidence that perceptual cues associated with homeostatic deficiency (e.g., words and visual images of dryness) influence self-regulation as if they were actual ones (Shalev, 2014, 2016; Halali et al., 2017). Likewise, neural processes involved in perceiving "real" events (Kosslyn et al., 2001) signaling danger or reward (Öhman and Mineka, 2001) and mental imagery areas are shared.

MCI AND THE CREATION OF EMOTION GOAL

To increase clarity of emotion, a moment-to-moment self-regulation should be based on tracking accessible multisensory data and creation of association between selective perceptual cues and individual need, suggesting that low emotional clarity results in self-regulation failure. While in predetermined situations, individuals are governed by automatic reflexive responses, if the person wants to make plausible decisions in unfamiliar circumstances, this will be achieved through a sort of integration of past experience and current alternatives in affective evaluation (Solms, 2014). Accordingly, emotional clarity enables complex organisms to identify, monitor, and handle deviations from homeostatic settlement points in uncertain contexts through integrative comprehension and voluntary action. Such change underpins learning from experience (Solms and Friston, 2018).

Although most emotion-regulating approaches address individual maladaptive feelings, MCI proposes treating emotion control as a function of contextual meaning. The general idea is that low emotional clarity caused low access to multisensory data. Therefore, awareness of sensation techniques may support the re-creation of meaning. To address this goal, the individual should first identify a situation or event that caused discomfort and invest selective attention in the physical sensations, associations, and mental images associated with this event. Techniques involve awareness-of-sensation to peripheral cues (e.g., mindfulness and experiential approach), enabling emergence of new details which may reconstruct the emotional experience. Gendlin (2012) suggested that a positive shift in psychological experience emerges from tracking the changes of bodily experience. As such, in the practice of focusing, the focuser assigns words, mental images, or phrases that express the present sensory experience. The focuser ranges from feeling to verbal association and returning to other physical sensations that arise during the process. After remaining throughout the interaction, a shift is generated and a greater understanding of the essence of the problem or action should be taken. Attention to perceptual cues results in the identification of a specific word or image that carries meaning or reveals an individual's unconscious personal purpose. Once relief has been achieved, the focuser labels items that she wants to memorize from the process or to formulate an integrative image of the experience.

Clinicians and neuropsychologists can train individuals to use awareness-of-sensation techniques to increase emotional clarity by integrating peripheral multisensory data and creating associations between perceptual cues, symbols, and the individual's personal needs to create an emotion regulation goal. Integrating peripheral cues in the moment-by-moment creation of meaning may bridge the gap between research on higher level emotion regulation (Suri et al., 2018) and various conditions that hinder personal goal accessibility. There is evidence that use of brief awareness of sensation enhances aspects of emotional processing such as emotional intensity, emotional memory, and emotional attention bias (Guendelman et al., 2017; Wu et al., 2019). Research indicates that mindfulness influences interpersonal emotional reactions through an experiential process, while altering the subjective and physiological experience of emotions, and also biasing interpersonal behavioral patterns (Grecucci et al., 2015). Likewise, there is evidence that when expecting negative images, prefrontal and right insular activations correlated negatively with trait-mindfulness, indicating that more attentive individuals need fewer regulatory resources to attenuate emotional arousal (Lutz et al., 2014). Future research should further study the relations between higher level individual goals and lower level interoceptive and exteroceptive signals to increase emotional clarity and to shed light on individual's needs. Integrating high-level processes and low-level signals may diagnose various psychological conditions and add to the state-of-the-art self-regulation research.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This research was funded by Ariel University Research Authority, grant number RA1900000095 to IS.

ACKNOWLEDGMENTS

The author thanks the editor and reviewers for their helpful comments.

REFERENCES

Ainley, V., and Tsakiris, M. (2013). Body conscious? Interoceptive awareness, measured by heartbeat perception, is negatively correlated with self-objectification. *PLoS One* 8:e55568. doi: 10.1371/journal.pone.0055568

Bargh, J., and Morsella, E. (2010). "Unconscious behavioral guidance systems" in *Then a miracle occurs: Focusing on behavior in social psychological theory and research*. eds. C. Agnew, D. E. Carlston, W. G. Graziano and J. R. Kelly (New York, NY: Oxford University Press), 89–118.

Bargh, J. A., and Shalev, I. (2012). The substitutability of physical and social warmth in daily life. *Emotion* 12, 154–162. doi: 10.1037/a0023527

Barrett, L. F. (2017). *How emotions are made: The secret life of the brain*. Houghton Mifflin Harcourt.

