Perception of emotion in psychiatric disorders: On the possible role of task, dynamics, and multimodality

Patricia Garrido-Vásquez a, Sarah Jessen a b & Sonja A. Kotz a

a Minerva Research Group “Neurocognition of Rhythm in Communication,” Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
b Cluster of Excellence “Languages of Emotion,” Free University of Berlin, Germany

Available online: 03 Oct 2011

To cite this article: Patricia Garrido-Vásquez , Sarah Jessen & Sonja A. Kotz (2011): Perception of emotion in psychiatric disorders: On the possible role of task, dynamics, and multimodality, Social Neuroscience, 6:5-6, 515-536

To link to this article: http://dx.doi.org/10.1080/17470919.2011.620771
Perception of emotion in psychiatric disorders: On the possible role of task, dynamics, and multimodality

Patricia Garrido-Vásquez¹*, Sarah Jessen¹,²*, and Sonja A. Kotz¹

¹Minerva Research Group “Neurocognition of Rhythm in Communication,” Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
²Cluster of Excellence “Languages of Emotion,” Free University of Berlin, Germany

Experimental evidence suggests an impairment in emotion perception in numerous psychiatric disorders. The results to date are primarily based on research using static displays of emotional facial expressions. However, our natural environment is dynamic and multimodal, comprising input from various communication channels such as facial expressions, emotional prosody, and emotional semantics, to name but a few. Thus, one critical open question is whether alterations in emotion perception in psychiatric populations are confirmed when testing patients in dynamic and multimodal naturalistic settings. Furthermore, the impact task demands may exert on results also needs to be reconsidered. Focusing on schizophrenia and depression, we review evidence on how emotions are perceived from faces and voices in these disorders and examine how experimental task demands, stimulus dynamics, and modality may affect study results.

Keywords: Emotion perception; Multimodal; Task effects; Depression; Schizophrenia.

Accurately perceiving others’ emotions is crucial for successful interpersonal interactions and may be a key factor in social deficits observed in many neuropsychiatric disorders. During recent decades, a great deal of evidence on emotion perception in psychiatric populations has accumulated. Even though this research has considerably advanced our understanding of how emotions are perceived in people with pathological mental states, we argue that inconsistencies in the results need to be viewed in light of the experimental procedures applied. As we will show, these procedures can be improved by considering higher ecological validity of stimuli, as, for instance, by using multimodal stimuli (i.e., the combination of face and voice). While an emotional expression in one modality is often enough to determine a person’s emotional state, these information sources normally occur simultaneously, providing seemingly redundant information. However, simultaneous perception of congruent emotional information is by no means redundant, but rather offers a number of benefits to the perceiver as will be outlined in this review. These benefits may be especially high when information-processing resources are limited, as is often apparent in psychiatric disorders. Along these lines, dynamic rather than static stimuli provide a more ecologically valid approach to test emotion perception. Furthermore, implicit rather than explicit task instructions may affect how emotional expressions are perceived as well as reduce confounds due to cognitive task demands.

*These authors contributed equally to the paper.

The authors would like to thank three anonymous reviewers for helpful comments on an earlier version of this manuscript. We also thank Rosie Wallis for carefully checking our English. We gratefully acknowledge the funding of the Canadian Institute of Health Research (CIHR-MPO:62867: Marc D. Pell and S.A.K.) and the German Science Foundation (DFG-FOR-499: S.A.K.).

© 2011 Psychology Press, an imprint of the Taylor & Francis Group, an Informa business

http://dx.doi.org/10.1080/17470919.2011.620771
The present review aims to promote a more naturalistic approach to emotion perception in patient investigations. To this end, we first present a general introduction to underline the benefits offered by multimodal, dynamic stimulus presentations and implicit task instructions. Second, we introduce models of how emotions are perceived from speech and faces. We then exemplify by focusing on two mental disorders, schizophrenia and major depression, and summarize what has been done so far to understand emotion perception in these two disorders. Finally, we draw disorder-specific and general conclusions from the current state of the art in research. The motivation for focusing on these specific disorders is that they are assumed to tap into different stages of emotion processing: While schizophrenia has been shown to already affect early sensory processes (Johnston, Stojanov, Devir, & Schall, 2005), major depression is considered to alter later processing stages (Gotlib & Joormann, 2010). Likewise, modality, dynamics, and task manipulations may have different effects in these two psychiatric populations, as we try to elucidate in the present review. Our literature review concerning schizophrenia and depression is limited to the perception of social emotional cues (i.e., vocal and facial expressions).

**MULTIMODAL EMOTION PROCESSING**

Several studies have shown facilitated emotional information processing when information is simultaneously presented in several modalities (Giard & Peronnet, 1999; Kreifelts, Ethofer, Grodd, Erb, & Wildgruber, 2007; Paulmann, Jessen, & Kotz, 2009). Emotions are recognized faster (Giard & Peronnet, 1999) and more accurately (Giard & Peronnet, 1999; Kreifelts et al., 2007) when simultaneously conveyed by face and voice rather than when presented unimodally.

Furthermore, multimodal perception allows fine-tuning of emotional expression that cannot be achieved unimodally. The voice can have a strong influence on how we perceive the face and vice versa (de Gelder & Vroomen, 2000). If, for instance, an ambiguous facial expression is presented with a happy voice, participants tend to perceive the facial expression as happy, while if the same facial expression is accompanied by sad prosody, the facial expression is also perceived as sad (de Gelder & Vroomen, 2000).

In addition, multimodal perception can ensure that our assessment of someone’s emotional state is clear if one of the modalities fails us. The less reliable information from one modality becomes, the more we benefit from information provided in another modality (Werner & Noppeney, 2010). Consider, for example, being in the dark when you can hardly see your conversational partner but very clearly hear his or her voice.

For these reasons, multisensory perception is an essential mechanism in emotion communication. Yet multisensory perception of emotions is largely understudied in the field of neuropsychiatric disorders. However, it seems crucial to study multisensory emotion perception in patients who suffer from emotional as well as social deficits. It is an open question whether multimodal perception offers the same benefit to these patients as it does to healthy people, or, in contrast, may hamper emotion perception.

**MODELS OF PROCESSING EMOTION FROM SPEECH AND FACES**

One important information source in emotional communication is a person’s voice. Three basic steps have been suggested for the processing of emotional speech (Schirmer & Kotz, 2006): a first sensory processing phase, followed by the integration of acoustic cues to form a salient percept, and finally cognitive processes operating on such percepts. In the first step, auditory input is processed in the primary auditory cortices, projecting to the bilateral superior temporal sulcus (STS). These primary features are then integrated in the superior temporal gyrus as well as sulcus, moving along to anterior portions of the STS. Based on these processes, different steps occur depending on the context, the task, and the exact specification of the stimulus. In emotional speech, emotional prosodic content is accompanied by semantic content. Thus, both types of emotion information need to be compared and aligned. This particular processing step engages the left inferior frontal gyrus. On the other hand, if participants have just to process and label prosodic information, the right inferior frontal gyrus, as well as the orbitofrontal cortex, seem to play an essential role (Schirmer & Kotz, 2006). A similar model was proposed by Wildgruber, Ackermann, Kreifelts, and Ethofer (2006), positing three steps: the extraction of supra-segmental cues, the representation of supra-segmental sequences, and the explicit judgment of emotional information. These steps correspond essentially to the three steps introduced above, and also involve mostly the same brain regions.

Regarding the processing of emotions from facial expressions, Adolphs (2002) proposes a two-step model, dividing the processing into a perception and a recognition part. However, the second step can be further divided into a recognition and a conceptual knowledge aspect, where the recognized emotion is...
associated with previously known information about the person and the general context. In the first, sensory step, simple as well as highly salient features are processed. This encoding encompasses subcortical regions such as the superior colliculus and the pulvinar nucleus of the thalamus as well as primary sensory cortices, such as the striate cortex. Following this first step, more detailed emotion-processing mechanisms that can be subsumed under the term recognition are employed. Relevant facial features are analyzed in visual association cortices. If motion information is contained, the middle and middle superior temporal areas are involved. This information is passed on to the fusiform and the superior temporal cortices, where emotional content and social relevance is processed. The amygdala and the orbitofrontal cortex are involved in then guiding various aspects of further emotion processing. First, feedback is directed back to regions implicated in earlier processing stages, allowing fine-tuned processing of the ongoing input. Second, connections to various cortical regions, such as the somatosensory and the prefrontal cortex, as well as the hippocampus, allow the perceived emotions to be put into emotional context based on previous knowledge. Finally, a simulation of another person’s emotional state is facilitated by connections to motor structures such as the basal ganglia and the frontal operculum, as well as brainstem nuclei, enabling us to “feel with” somebody (Adolphs, 2002). Besides these elaborate emotion-processing pathways, the pulvinar-amygdalar pathway provides a second, faster route for processing emotion information (Adolphs, 2002; Vuilleumier & Pourtois, 2007). Here, coarse stimulus features, in particular, are used to quickly determine potentially dangerous situations in order to allow a fast and adequate response.

