Shorter communication

Recognition and evaluation of emotional prosody in individuals with generalized social phobia: A pilot study

Susanne Quadflieg\textsuperscript{a}, Beate Wendt\textsuperscript{b}, Alexander Mohr\textsuperscript{a}, Wolfgang H.R. Miltner\textsuperscript{a}, Thomas Straube\textsuperscript{a,\ast}

\textsuperscript{a}Friedrich-Schiller-University Jena, Germany
\textsuperscript{b}Leibniz Institute for Neurobiology, Magdeburg, Germany

Received 18 September 2006; received in revised form 22 June 2007; accepted 6 August 2007

Abstract

Studies using facial emotional expressions as stimuli partially support the assumption of biased processing of social signals in social phobia. This pilot study explored for the first time whether individuals with social phobia display a processing bias towards emotional prosody. Fifteen individuals with generalized social phobia and fifteen healthy controls (HC) matched for gender, age, and education completed a recognition test consisting of meaningless utterances spoken in a neutral, angry, sad, fearful, disgusted or happy tone of voice. Participants also evaluated the stimuli with regard to valence and arousal. While these ratings did not differ significantly between groups, analysis of the recognition test revealed enhanced identification of sad and fearful voices and decreased identification of happy voices in individuals with social phobia compared with HC. The two groups did not differ in their processing of neutral, disgust, and anger prosody.

Keywords: Emotional facial expressions; Emotional prosody; Information-processing bias; Social phobia

Introduction

Etiological models as well as an extensive body of research suggest that the maintenance of social phobia might be due to cognitive biases in the processing of social information (e.g. Clark & Wells, 1995; Heinrichs & Hofmann, 2001). In particular, social anxiety has been associated with an increased sensitivity towards signals of negative evaluation by others, even when these signals are ambiguous (Stopa & Clark, 2000; Yoon & Zimbarg, 2007). Consistent with this notion, hypersensitivity to negative social signals in social phobia has been demonstrated across studies of attention, memory, interpretation, and/or judgment (for a review, see Heinrichs & Hofmann, 2001), by means of changed explicit emotional evaluation of stimuli (Straube, Kolassa, Mentzel, & Miltner, 2004; Straube, Mentzel, & Miltner, 2005), and/or by increased brain activation (e.g. Amir et al., 2005; Phan, Fitzgerald, Nathan, & Tancer, 2006; Stein, Goldin, Sareen, Zorrilla, & Brown, 2002; Straube et al., 2004, 2005).

\textsuperscript{\ast}Corresponding author.

E-mail address: straube@biopsy.uni-jena.de (T. Straube).

0005-7967/S - see front matter \textcopyright 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.brat.2007.08.003
Importantly, the majority of studies on social phobia have utilized emotional facial expressions (EFE) as stimuli conveying negative social signals. For example, functional neuroimaging studies showing increased brain activation to facial expressions of anger, contempt, disgust, and fear in socially phobic compared with nonphobic individuals support the assumption of enhanced processing of negative social signals in social phobia (Amir et al., 2005; Phan et al., 2006; Stein et al., 2002; Straube et al., 2004, 2005). Behavioral studies also indicate that social phobic individuals show increased attentional (Eastwood et al., 2005; Mogg, Philippot, & Bradley, 2004; Pishyar, Harris, & Menzies, 2004) and biased memory processing (Coles & Heimberg, 2005; Foa, Gilboa-Schechtman, Amir, & Freshman, 2000) of negative EFE (but see Chen, Ehlers, Clark, & Mansell, 2002; Pérez-López & Woody, 2001, for an opposing view) in contrast to nonphobic controls. Furthermore, enhanced recognition of negative facial cues in socially anxious compared to less anxious controls has been observed (Joormann & Gotlib, 2006; Winton, Clark, & Edelmann, 1995; but see Philippot & Douilliez, 2005). Remarkably though, most studies have failed to find any evidence that individuals with social phobia differ from normal controls in the explicit evaluation of negative EFE on dimensions such as valence or arousal (Amir et al., 2005; Merckelbach, Van Hout, Van den Hout, & Mersch, 1989; Philippot & Douilliez, 2005; Stein et al., 2002; but see Straube et al., 2004, 2005). Moreover, findings obtained with negative EFE have recently been complemented by studies showing that positive EFE can also elicit biased processing in social phobia (Phan et al., 2006; Straube et al., 2005). Although socially phobic individuals seem to evaluate happy facial expressions as pleasant (Philippot & Douilliez, 2005) and even more arousing than nonphobic controls (Straube et al., 2005), they take longer to recognize happy EFE than normal controls (Silvia, Allan, Beauchamp, Maschauer, & Workman, 2006), and lack a positivity bias towards happy faces shown by healthy individuals (Coles & Heimberg, 2005, for memory bias; Eastwood et al., 2005; Pishyar et al., 2004, for attentional bias).