Barrett, L. F., Gross, J., Christensen, T. C., and Benvenuto, M. (2001). Knowing what you're feeling and knowing what to do about it: mapping the relation between emotion differentiation and emotion regulation. *Cognit. Emot.* 15, 713–724. doi: 10.1080/02699930143000239

Barsalou, L. W. (1999). Perceptual symbol systems. *Behav. Brain Sci.* 22, 577–609. doi: 10.1017/S0140525X99002149

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645. doi: 10.1146/annurev.psych.59.103006.093639

Bechara, A., and Damasio, A. R. (2005). The somatic marker hypothesis: a neural theory of economic decision. *Games Econ. Behav.* 52, 336–372. doi: 10.1016/j.geb.2004.06.010

Boden, M. T., and Thompson, R. J. (2015). Facets of emotional awareness and associations with emotion regulation and depression. *Emotion* 15, 399–410. doi: 10.1037/emo0000057

Boden, M. T., Westermann, S., McRae, K., Kuo, J., Alvarez, J., Kulkarni, M. R., et al. (2013). Emotion regulation and posttraumatic stress disorder: a prospective investigation. *J. Soc. Clin. Psychol.* 32, 296–314. doi: 10.1521/jscp.2013.32.3.296

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 36, 181–204. doi: 10.1017/S0140525X12000477

Cole, P. M., Michel, M. K., and Teti, L. O. D. (1994). The development of emotion regulation and dysregulation: a clinical perspective. *Monogr. Soc. Res. Child Dev.* 59, 73–100. doi: 10.1111/j.1540-5834.1994.tb01278.x

Craig, A. D. (2004). Human feelings: why are some more aware than others? *Trends Cogn. Sci.* 8, 239–241. doi: 10.1016/j.tics.2004.04.004

Damasio, A. R. (1999). The feeling of what happens: body and emotion in the making of consciousness. *Nature* 401. doi: 10.1176/appi.ps.51.12.1579

Damasio, A. R. (1994). *Descartes' error: Emotion, rationality and the human brain*. New York: G. P. Putnam's Sons.

Davidson, R. J., Jackson, D. C., and Kalin, N. H. (2000). Emotion, plasticity, context, and regulation: perspectives from affective neuroscience. *Psychol. Bull.* 126, 890–909. doi: 10.1037/0033-2909.126.6.890

Eitam, B., and Higgins, E. T. (2010). Motivation in mental accessibility: relevance of a representation (ROAR) as a new framework. *Soc. Personal. Psychol. Compass* 4, 951–967. doi: 10.1111/j.1751-9004.2010.00309.x

Fishbach, A., and Ferguson, M. J. (2007). "The goal construct in social psychology" in *Social psychology: Handbook of basic principles*. *2nd Edn.* New York, US: Guilford Press, 490–515.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787

Garfinkel, S. N., Seth, A. K., Barrett, A. B., Suzuki, K., and Critchley, H. D. (2015). Knowing your own heart: distinguishing interoceptive accuracy from interoceptive awareness. *Biol. Psychol.* 104, 65–74. doi: 10.1016/j.biopsycho.2014.11.004

Gendlin, E. T. (2012). *Focusing-oriented psychotherapy: A manual of the experiential method*. Guilford Press.

Givon-Benjio, N., and Okon-Singer, H. (2020). Biased estimations of interpersonal distance in non-clinical social anxiety. *J. Anxiety Disord.* 69:102171. doi: 10.1016/j.janxdis.2019.102171

Gratz, K. L., and Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: development, factor structure, and initial validation of the difficulties in emotion regulation scale. *J. Psychopathol. Behav. Assess.* 26, 41–54. doi: 10.1023/B:JOBA.0000007455.08539.94

Gratz, K. L., and Roemer, L. (2008). The relationship between emotion dysregulation and deliberate self-harm among female undergraduate students at an urban commuter university. *Cogn. Behav. Ther.* 37, 14–25. doi: 10.1080/16506070701819524

Grecucci, A., De Pisapia, N., Thero, D. K., Paladino, M. P., Venuti, P., and Job, R. (2015). Baseline and strategic effects behind mindful emotion regulation: behavioral and physiological investigation. *PLoS One* 10:e0116541. doi: 10.1371/journal.pone.0116541

Gross, J. J. (1998). The emerging field of emotion regulation: an integrative review. *Rev. Gen. Psychol.* 2, 271–299. doi: 10.1037/1089-2680.2.3.271

Gross, J. J. (2015). Emotion regulation: current status and future prospects. *Psychol. Inq.* 26, 1–26. doi: 10.1080/1047840X.2014.940781

