J. Behav. Ther. & Exp. Psychiat. 57 (2017) 206–211

Contents lists available at ScienceDirect

Journal of Behavior Therapy and Experimental Psychiatry

journal homepage: www.elsevier.com/locate/jbtep

Overestimation of threat from neutral faces and voices in social anxiety

Virginie Peschard*, Pierre Philippot

Laboratory for Experimental Psychopathology, Psychological Sciences Research Institute, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

ARTICLE INFO

Article history:
Received 28 February 2017
Received in revised form
8 June 2017
Accepted 21 June 2017
Available online 23 June 2017

Keywords:
Social anxiety
Emotional facial expressions
Emotional prosody
Misinterpretation
Audio-visual integration

ABSTRACT

Background and objectives: Social anxiety (SA) is associated with a tendency to interpret social information in a more threatening manner. Most of the research in SA has focused on unimodal exploration (mostly based on facial expressions), thus neglecting the ubiquity of cross-modality. To fill this gap, the present study sought to explore whether SA influences the interpretation of facial and vocal expressions presented separately or jointly.
Methods: Twenty-five high socially anxious (HSA) and 29 low socially anxious (LSA) participants completed a forced two-choice emotion identification task consisting of angry and neutral expressions conveyed by faces, voices or combined faces and voices. Participants had to identify the emotion (angry or neutral) of the presented cues as quickly and precisely as possible.
Results: Our results showed that, compared to LSA, HSA individuals show a higher propensity to misattribute anger to neutral expressions independent of cue modality and despite preserved decoding accuracy. We also found a cross-modal facilitation effect at the level of accuracy (i.e., higher accuracy in the bimodal condition compared to unimodal ones). However, such effect was not moderated by SA.
Limitations: Although the HSA group showed clinical cut-off scores at the Liebowitz Social Anxiety Scale, one limitation is that we did not administer diagnostic interviews. Upcoming studies may want to test whether these results can be generalized to a clinical population.
Conclusions: These findings highlight the usefulness of a cross-modal perspective to probe the specificity of biases in SA.

Social anxiety (SA) is characterized by fear and avoidance of social or performance situations in which negative evaluation by others may occur. Even at subclinical levels, SA is associated with adverse outcomes, including significant psychosocial impairments, reduced quality of life and increased comorbidity risk (Fehm, Beesdo, Jacobi, & Fiedler, 2008; Filho et al., 2010). Cognitive models of SA (Clark & Wells, 1995; Rapee & Heimberg, 1997) suggest that information-processing biases play a key role in the maintenance of SA. In particular, SA is related to a tendency to interpret benign or ambiguous social information in a more threatening manner. Consistent with these predictions, several studies using socially relevant verbal stimuli (e.g., descriptions of social scenarios) have shown that high socially anxious (HSA) individuals interpret ambiguous social information more negatively than low socially anxious (LSA) ones (for a review see Mobini, Reynolds, & MacKintosh, 2013).

Another line of research has explored interpretation biases by means of nonverbal information, with a main focus on emotional facial expressions (for reviews see Gilboa-Schechtman & Shachar-Lavie, 2013; Staugaard, 2010). Most studies showed that HSA individuals do not differ from LSA individuals in their ability to accurately decode facial expressions (Bell et al., 2011; Heuer, Lange, Isaac, Rinck, & Becker, 2010; Philippot & Douilliez, 2005). Other works have reported that HSA individuals may present an increased sensitivity to threat-related—morphed or blended—facial expressions as compared to LSA individuals (Gutiérrez-Garcia & Calvo, 2017; Gutiérrez-García & Calvo, 2014; Joormann & Gotlib, 2006; Yoon, Yang, Chong, & Oh, 2014). It has also been shown that ambiguous emotional faces are more likely to be misclassified as expressing threatening emotions (Bell et al., 2011; Heuer et al., 2010; Yoon et al., 2014) or to be interpreted in a less benign way

* Corresponding author. Present address: Laboratory of Emotion processing, Department of Psychology and the Gonda Brain Science Center, Bar-Ilan University, Ramat-Gan, 52900, Israel.
E-mail address: peschardv@gmail.com (V. Peschard).

http://dx.doi.org/10.1016/j.jbtep.2017.06.003
0005-7916/

(Gutiérrez-Garcia & Calvo, 2016; Gutiérrez-García & Calvo, 2014) by HSA individuals relative to LSA ones. Furthermore, higher degrees of social anxiety have been related to more misclassifications of emotionally neutral faces as angry (Bell et al., 2011; Gutiérrez-Garcia & Calvo, 2017; Mohlman, Carmin, & Price, 2007). Relatedly, studies using implicit tasks (Yoon & Zinbarg, 2007, 2008) suggest that HSA, compared to LSA, individuals interpret neutral faces in a more negative direction. Altogether, these findings indicate that HSA and LSA individuals are unlikely to differ in their general ability to accurately decode facial expressions but they may differ in their sensitivity to threat-related facial cues, leading to their overestimation and misclassification.