Overall, processing of auditory as well as visual information can thus be divided into three essential steps: namely, early feature processing, integration of these features, and finally evaluation of the percept. Different regions in the brain have been proposed to be involved in the integration of auditory and visual emotional information, such as the right posterior STS (Ethofer, Poutois, & Wildgruber, 2006), the right posterior insula (Ethofer et al., 2006), the amygdala (Ethofer et al., 2006; de Gelder, Vroomen, & Pourtois, 2004; O’Doherty, Rolls, & Kringelbach, 2004), and the orbitofrontal cortex (O’Doherty et al., 2004). However, no specific model exists which accounts for the integration of multisensory emotional information, and it remains unclear to what extent an integration process is emotion specific or, rather, draws on mechanisms involved in multisensory perception, such as audiovisual speech perception, in general.

**DYNAMIC VERSUS STATIC EMOTIONAL STIMULI**

While speech and facial expressions are both inherently dynamic, facial expressions are often tested in static displays. However, our environment is dynamic, and the use of dynamic stimuli in emotion research is hence a much more naturalistic approach for studying emotion perception than from photos or drawings of facial expressions.

In fact, dynamic face stimuli are more easily recognized, especially in the case of non-prototypical emotion displays (Ambadar, Schooler, & Cohn, 2005; Bould & Morris, 2008), and emotional deviants are easier to detect in a dynamic visual search task (Horstmann & Ansorge, 2009). The advantage of dynamic stimuli is further underlined in a steady-state, visual evoked-potential study (Mayes, Pipingas, Silberstein, & Johnston, 2009). This method, in which stimuli are presented together with a flicker while the electroencephalogram (EEG) is recorded, provides information on cortical activation patterns and their latencies, as well as on inhibitory and excitatory processes. Applying this method, the authors found that processing of dynamic facial stimuli proved to be faster and yielded more activity at temporal electrodes and less activity at frontal electrodes than static ones. This may indicate more efficient processing of dynamic stimuli (Mayes et al., 2009). Furthermore, neuroimaging studies report more extended activation patterns for dynamic than for static facial stimuli. This concerns activation of the middle temporal gyrus, inferior and superior temporal gyri, visual areas, most notably the middle occipital and fusiform gyri, and frontal regions, especially the inferior frontal gyrus (Kilts, Egan, Gideon, Ely, & Hoffman, 2003; LaBar, Crupain, Voyeradic, & McCarthy, 2003; Sato, Kochiyama, Yoshikawa, Naito, & Matsumura, 2004; Trautmann, Fehr, & Herrmann, 2009). Increased middle occipital activation, however, is also observed when viewing dynamic mosaics (Sato et al., 2004) whereas the effects in temporal brain regions are connected to the social relevance of stimuli.

**EXPLICIT VERSUS IMPLICIT TASK INSTRUCTIONS**

In emotion perception experiments, task instructions can be generally divided into two categories: explicit and implicit. Explicit tasks are aimed at the emotional content of stimuli and comprise, among others, emotion categorization (assigning a verbal emotion label to a stimulus), matching of emotional stimuli.
in terms of the emotion they express, or valence ratings. There is also a wealth of implicit tasks that are not directed toward the emotionality of a stimulus. An example for an implicit measure is the dot-probe paradigm, in which a neutral and an emotional stimulus are presented simultaneously at different spatial locations and one stimulus is subsequently replaced by a dot probe to which subjects have to react as quickly as possible. Response latencies in this task can reveal attentional biases, as reactions should be faster when the probe appears in a previously attended location. Other examples of implicit tasks are decisions on physical stimulus characteristics, such as color, or gender decisions.

Some studies have directly compared how identical emotional stimulus material is processed under different task instructions. For example, event-related potential (ERP) studies have shown that explicit and implicit task instructions lead to quantitative differences in face processing. Some researchers reported different scalp distributions and latencies of an N400-like component when comparing identity- and emotion-matching tasks, as well as longer reaction times and higher error rates in an explicit task (Bobes, Martin, Olavres, & Valdés-Sosa, 2000; Münte et al., 1998), supporting the view that information processing is facilitated under implicit task instructions.

In neuroimaging research, increased amygdala activations in explicit compared to implicit tasks (Gorno-Tempini et al., 2001; R. C. Gur et al., 2002) and also the opposite pattern (Critchley et al., 2000) have been reported. In fact, the amygdala seems to be more sensitive to the absence of a task (i.e., passive viewing) than to differences between implicit and explicit tasks (Costafreda, Brammer, David, & Fu, 2008; but see Fusar-Poli et al., 2009). Furthermore, an absence of task-modulated amygdala effects but heightened responses in the somatosensory cortex have been reported during explicit face processing (Winston, O’Doherty, & Dolan, 2003).

ERP studies on emotional prosody perception also suggest task-dependent differences. Wambacq, Shea-Miller, and Abubakr (2004) reported an early differentiation (P200) between neutral and negative emotional prosody only when emotion was not task-relevant, while the explicit condition yielded a similar but later effect around 360 ms after stimulus onset. In a cross-modal priming study, Schirmer, Kotz, and Friederici (2002) presented an emotional-prosodic stimulus followed by a visually presented word or non-word, with participants required to make a lexical decision. On incongruent trials, the authors found an N400 effect that differed between women and men as a function of the interstimulus interval. In contrast, when the task was to determine the valence of a word, no sex effects were found (Schirmer, Kotz, & Friederici, 2005). Kotz and Paulmann (2007) reported comparable effects when assessing ERP responses to emotional-prosodic and combined prosodic-semantic expectancy violations under implicit and explicit processing instructions. However, ERP amplitudes were overall more positive-going and larger in the explicit task.

Brain activation patterns revealed by neuroimaging methods during emotional prosody processing also depend considerably on study designs, among them task instructions (Kotz, Meyer, & Paulmann, 2006). In fMRI studies, explicit versus implicit emotional prosody processing has been related to increased orbitofrontal and inferior frontal activations (Ethofer et al., 2009; Imaizumi et al., 1997; Wildgruber et al., 2004, 2005). Enhanced activations in the right superior temporal region have also been observed under explicit versus implicit task instructions when processing emotional prosody (Ethofer et al., 2009; Wildgruber et al., 2005). To sum up, task demands appear to have a considerable impact on study results, both at the neural level and behaviorally. In many cases, reaction times and/or error rates are lower for implicit compared to explicit tasks, for face and prosody processing (Critchley et al., 2000; Ethofer et al., 2009; Gorno-Tempini et al., 2001; Kotz & Paulmann, 2007; Münte et al., 1998; Wildgruber et al., 2005). However, depending on the nature of the task assigned, an implicit task might be more difficult than one that is explicit (R. C. Gur et al., 2002). Thus, a simple distinction between implicit and explicit task dimensions may be too coarse, but choosing a simple implicit task could significantly reduce task demands. This may be a critical issue when testing patients who suffer from executive dysfunction. High task demands, as often occur in explicit tasks, can confound executive dysfunction with emotional processing deficits. This could lead to false assumptions about emotional processing deficits in patients (e.g., Adolphs, Schul, & Tranel, 1998). On the other hand, a very simple task, or even more so, the absence of a task, may lead to increased distraction in psychiatric populations and could also alter study results. These facts need to be considered when designing patient studies or interpreting their outcomes.

**SCHIZOPHRENIA**

Schizophrenia is a complex neuropsychiatric disorder comprising numerous symptoms such as hallucinations, delusions, incoherent thought, and blunting of
affect, which can occur in basically any combination. While disturbances in the dopaminergic system have been commonly associated with schizophrenia, other pathological changes in the brain also seem to play an important role in the neuronal basis of schizophrenia.

Among the numerous deficits associated with schizophrenia, one often observes abnormalities in early sensory processing and deficits in emotion perception. It has been frequently reported that patients with schizophrenia seem to be unable to filter out relevant sensory information from irrelevant information in a stream of complex sensory input. This phenomenon has been described as a deficit in sensory gating (Freedman et al., 1987), and is also reflected in the electrophysiological brain response. When healthy people hear two consecutive sounds, an early auditory evoked potential (P50) in response to the second sound is reduced. This is commonly interpreted as reflecting the above-mentioned gating mechanism. In individuals with schizophrenia, no such reduction can be observed (Bramon, Rabe-Hesketh, Sham, Murray, & Frangou, 2004). Hence, schizophrenic patients seem to process sensory information in a different way than healthy controls.