Gross, J. J., and Jazaieri, H. (2014). Emotion, emotion regulation, and psychopathology: an affective science perspective. *Clin. Psychol. Sci.* 2, 387–401. doi: 10.1177/2167702614536164

Guendelman, S., Medeiros, S., and Rampes, H. (2017). Mindfulness and emotion regulation: insights from neurobiological, psychological, and clinical studies. *Front. Psychol.* 8:220. doi: 10.3389/fpsyg.2017.00220

Gyurak, A., Gross, J. J., and Etkin, A. (2011). Explicit and implicit emotion regulation: a dual-process framework. *Cognit. Emot.* 25, 400–412. doi: 10.1080/02699931.2010.544160

Halali, E., Meiran, N., and Shalev, I. (2017). Keep it cool: temperature priming effect on cognitive control. *Psychol. Res.* 81, 343–354. doi: 10.1007/s00426-016-0753-6

Holmes, E. A., and Mathews, A. (2005). Mental imagery and emotion: a special relationship? *Emotion* 5, 489–497. doi: 10.1037/1528-3542.5.4.489

Hommel, B., Moors, A., Sander, D., and Deonna, J. (2017). Emotion meets action: towards an integration of research and theory. *Emot. Rev.* 9, 295–298. doi: 10.1177/1754073916689379

Hutchinson, J. B., and Barrett, L. F. (2019). The power of predictions: an emerging paradigm for psychological research. *Curr. Dir. Psychol. Sci.* 28, 280–291. doi: 10.1177/0963721419831992

James, W., and Dennis, W. (1884). "What is emotion?" in *Readings in the history of psychology*. ed. W. Dennis (New York: Appleton-Century-Crofts), 290.

Khalsa, S. S., Adolphs, R., Cameron, O. G., Critchley, H. D., Davenport, P. W., Feinstein, J. S., et al. (2018). Interoception and mental health: a roadmap. *Biol. Psychiatry. Cogn. Neurosci. Neuroimaging* 3, 501–513. doi: 10.1016/j.bpsc.2017.12.004

Koole, S. L., and Rothermund, K. (2011). "I feel better but I don't know why": the psychology of implicit emotion regulation. *Cognit. Emot.* 25, 389–399. doi: 10.1080/02699931.2010.550505

Kosslyn, S. M., Ganis, G., and Thompson, W. L. (2001). Neural foundations of imagery. *Nat. Rev. Neurosci.* 2, 635–642. doi: 10.1038/35090055

Kruglanski, A. W., Shah, J. Y., Pierro, A., and Mannetti, L. (2002). When similarity breeds content: need for closure and the allure of homogeneous and self-resembling groups. *J. Pers. Soc. Psychol.* 83, 648–662. doi: 10.1037/0022-3514.83.3.648

Lane, R. D., Nadel, L., and Kaszniak, A. W. (2000). "Epilogue: The future of emotion research from the perspective of cognitive neuroscience" in *Cognitive neuroscience of emotion*. eds. R. D. Lane and L. Nadel (New York: Oxford University Press), 407–410.

Linehan, M. M. (1993). *Skills training manual for treating borderline personality disorder*. New York, US: Guilford Press.

Loersch, C., and Payne, B. K. (2011). The situated inference model: an integrative account of the effects of primes on perception, behavior, and motivation. *Perspect. Psychol. Sci.* 6, 234–252. doi: 10.1177/1745691611406921

Longhurst, J. (2011). "Regulation of autonomic function by visceral and somatic afferents" in *Central regulation of autonomic functions. 2nd Edn*. eds. I. J. Llewellyn-Smith and A. J. M. Verberne (New York, NY, USA: Oxford University Press), 161–179.

Lutz, J., Herwig, U., Opialla, S., Hittmeyer, A., Jäncke, L., Rufer, M., et al. (2014). Mindfulness and emotion regulation—an fMRI study. *Soc. Cogn. Affect. Neurosci.* 9, 776–785. doi: 10.1093/scan/nst043

Meier, B. P., Schnall, S., Schwarz, N., and Bargh, J. A. (2012). Embodiment in social psychology. *Top. Cogn. Sci.* 4, 705–716. doi: 10.1111/j.1756-8765.2012.01212.x

Melnick, S. M., and Hinshaw, S. P. (2000). Emotion regulation and parenting in AD/HD and comparison boys: linkages with social behaviors and peer preference. *J. Abnorm. Child Psychol.* 28, 73–86. doi: 10.1023/A:1005174102794

Moors, A., Boddez, Y., and De Houwer, J. (2017). The power of goal-directed processes in the causation of emotional and other actions. *Emot. Rev.* 9, 310–318. doi: 10.1177/1754073916669595

Murphy, J., Catmur, C., and Bird, G. (2018). Alexithymia is associated with a multidomain, multidimensional failure of interoception: evidence from novel tests. *J. Exp. Psychol. Gen.* 147, 398–408. doi: 10.1037/xge0000366

Nemiah, J. C., Freyberger, H. J., and Sifneos, P. E. (1976). "Alexithymia: A view of the psychosomatic process" in *Modern trends in psychosomatic medicine. Vol. 3*. ed. O. W. Hill (London: Buttersworth), 430–439.

