



Towards a cross-modal perspective of emotional perception in social anxiety: review and future directions

Virginie Peschard *, Pierre Maurage and Pierre Philippot

Laboratory for Experimental Psychopathology, Institute of Psychology, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

Edited by:

Quincy Wong, Macquarie University, Australia

Reviewed by:

Matthias J. Wieser, University of Würzburg, Germany
Eva Gilboa-Schechtman, Bar Ilan University, Israel

*Correspondence:

Virginie Peschard, Laboratory for Experimental Psychopathology, Institute of Psychology, Université Catholique de Louvain, Place du Cardinal Mercier 10, B-1348 Louvain-la-Neuve, Belgium
e-mail: virginie.peschard@uclouvain.be

The excessive fear of being negatively evaluated constitutes a central component of social anxiety (SA). Models posit that selective attention to threat and biased interpretations of ambiguous stimuli contribute to the maintenance of this psychopathology. There is strong support for the existence of processing biases but most of the available evidence comes from face research. Emotions are, however, not only conveyed through facial cues, but also through other channels, such as vocal and postural cues. These non-facial cues have yet received much less attention. We therefore plead for a cross-modal investigation of biases in SA. We argue that the inclusion of new modalities may be an efficient research tool to (1) address the specificity or generalizability of these biases; (2) offer an insight into the potential influence of SA on cross-modal processes; (3) operationalize emotional ambiguity by manipulating cross-modal emotional congruency; (4) inform the debate about the role of top-down and bottom-up factors in biasing attention; and (5) probe the cross-modal generalizability of cognitive training. Theoretical and clinical implications as well as potential fruitful avenues for research are discussed.

Keywords: cross-modality, emotion, social anxiety, face, voice

INTRODUCTION

Influential models of social anxiety (SA) implicate cognitive biases as maintaining factors (Clark and Wells, 1995; Rapee and Heimberg, 1997). The existing evidence concerning biases in SA has largely relied on faces (for a review, Staugaard, 2010). Particularly, there is strong support for attentional biases (AB) towards facial stimuli among high socially anxious (HSA) individuals. While some studies indicated a facilitated attention to threatening faces (Mogg et al., 2004; Pishyar et al., 2004), others demonstrated difficulties in disengaging attention from these cues (Buckner et al., 2010; Schofield et al., 2012). Significant efforts have also been directed at understanding the effect of SA on the interpretation of faces, but have yielded mixed results, possibly due to methodological differences in dependent variables, stimuli and tasks. While several studies indicate that SA modulates the interpretation of emotional facial expressions (e.g., ratings of the emotional cost for interacting with the expressor: Schofield et al., 2007; Douilliez et al., 2012), other studies did not find any differences between HSA and controls (e.g., disapproval ratings: Douilliez and Philippot, 2003; decoding accuracy: Philippot and Douilliez, 2005).

To date, evidence linking SA to cognitive biases provided much information about how HSA individuals process faces. However, conclusions from these studies are limited to the processing of faces. Further, some questions are still controversial, in part due to the inherent methodological limitations of face research. Social interactions mobilize multiple channels, including speech style, facial expressions, postures, gestures, and tone of voice. Focusing research solely on faces raises the risk of overlooking

other channels that are heavily implicated in social interactions. We argue that the investigation of SA-related biases needs to be extended to a multi-modal approach (as also suggested by Gilboa-Schechtman and Shachar-Lavie, 2013; Schulz et al., 2013), including the modalities that are most important in social interaction: vision and hearing. The use of cross-modal paradigms will allow the re-evaluation of studies using uni-modal stimuli, which could underestimate the cognitive biases present in real life. To support this statement, we developed several arguments based on empirical evidence, with the aim of identifying useful avenues for future research.

ARGUMENTS

INCLUDING EMOTIONAL PROSODY TO PROBE THE GENERALIZABILITY OF COGNITIVE BIASES IN SOCIAL ANXIETY

Emotional prosody refers to all changes in acoustic parameters, such as intonation, amplitude, envelope, tempo, rhythm and voice quality during emotional episodes (Grandjean et al., 2006). It is a powerful communication tool transmitting paralinguistic information, and notably the speaker's emotional state (Belin et al., 2004). Research that neglects the latter channel ignores crucial information for interpersonal interactions. To document its relevance, we will review research on the modulation of attention and emotional judgments by prosody.