At the same time, patients with schizophrenia appear to have difficulty in processing emotional information (Chan, Li, Cheung, & Gong, 2010). However, emotional information is usually presented in one or the other sensory modality, raising the question of to what degree emotion perception in patients with schizophrenia is influenced by a general sensory processing deficit, and what role different modalities play in emotion perception. Do the reported deficits in emotion perception originate from an early, sensory deficit, from a deficit in early emotion processing, or rather at a later stage in emotion processing? Furthermore, differences at several of these stages are conceivable, and may have a differential effect dependent on the precise experimental setup. An overview of selected emotion perception studies in schizophrenia is provided in Table 1.

### Schizophrenia: Explicit versus implicit tasks

Various tasks, such as emotion identification (e.g., Bach et al., 2009; Hempel, Hempel, Schönknecht, Stippich, & Schröder, 2003; Kucharska-Pietura, David, Masiak, & Phillips, 2005; Quintana, Wong, Ortiz-Portillo, Marder, & Mazzotti, 2003) and emotion matching (e.g., Hempel et al., 2003; Martin, Baudouin, Tiberghien, & Franck, 2005; Quintana et al., 2003; Salgado-Pineda, Fakra, Delaveau, Hariri, & Blin, 2010), have been employed to investigate explicit emotion perception in patients with schizophrenia. While these tasks themselves differ in that one requires explicit labeling while the other only requires a comparison between a number of stimuli, a recent meta-analysis by Kohler, Walker, Martin, Healey, and Moberg (2010) shows that no difference in impairment can be found by these tasks. This suggests that the observed deficits, also in emotion-identification tasks, are likely to arise from emotional or at least stimulus-processing deficits rather than an impairment in correctly labeling emotions. Regarding implicit emotion processing, gender decision tasks (Johnston et al., 2005; Lepage et al., 2011; Williams et al., 2007), identity matching (Martin et al., 2005; Quintana et al., 2003), and age discrimination (R. E. Gur et al., 2002) have been used primarily in imaging studies, while behavioral studies have focused on the priming paradigm (Höschel & Irle, 2001; Rauch et al., 2010) or the Stroop task (Roux, Christophe, & Passerieux, 2010).

Considering behavioral studies, large differences in emotion processing between patients and controls are observed in explicit and implicit tasks. For instance, in an emotion-matching and also an identity-matching task with patients, Martin et al. (2005) reported a performance decrease that was especially severe if both emotion and identity varied. Interestingly, this suggests that patients had difficulty in processing only one stimulus dimension while ignoring the other. Furthermore, the performance deficit was larger in the emotion than in the identity task. A similar picture, namely a deficit in both tasks, albeit larger in an emotion-identification task, was also found by several other authors (e.g., Sachs, Steger-Wuchse, Kryspin-Exner, Gur, & Katschnig, 2004). While these results suggest a combination of general and emotion-specific impairments, other studies provide results pointing clearly in one or the other direction. Kerr and Neale (1993) reported no emotion-specific processing deficit, while Edwards, Pattison, Jackson, and Wales (2001) described a selective impairment in emotional processing. Evidence of both a general and a specific impairment, has been reported for auditory and visual stimuli (Kerr & Neale, 1993; Edwards et al., 2001).

While the pattern of results remains unclear for studies employing explicit tasks, a clearer picture emerges when considering implicit processing tasks. In an affective priming paradigm, Höschel and Irle (2001) demonstrated that both healthy controls and patients with schizophrenia show increased priming by negative stimuli. However, this effect was even more strongly increased in the patient group, suggesting enhanced automatic processing of (negative)
## TABLE 1
Comparison of selected studies investigating emotion perception in schizophrenia.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Participants</th>
<th>Medication</th>
<th>Emotions</th>
<th>Method</th>
<th>Anger</th>
<th>Fear</th>
<th>Sadness</th>
<th>Happiness</th>
<th>Disgust</th>
<th>Neutral</th>
<th>Other</th>
<th>Task Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>de Gelder et al. (2005)</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>0</td>
<td>Behav</td>
<td>AV</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>e: Weaker influence of voice but stronger influence of face</td>
</tr>
<tr>
<td>de Jong et al. (2009)</td>
<td>55</td>
<td>50</td>
<td>52</td>
<td>0</td>
<td>Behav</td>
<td>AV</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>e: Weaker influence of face</td>
</tr>
<tr>
<td>de Jong et al. (2010)</td>
<td>55</td>
<td>50</td>
<td>52</td>
<td>0</td>
<td>Behav</td>
<td>AV</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>e, i: Stronger integration when auditory distractor present, no impairment by visual distractor</td>
</tr>
<tr>
<td>Linden et al. (2010)</td>
<td>34</td>
<td>34</td>
<td>30</td>
<td>4</td>
<td>Behav</td>
<td>V</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td>e, i: No difference for i task but impairment for e task</td>
</tr>
<tr>
<td>Suslow et al. (2005)</td>
<td>88</td>
<td>30</td>
<td>88</td>
<td>0</td>
<td>Behav</td>
<td>V</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>e: Differences only for patients with affective symptoms</td>
</tr>
<tr>
<td>Tomlinson et al. (2006)</td>
<td>16</td>
<td>24</td>
<td>?</td>
<td>?</td>
<td>Behav</td>
<td>V</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>e: Motion benefit for both groups, but smaller for patients</td>
</tr>
<tr>
<td>van den Stock et al.</td>
<td>31</td>
<td>21</td>
<td>28</td>
<td>3</td>
<td>Behav</td>
<td>V</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td>e: Deficits for perception of body expressions</td>
</tr>
<tr>
<td>Das et al. (2007)</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>0</td>
<td>Behav</td>
<td>AV</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>e: Increased audiovisual integration</td>
</tr>
<tr>
<td>Johnston et al. (2005)</td>
<td>11</td>
<td>15</td>
<td>11</td>
<td>0</td>
<td>fMRI</td>
<td>V</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>p: Reduced activation/differences in connectivity of AMG</td>
</tr>
<tr>
<td>Salgado-Pineda et al.</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>0</td>
<td>fMRI</td>
<td>V</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>e, i: ERP: reduction of VPP and P3; fMRI: reduced activity of FFG</td>
</tr>
<tr>
<td>An et al. (2003)</td>
<td>20</td>
<td>20</td>
<td>15</td>
<td>11</td>
<td>EEG</td>
<td>V</td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
<td>+</td>
<td>i: Reduced P3</td>
</tr>
<tr>
<td>Turetsky et al. (2007)</td>
<td>16</td>
<td>16</td>
<td>10</td>
<td>6</td>
<td>EEG</td>
<td>V</td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
<td>+</td>
<td>e: Larger N170</td>
</tr>
<tr>
<td>Wynn et al. (2008)</td>
<td>26</td>
<td>27</td>
<td>26</td>
<td>0</td>
<td>EEG</td>
<td>V</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>e, i: Reduced N250</td>
</tr>
</tbody>
</table>

Notes: H. Controls: healthy controls; Behav: behavioral; V: visual; A: auditory (prosody); e: explicit; i: implicit; p: passive viewing; FFG: fusiform gyrus; AMG: amygdala. 1: 53 paranoid, 2 residual; 2: 30 schizophrenic patients with flat affect, 30 schizophrenic patients with anhedonia, 28 schizophrenic patients without affective symptoms; 3: point-light faces; 4: body language; 5: 16 paranoid, 4 undifferentiated.
emotional information. In contrast, other studies did not find any difference in implicit emotion processing (Demily et al., 2010; Linden et al., 2010; Roux et al., 2010), or found an impairment only for patients suffering from negative affective symptoms (Suslow, Droste, Roestel, & Arolt, 2005).

In sum, behavioral studies yield a rather heterogeneous picture regarding a possible emotion-processing deficit in schizophrenia. The fact that implicit emotion processing seems less affected than explicit processing may suggest that pre-attentive processing is relatively spared. The reported deficits in explicit tasks may arise from rather late evaluative processing. Furthermore, the ability to correctly discriminate emotional facial expressions under explicit conditions correlates with general cognitive performance in individuals with schizophrenia (Kohler, Bilker, Hagendoorn, Gur, & Gür, 2000; Sachs et al., 2004), suggesting a link between the performance in emotion discrimination and other tasks.