Surprisingly, assumptions about cognitive biases towards the processing of social cues in social phobia have hardly been investigated with expressive social signals other than facial expressions. One should expect, however, to see disorder-associated cognitive biases towards such cues across a range of stimuli and modalities. For example, it is unknown whether a biased processing of emotional signals in social phobia is also displayed in response to prosodic components of human speech. Due to typical acoustic patterns based on pitch contour, fundamental frequency, tempo, loudness, and stress, a reliable recognition of emotions from speech can be accomplished (Banse & Scherer, 1996). In consequence, prosodic features can reveal a speaker’s emotional state and intention towards others and provide powerful signals of social approval and disapproval. Recently, it has been emphasized that the recognition of emotional prosody is impaired in various neurological and psychiatric disorders (see Bostanov & Kotchoubey, 2004). Psychopathic individuals, for example, have been observed to be restricted in their recognition of fearful emotional prosody compared with healthy controls (HC) (Blair et al., 2002). Depressive individuals, in contrast, have been shown to suffer from a general impairment of prosody recognition across different types of emotional categories (Luck & Dowrick, 2004). Although of high relevance, the response of individuals with social phobia towards cues of emotional prosody has not yet been the focus of investigations. Rather, a review article on cognitive biases in social phobia has emphasized the fact that studies on auditory processing biases in social phobia are seriously lacking (Heinrichs & Hofmann, 2001). Thus, it remains to be examined whether attentional, mnemonic, and recognition biases, or changed explicit valence and arousal ratings towards emotional prosody may be found in individuals with social phobia compared with HC.

Therefore, this pilot study aimed to explore the processing of emotional prosody in individuals with social phobia. Specifically, we investigated whether individuals suffering from generalized social phobia (GSP) display a recognition bias for emotional voices compared with HC. According to previous work on the presence of a recognition bias for facial expressions in GSP, we examined whether individuals with GSP would show improved recognition of negative prosodic stimuli. Based on current theories proposing an increased sensitivity for social signals of rejection and disapproval in social phobia in particular, we expected that such a recognition bias should be found especially for anger prosody. Furthermore, we explored whether individuals with GSP would evaluate emotional prosody differently with regard to arousal and valence as compared with HC. We predicted that individuals with GSP would experience negative—especially anger—prosody as more negative and more arousing than HC.
Method

Participants

Fifteen Caucasian participants diagnosed with GSP (SP) and fifteen Caucasian volunteers (HC) free of any psychiatric symptoms (besides one individual who suffered from a mild phobia of heights) took part in the experiment. SP and HC were matched for age [HC, M = 23.93 vs. SP, M = 23.27, t(28) = .35, p > ns], gender (seven men and eight women per group), and level of education (student population only). All participants provided written informed consent for this study, which was approved by the ethics committee of the University of Jena. For participation in the study, 6 €/h was paid to all individuals.

Participants were recruited by public announcement to the university student population and were required to be right-handed as determined by the Edinburgh Handedness Inventory (Oldfield, 1971). All participants underwent a structured psychiatric clinical interview for DSM-IV Axis I disorders by an experienced interviewer (SCID, Wittchen, Zaudig, & Fydrich, 1997). Exclusion criteria were (1) a diagnosis of panic disorder or agoraphobia, obsessive–compulsive disorder, current alcohol or substance abuse, psychotic disorder or dementia; (2) a history of seizures or head injury with loss of consciousness; (3) a severe uncontrollable medical condition; (4) the use of any psychotropic medication during the preceding 6 months; and (5) primary or secondary major depression. Individuals with GSP were diagnosed as primary social phobics of the generalized subtype according to the DSM-IV criteria. About 10% of all invited and/or diagnosed subjects fulfilled the inclusion criteria. The present 15 participants suffered from moderate to severe GSP, and considered themselves anxious in most social situations, including at least one performance situation and one situation of social interaction. Participants with GSP did not meet the criteria for any other Axis I diagnosis but specific phobia (which happened to be the case in 5 participants).