Selective attention to emotional prosody

Efficient detection of salient or goal-relevant stimuli is essential to adjust behaviors accordingly. Given the limited processing capacity of our brain, mechanisms of attention play a critical

role in selecting most important information from the myriad of sensory inputs. In the competition for processing resources, emotions have been shown to modulate attention (Vuilleumier et al., 2004; Vuilleumier, 2005). To date, the effect of emotional prosody on attention has been mostly assessed during dichotic listening or during the variation of feature-based attention.

The dichotic-listening technique is an attentional filtering task that assesses the ability to suppress or ignore distractors co-occurring with targets. Dichotic-listening investigations typically involve the simultaneous presentation of lateralized male and female voices with identical or different emotional prosody. Participants are requested to focus their attention on one ear and to determine the gender of the speaker on the attended ear. Recently, Aue et al. (2011) reported that, compared to neutral prosody, angry prosody attracts attention and induces behavioral and physiological changes (e.g., increased forehead temperature) with or without voluntary attention. Moreover, neuroimaging studies indicated greater activation for angry relative to neutral prosody in the superior temporal sulcus (Grandjean et al., 2005; Sander et al., 2005) and the amygdala (Sander et al., 2005) irrespective of the focus of attention. These findings suggest that threatening voices might be processed automatically by specific brain regions (but see Mothes-Lasch et al., 2011).

In addition to dichotic-listening methods, several studies (Quadflieg et al., 2008; Ethofer et al., 2009) investigated whether brain responses to angry compared to neutral prosody are modulated by variations in feature-based auditory attention. For example, Quadflieg et al. (2008) examined brain responses to neutral and angry voices while control and HSA subjects judged either the emotion or the gender of the voice. This study confirmed the findings of Sander et al. (2005) showing stronger activation for angry than neutral prosody in amygdala regardless of the task and in orbitofrontal cortex (OFC) during task-relevant as compared to task-irrelevant emotional prosody processing. Additionally, their results indicated that compared to controls, HSA individuals exhibit stronger right OFC response to angry versus neutral prosody regardless of the focus of attention. These findings suggest that the OFC might be implicated in biased processing of threatening prosody in SA.

To conclude, few studies have explored the implicit and explicit processing of emotional prosody via uni-modal attentional distraction from emotion. The lack of studies examining attention to prosodic information in the general population as well as in socially anxious samples is surprising, since the exploration of these processes could contribute to new insights into the attentional processing of emotional information. The above mentioned paradigms offer an interesting opportunity to provide evidence from the auditory modality that might be congruent or incongruent with the evidence accumulated in the visual domain.

Interpretation of emotional prosody

Other studies have focused on the interpretation of affective signals conveyed by faces or voices. These abilities have been increasingly studied in several psychopathologies, including alcohol-dependence (Maurage et al., 2009; Kornreich et al., 2013), depression (Naranjo et al., 2011) and bipolar disorder (Van Rheenen and Rossell, 2013).

Despite this growing interest, we found only one study (Quadflieg et al., 2007) that probed the presence of biases in the interpretation of emotional prosody in SA. Findings indicated that compared to controls, HSA participants present higher correct identification rates for fearful and sad prosody than controls, but conversely show impaired performances for happy prosody. Surprisingly, there were no differences between groups for neutral, anger and disgust prosody, as well as with regard to valence and arousal ratings for any prosody. These findings suggest that HSA individuals interpret prosody in a different manner than low socially anxious (LSA) individuals. However, it should be noted that this observation is at odds with theoretical predictions of a threat-specific bias, since fearful and sad expressions do not specifically indicate a social threat as would angry expressions do, thereby highlighting the importance of further investigations.

Summary

The lack of studies on emotional prosody in SA is problematic, since a threatening voice is a clear sign of danger and therefore a good candidate for capturing the attention of HSA individuals and eliciting biased interpretations. The study of emotional prosody constitutes a promising tool to investigate the cognitive biases in SA more completely. Presently, it is unclear whether these biases, which are repeatedly described in SA for visual processing, are similar in the auditory channel. Yet, the few existing data suggest some particularities in the processing of emotional prosody by HSA individuals. In addition to emotional prosody, other affective stimuli could be useful to probe the generalizability of cognitive biases in SA, notably body language (for an illustration in depression see Loi et al., 2013).