Niedenthal, P. M. (2007). Embodying emotion. *Science* 316, 1002–1005. doi: 10.1126/science.1136930

Ochsner, K. N., and Gross, J. J. (2005). The cognitive control of emotion. *Trends Cogn. Sci.* 9, 242–249. doi: 10.1016/j.tics.2005.03.010

Öhman, A., and Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychol. Rev.* 108, 483–522. doi: 10.1037/0033-295X.108.3.483

Panasiti, M. S., Ponsi, G., Monachesi, B., Lorenzini, L., Panasiti, V., and Aglioti, S. M. (2019). Cognitive load and emotional processing in psoriasis: a thermal imaging study. *Exp. Brain Res.* 237, 211–222. doi: 10.1007/s00221-018-5416-y

Rao, R. P. N., 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

Rescorla, R. A. (1985). "Pavlovian conditioning analogues to Gestalt perceptual principles" in *Affect, conditioning, and cognition: Essays on the determinants of behavior*. eds. F. R. Brush and J. B. Overmier (Hillsdale, NJ: Erlbaum), 113–130.

Salovey, P., Mayer, J. D., Golman, S. L., Turvey, C., and Palfai, T. P. (1995). "Emotional attention, clarity, and repair: Exploring emotional intelligence using the Trait Meta-Mood Scale" in *Emotion, disclosure, and health*. ed. J. W. Pennebaker (Washington, DC: American Psychological Association), 125–154.

Schwarz, N. (1990). "Feelings as information: Informational and motivational functions of affective states" in *Handbook of motivation and cognition. Foundations of social behavior. Vol. 2*. eds. E. T. Higgins and R. M. Sorrentino (New York, NY: Guilford Press), 527–561.

Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends Cogn. Sci.* 17, 565–573. doi: 10.1016/j.tics.2013.09.007

Shalev, I. (2014). Implicit energy loss: embodied dryness cues influence vitality and depletion. *J. Consum. Psychol.* 24, 260–270. doi: 10.1016/j.jcps.2013.09.011

Shalev, I. (2015). The architecture of embodied cue integration: insight from the "motivation as cognition" perspective. *Front. Psychol.* 6:658. doi: 10.3389/fpsyg.2015.00658

Shalev, I. (2016). Pictorial and mental arid landscape images reduce the motivation to change negative habits. *J. Environ. Psychol.* 45, 30–39. doi: 10.1016/j.jenvp.2015.11.005

Shalev, I. (2018). Using motivated cue integration theory to understand a moment-by-moment transformative change: a new look at the focusing technique. *Front. Hum. Neurosci.* 12:307. doi: 10.3389/fnhum.2018.00307

Shalev, I. (2019). Motivated cue integration in alexithymia: improving interoception and emotion information processing by awareness-of-sensation techniques. *Front. Psychiatry* 10:329. doi: 10.3389/fpsyt.2019.00329

Shalev, I., and Bargh, J. (2015). On the association between loneliness and physical warmth-seeking through bathing: reply to Donnellan et al. (2014) and three further replications of Bargh and Shalev (2012) study 1. *Emotion* 15, 120–123. doi: 10.1037/emo0000014

Smith, R., and Lane, R. D. (2015). The neural basis of one's own conscious and unconscious emotional states. *Neurosci. Biobehav. Rev.* 57, 1–29. doi: 10.1016/j.neubiorev.2015.08.003

Smith, R., Thayer, J. F., Khalsa, S. S., and Lane, R. D. (2017). The hierarchical basis of neurovisceral integration. *Neurosci. Biobehav. Rev.* 75, 274–296. doi: 10.1016/j.neubiorev.2017.02.003

Solms, M. (2014). A neuropsychoanalytical approach to the hard problem of consciousness. *J. Integr. Neurosci.* 13, 173–185. doi: 10.1142/S0219635214400032

Solms, M., and Friston, K. (2018). How and why consciousness arises: some considerations from physics and physiology. *J. Conscious. Stud.* 25, 202–238.