In order to shed light on the underlying mechanisms and potential deficits of emotion perception, numerous neuroimaging studies have been conducted in recent years. Overall, the results suggest deficits at multiple levels. Several studies have shown that patients with schizophrenia show a decrease in activation in the basal-limbic system, including the amygdala (R. E. Gur et al., 2002; Hempel et al., 2003; Johnston et al., 2005; Leitman et al., 2007; Schneider et al., 1998; Williams et al., 2007), the hippocampus (R. E. Gur et al., 2002; Hempel et al., 2003), the anterior cingulate cortex (Hempe et al., 2003; Williams et al., 2007), and the medial prefrontal cortex (MPFC) (Das et al., 2007; Williams et al., 2007). Based on these findings, Williams and colleagues suggest a breakdown in the amygdala–MPFC connection to be related to the observed deficits in the processing of emotional, especially of fearful, stimuli. As abnormalities in functional integration of these regions were observed irrespective of whether fearful stimuli were processed consciously or subconsciously, this suggests deficits in early, fast processing as well as disturbed frontal control mechanisms affecting late evaluative processing.

An interesting differentiation between explicit and implicit tasks can be seen when considering a meta-analysis by Li, Chan, McAlonan, and Gong (2010), who report an overall decrease in activation particularly in the amygdala and in the fusiform gyrus in emotion perception. However, while changes in amygdala activation were observed irrespective of task, the fusiform gyrus was less activated only in explicit tasks. Similar results are reported in a recent study by Quintana et al. (2011), who report underactivation in the fusiform gyrus only when attention is directed to emotional features of a stimulus. These findings further support the notion that at least two separate systems are impaired in emotion processing in patients with schizophrenia: a fast, pre-attentive system, involving the amygdala and its surrounding network, and, at least in visual emotion perception, an attention-modulated system, which also seems deficient but is only involved when participants have to consciously process facial features.

Some studies also report differences in sensory areas, such as the middle occipital gyrus (Johnston et al., 2005), in visual emotion perception, both in explicit and implicit tasks, and primary auditory areas for explicit emotion perception from voices (Leitman et al., 2007). These findings suggest a third aspect that may also affect emotion processing at later stages, namely deficits in an early encoding phase.

A differentiation between an early, purely sensory, and later, more evaluative processing stages is also in line with electrophysiological evidence specifically investigating the time course of visual emotion perception. In addition to an fMRI study, Johnston et al. (2005) conducted an ERP study with the same patient group and paradigm. They reported a reduced vertex positive potential (VPP) amplitude in patients, associated with deficits in the encoding of emotion information. Furthermore, they observed differences in the P3a, which they interpreted as being associated with the encoding deficits at earlier processing stages. Hence, these results support a more general view in which the encoding of facial features is disturbed, irrespective of specific emotional content. Similar results were observed by Turetsky et al. (2007), who reported N170 amplitude differences in an emotion-recognition task as indicative of deficient structural encoding, which also affects differences at the P3. Support for the hypothesis that early processing deficits are at least partly responsible for emotion-processing deficits comes from a study by Kee, Kern, and Green (1998), who found a strong correlation with deficits in visual (and auditory) emotion recognition and measures of neurocognitive function assessing early deficits in perceptual processing. Other studies, however, did not report any N170 differences, but only differences at later emotion processing stages, irrespective of task specification (An et al., 2003; Wynn, Lee, Horan, & Green, 2008).

Overall, behavioral and neuroimaging results suggest deficits in at least two different sub-processes of emotion processing. Explicit and implicit tasks lead to differences in performance and point to separate brain structures underlying the behaviorally observed...
deficits. Hence, they provide a valuable tool in assessing the interplay between the observed networks.

**Schizophrenia: Dynamic versus static stimuli**

Another potentially important factor influencing visual emotion perception concerns the role of dynamic information; what happens if stimuli are not static pictures but dynamic videos, a situation much more comparable to emotion perception in an everyday life? It has been reported that healthy controls show an improvement in emotion recognition for dynamic over static point-light displays (Atkinson, Dittrich, Gemmell, & Young, 2004). Furthermore, imaging studies have shown that static stimuli activate different brain regions than dynamic stimuli (Kilts et al., 2003). In particular, static stimuli activate areas associated with motor imagery, such as motor and parietal areas (Decety & Grèzes, 1999), suggesting that static stimuli require motion simulation to be processed completely. Dynamic stimuli therefore present an interesting case when we investigate emotion perception in patients with schizophrenia. On the one hand, we could expect a decrease in performance, as the processing of dynamic information is an additional processing load and information needs to be integrated correctly in order to be beneficial rather than detrimental. Indeed, Archer, Hay, and Young (1994) showed that patients performed worse in an emotion-identification task when stimuli were dynamic rather than static. On the other hand, one can argue that dynamic stimuli are more natural, utilize different brain networks, and lead to a processing benefit in healthy participants, and hence may also be beneficial to patients. Evidence in support of this view comes from a study by Tomlinson, Jones, Johnston, Meaden, and Wink (2006), who observed that patients with schizophrenia showed an overall worse performance in emotion identification, but nevertheless improved when dynamic stimuli were presented. In contrast to Archer et al. (1994), Tomlinson et al. (2006) used point-light stimuli and were thus able to investigate the perception of motion information selectively. Therefore, motion itself seems to provide a benefit to patients with schizophrenia, and the deficits observed by Archer and colleagues are likely to arise from a different stimulus aspect. Determining the interplay between different aspects, and establishing which aspects prove beneficial and which detrimental is particularly necessary when investigating early, sensory-driven emotion perception.

**Schizophrenia and multimodal emotion perception**

One aspect that needs investigation is the perception of emotional information from multiple sensory modalities at the same time. On the one hand, multimodal presentation of congruent emotional information in healthy controls usually leads to facilitated perception (de Gelder & Vroomen, 2000); on the other hand, it increases the processing demands in terms of sorting out which piece of information from one processing stream is associated with which piece from the other processing stream. As patients with schizophrenia have trouble integrating different sensory streams (Magnée, Oranje, van Engeland, Kahn, & Kemner, 2009), multimodal perception may therefore do more harm than good. Previous studies have shown that patients with schizophrenia seem to be able to accomplish audiovisual integration for very simple sensory input but show strong deficits in more complex, social situations, such as audiovisual speech perception (de Gelder, Vroomen, Annen, Masthof, & Hodiamont, 2003).

A few studies have investigated the multimodal perception of emotional information in patients with schizophrenia (de Jong, Hodiamont, & de Gelder, 2010; de Jong, Hodiamont, van den Stock, & de Gelder, 2009; de Gelder et al., 2005; van den Stock, de Jong, Hodiamont, & de Gelder, in press). Using photos of emotional facial expressions accompanied by matching or mismatching prosodic information, all these studies describe an anomalous integration between auditory and visual information, but no emotion-specific effects. Regarding the exact nature of this anomalous integration, the evidence provided by the studies is inconclusive. When patients are instructed to determine the emotion conveyed by a facial expression and ignore simultaneously presented auditory information, reduced integration compared to healthy controls is observed (de Gelder et al., 2005). When participants are instructed to attend to the auditory modality while ignoring the visual information, the reverse pattern is seen; patients with schizophrenia then show stronger integration than healthy controls (de Gelder et al., 2005). In the study by de Jong et al. (2009), a slightly modified paradigm was used; again participants were instructed to attend to the auditory modality while ignoring the visual information. This time, decreased integration was found. If, however, an auditory distractor was included in the paradigm, integration increased in patients while it decreased in healthy controls (de Jong et al., 2010). In a fourth study by the same group, increased integration of emotional body language and emotional vocalizations...
was observed in patients instructed to pay attention to visual information (van den Stock et al., in press).

Overall, the studies described here seem to show that there is an abnormality in the multimodal integration of emotional information, but its exact nature is still unclear. As no comparison to neutral stimuli was made, it is still an open question whether the observed effects are specific to emotions, or are deficits in integration per se.