PROVIDING INSIGHTS ABOUT THE POTENTIAL INFLUENCE OF SOCIAL ANXIETY ON THE INTERACTIONS BETWEEN MODALITIES

Audio-visual integration

A specific line of research addresses the ability of humans to integrate co-occurring sources of facial and vocal affective information. In natural environment, humans are immersed in a stream of stimulations from multiple modalities. The ability to integrate these multimodal inputs allows for an unified and coherent representation of the world and for taking advantage of non-redundant and complementary information from a single modality (Ernst and Bühlhoff, 2004). The multimodal integration of affective facial and vocal expressions has led to a growing interest in the literature (for a review, Campanella and Belin, 2007). It has been demonstrated that congruency in the facial and vocal expression of emotion facilitates their identification compared to an uni-modal (i.e., face or voice presented in isolation) source of information (e.g., Collignon et al., 2008). Interestingly, integrative processes have been shown to be altered during the emotional perception of facial and vocal expressions in psychopathological populations, such as in alcohol-dependent subjects (Maurage et al., 2007, 2008, 2013). Specifically, alcohol-dependent individuals do not only suffer from a deficit in decoding facial and vocal expressions, but they also present a specific deficit in integrating messages conveyed by these two modalities. Hence, their resulting impairment is not just the sum of impairments in each modality,

but it is further aggravated by a difficulty in integrating these modalities.

To our knowledge, no study has investigated the effect of SA on the ability to decode emotions in audio-visual modality, and the possible deficit in integrating these two modalities. This issue is important, as it would suggest that the total deficit in emotional information processing by HSA individuals would not be the addition of the deficits in each modality, but would be even more important, given the over-added integration deficit. Hence, the closer a paradigm would be to a real-life multi-sensory situation, the more pronounced might be the biases. Consequently, earlier uni-modal studies might have underestimated the extent of these biases.

Cross-modal attention

A second line of research has investigated how signals from different modalities influence each other in capturing attention. It has been shown that emotional prosody can serve as an exogenous cue to orient attention towards relevant visual events. Using a cross-modal adaptation of the dot-probe task, Brosch et al. (2008) showed decreased response times to non-emotional visual targets preceded by angry prosody compared to targets preceded by neutral prosody. Brosch et al. (2009) replicated and extended these behavioral findings by showing an amplification of the P1 (an electrophysiological component indexing early visual processing) for visual targets occurring at the spatial location of angry as compared to neutral prosody. These results suggest that emotional attention can operate across modalities because auditory stimuli can enhance early visual processing stages.

Several studies similarly demonstrated that emotional stimuli in one modality influence the processing of emotional information in another modality. For example, emotional prosody can facilitate attention to emotionally congruent facial expressions in visual search (Paulmann et al., 2012; Rigoulot and Pell, 2012) and in cross-modal priming tasks (Pell, 2005a,b; Paulmann and Pell, 2010). Other studies revealed that the judgment of emotional prosody is biased by a concurrent emotional face despite the instruction to ignore this channel (de Gelder and Vroomen, 2000; Vroomen et al., 2001). The reverse effect has also been observed, showing that emotional prosody biases the judgment of the emotion expressed in the face (de Gelder and Vroomen, 2000). These studies suggest that audio-visual integration of emotional signals may be an automatic and mandatory process, as this effect seems to arise independently of voluntary attentional factors (de Gelder and Vroomen, 2000; Vroomen et al., 2001) and of the awareness of the face (de Gelder et al., 2002).

Based on this line of research, one would want to investigate whether such automatic control of attention across modalities is modulated by SA. Such research could help identifying the origin of the SA biases on the top-down—bottom-up continuum. One could also hypothesize that HSA individuals could be more influenced than LSA individuals by cross-modal interference, if that interference can be interpreted as a social threat. These kind of studies need still to be conducted. The results obtained in healthy populations also raise the question of how conflicting emotional information is processed by HSA individuals. This topic will be developed in the next section.