In addition to the interview, all individuals completed two self-report measures: the Social Phobia Inventory (SPIN, German version, Stangier & Steffens, 2001) and the Beck Depression Inventory (BDI, German version, Hautzinger, Bailer, Worall, & Keller, 1995). A between-group comparison regarding the SPIN scores confirmed that SP expressed stronger social fears than HC [38.73 (5.42) vs. 9.00 (5.48), t(28) = 14.95, p < .05]. On the BDI, SP scored higher than HC [11.2 (5.28) vs. 4.13 (4.66), t(28) = 3.89, p < .05], but far below clinical cut-off scores. In order to ensure comparability of our sample with earlier studies, individuals with GSP were also administered the Liebowitz Social Anxiety Scale (LSAS, German version, Stangier & Heidenreich, 1996). Scores ranged from 41 to 89, with an average score of 62.87.

Stimuli, tasks and dependent measures

Stimuli for the emotional prosody recognition task were taken from the “Magdeburger Prosodie-Korpus” constructed and validated by Wendt and Scheich (2002). Stimuli comprised 384 pseudo-words without semantic meaning spoken in a happy, sad, fearful, angry, disgusted, or neutral tone of voice. For each of the six emotional categories 32 pseudo-words were spoken by an actress and an actor, resulting in 64 pseudo-word trials per emotional category. Although utterances with different kinds of emotional prosody were presented in a randomized fashion, the order of presentation was fixed and therefore identical for all participants.

Participants were instructed to listen to the pseudo-words and to indicate on a questionnaire which of the six emotional categories an utterance expressed. For a subset of stimuli, participants were prompted to complete the recognition task first and, additionally, to evaluate the utterances according to their valence and arousal afterwards. The evaluation scales provided for these trials displayed seven grades of valence (1 = very unpleasant to 7 = very pleasant) and arousal (1 = not arousing to 7 = very arousing). These additional evaluation ratings were required on twenty trials per emotional category (only 19 instead of 20 items were prompted in the case of happy utterances due to an error in protocol preparation). All participants received the evaluation prompt on identical trials. The task was split into 4 blocks of 96 trials each. The blocks were equalized on the number of stimuli from each emotional category. In consideration of the test length and possible exhaustion phenomena, participants were prompted for evaluation judgments on 36 trials in the first two blocks and on 24 trials in the third and fourth blocks. Since the prosodic stimuli for which evaluation judgments were requested had been selected randomly, the evaluation prompts occurred in varying intervals
throughout the blocks. Between each block, participants had a break of approximately 5 min. Time between pseudo-word presentations was 8 s by default, but participants were allowed to ask to pause the CD while completing valence and arousal ratings. Time required beyond the default setting was not measured. Stimuli were presented with a portable Panasonic CD-player.

In order to become familiar with the task, participants were exposed to 12 training trials (2 per emotional category) not included in the experiment proper. During these practice trials no feedback was given, in order to avoid participants answering according to what they have “learned” during practice to be considered the accurate answer. Rather, participants were assured to respond according to their immediate, idiosyncratic impression. Participants themselves adjusted the level of loudness of the voices during these trials.

Data were analyzed by means of analysis of variance (ANOVA), correlational analysis, and t-tests using SPSS (Version 12; SPSS, Inc., Chicago). Greenhouse–Geisser corrections were used when appropriate to correct violations of sphericity. To strike a balance between type 1 and 2 errors in the context of small sample size, a probability value of .05 was considered as statistically significant in the overall and follow-up analysis. All data are expressed as mean (M) ± standard deviation (SD).