The extensive exploration of emotional facial expressions (EFE) processing in SA contrasts with the limited number of studies on nonverbal auditory cues, such as emotional prosody. One of the first behavioral explorations (Quadflieg, Wendt, Mohr, Miltner, & Straube, 2007) revealed that individuals with generalized social phobia more frequently labelled fearful and sad emotional prosodies correctly but were less accurate for happy prosodies. Surprisingly, the two groups did not differ in their identifications of neutral, angry or disgusted prosody nor in their rating of valence or arousal for any prosody. In a follow-up fMRI study asking social phobia and control participants to label the emotion and gender of angry or neutral prosodies (Quadflieg, Mohr, Mentzel, Miltner, & Straube, 2008), limbic and fronto-temporal activation was greater in response to angry relative to neutral prosodies. While control participants exhibited increased activation in their right orbitofrontal cortex to angry relative to neutral prosody if the emotional prosody was task-relevant, individuals with generalized social phobia had this activation pattern, regardless of whether the prosodies were task-relevant. Yet, there were no group differences with respect to accuracy and speed. Combined, these results suggest that social phobic individuals may process emotional prosody differently than their non-anxious peers. However, given the small number of studies, future investigations are clearly needed.

Two significant limitations of the aforementioned literature include the prevailing use of EFEs and one sensory modality at a time. Notwithstanding the fact that any channel can be informative on its own, facial and vocal emotional expressions hardly ever occur in isolation in everyday situations. The integration of these co-occurring sources of information into a unified coherent percept facilitates social comprehension. More specifically, several studies conducted in the general populations (Collignon et al., 2008; Kreifelts, Ethofer, Grodd, Erb, & Wildgruber, 2007) have demonstrated that (congruent) audiovisual emotional cues, such as pairs of facial expressions and emotional prosodies, are more accurately and/or quickly identified as compared to either unimodal stimuli. To our knowledge, there is currently no study examining the effect of SA on the integration of facial and vocal emotional cues. This is surprising considering that perception and production of emotions are routinely based on multiple sensory channels. Recent reviews (Gilboa-Schechtman & Shachar-Lavie, 2013; Peschard, Maurage, & Philippot, 2014; Schulz, Mothes-Lasch, & Straube, 2013) have highlighted the potential value of using a cross-modal perspective in furthering our understanding of SA-related processing biases. In particular, the inclusion of cross-modal cues – especially face and voice – may shed new light on how socially anxious individuals interpret interpersonal information in situations that are closer to real-life. So far, one cannot rule out the possibility that cross-modal interactions between, and integration of, emotional faces and voices could be moderated by SA in the presence (or absence) of biases in the unimodal (face or voice) processing of emotions.

The present study therefore sought to extend the research scope by examining the effect of SA on the interpretation of facial and vocal expressions presented either in isolation or in combination. To this end, participants completed a forced two-choice emotion identification task consisting of angry and neutral expressions conveyed by faces (i.e., unimodal face condition), voices (i.e., unimodal voice condition) or combined faces and voices (bimodal face-voice condition). Participants were asked to identify the emotion (angry or neutral) of the presented cues as quickly and precisely as possible. The outcome measures included response accuracy, reaction times (RTs) for correct responses, and misclassification rates.

Given the results of previous studies examining EFEs decoding (Staugaard, 2010) and the few available data on emotional prosody processing in SA (Quadflieg et al., 2008, 2007), we did not expect to observe a substantial difference between HSA and LSA individuals with respect to decoding accuracy. However, based on prior face-based investigations (Bell et al., 2011; Gutiérrez-Garcia & Calvo, 2017; Mohlman et al., 2007), we predicted that HSA individuals would be more prone to misclassify (ambiguous) neutral faces as expressing anger. Lastly, we expected to identify the classical cross-modal facilitation effect for all participants (e.g., Collignon et al., 2008), such as mirrored by higher accuracy and/or faster RTs in the bimodal face-voice condition as compared to both unimodal conditions. As this study was the first to explore audio-visual processing as a function of SA, we did not formulate hypotheses regarding group differences.

1. Method

1.1. Participants

Sixty-one French-speaking volunteers (51 women, 10 men; mean age = 22.77, SD = 4.87) took part in the experiment. They were selected from a pool of 195 students studying at the University, based on their scores on the French validated version of Liebowitz Social Anxiety Scale (Liebowitz, 1987). HSA individuals were defined as those scoring 80 or more on the LSAS and LSA individuals were defined as those scoring 45 or below. To ensure that the screening procedure was effective, participants completed the LSAS again on the day of the experiment. Six participants had to be excluded because they did not meet the LSAS criteria anymore. In addition, one participant was discarded from the analyses due to an abnormally low level of accuracy (see data preparation), so that the final sample consisted of 54 participants (25 participants in the HSA group and 29 participants in the LSA group). Each participant reported normal auditory perception and normal or corrected-to-normal vision. The study was approved by the ethical committee of the psychology department.

1.2. Questionnaires

1.2.1. Social anxiety scales

In addition to the LSAS, participants completed the validated French version of the following self-report questionnaires: the Fear of Negative Evaluation (FNE, Watson & Friend, 1969), the short version of the Personal Report of Confidence as Speaker (PRCS, Paul, 1966), and the Self-Beliefs Related to Social Anxiety Scale (SBSA, Wong & Moulds, 2009). The LSAS is a 24-item scale that assesses anxiety and avoidance of social interaction and performance situations. The FNE is a 30-item questionnaire that measures a person's apprehension about being negatively evaluated. The PRCS is a 12-item questionnaire that measures a person's fear of public speaking. Finally, the SBSA is a 15-item self-report questionnaire assessing the three categories of self-beliefs postulated in Clark and Wells' (1995) model of social anxiety. The internal reliabilities of the LSAS, FNE, PRCS, and SBSA were 0.98, 0.78, 0.85 and 0.92 respectively.