MANIPULATING THE CROSS-MODAL EMOTIONAL CONGRUENCY AS A TOOL TO OPERATIONALIZE AMBIGUITY

In the environment, we frequently encounter conflicting situations in which two modalities convey incongruent information (De Gelder and Bertelson, 2003). As mentioned, the categorization of emotional stimuli is affected by incongruent information provided by the second channel in cross-modal situations. Few studies have investigated such cross-modal incongruence effects among psychopathological populations. Some studies have described disturbed cross-modal integration of emotional faces and voices in schizophrenia (de Gelder et al., 2005; de Jong et al., 2009). However, no study has explored the effect of SA on the ability to decode incongruent emotional faces and voices. Yet, in real-life conditions, conversational partner often do not provide direct unambiguous feedback about their approval or disapproval. Such ambiguity leaves room for the socially anxious' tendency to interpret responses as signs of negative evaluation. Recently, Koizumi et al. (2011) used a cross-modal bias paradigm (Bertelson and De Gelder, 2004) that included emotionally congruent or incongruent voice-face pairs. Participants had to decode the emotion displayed in one of the two channels (e.g., face) while ignoring the other (e.g., voice). Results indicate that individuals with heightened trait anxiety were likely to interpret the stimuli more negatively, putting more weight on the to-be-ignored angry faces or voices. As a consequence, manipulating emotional congruency across modalities can be a powerful way to examine the impact of ambiguity on the judgment of social information and to renew the exploration of biases in SA.

INFORMING DEBATE ABOUT THE ROLE OF TOP-DOWN AND BOTTOM-UP FACTORS IN BIASING ATTENTION TO THREAT

Different models of anxiety have questioned the balance between bottom-up and top-down attention to explain cognitive biases. First, Bishop (2007) proposes that anxiety leads to AB by amplifying amygdala responsiveness to threat and/or by impairing the recruitment of top-down attention control, particularly under conditions of low perceptual load. In the same vein, the attentional control theory (Eysenck et al., 2007) and recent developments (e.g., Berggren and Derakshan, 2013; Berggren et al., 2013) suggest that individuals reporting high trait anxiety have to engage a greater amount of attentional control under low cognitive load (thereby reducing efficiency) to attain the level of performance achieved by low-anxious individuals. However, high cognitive load can disrupt performance in tasks requiring attentional control particularly in high anxious individuals. Finally, Hirsch and Mathews (2012) propose that high levels of anxiety are characterized by an imbalance between (weak) top/down and (strong) bottom/up attentional processes, the latter being automatically fueled by threat.

While behavioral studies demonstrated a rapid orientation towards threatening faces (Mogg et al., 2004; Pishyar et al., 2004), neuroimaging studies showed increased amygdala response, exaggerated negative emotion reactivity, and reduced cognitive regulation-related neural activation to faces in SA (Goldin et al., 2009; Ball et al., 2012). An increased vigilance for faces, indexed by enhanced P1, is also well documented in SA

(Rossignol et al., 2012; Peschard et al., 2013). Nevertheless, most of this research is limited to visual stimuli and therefore prevents us from drawing firm conclusions about the implication of top-down and bottom-up factors in the generation of cognitive biases. Investigating the presence of biases across modalities offers an interesting paradigm to provide an insight into the contribution of top-down and bottom-up influences. Indeed, if a bias is generated at an early perceptual level, and thus nested in a specific modality, it is unlikely that the same bias would be reproduced in all other modalities. Consequently, the absence of generalization of a cognitive bias across modalities would support the notion that this bias is yielded by bottom-up processes, whereas its presence across modalities would rather support the notion of a top-down influence. As far as we know, no study has yet explored these integrative processes in SA, thus stressing the need to initiate this field of research.

THE CROSS-MODAL GENERALIZABILITY OF COGNITIVE TRAINING

Recent studies have shown that training HSA individuals to attend to non-threatening stimuli reduces AB, which in turn diminishes anxiety (Amir et al., 2008; Heeren et al., 2012b). It has also been demonstrated that inducing AB for threat induces anxiety (Heeren et al., 2012a). These findings support the proposal that AB to threat play a causal role in the maintenance and the development of SA. However, previous research has left unaddressed several important issues both at the fundamental and clinical level. First, there is a need to obtain a more ecological and complete AB evaluation before AB training. It should be established whether similar AB are present across modalities (as posed by theoretical models) or whether they are proper to a specific modality, hence suggesting retraining in that specific modality. Moreover if research findings show that AB appear across modalities, a crucial question would be whether training in one modality would transfer its effects to other modalities. This cross-modal perspective can offer an interesting paradigm to disentangle top-down and bottom-up determinations of AB. Finally, this perspective could lead to innovative AB training based on the combination of different modalities.

