



Available online at www.sciencedirect.com



Schizophrenia Research 72 (2005) 195–203

**SCHIZOPHRENIA
RESEARCH**

www.elsevier.com/locate/schres

Multisensory integration of emotional faces and voices in schizophrenics

Beatrice de Gelder^{a,b,*}, Jean Vroomen^a, Sjakko J. de Jong^c, Erik D. Masthoff^c, Fons J. Trompenaars^c, Paul Hodiamont^{a,c}

^a Cognitive Neuroscience Laboratory, Tilburg University, P.O. Box 90153, The Netherlands

^b MGH-NMR Center, Harvard Medical School, Building 36, First Street, Charlestown, MA 02129, USA

^c GGZ, Midden Brabant, The Netherlands

Received 3 October 2003; received in revised form 3 February 2004; accepted 6 February 2004

Available online 23 April 2004

Abstract

In their natural environment, organisms receive information through multiple sensory channels and these inputs from different sensory systems are routinely combined into integrated percepts. Previously, we reported that in a population of schizophrenics, deficits in audiovisual integration were observed for complex stimuli (auditory and visual syllables), but not for more simple ones (beeps and light flashes). Here, we investigated multisensory integration of emotional information in a group of schizophrenic patients. In Experiment 1, we found a reduced effect of an emotional voice on the categorization of a facial expression. In Experiment 2, the reverse test situation was presented, and, here, we observed an exaggerated effect of a face expression on the categorization of an emotional voice. Results are discussed in the light of current models of multisensory integration and their relevance for schizophrenia.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Schizophrenia; Face expression; Emotional voice; Affective prosody; Cross-modal integration; Audio–visual perception

1. Introduction

Classical descriptions of cognitive deficits in schizophrenia include affective disturbances like abnormal expression and experience of emotional states, loss of feeling of presence and disturbances in the experience of personal identity and of self. But only a limited subset of these deficits have so far been addressed experimentally. Over the last decade, a

number of studies have substantiated this clinical picture and consistently reported that schizophrenics have difficulty with the recognition of emotions (for an overview, see Kohler et al., 2000). The area that has received most attention is that of understanding emotions presented in facial expressions (Borod et al., 1993; Feinberg et al., 1986; Mueser et al., 1996; Streit et al., 1997; Wölwer et al., 1996; see Mandal et al., 1998 for an overview). A few studies have also investigated emotion perception in written or in spoken language. For example, in studies focussing on the auditory modality, deficits have been reported in the categorisation of emotional voices (Ross et al.,

* Corresponding author. MGH-NMR Center, Harvard Medical School, Building 36, First Street, Charlestown, MA 02129, USA.

E-mail address: degelder@nmr.mgh.harvard.edu (B. de Gelder).

2001; see [de Gelder et al., \(2003\)](#), for further discussion). Researchers on schizophrenia have also raised the issue of a combined deficit and looked for correlations between deficits in hearing and in seeing emotions. Following the approach of cognitive neuropsychologists, they were interested in finding out whether there existed an underlying emotion deficit that would be independent of the modality of stimulus presentation and that would correspond to a single processing deficit manifesting itself in multiple channels ([Borod et al., 1998](#)).

The present study deals with the combined perception of emotions in the face and in the voice. As far as the study of cognitive deficits in schizophrenia is concerned, this is a novel area of research. We are interested not so much in deficits in processing emotional stimuli in the visual and/or auditory modality but in how the perceptual system combines what it hears and sees in the service of adaptive behaviour and whether this multisensory integration process is impaired in schizophrenia. Some well-known examples of multisensory integration processes are the influence of a visual stimulus on spatial localization of an auditory one (ventriloquism), the reciprocal influence of auditory and visual information in spoken language perception (leading to the so-called McGurk effect), and of emotion (see [Bertelson and de Gelder, in press](#); [de Gelder and Bertelson, 2003](#) for an overview and discussion of the main issues). When the focus is on multisensory perception, the central question is not really whether there is a similarity between two sensory modalities each considered on their own, but whether two sensory modalities interact when information is presented in two sensory modalities at the same time. This issue thus cannot be answered by asking whether a deficit in one modality (like, for example, an impairment in recognizing emotion in the face) is also observed in another modality (like, for example, an impairment in recognizing emotion in the voice). Instead of presenting tasks in each sensory modality and examining correlations in performance, one needs to use paradigms that provide a window on intersensory processes taking place in the course of perception.