**Emotion perception in schizophrenia: Appraisal and outlook**

Task, stimulus dynamics, and multimodality are three factors that are likely to influence emotion perception in patients with schizophrenia. The heterogeneous pattern of results in explicit tasks and the relatively better performance in implicit tasks suggest several underlying mechanisms that may be differentially impaired. If this observation of intact implicit processing can be confirmed, it would necessitate a new discussion about what exactly emotion discrimination tasks are measuring, and whether it can still be claimed that patients with schizophrenia are impaired in emotion perception per se. This assumption is further supported by brain-imaging results; at least two separate aspects, namely an early sensory and a late cognitive processing aspect, seem to be affected to different degrees. Further light can be cast on the interaction and impairment of these mechanisms by investigating dynamic and multimodal stimuli, two features that in healthy participants lead to facilitated processing. Movement information seems to result in a processing benefit, providing evidence of an intact integration of motion information, possibly at early processing steps. Furthermore, multimodal perception seems to be affected. However, the exact extent and emotion specificity remain unclear.

In summary, it is of relevance for future studies to disentangle these different components in order to gain a clearer understanding of emotion-processing deficits in patients with schizophrenia.

**DEPRESSION**

Major depressive disorder (MDD) is one of the most prevalent neuropsychiatric disorders in Western society. For example, around one-sixth of the American population suffers from MDD at least once during life (Kessler et al., 2003). There is evidence that MDD leads to biased emotional processing, which is a maintaining factor of the disorder and may also be present in people at risk or in remission. Understanding how emotion is processed in MDD or MDD-prone individuals could thus be important in preventing and treating the disorder.

It should first be considered that depression severely alters cognitive functions. Problems in thinking and concentration are among the diagnostic criteria for MDD, according to the *Diagnostic and Statistic Manual of Mental Disorders* (DSM-IV, American Psychiatric Association, 1994). MDD patients exhibit slowed information processing, as evidenced by reaction times (e.g., Leyman, de Raedt, Schacht, & Koster, 2007) and enhanced latencies in the ERP, such as the P300 (Vandoolaeghe, Hunsel, Nuyten, & Maes, 1998). MDD patients are also more interference prone than healthy controls, as most clearly manifested in increased error rates (Elliott et al., 1997). These deficits need to be taken into account when running experiments with this patient population, as they could considerably influence study results.

Generally, there are three views on emotion processing in depression:

1. Negative potentiation posits that MDD patients are biased toward negatively valenced stimuli (e.g., Beck, 1967).
2. Positive attenuation considers that processing of positively valenced information is impaired in MDD (e.g., Clark & Watson, 1991).
3. General blunting of emotional processing irrespective of valence is also suggested (emotion context insensitivity; Rottenberg & Gotlib, 2004).

Importantly, points (1) and (2) do not exclude each other and can be present in combination, while (3) is in line with (2), but additionally posits blunted processing of negatively valenced stimuli. An overview of selected studies assessing emotion perception in MDD is provided in Table 2.

**Depression: Explicit versus implicit tasks**

A negativity bias in MDD is widely supported by studies applying explicit categorization of facial emotion, as patients tend to classify neutral or ambiguous facial expressions as sad or negative (Douglas & Porter, 2010; Gollan, Pane, McCloskey, & Coccaro, 2008; R. C. Gur et al., 1992; Hale, 1998; Hale, Jansen, Bouhyus, & van den Hoofdakker, 1998; Leppänen, Milders, Bell, Terriere, & Hietanen, 2004; Luck & Dowrick, 2004; Naranjo et al., 2010), and happy faces as neutral (R. C. Gur et al., 1992; Surguladze et al.,...
TABLE 2

Comparison of selected studies testing emotion perception in major depression. Note that due to a very high number of facial emotion categorization studies in the literature, they are treated exclusively in the text. In the case of treatment studies, only results at baseline are reported.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Patients</th>
<th>Controls</th>
<th>Medication</th>
<th>Method</th>
<th>Stimuli</th>
<th>Anger</th>
<th>Fear</th>
<th>Sadness</th>
<th>Happiness</th>
<th>Disgust</th>
<th>Neutral</th>
<th>Other</th>
<th>Task</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fritzsche et al. (2010)</td>
<td>20</td>
<td>20</td>
<td>2</td>
<td>Behav</td>
<td>V</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>i Attentional bias toward sad and away from happy faces</td>
<td></td>
</tr>
<tr>
<td>Goeleven et al. (2006)</td>
<td>20</td>
<td>20</td>
<td>?</td>
<td>Behav</td>
<td>V</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>e Reduced negative priming of sad faces, indicating reduced inhibition</td>
<td></td>
</tr>
<tr>
<td>Gotlib, Kasch et al. (2004)</td>
<td>88</td>
<td>55</td>
<td>?</td>
<td>Behav</td>
<td>V</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>i Attentional bias toward sad faces</td>
<td></td>
</tr>
<tr>
<td>Gotlib, Kransnoperova et al. (2004)</td>
<td>19</td>
<td>16</td>
<td>8</td>
<td>Behav</td>
<td>V</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>i Attentional bias toward sad faces</td>
<td></td>
</tr>
<tr>
<td>Joormann et al. (2007)</td>
<td>26</td>
<td>19</td>
<td>16</td>
<td>Behav</td>
<td>V</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>i Attentional bias toward sad and away from happy faces</td>
<td></td>
</tr>
<tr>
<td>Kan et al. (2004)</td>
<td>16</td>
<td>20</td>
<td>15</td>
<td>Behav</td>
<td>A</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td>e Tendency to categorize surprise as negative</td>
<td></td>
</tr>
<tr>
<td>Leyman et al. (2007)</td>
<td>20</td>
<td>20</td>
<td>16</td>
<td>Behav</td>
<td>V</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>i Reduced attentional disengagement from angry faces</td>
<td></td>
</tr>
<tr>
<td>Luck et al. (2004)</td>
<td>49</td>
<td>30</td>
<td>0</td>
<td>Behav</td>
<td>A</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>e More negative classifications</td>
<td></td>
</tr>
<tr>
<td>Naranjo et al. (2010)</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>Behav</td>
<td>A</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td>e Tendency to categorize neutral and surprise as negative</td>
<td></td>
</tr>
<tr>
<td>Uekermann et al. (2008)</td>
<td>29</td>
<td>29</td>
<td>12</td>
<td>Behav</td>
<td>A</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>e Impaired for all subtests except for stimuli with congruent semantics and prosody; all categories except sadness affected</td>
<td></td>
</tr>
<tr>
<td>Almeida et al. (2010)</td>
<td>15</td>
<td>15</td>
<td>13</td>
<td>fMRI</td>
<td>V</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>e No group differences</td>
<td></td>
</tr>
<tr>
<td>Dannlowski et al. (2008)</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td>fMRI</td>
<td>V(^1)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>p No group differences; increased AMG response in risk allele carriers</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Subj.</td>
<td>Ctrl.</td>
<td>Emo.</td>
<td>Task</td>
<td>Modality</td>
<td>Status</td>
<td>Notes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------</td>
<td>-------</td>
<td>------</td>
<td>-----------------------</td>
<td>-----------</td>
<td>---------</td>
<td>-----------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frodl et al. (2009)</td>
<td>12</td>
<td>12</td>
<td>8</td>
<td>4</td>
<td>fMRI V</td>
<td>+</td>
<td>Higher neural response, reduced with explicit task</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fu et al. (2004)</td>
<td>19</td>
<td>19</td>
<td>0</td>
<td>19</td>
<td>fMRI V^2</td>
<td>+</td>
<td>Stronger neural response, slower RT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fu et al. (2007)</td>
<td>19</td>
<td>19</td>
<td>0</td>
<td>19</td>
<td>fMRI V^2</td>
<td>+</td>
<td>Range of neural response to different intensities reduced in limbic-subcortical and visual areas higher overall AMG-hippocampus activity, stronger reaction to intensity increase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fu et al. (2008)</td>
<td>16</td>
<td>16</td>
<td>0</td>
<td>16</td>
<td>fMRI V^2</td>
<td>+</td>
<td>Stronger neural response to sad and happy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gotlib et al. (2005)</td>
<td>18</td>
<td>18</td>
<td>9</td>
<td>9</td>
<td>fMRI V</td>
<td>+</td>
<td>Weaker neural response to all faces, subcortically and frontally</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lawrence et al. (2004)</td>
<td>9</td>
<td>11</td>
<td>9</td>
<td>0</td>
<td>fMRI V^2</td>
<td>+</td>
<td>Weaker neural response, subcortically and frontally</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee et al. (2008)</td>
<td>21</td>
<td>15</td>
<td>10</td>
<td>11</td>
<td>fMRI V</td>
<td>+</td>
<td>Greater AMG activation to emotional faces</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sheline et al. (2008)</td>
<td>11</td>
<td>11</td>
<td>0</td>
<td>11</td>
<td>fMRI V^1</td>
<td>+</td>
<td>Happy; reduced activation in FFG and putamen; sad; increased activation in FFG, putamen, parahippocampal gyrus/AMG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surguladze et al. (2005)</td>
<td>16</td>
<td>14</td>
<td>16</td>
<td>0</td>
<td>fMRI V^2</td>
<td>+</td>
<td>Stronger AMG reaction to sad primes; reduced to happy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suslow et al. (2010)</td>
<td>30</td>
<td>26</td>
<td>30</td>
<td>0</td>
<td>fMRI V^1</td>
<td>+</td>
<td>Increased activation in AMG, IFG, and ACC to labeling task</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>van Wingen et al. (2010)</td>
<td>18</td>
<td>30</td>
<td>0</td>
<td>18</td>
<td>fMRI V</td>
<td>+</td>
<td>Stronger AMG reaction to sad primes; reduced to happy; faster RT to sad primes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Victor et al. (2010)</td>
<td>22</td>
<td>25</td>
<td>0</td>
<td>22</td>
<td>fMRI V^1</td>
<td>+</td>
<td>No difference between neutral and emotional stimuli from around 220 ms after onset</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chang et al. (2011)</td>
<td>15</td>
<td>15</td>
<td>0</td>
<td>15</td>
<td>EEG V</td>
<td>+</td>
<td>Enhanced P1 to sad face preceded by sad target; no negative priming for sad faces</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dai et al. (2011)</td>
<td>19</td>
<td>20</td>
<td>?</td>
<td>?</td>
<td>EEG/Behav</td>
<td>+</td>
<td>Abolished P300 difference between previously seen sad and happy faces</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deklin et al. (2001)</td>
<td>19</td>
<td>15</td>
<td>14</td>
<td>5</td>
<td>EEG V</td>
<td>+</td>
<td>No slow wave amplitude reduction for sad faces</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deveney &amp; Deklin (2004)</td>
<td>17</td>
<td>17</td>
<td>6</td>
<td>11</td>
<td>EEG V</td>
<td>+</td>
<td>No difference of emotional versus neutral faces in late positive potential</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foti et al. (2010)</td>
<td>19</td>
<td>25</td>
<td>0</td>
<td>19</td>
<td>EEG V</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: H. Controls; Healthy controls; Behav: behavioral; V: visual; A: auditory (prosody); e: explicit; i: implicit; p: passive viewing; ACC: anterior cingulate cortex; IFG: inferior frontal gyrus; FFG: fusiform gyrus; AMG: amygdala; RT: reaction time. 1: subliminally presented, masked faces; 2: morphed to different emotional intensities; 3: two explicit tasks: matching of facial expressions, matching of facial expression to verbal label.
In the latter study by Surguladze et al. (2004), this was limited to happy expressions of 50% intensity, and only at longer stimulus durations (2000 ms instead of 100 ms). Thus, biases in MDD may benefit from long presentation times and non-prototypical emotion expressions providing more scope for an individual’s own interpretations. This speaks in favor of biases at later, evaluative processing stages. This evaluative bias fits with the finding that morphed sad faces need less intensity to be categorized as sad by MDD patients, while the intensity of happy facial expressions needs to be higher in order to be categorized as happy (Joormann & Gotlib, 2006). Even healthy subjects have been shown to perceive more sadness or rejection in ambiguous line drawings of faces after depressive mood induction (Bouhuys, Bloem, & Groothuis, 1995). Thus, evaluative biases can also occur in the absence of MDD; mood itself may play an essential role. A positivity bias in controls not present in patients has also been reported, reflected by an increased tendency of controls to classify neutral faces as happy (Douglas & Porter, 2010; Gollan et al., 2008; Surguladze et al., 2004).