CONCLUSION

We developed several arguments pleading for a cross-modal perspective in the investigation of biases in SA. In addition to the gain of a more complete and ecological picture of cognitive biases, a cross-modal perspective opens up new possibilities for understanding fundamental processes underlying biases in SA. This perspective might help to determine the stage of processing at which these biases occur. In this contribution, we mainly focused on auditory and visual modalities. However, signals from other modalities, like olfaction, could also influence information processing and should thus be considered in psychopathological research (Maurage et al., 2014). Recently, Adolph et al. (2013) have reported that HSA individuals might be particularly sensitive to chemosensory contextual social information during the processing of anxious facial expressions. This outlines the usefulness to exploring cross-modal processing in order to precisely describe cognitive biases in SA.

ACKNOWLEDGMENTS

The authors are funded by the Belgian Fund for Scientific Research (F.N.R.S., Belgium). The authors also appreciate helpful comments of Alexandre Heeren, Magali Lahaye, Vincent Dethier, and Anne Kever on earlier drafts of this paper.

REFERENCES

- Adolph, D., Meister, L., and Pause, B. M. (2013). Context counts! social anxiety modulates the processing of fearful faces in the context of chemosensory anxiety signals. *Front. Hum. Neurosci.* 7:283. doi: 10.3389/fnhum.2013.00283
- Amir, N., Weber, G., Beard, C., Bomyea, J., and Taylor, C. T. (2008). The effect of a single-session attention modification program on response to a public-speaking challenge in socially anxious individuals. *J. Abnorm. Psychol.* 117, 860–868. doi: 10.1037/a0013445
- Aue, T., Cuny, C., Sander, D., and Grandjean, D. (2011). Peripheral responses to attended and unattended angry prosody: a dichotic listening paradigm. *Psychophysiology* 48, 385–392. doi: 10.1111/j.1469-8986.2010.01064.x
- Ball, T. M., Sullivan, S., Flagan, T., Hitchcock, C. A., Simmons, A., Paulus, M. P., et al. (2012). Selective effects of social anxiety, anxiety sensitivity, and negative affectivity on the neural bases of emotional face processing. *Neuroimage* 59, 1879–1887. doi: 10.1016/j.neuroimage.2011.08.074
- Belin, P., Fecteau, S., and Bedard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends Cogn. Sci.* 8, 129–135. doi: 10.1016/j.tics.2004.01.008
- Berggren, N., and Derakshan, N. (2013). Attentional control deficits in trait anxiety: why you see them and why you don't. *Biol. Psychol.* 92, 440–446. doi: 10.1016/j.biopsycho.2012.03.007
- Berggren, N., Richards, A., Taylor, J., and Derakshan, N. (2013). Affective attention under cognitive load: reduced emotional biases but emergent anxiety-related costs to inhibitory control. *Front. Hum. Neurosci.* 7:188. doi: 10.3389/fnhum.2013.00188
- Bertelson, P., and De Gelder, B. (2004). "The psychology of multimodal perception," in *Crossmodal Space and Crossmodal Attention*, ed J. Driver (Oxford: Oxford University Press), 151–177.
- Bishop, S. J. (2007). Neurocognitive mechanisms of anxiety: an integrative account. *Trends Cogn. Sci.* 11, 307–316. doi: 10.1016/j.tics.2007.05.008
- Brosch, T., Grandjean, D., Sander, D., and Scherer, K. R. (2008). Behold the voice of wrath: cross-modal modulation of visual attention by anger prosody. *Cognition* 106, 1497–1503. doi: 10.1016/j.cognition.2007.05.011
- Brosch, T., Grandjean, D., Sander, D., and Scherer, K. R. (2009). Cross-modal emotional attention: emotional voices modulate early stages of visual processing. *J. Cogn. Neurosci.* 21, 1670–1679. doi: 10.1162/jocn.2009.21110
- Buckner, J. D., Maner, J. K., and Schmidt, N. B. (2010). Difficulty disengaging attention from social threat in social anxiety. *Cognit. Ther. Res.* 34, 99–105. doi: 10.1007/s10608-008-9205-y
- Campanella, S., and Belin, P. (2007). Integrating face and voice in person perception. *Trends. Cogn. Sci.* 11, 535–543. doi: 10.1016/j.tics.2007.10.001
- Clark, D. M., and Wells, A. (1995). "A cognitive model of social phobia," in *Diagnosis, Assessment and Treatment*, eds R. G. Heimberg, M. R. Liebowitz, D. A. Hope and R. Schneider (New York: Guilford), 69–93.
- Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., et al. (2008). Audio-visual integration of emotion expression. *Brain Res.* 1242, 126–135. doi: 10.1016/j.brainres.2008.04.023
- De Gelder, B., and Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends Cogn. Sci.* 7, 460–467. doi: 10.1016/j.tics.2003.08.014
- de Jong, J. J., Hodiamont, P. P., Van den Stock, J., and de Gelder, B. (2009). Audiovisual emotion recognition in schizophrenia: reduced integration of facial and vocal affect. *Schizophr. Res.* 107, 286–293. doi: 10.1016/j.schres.2008.10.001
- de Gelder, B., Pourtois, G., and Weiskrantz, L. (2002). Fear recognition in the voice is modulated by unconsciously recognized facial expressions but not by unconsciously recognized affective pictures. *Proc. Natl. Acad. Sci. U S A* 99, 4121–4126. doi: 10.1073/pnas.062018499
- de Gelder, B., and Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cogn. Emot.* 14, 289–311. doi: 10.1080/026999300378824
- de Gelder, B., Vroomen, J., De Jong, S. J., Masthoff, E. D., Trompenaars, F. J., and Hodiamont, P. (2005). Multisensory integration of emotional faces and voices in schizophrenics. *Schizophr. Res.* 72, 195–203. doi: 10.1016/j.schres.2004.02.013