Only a few studies have looked at audio–visual integration in schizophrenia. A number of neurological reports found impairments in audio–visual integration using crossmodal matching tasks. [Heinrichs](#)

and [Buchanan \(1988\)](#) reviewed several neurological studies of abnormal signs of multisensory integration on clinical neurological examinations. They concluded that basic mechanisms of sensory processing appear to be intact in schizophrenic patients, but that impairments existed in three higher-order functional areas: the co-ordination of motor activity, the sequencing of motor patterns and the integration of more complex sensory units. Later results were based on crossmodal matching. [Ismail et al. \(1998\)](#) investigated the prevalence and type of neurological abnormalities in schizophrenic patients and their non-psychotic siblings. Both patients and siblings scored worse than comparison subjects on integration of higher sensory functions. [Ross et al. \(1998\)](#) found an association between poor sensory integration and eye tracking disorder. They proposed that a circuit, which includes the posterior parietal cortex, subserved both smooth-pursuit eye movements and audio–visual integration.

In previous studies, we investigated multisensory processes in schizophrenia by focussing on two well-known phenomena from the field of multisensory perception: ventriloquism (i.e., the crossmodal bias of the location of a sound towards a spatially displaced light) and audio–visual speech perception (the McGurk effect). Our findings indicated that there was no integration deficit with respect to the first, but a marked impairment for the second ([de Gelder et al., 2003](#)). In the present study, we examined the reactions of schizophrenic patients in a new situation recently explored in our laboratory, that of multisensory perception of emotion ([de Gelder et al., 1995](#); [Massaro and Egan, 1996](#); [de Gelder and Vroomen, 2000](#); [Dolan et al., 2001](#)). Our experiments with normal participants had established that an emotional expression in the voice influenced categorization of a face expression ([de Gelder et al., 1995](#)). In the reverse situation, in which subjects were asked to ignore the face but to categorize the emotional expression of the voice, we found a similar effect this time from the face expression on the recognition of emotional expression in the voice ([de Gelder and Vroomen, 2000](#), Experiment 3). Participants judging the emotional tone of a voice (happy or fearful) were systematically affected by the expression of a face (e.g., the voice was judged as less fearful when a smiling face was seen). Since this first report, we have applied a number of different methods (behavioural, neurophysiological, TMS and

brain imaging) to further analyse several aspects of this integration process, like its time course, the role of attention and of awareness in generation integration, and the neuroanatomical substrates (de Gelder et al., 1999; Dolan et al., 2001; Pourtois et al., 2000, 2002; Vroomen et al., 2001).

In an earlier study on audiovisual integration deficits in schizophrenia, we found that crossmodal spatial integration was normal when the stimuli consisted of simple tones and light flashes, but a deficit was revealed when more complex linguistic stimuli were used (de Gelder et al., 2003). Here, our question was whether the basic process of multisensory integration of emotions is intact in schizophrenics. We conducted two experiments adapted from the ones used with normal subjects. We investigated both the influence of an emotional tone of voice on face categorisation, and that of a facial expression on the categorisation of a voice.

2. Experiment 1. The effect of an emotional tone of voice on categorization of facial expression

The goal of the first experiment was to examine whether in a situation of bi-sensory input about emotional state information from the voice influences decisions about the facial expression. On each bimodal trial, a still photograph of a face taken from an 11-step continuum extending between two natural tokens expressing respectively sadness or happiness was presented on the PC screen. At the same time, participants heard a voice speaking a short sentence in one of two same emotional tones. The subjects were asked to indicate, by pressing one of two keys, whether the face was happy or sad, and to ignore the voice.