Results

Recognition rates

For each individual, the percentage of correct answers of the 64 trials in each emotional category was computed. Table 1 depicts the two groups’ average scores of correctly identified trials for each emotional category included in the experiment. A 2 (groups: SP vs. HC) by 6 (emotional prosody: recognition rates for neutral, happy, angry, sad, fearful, and disgusted utterances) repeated-measures ANOVA was conducted with clinical status as the between-subjects variable and emotion recognition across categories as the within-subject variable. A significant main effect of emotion \(F(3.31, 92.55) = 18.11, p < .05\] and the marginal significant main effect of group \(F(1, 28) = 3.98, p = .056\] were qualified by the expected significant group \(C2\) emotion interaction \(F(3.31, 92.55) = 3.46, p < .05\].

To clarify for which emotions significant differences between groups existed, follow-up univariate analyses of variance for the 6 emotional categories were conducted. Individuals with GSP identified significantly less frequently utterances with happy prosody \(F(1, 28) = 6.08, p < .05\], but recognized more frequently utterances with fearful \(F(1, 28) = 5.23, p < .05\] and sad \(F(1, 28) = 5.37, p < .05\] emotional prosody as compared with HC. SP and HC did not differ significantly in their recognition of utterances with neutral \(F(1, 28) = 2.92, \text{ns}\], angry \(F(1, 28) = .19, \text{ns}\], or disgusted \(F(1, 28) = .88, \text{ns}\) prosody.

Furthermore, correlational analyses were carried out to examine the relations between the recognition rates and scores of the BDI. The only significant correlation was found for HC: Utterances with angry prosody were identified significantly more frequently the less pronounced symptoms of depression were \(r = -.56, p < .05\]. In addition, there was a tendency of an inverse relation between recognizing happy prosody and BDI scores in both groups (HC, \(r = -.45, p = .10\); SP, \(r = -.42, p = .12\]).

Given those correlative results and the fact that groups differed significantly in their BDI scores, ANOVAS for happy, sad, anger, and fearful prosody with the BDI scores as a covariate were computed. These analyses showed that the observed significant difference in recognition rates between groups remained stable for sad

<table>
<thead>
<tr>
<th>Happy</th>
<th>Neutral</th>
<th>Sad</th>
<th>Disgusted</th>
<th>Fearful</th>
<th>Angry</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>HC</td>
<td>75.63</td>
<td>67.61</td>
<td>69.27</td>
<td>65.94</td>
<td>71.88</td>
<td>94.79</td>
</tr>
<tr>
<td></td>
<td>(13.53)</td>
<td>(19.44)</td>
<td>(10.10)</td>
<td>(14.10)</td>
<td>(8.90)</td>
<td>(4.03)</td>
</tr>
<tr>
<td>SP</td>
<td>64.79</td>
<td>77.92</td>
<td>79.27</td>
<td>70.62</td>
<td>80.21</td>
<td>93.86</td>
</tr>
<tr>
<td></td>
<td>(10.33)</td>
<td>(12.98)</td>
<td>(13.30)</td>
<td>(13.26)</td>
<td>(10.96)</td>
<td>(7.27)</td>
</tr>
</tbody>
</table>

Note: Standard deviations are given in parentheses.
[F(1, 27) = 4.47, p < .05] and fearful [F(1, 27) = 4.46, p < .05] but not for happy [F(1, 27) = .50, ns] prosody. Again, there was no significant effect for anger prosody.

To clarify where the significant differences in recognition rates for fearful, sad, and happy utterances came from, erroneous trials were examined in more detail for these three emotions. Both groups made similar errors of misidentification in response to the stimuli. With regard to fearful prosody, HC and individuals with GSP similarly misidentified fearful utterances most often as sad [HC = 15.42% vs. SP = 10.63%; t(28) = 1.35, p > .05] or as happy prosody [HC = 6.98% vs. SP = 4.06%; t(28) = 1.63, p > .05]. With regard to sad prosody, HC and SP misidentified sad trials most often as expressing neutral or happy emotions. However, while HC and individuals with GSP did not differ in their misidentification of sad utterances as being neutral [HC: 17.40% vs. SP: 14.58%; t(28) = .63, p > .05], individuals with GSP were less likely to misidentify them as happy compared with HC [HC: 7.71% vs. SP: 2.61%; t(28) = 2.26, p < .05]. Furthermore, happy utterances were most often misidentified by both groups as neutral, however, more so by socially phobic compared with nonphobic individuals (HC: 20.84% vs. SP: 32.40%; t(28) = -2.50, p < .05).