- Douilliez, C., and Philippot, P. (2003). Biais dans l'évaluation volontaire de stimuli verbaux et non-verbaux: effet de l'anxiété sociale. *Francophone J. Clin. Behav. Cogn.* 8, 12–18.
- Douilliez, C., Yzerbyt, V., Gilboa-Schechtman, E., and Philippot, P. (2012). Social anxiety biases the evaluation of facial displays: evidence from single face and multi-facial stimuli. *Cogn. Emot.* 26, 1107–1115. doi: 10.1080/02699931.2011.632494
- Ernst, M. O., and Bühlhoff, H. H. (2004). Merging the senses into a robust percept. *Trends Cogn. Sci.* 8, 162–169. doi: 10.1016/j.tics.2004.02.002
- Ethofer, T., Kreifelts, B., Wiethoff, S., Wolf, J., Grodd, W., Vuilleumier, P., et al. (2009). Differential influences of emotion, task and novelty on brain regions underlying the processing of speech melody. *J. Cogn. Neurosci.* 21, 1255–1268. doi: 10.1162/jocn.2009.21099
- Eysenck, M. W., Derakshan, N., Santos, R., and Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion* 7, 336–353. doi: 10.1037/1528-3542.7.2.336
- Gilboa-Schechtman, E., and Shachar-Lavie, I. (2013). More than a face: a unified theoretical perspective on nonverbal social cue processing in social anxiety. *Front. Hum. Neurosci.* 7:904. doi: 10.3389/fnhum.2013.00904
- Goldin, P. R., Manber, T., Hakimi, S., Canli, T., and Gross, J. J. (2009). Neural bases of social anxiety disorder: emotional reactivity and cognitive regulation during social and physical threat. *Arch. Gen. Psychiatry* 66, 170–180. doi: 10.1001/archgenpsychiatry.2008.525
- Grandjean, D., Banziger, T., and Scherer, K. R. (2006). Intonation as an interface between language and affect. *Prog. Brain Res.* 156, 235–247. doi: 10.1016/s0079-6123(06)56012-1
- Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., et al. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nat. Neurosci.* 8, 145–146. doi: 10.1038/nn1392
- Heeren, A., Peschard, V., and Philippot, P. (2012a). The causal role of attentional bias for threat cues in social anxiety: a test on a cyber-ostracism task. *Cognit. Ther. Res.* 36, 512–521. doi: 10.1007/s10608-011-9394-7
- Heeren, A., Reese, H. E., McNally, R. J., and Philippot, P. (2012b). Attention training toward and away from threat in social phobia: effects on subjective, behavioral and physiological measures of anxiety. *Behav. Res. Ther.* 50, 30–39. doi: 10.1016/j.brat.2011.10.005
- Hirsch, C. R., and Mathews, A. (2012). A cognitive model of pathological worry. *Behav. Res. Ther.* 50, 636–646. doi: 10.1016/j.brat.2012.06.007
- Koizumi, A., Tanaka, A., Imai, H., Hiramatsu, S., Hiramoto, E., Sato, T., et al. (2011). The effects of anxiety on the interpretation of emotion in the face-voice pairs. *Exp. Brain Res.* 213, 275–282. doi: 10.1007/s00221-011-2668-1
- Kornreich, C., Breviers, D., Canivet, D., Ermer, E., Naranjo, C., Constant, E., et al. (2013). Impaired processing of emotion in music, faces and voices supports a generalized emotional decoding deficit in alcoholism. *Addiction* 108, 80–88. doi: 10.1111/j.1360-0443.2012.03995.x
- Loi, F., Vaidya, J. G., and Paradiso, S. (2013). Recognition of emotion from body language among patients with unipolar depression. *Psychiatry Res.* 209, 40–49. doi: 10.1016/j.psychres.2013.03.001