Results from negative priming, a method in which a former distractor becomes task-relevant in a subsequent trial, revealed deficient inhibition in MDD patients when the distractor was a sad face (Dai, Feng, & Koster, 2011; Goeleven, de Raedt, Baert, & Koster, 2006). This suggests that patients allocated attention to the sad face even though it had to be ignored, while the happy face had to be categorized.

The negative bias in MDD previously outlined for faces is also reported for the categorization of vocal stimuli (Kan, Mimura, Kamijima, & Kawamura, 2004; Luck & Dowrick, 2004; Naranjo et al., 2010). For example, MDD patients judge negatively valenced stimuli as more negative than healthy controls (Naranjo et al., 2010). Importantly, stimulus characteristics may play a pivotal role in MDD. Uekermann, Abdel-Hamid, Lehmkämper, Vollmoeller, and Daum (2008) tested emotional prosody categorization and found broad impairments in MDD patients for almost all prosodic stimuli (with neutral semantics and with mismatching semantics). However, no group effects were found when semantics matched the emotional tone conveyed in a sentence. Additionally, the participants’ executive functions were related to task performance. Thus, convergent semantic information may largely reduce between-group differences. Reduced impact on executive functions and richer information availability could account for this effect. Likewise, in a study by Emerson, Harrison, and Everhart (1999) examining schoolboys, it was shown that although all participants were better at emotional categorization when semantics and prosody of a sentence matched, the performance decline for mismatching sentences was much more pronounced in depressed than in healthy participants.

Preliminary ERP data from our laboratory on a group of elderly participants without clinically relevant depression symptoms indicate that increased depression scores affect the integration of emotional-prosodic and semantic information, as reflected in smaller differences in the ERP response to pure prosodic (unintelligible) and normal speech. Based on these results, one may argue that depressive symptoms hamper the integration of information channels, and that the benefit MDD patients showed for semantically and prosodically congruent stimuli in Uekermann et al. (2008) rather reflects richer information availability and/or the absence of distracting information than integration per se.

Taken together, the results on emotional prosody perception in depression confirm the presence of a negative evaluation bias and thus extend the findings from facial expressions to speech, although stimulus characteristics may play an additional role in these findings.

Apart from explicit tasks, implicit methods have also been used at the behavioral level, most notably the dot-probe task (described in the introduction). At face presentation times of 1000 ms, an attentional bias toward sad faces has been reported for MDD patients (Fritzsche et al., 2010; Gotlib, Kasch et al., 2004; Gotlib, Krasnoperova, Yue, & Joormann, 2004; Joormann & Gotlib, 2007), and this bias was preserved in remitted patients (Fritzsche et al., 2010; Joormann & Gotlib, 2007). Even in never-depressed daughters of depressed mothers, this effect was found after negative mood induction (Joormann, Talbot, & Gotlib, 2007; Kujawa et al., 2011). A related method, spatial cueing, revealed prolonged attending to angry versus neutral faces in MDD compared to controls, suggesting that the bias extends beyond expressions of sadness when participants are instructed to attend to the faces (Leyman et al., 2007). By contrast, in dot-probe studies including anger, the bias appears to be sadness-specific (Gotlib, Krasnoperova et al., 2004; Gotlib, Kasch et al., 2004); thus, instructions that direct attention may be significant. Some dot-probe studies also report a positivity bias in healthy controls not present in MDD or MDD-prone individuals (Fritzsche et al., 2010; Joormann & Gotlib, 2007; Joormann et al., 2007). These studies demonstrate that depression leads to preferred attention to sad, or more generally, negatively valenced faces and away from positive expressions
at late processing stages, even when emotion is not task-relevant.

In short, both explicit and implicit behavioral studies revealed biased emotion processing in depression, and not only evaluative but also attentional biases are supported by the literature.

Electrophysiology has also yielded interesting results: Dai et al. (2011) reported abnormal ERP responses to sad faces in MDD already occurring around 100 ms after stimulus onset when the face was preceded by a sad target that had to be categorized. Thus, after judging a face as sad, this kind of stimulus is processed abnormally in MDD, as can be observed during early sensory processing. Importantly, sad faces preceded by a happy target did not provoke such early effects, speaking against an early sensory deficit or bias. In an electromyographic study (Sloan, Bradley, Dimoulas, & Lang, 2002), students who were dysphoric revealed reduced activity of the zygomaticus, a facial muscle which responds to positive stimuli, during the explicit categorization of happy faces. Instead, they showed more corrugator activity for happy faces, which is normally correlated with processing negative stimuli. In fact, positive and negative facial expressions were processed as if they were all negative in dysphorics, even though categorization performance was comparable to controls, and thus there was no evaluative bias. This study, which temporally covers the seconds range, is complemented by an ERP study with millisecond resolution, in which students who were depressed exhibited reduced processing especially of mildly happy faces, starting around 350 ms after face onset (Cavanagh & Geisler, 2006).