Valence and arousal ratings

For each individual, the average rating on arousal and valence scores per prosody was computed, regardless of the participant’s decision on the recognition task for these trials. Therefore, each participant had one aggregated value for arousal and valence for each of the six emotional categories. Missing values constituted less than 5.5% and were substituted by the average value. Table 2 displays the mean values of valence and arousal ratings for each emotional category and for each group.

For valence and arousal ratings, two 2 (group: SP vs. HC) by 6 (emotional prosody: ratings for neutral, happy, angry, sad, fearful, and disgusted utterances) repeated-measures ANOVA were computed separately. Each revealed only a main effect of emotion [valence, F(5, 24) = 41.5, p < .05; arousal, F(5, 24) = 12.70, p < .05]. In sum, the profiles of evaluation were modulated by the type of emotion of the presented stimuli, but not by participants’ clinical status. Correlational analysis on the evaluative ratings and participants’ BDI scores revealed only one significant result: HC rated neutral utterances less positive on the valence scale the more symptoms of depression they expressed in the BDI. Including the BDI score as covariate in the ANOVA did not affect the nonsignificant outcome regarding group differences in valence and arousal ratings.

Discussion

This pilot study explored how individuals with GSP recognize and evaluate emotional prosody. Compared with HC, SP more often correctly identified fearful and sad voices, but displayed a decreased recognition of happy voices. When participants’ BDI scores were considered as a covariate in the analysis, significant differences in prosody recognition remained for fearful and sad voices across groups. Furthermore, the two groups did not differ in the recognition of neutral, disgust, and anger prosody, and showed no differences with regard to valence and arousal ratings for any prosody. These findings suggest that differences in the

Table 2
Mean values of valence and arousal ratings for each emotional category according to groups

<table>
<thead>
<tr>
<th></th>
<th>Happy</th>
<th>Neutral</th>
<th>Sad</th>
<th>Disgusted</th>
<th>Fearful</th>
<th>Angry</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Valence</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HC</td>
<td>4.77</td>
<td>3.89</td>
<td>3.38</td>
<td>2.99</td>
<td>2.80</td>
<td>2.65</td>
</tr>
<tr>
<td></td>
<td>(.44)</td>
<td>(.44)</td>
<td>(.58)</td>
<td>(.88)</td>
<td>(.46)</td>
<td>(.176)</td>
</tr>
<tr>
<td>SP</td>
<td>4.58</td>
<td>3.94</td>
<td>3.31</td>
<td>2.74</td>
<td>2.75</td>
<td>2.36</td>
</tr>
<tr>
<td></td>
<td>(.44)</td>
<td>(.44)</td>
<td>(.94)</td>
<td>(.65)</td>
<td>(.73)</td>
<td>(.88)</td>
</tr>
<tr>
<td><strong>Arousal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HC</td>
<td>1.90</td>
<td>1.60</td>
<td>1.97</td>
<td>2.51</td>
<td>2.66</td>
<td>3.56</td>
</tr>
<tr>
<td></td>
<td>(.96)</td>
<td>(.81)</td>
<td>(.99)</td>
<td>(1.28)</td>
<td>(1.35)</td>
<td>(1.64)</td>
</tr>
<tr>
<td>SP</td>
<td>1.82</td>
<td>1.58</td>
<td>2.42</td>
<td>2.92</td>
<td>3.11</td>
<td>3.80</td>
</tr>
<tr>
<td></td>
<td>(.73)</td>
<td>(.55)</td>
<td>(1.25)</td>
<td>(1.54)</td>
<td>(1.60)</td>
<td>(1.84)</td>
</tr>
</tbody>
</table>

*Note: Standard deviations are given in parentheses.*
recognition of emotional prosody between socially phobic and nonphobic individuals depend on the kind of emotion expressed in the voice. Since it is possible that the small sample size has compromised the power of the present study to detect further differences between groups, however, the preliminary nature of these results needs to be emphasized.