- Maurage, P., Campanella, S., Philippot, P., Charest, I., Martin, S., and de Timary, P. (2009). Impaired emotional facial expression decoding in alcoholism is also present for emotional prosody and body postures. *Alcohol Alcohol.* 44, 476–485. doi: 10.1093/alcac/agg037
- Maurage, P., Campanella, S., Philippot, P., Pham, T. H., and Joassin, F. (2007). The crossmodal facilitation effect is disrupted in alcoholism: a study with emotional stimuli. *Alcohol Alcohol.* 42, 552–559. doi: 10.1093/alcac/agm134
- Maurage, P., Joassin, F., Pesenti, M., Grandin, C., Heeren, A., Philippot, P., et al. (2013). The neural network sustaining crossmodal integration is impaired in alcohol-dependence: an fMRI study. *Cortex* 49, 1610–1626. doi: 10.1016/j.cortex.2012.04.012
- Maurage, P., Philippot, P., Joassin, F., Pauwels, L., Pham, T., Prieto, E. A., et al. (2008). The auditory-visual integration of anger is impaired in alcoholism: an event-related potentials study. *J. Psychiatry Neurosci.* 33, 111–122.
- Maurage, P., Rombaux, P., and de Timary, P. (2014). Olfaction in alcohol-dependence: a neglected yet promising research field. *Front. Psychol.* 4:1007. doi: 10.3389/fpsyg.2013.01007
- Mogg, K., Philippot, P., and Bradley, B. P. (2004). Selective attention to angry faces in clinical social phobia. *J. Abnorm. Psychol.* 113, 160–165. doi: 10.1037/0021-843x.113.1.160
- Mothes-Lasch, M., Mentzel, H. J., Miltner, W. H. R., and Straube, T. (2011). Visual attention modulates brain activation to angry voices. *J. Neurosci.* 31, 9594–9598. doi: 10.1523/JNEUROSCI.6665-10.2011
- Naranjo, C., Kornreich, C., Campanella, S., Noël, X., Vandriette, Y., Gillain, B., et al. (2011). Major depression is associated with impaired processing of emotion in music as well as in facial and vocal stimuli. *J. Affect. Disord.* 128, 243–251. doi: 10.1016/j.jad.2010.06.039
- Paulmann, S., and Pell, M. D. (2010). Contextual influences of emotional speech prosody on face processing: how much is enough? *Cogn. Affect. Behav. Neurosci.* 10, 230–242. doi: 10.3758/CABN.10.2.230
- Paulmann, S., Titone, D., and Pell, M. D. (2012). How emotional prosody guides your way: evidence from eye movements. *Speech Commun.* 54, 92–107. doi: 10.1016/j.specom.2011.07.004
- Pell, M. D. (2005a). Nonverbal emotion priming: evidence from the 'facial affect decision task'. *J. Nonverbal. Behav.* 29, 45–73. doi: 10.1007/s10919-004-0889-8
- Pell, M. D. (2005b). Prosody-face interactions in emotional processing as revealed by the facial affect decision task. *J. Nonverbal Behav.* 29, 193–215. doi: 10.1007/s10919-005-7720-z
- Peschard, V., Philippot, P., Joassin, F., and Rossignol, M. (2013). The impact of the stimulus features and task instructions on facial processing in social anxiety: an ERP investigation. *Biol. Psychol.* 93, 88–96. doi: 10.1016/j.biopsycho.2013.01.009
- Philippot, P., and Douilliez, C. (2005). Social phobics do not misinterpret facial expression of emotion. *Behav. Res. Ther.* 43, 639–652. doi: 10.1016/j.brat.2004.05.005
- Pishyar, R., Harris, L. M., and Menzies, R. G. (2004). Attentional bias for words and faces in social anxiety. *Anxiety Stress Coping* 17, 23–36. doi: 10.1080/10615800310001601458
- Quadflieg, S., Mohr, A., Mentzel, H. J., Miltner, W. H. R., and Straube, T. (2008). Modulation of the neural network involved in the processing of anger prosody: the role of task-relevance and social phobia. *Biol. Psychol.* 78, 129–137. doi: 10.1016/j.biopsycho.2008.01.014