Taken together, results from electrophysiological studies using explicit tasks indicate that subclinical depression leads to reduced positive reactivity, and that sad stimuli are processed abnormally in MDD when attention has previously been drawn to them. The time course suggests that biases already emerge before evaluative decisions, but no early, sensory processing deficits are confirmed.

ERP studies applying implicit tasks also support altered processing of facial emotion in MDD. Deldin, Keller, Gergen, and Miller (2000) observed a generally attenuated N200 component at right-posterior electrode sites in MDD, an effect which was strongest in response to happy facial expressions. In another study, MDD patients did not differ in their P300 response to previously seen happy versus sad faces while controls did (Deldin, Keller, Gergen, & Miller, 2001). The mismatch negativity (MMN), a measure of pre-attentive change detection, was not evident in MDD patients from around 220 ms post-stimulus onset onward, as their ERP response to sad, happy, and neutral schematic faces did not differ. However, the MMN was present before this time window, as neutral faces initially elicited brain responses that differed from the two emotional categories (Chang, Xu, Shi, Zhang, & Zhao, 2010). This means that early perceptual encoding including the face-related N170 component may be largely intact, and that blunting takes effect at a stage in which emotional significance is supposed to be extracted. Foti, Olvet, Klein, and Hajcak (2010) reported a reduced differentiation of fearful and angry faces from neutral ones in the ERP in MDD at a late processing stage (late positive potential). This further supports the view of emotional blunting in MDD, although it may also reflect the fact that MDD patients are more likely to perceive neutral faces as negative. Blunting is also supported in a study by Kayser, Bruder, Tenke, Stewart, and Quitkin (2000), who presented unpleasant (wounded) and neutral (intact) faces and reported a lack of a differentiation of these two stimulus types in MDD in a comparable time window, suggesting generally reduced emotional reactivity. Deveney and Deldin (2004) assessed memory retention of emotional faces during several seconds (slow-wave). While this component was increasingly reduced for sad relative to neutral and happy faces in healthy participants, no differences emerged in the depressed group, suggesting that MDD patients fail to disengage from sad faces over time. This is nicely complemented by an fMRI study on memory retention of positively and negatively valenced pictures: Successful retention of the latter was associated with enhanced amygdala activation during encoding in MDD patients (Hamilton & Gotlib, 2008); thus, amygdala hyperactivity might play a role in the enhanced retention of negative stimuli.

To sum up, electrophysiological studies indicate that a few hundred milliseconds from the onset of a face, emotion perception is blunted in MDD, as reflected in reduced neural differentiation between positive, negative, and neutral faces. Negative potentiation, in contrast, seems to occur in the seconds range and may reflect a lack of attentional deployment from negative stimuli. This can be observed with explicit and implicit tasks.

In fMRI studies, there is some indication that explicit compared to implicit tasks may reduce group effects at the neural level. Studies using implicit tasks or passive viewing provide ample evidence of negative potentiation—that is, heightened neural responses in MDD compared to controls when individuals are presented with negative facial expressions (Frodl et al., 2009; Fu et al., 2004, 2008; Gotlib et al., 2005; Sheline et al., 2001; Surguladze et al., 2005; Victor, Furey, Fromm, Ohman, & Drevets, 2010, but see Lawrence...
et al., 2004; Lee et al., 2008). This extends to remitted patients (Neumeister et al., 2006) and individuals at genetic risk of MDD (Dannlowski et al., 2008; Hariri et al., 2005). Positive attenuation has also been corroborated by neuroimaging studies using implicit tasks (Fu et al., 2007; Lawrence et al., 2004; Surguladze et al., 2005; Victor et al., 2010).

In contrast, a study using emotional categorization reported no significant activation differences between MDD and control participants, for either positive or negative emotional faces (Almeida, Versace, Hassel, Kupfer, & Phillips, 2010). This fits well with Frodl et al. (2009), who observed that negative potentiation was greater when using a gender matching instead of an emotion matching task. It is also in line with a study by Monk et al. (2008) comparing passive viewing with different tasks in adolescents at risk of depression. While the data obtained from passive viewing support negative potentiation and positive attenuation, the data from the constrained attention conditions do not. The findings from Monk et al. also suggest that a challenging implicit task may diminish group differences. Thus, what has to be evaluated is probably not the explicit-implicit distinction but rather the potential cognitive load introduced by the task. Stronger frontal activations in the risk compared to the non-risk group when performing the tasks support this notion (Monk et al., 2008). Along these lines, an easy explicit task may also give rise to negative potentiation in MDD, compared to a more challenging one (van Wingen et al., 2011).

The amygdala, in particular, seems to play a central role in MDD, as most of the studies reporting negative potentiation find enhanced activations in this structure. Moreover, Dannlowski et al. (2007) observed that amygdala activation to sad and angry faces correlated positively with behaviorally manifested negative biases in MDD. Connectivity studies with depressed individuals have shown abnormal interactions of prefrontal regions with brain structures attributed to emotional processing. More specifically, connections between the amygdala and prefrontal cortex appear to be impaired in MDD (Dannlowski et al., 2009; Siegle, Thompson, Carter, Steinhauer, & Thase, 2007), a connection which has been implicated in successful emotion regulation (Banks, Eddy, Angstadt, Nathan, & Phan, 2007). Low cognitive load, as in easy implicit tasks, may exacerbate the impact of this disturbed circuit, enabling excessive amygdala activations in response to negative stimuli. Findings from the cognitive domain indicate that tasks which strongly bind attention and suppress rumination, a core feature of the disorder, have been associated with reduced between-group differences (for a review, see Gotlib & Joormann, 2010).

Depression: Dynamic versus static stimuli

Categorization studies using dynamic facial expressions have yielded promising results. Kan et al. (2004) tested videos of six emotions (happiness, surprise, anger, disgust, fear, and sadness) and found no group effects in recognition performance, a result which is in line with a recent study (Schaefer, Baumann, Rich, Luckenbaugh, & Zarate, 2010). One possible explanation for these results is that the information provided by dynamic stimuli is richer than from static displays and makes it easier for patients to recognize them, as discussed by the authors of both studies. Another explanation is that emotional information provided in moving stimuli that is continuously changing prevents MDD patients from focusing on a facial expression and starting to ruminate about it or project negativity onto it. This suggestion is in line with intact early processing of emotional cues and biases at later, more cognitive processing stages. In any case, these study results are promising and will have to be corroborated by further investigations.

However, as outlined in the previous section, impairments and negative biases have been reported in studies on emotional prosody categorization in MDD (Kan et al., 2004; Luck & Dowrick, 2004; Naranjo et al., 2010; Uekermann et al., 2008), and prosody is also dynamic in nature. In the realm of an apparent dissociation between dynamic speech and face stimuli, it must be considered that recognition performance is generally higher for emotions conveyed by faces than by voices (e.g., Kan et al., 2004). Likewise, cognitive impairment associated with MDD could potentiate challenging task demands in explicit tasks (i.e., prosody categorization), while congruent semantic information may reduce cognitive demands and decrease group differences (Emerson et al., 1999; Uekermann et al., 2008).

To sum up, dynamic facial expressions, as well as prosodic stimuli with matching semantics, have led to promising results in MDD. However, even in the absence of behavioral differences, there may still be alterations at the neural level (e.g., Sloan et al., 2002), as explicit categorization is not very informative about earlier processing steps. More research is needed to shed light on these open issues.
Depression and multimodal emotion perception

To our knowledge, so far there are no studies on multimodal emotion perception in MDD patients. There is, however, one ERP experiment by Campanella et al. (2010) that assesses a group of students displaying elevated but subclinical anxiety and depression scores. In an emotional target-detection oddball paradigm, they reported a reduced P300 amplitude when compared with students with low scores in these measures. Interestingly, this was the case only in the multimodal (prosodic speech cue and static facial expression) condition irrespective of emotion, and not in the unimodal conditions. Moreover, the P300 amplitude was negatively correlated with depression scores. This suggests that under multimodal input conditions, emotion-processing deficits may be more likely to take effect than under unimodal input. However, this is but one experiment in a subclinical population, and one can only speculate about multimodal emotion processing in MDD. Furthermore, in the experiment, a static facial expression was combined with a prosodic stimulus. Even though these were matched in terms of the emotion they expressed, the static-dynamic combination represents a mismatch, and mismatching information may affect depressive individuals more strongly than controls (Emerson et al., 1999; Uekermann et al., 2008). As long as multimodal information is congruent, MDD patients should benefit from it, at least when it comes to explicit recognition tasks, while neural correlates of multimodal stimuli still need to be elucidated in MDD.