Although preliminary, the present finding of increased recognition of sad and fearful prosody in social phobia supports previous research which has found biased processing of facial expressions indicating fear (Phan et al., 2006) and sadness (Horley, Williams, Gonsalvez, & Gordon, 2003) in social phobia. Moreover, the present results complement earlier observations that mental disorders can be associated with an altered capacity to comprehend emotional prosody. In contrast to observations of impaired responsiveness associated with psychopathy (Blair et al., 2002) or depression (Luck & Dowrick, 2004), the current findings indicate that a consequence of disorder-associated changes in prosody perception might also be an increased responsiveness towards emotional prosody. However, it remains uncertain whether the observed processing bias is specifically associated with social phobia or can also be demonstrated for other anxiety or psychiatric disorders since expressions of fear and sadness do not specifically imply a social threat or rejection component as would expressions of anger do.

We found no recognition bias in response to utterances expressing anger or disgust in individuals with social phobia. The absence of a significant between-group difference for angry utterances might partially be due to a ceiling effect. The amount of correctly identified utterances expressing anger was above 93% for both groups. Thus, the groups’ high recognition rates left little space for individuals with GSP to exceed the performance of HC. To reduce ceiling effects, future investigations on the processing of emotional prosody in social phobia might use more ambiguous anger stimuli or consider the use of response times instead of recognition rates as the critical measure.

The lack of difference between groups for utterances expressing disgust is difficult to interpret. Although disgust is considered to be a negatively valenced emotion, it is often used for the characterization of contaminated objects instead of people. In consequence, its meaning as a cue of disapproval, rejection, or discomfort in social interactions is controversial. Recently, a functional brain imaging experiment revealed that, compared with HC, individuals with GSP show increased neural processing of facial expressions of disgust (Amir et al., 2005). However, in agreement with our study, the observed enhanced neural activity in SP was not reflected in a more negative evaluation of facial expressions of disgust. Replication of these findings is needed before firm conclusions can be drawn.

Our finding that individuals with GSP made more errors than HC in recognizing happy prosody supports the idea that social phobia might be characterized by a bias with regard to not only negative but also positive emotional cues (Coles & Heimberg, 2005; Eastwood et al., 2005; Phan et al., 2006; Pishyar et al., 2004; Silvia et al., 2006; Straube et al., 2005). A misidentification of positive emotional signals might be associated with decreased experience of social approval in everyday interactions in social phobia. However, the decreased recognition of positive emotional signals in social phobia correlated with participants’ concurrent mild depressive symptoms in the current study. Such a relation was not found for negative emotional signals. Thus, the influence of comorbid depressive symptoms on information processing in social phobia might depend on the valence of the expressed emotions (but see Musa, Lépine, Clark, Mansell, & Ehlers, 2003).

The present study did not reveal any group difference with regard to valence and arousal ratings for any of the emotional categories considered. A lack of difference in explicit evaluation of EFE has been observed before (Merckelbach, Van Hout, Van den Hout, & Mersch, 1989; Philippot & Douilliez, 2005; Stein et al., 2002; but see Straube et al., 2004, 2005). In the current study, the absence of group differences might also be due to a lack of power since the arousal data suggest that SP experienced more arousal than HC to emotional stimuli of negative, but not of neutral or positive, valence. Furthermore, recruitment from a student population together with the need to be medication free might also have contributed to decreased power of the present results. The BDI scores are low for a group that bears a diagnosis of GSP. Therefore, there is a strong need for further studies on the processing of emotional prosody with increased samples derived from clinical settings.

Besides its limitations, this pilot study provides further evidence for biases in emotional processing in social phobia, using prosodic rather than facial information. Future research might not only target the recognition and evaluation of emotional prosody in social phobia, but also examine whether evidence for attentional or
mnemonic biases in prosody perception can be found. Altered processing of facial expressions, emotional prosody, gestures, and/or verbal cues might amplify the development of pathological fears towards embarrassment and rejection by others. Therefore, increasing our understanding of biased processing of emotional cues in social phobia across a wide range of ecologically valid stimuli poses a major challenge.

Acknowledgements

The study was supported by grants from the Deutsche Forschungsgemeinschaft and the federal state of Thuringia and the University of Jena awarded to Thomas Straube and Wolfgang H.R. Miltner.

References