Spratling, M. W. (2017). A hierarchical predictive coding model of object recognition in natural images. *Cogn. Comput.* 9, 151–167. doi: 10.1007/s12559-016-9445-1

Strominger, N. L., Demarest, R. J., and Laemle, L. B. (2012). "Reflexes and muscle tone" in *Noback's human nervous system, Seventh edition: Structure and function*. Springer Science and Business Media. Totowa, NJ: Humana Press, 143–154.

Suri, G., Sheppes, G., Young, G., Abraham, D., McRae, K., and Gross, J. J. (2018). Emotion regulation choice: the role of environmental affordances. *Cognit. Emot.* 32, 963–971. doi: 10.1080/02699931.2017.1371003

Tajadura-Jiménez, A., Longo, M. R., Coleman, R., and Tsakiris, M. (2012). The person in the mirror: using the enfacement illusion to investigate the experiential structure of self-identification. *Conscious. Cogn.* 21, 1725–1738. doi: 10.1016/j.concog.2012.10.004

Tajadura-Jiménez, A., and Tsakiris, M. (2014). Balancing the "inner" and the "outer" self: interoceptive sensitivity modulates self–other boundaries. *J. Exp. Psychol. Gen.* 143:736. doi: 10.1037/a0033171

Tamir, M., and Bigman, Y. E. (2018). Expectations influence how emotions shape behavior. *Emotion* 18, 15–25. doi: 10.1037/emo0000351

Tamir, M., and Ford, B. Q. (2009). Choosing to be afraid: preferences for fear as a function of goal pursuit. *Emotion* 9, 488–497. doi: 10.1037/a0015882

Tamir, M., and Ford, B. Q. (2012). When feeling bad is expected to be good: emotion regulation and outcome expectancies in social conflicts. *Emotion* 12, 807–816. doi: 10.1037/a0024443

Tamir, M., Mitchell, C., and Gross, J. J. (2008). Hedonic and instrumental motives in anger regulation: research report. *Psychol. Sci.* 19, 324–328. doi: 10.1111/j.1467-9280.2008.02088.x

Vine, V., and Aldao, A. (2014). Impaired emotional clarity and psychopathology: a transdiagnostic deficit with symptom-specific pathways through emotion regulation. *J. Soc. Clin. Psychol.* 33, 319–342. doi: 10.1521/jscp.2014.33.4.319

Williams, D. P., Cash, C., Rankin, C., Bernardi, A., Koenig, J., and Thayer, J. F. (2015). Resting heart rate variability predicts self-reported difficulties in emotion regulation: a focus on different facets of emotion regulation. *Front. Psychol.* 6:261. doi: 10.3389/fpsyg.2015.00261

Williams, L. E., Huang, J. Y., and Bargh, J. A. (2009). The scaffolded mind: higher mental processes are grounded in early experience of the physical world. *Eur. J. Soc. Psychol.* 39, 1257–1267. doi: 10.1002/ejsp.665

Winkielman, P., Niedenthal, P., Wielgosz, J., Eelen, J., and Kavanagh, L. C. (2015). "Embodiment of cognition and emotion" in *APA handbook of personality and social psychology. Vol. 1*. eds. M. Mikulincer and P. R. Shaver (Washington, DC: American Psychological Association), 151–175.

Wu, R., Liu, L. -L., Zhu, H., Su, W. -J., Cao, Z. -Y., Zhong, S. -Y., et al. (2019). Brief mindfulness meditation improves emotion processing. *Front. Neurosci.* 13:1074. doi: 10.3389/fnins.2019.01074

Zajonc, R. B. (1980). Feeling and thinking: preferences need no inferences. *Am. Psychol.* 35, 151–175. doi: 10.1037/0003-066X.35.2.151

Zamariola, G., Frost, N., Van Oost, A., Corneille, O., and Luminet, O. (2019). Relationship between interoception and emotion regulation: new evidence from mixed methods. *J. Affect. Disord.* 246, 480–485. doi: 10.1016/j.jad.2018.12.101

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

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read
for greatest visibility
and readership

FAST PUBLICATION
Around 90 days
from submission
to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative,
and constructive
peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers
acknowledged by name
on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data
and methods to enhance
research reproducibility

DIGITAL PUBLISHING
Articles designed
for optimal readership
across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics
track visibility across
digital media

EXTENSIVE PROMOTION
Marketing
and promotion
of impactful research

LOOP RESEARCH NETWORK
Our network
increases your
article's readership