Emotion perception in depression: Appraisal and outlook

While a considerable amount of work targeting facial emotion processing in MDD has accumulated in both implicit and explicit tasks, studies on emotional prosody or multimodal emotion displays are largely or completely missing to date. Furthermore, face stimuli have been presented statically rather than dynamically in the vast majority of studies.

As outlined in the introduction, emotion processing is not a unitary process but rather involves different processing steps in which different tasks are accomplished and different neural correlates may be engaged. Studies with MDD patients have helped to shape our understanding of how depression may influence emotion processing from faces. ERP studies indicate that processing alterations in MDD may start around a few hundred milliseconds after stimulus onset, where there is evidence of general blunting, as neural response patterns fail to distinguish stimuli according to their emotional content (Chang et al., 2010; Foti et al., 2010). At around one second after face onset, patients have directed their attention to sad stimuli (Fritzsche et al., 2010; Gotlib, Krasnoperova, et al., 2004) and away from happy stimuli (Joormann & Gotlib, 2007; Joormann et al., 2007) even when faces are not task-relevant. This attentional focus then persists, as suggested by two ERP studies looking at late processing stages (Dai et al., 2011; Deveney & Deldin, 2004). When prompted to attend to the emotionality of a face, happy expressions may be processed as if they were negative (Sloan et al., 2002), a result which could explain biases evident in explicit categorization experiments. Thus, it seems that depression does not lead to preferential attention to negative or attention deployment from positive stimuli, but once negative stimuli have captured a patient’s attention, the patient cannot disengage from them (Gotlib & Joormann, 2010). This may be especially the case when an experimental task is not challenging (Frodl et al., 2009) and does not provide sufficient distraction to suppress rumination.

Even though there is no convincing evidence so far that MDD biases early perceptual encoding of facial emotion, support for early alterations has been provided by three recent fMRI studies using subliminally presented, masked face primes, which revealed group-related activation differences in the amygdala (Sheline et al., 2001; Suslow et al., 2010; Victor et al., 2010). Another study reported similar results only when dividing the sample into genetic risk and non-risk groups rather than diagnostic groups (Dannlowski et al., 2008). The effects may correspond to the fast emotional face-processing pathway directly feeding into the amygdala (Adolphs, 2002; Vuilleumier & Pourtois, 2007). However, neutral faces were used to mask the emotional faces in these studies. Due to the low temporal resolution of fMRI, it cannot be ruled out that the activations reflect the processing of the neutral mask, influenced by emotional prime valence. Experiments using non-facial masks and the application of techniques providing higher temporal resolution, such as EEG, may be useful tools to further address this issue. ERP studies, despite their excellent temporal resolution, have so far failed to find alterations in early sensory processing. In sum, at this point, we do not know whether depression affects early sensory processing. Evidence on later processing stages is much clearer, although the fact that the vast majority of studies are based on static face stimuli must be considered.
GENERAL DISCUSSION

Accurately perceiving others’ emotions is an essential component of successful social interaction, a skill, which is impaired in many if not all psychiatric disorders. However, emotion perception is not a single-step process but rather a complex mechanism involving several interacting subcomponents.

As elaborated in the introduction, at least three such substeps can be distinguished: (1) sensory processing, (2) integration of sensory cues, and (3) evaluation of the perceived cues (Adolphs, 2002; Schirmer & Kotz, 2006). In principle, all of these steps are vulnerable to malfunction, raising the question of at which processing stage the observed deficits in a given neuropsychiatric disorder may occur.

Here, schizophrenia and depression are of particular interest. While a common finding in patients with schizophrenia is the presence of early sensory processing deficits, depression is characterized by deficits at later, cognitive processing stages. This dissociation in general deficits suggests that the processing of emotional information expressed in faces and voices may also result in specific processing deficits. Thus, while at first glance, emotion perception appears disturbed in both disorders, the underlying mechanisms leading to these alterations may be different. It is therefore conceivable that deficits arise in several processing steps, or from a different mechanism altogether, such as cognitive deficits affecting response behavior. Furthermore, it is open to discussion to what extent observed deficits are modality-specific or supramodal, and whether multimodal emotion perception can alleviate symptoms or, in contrast, lead to a decrease in performance. All these questions cannot be fully answered by the current state of the art in psychiatric research on emotion expressions, but the evidence reviewed here offers the possibility of some justified speculation, motivating further research.

Current studies suggest that emotion-processing deficits occur at various processing levels. For schizophrenia, in particular, studies show very early processing differences (Johnston et al., 2005; Leitman et al., 2007), suggesting that deficits already occur at the level of sensory encoding. Considering the well-known problems in sensory gating that are often described as one characteristic finding in patients with schizophrenia, this raises the question of to what extent these deficits are emotion-specific or rather a result of pathological early sensory encoding per se. In depression, evidence on deficits in early perceptual encoding is, to date, not convincing (e.g., Gotlib & Joormann, 2010).

Regarding later processing stages, emotional processing deficits have been reported for both patient populations (An et al., 2003; Dai et al., 2011; Deveney & Deldin, 2004; Wynn et al., 2008). While emotional valence appears to play a role in depression, as positively valenced material is attenuated and negatively valenced material is processed in an enhanced manner, no such distinction can be described for schizophrenic patients. These patterns suggest that the observed deficits in emotion processing arise from different underlying mechanisms. The observations regarding MDD speak in favor of an emotional bias, as processing differences are found at later stages and appear to be valence-driven. In schizophrenic patients, however, the picture is not that clear. While there is evidence for deficits at later evaluative steps, disturbances in early sensory processes are also commonly observed. On the one hand, this complex pattern may indicate that two separate processing mechanisms are affected. On the other hand, the late processing deficits could also be interpreted as a consequence of deficits at earlier stages. Thus, alterations in emotion processing per se seem more likely in depression than in schizophrenia.

To shed more light on emotional processing in these two disorders, it seems important to examine more closely several factors influencing emotion perception. Here, we focused on the three aspects—task, stimulus dynamics, and multimodality—as they seem to be especially relevant to emotion perception (e.g., Atkinson et al., 2004; Giard & Peronnet, 1999). Distinguishing between explicit and implicit tasks is important when investigating patient populations, as explicit task instructions may introduce differences between patient and healthy populations that are unrelated to emotional processing, but rather reflect cognitive deficits. In fact, schizophrenia patients generally perform worse in emotional as well as control tasks, suggesting a strong effect of cognitive factors (Kerr & Neale, 1993; Pomarol-Clotet et al., 2010). This is supported by the finding that difficulties are more pronounced in explicit than in implicit emotional processing tasks (Linden et al., 2010; Suslow et al., 2005). Patients with depression seem to be less affected by task settings and may even benefit from challenging tasks at the neural level (Frodl et al., 2009).

A second aspect to be taken into consideration is multisensory emotion perception. In healthy participants, providing congruent emotional information via several modalities usually leads to facilitated processing (Giard & Peronnet, 1999; Paulmann et al., 2009). Can similar benefits be observed in patient populations? Or do deficits rather increase, as multisensory
stems are more complex than unsensory, thus requiring more elaborate processing? While investigating multisensory emotion perception may thus provide crucial insights into the relation between sensory and affective deficits, this dimension has hardly been investigated in patient populations. While no studies are known that have investigated multimodal emotion perception in MDD patients, a few studies have addressed this issue at the behavioral level in patients with schizophrenia (de Gelder et al., 2005; de Jong et al., 2009, 2010; van den Stock et al., in press). However, no clear picture results from these studies with respect to whether these patients can benefit from the multimodal perception of emotional information. For both disorders, it remains an open question whether multimodal emotional input leads to improved processing or introduces new problems.

A last aspect, also improving emotion perception in healthy controls, is the use of dynamic visual stimuli. Dynamic visual stimuli have higher ecological validity and provide more information than static pictures. In relation to the face stimuli used, it has been shown that, just like healthy controls, patient groups benefit from dynamic information (Atkinson et al., 2004; Kan et al., 2004; Schaefer et al., 2010; Tomlinson et al., 2006). Nevertheless, they are not widely used but should be considered for further research aiming at more naturalistic stimulus displays.

To conclude, emotion perception in patient populations is influenced by numerous factors, such as task specificity and stimulus dynamics. However, it seems that patients with schizophrenia and depression are affected differentially; while in schizophrenic patients, emotion-processing deficits occur at early as well as later stages in processing in patients with schizophrenia, MDD patients show impairments primarily at later evaluative stages. Acknowledging this fact could not only affect our understanding of a given disorder but also play a crucial role in the correct assessment and intervention.

REFERENCES