- Quadflieg, S., Wendt, B., Mohr, A., Miltner, W. H. R., and Straube, T. (2007). Recognition and evaluation of emotional prosody in individuals with generalized social phobia: a pilot study. *Behav. Res. Ther.* 45, 3096–3103. doi: 10.1016/j.brat.2007.08.003
- Rapee, R. M., and Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behav. Res. Ther.* 35, 741–756. doi: 10.1016/s0005-7967(97)00022-3
- Rigoulot, S., and Pell, M. D. (2012). Seeing emotion with your ears: emotional prosody implicitly guides visual attention to faces. *PLoS One* 7:e30740. doi: 10.1371/journal.pone.0030740
- Rossignol, M., Philippot, P., Bissot, C., Rigoulot, S., and Campanella, S. (2012). Electrophysiological correlates of enhanced perceptual processes and attentional capture by emotional faces in social anxiety. *Brain Res.* 1460, 50–62. doi: 10.1016/j.brainres.2012.04.034
- 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
- Schofield, C. A., Coles, M. E., and Gibb, B. E. (2007). Social anxiety and interpretation biases for facial displays of emotion: emotion detection and ratings of social cost. *Behav. Res. Ther.* 45, 2950–2963. doi: 10.1016/j.brat.2007.08.006
- Schofield, C. A., Johnson, A. L., Inhoff, A. W., and Coles, M. E. (2012). Social anxiety and difficulty disengaging threat: evidence from eye-tracking. *Cogn. Emot.* 26, 300–311. doi: 10.1080/02699931.2011.602050
- Schulz, C., Mothes-Lasch, M., and Straube, T. (2013). Automatic neural processing of disorder-related stimuli in social anxiety disorder: faces and more. *Front. Psychol.* 4:282. doi: 10.3389/fpsyg.2013.00282
- Staugaard, S. R. (2010). Threatening faces and social anxiety: a literature review. *Clin. Psychol. Rev.* 30, 669–690. doi: 10.1016/j.cpr.2010.05.001
- Van Rhenen, T. E., and Rossell, S. L. (2013). Is the non-verbal behavioural emotion-processing profile of bipolar disorder impaired? A critical review. *Acta Psychiatr. Scand.* 128, 163–178. doi: 10.1111/acps.12125
- Vroomen, J., Driver, J., and de Gelder, B. (2001). Is cross-modal integration of emotional expressions independent of attentional resources? *Cogn. Affect. Behav. Neurosci.* 1, 382–387. doi: 10.3758/cabn.1.4.382

- Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends Cogn. Sci.* 9, 585–594. doi: 10.1016/j.tics.2005.10.011
- Vuilleumier, P., Armony, J. L., and Dolan, R. J. (2004). “Reciprocal links between emotion and attention,” in *Human Brain Function*, eds R. S. J. Frackowiak, K. J. Friston, C. D. Frith, R. J. Dolan, C. J. Price, S. Zeki, et al. (San Diego: Academic Press), 419–444.

Conflict of Interest Statement: All authors report no competing financial interests or potential conflicts of interest, and no financial relationships with commercial interests. The authors are funded by the Belgian Fund for Scientific Research (F.N.R.S., Belgium), but this fund did not exert any editorial direction or censorship on any part of this article.

Received: 12 July 2013; accepted: 30 April 2014; published online: 15 May 2014.

Citation: Peschard V, Maurage P and Philippot P (2014) Towards a cross-modal perspective of emotional perception in social anxiety: review and future directions. *Front. Hum. Neurosci.* 8:322. doi: 10.3389/fnhum.2014.00322

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

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