2.1. Method

2.1.1. Participants

A group of 13 schizophrenic patients (11 male and 2 female) was compared with a group of 13 normal controls matched for gender, age and socio-economic status. Mean age of the patients was 37 years (range=27–51). Only patients meeting the criteria for schizophrenia set by the DSM-IV (Bertelson and de Gelder, in press) were included. All patients were

under treatment at the local hospital. Diagnosis was established with the Schedules for Clinical Assessment in Neuropsychiatry (SCAN), a standardised interview for diagnosing axis I disorders, conducted by two trained interviewers. All patients were on antipsychotic medication, receiving benztioxasol, thioxanthene, thienobenzothiazepine, butyrophenon or dibenzothiazepine.

The control group was also screened with the SCAN. Subjects with a psychiatric disorder, a brain dysfunction or a genetic predisposition for schizophrenia were excluded from participation. All participants were paid for participation.

2.1.2. Materials and procedure

Visual materials consisted of 11 black-and-white photographs making a continuum between a sad and a happy expression that were used in an earlier study (de Gelder and Vroomen, 2000). The two end-photographs were of a male model from the Ekman and Friesen (1976) series, posing a sad and a happy expression. Nine intermediate faces were obtained by a morphing procedure. Each photograph occupied a 9.5×6.5 cm rectangle on the computer screen, which at the mean viewing distance of 60 cm corresponding to a visual angle of $10.0 \times 6.8^\circ$.

2.1.3. Auditory material

A semantically neutral sentence (“Zijn vriendin kwam met het vliegtuig”, which means “His girlfriend came by plane”) was spoken by a professional male Dutch actor who had been instructed to pronounce it once ‘as if he was happy’ and the other time ‘as if he was sad’. The sentences were recorded on Digital Audio Tape. The duration of the happy utterance was 1.78 s, with a mean F_0 of 205 Hz (S.D.=39.3); the sad utterance had a duration of 2.12 s, with a mean F_0 of 170 Hz (S.D.=19.2). Further acoustic details can be found in de Gelder and Vroomen (2000, Experiment 2).

2.1.4. Procedure

Subjects were tested individually in a quiet room. After they pushed a button to initiate a trial, stimuli were presented 750 ms later. On *visual-only trials*, one of the 11 photographs was shown for 500 ms without any auditory accompaniment. On *bimodal trials*, the auditory stimulus (one of the two utterances) started and 300 ms after onset, a facial expression appeared

on the screen for 500 ms. This procedure was determined by the fact that we needed to combine a short visual presentation with a much longer auditory fragment and by the fact that the initial segment of the auditory fragment does not contain emotional information as this tends to be presented in a cumulative fashion over the duration of the fragment. Subjects pressed as fast as possible one of two keys to indicate whether the face was happy or sad. Testing was divided into five blocks of 33 trials each, preceded by a short warming-up period of four trials. Each block consisted of the 22 possible bimodal trials (11 faces \times 2 utterances), and 11 visual trials, all presented in random order.

Subjects were seated at a 60-cm distance from the computer screen. They were informed of the different types of trials, which were demonstrated during a practice phase. They were instructed to ignore the voice and to watch the screen in order to judge whether the face was sad or happy. Subjects pressed one of two response keys labelled ‘sad’ and ‘happy’. The computer recorded responses.

2.2. Results

Fig. 1 presents the proportion of ‘sad’ responses as a function of facial expression separately for each voice and group. **Fig. 1**, the auditory-presented ‘happy’ sentence shifted the identification function towards the ‘happy’ end of the continuum, and the

‘sad’ utterance shifted it to the ‘sad’ end if compared to the visual-only trials (see **Fig. 1**). This crossmodal bias effect of a voice on a face was larger in normal subjects than in schizophrenics. An ANOVA on the proportion of sad responses was performed with group (normal or schizophrenic) as between-subjects factor and voice (happy, sad or no) and face as within-subjects factors. As expected, the average proportion of ‘sad’ responses increased as the face moved from the happy to the sad end of the continuum, $F(10,240)=304.99, p < 0.001$. The effect of the Voice was significant, $F(2,48)=15.46, p < 0.001$, as was the interaction between Face and Voice, $F(20,480)=4.24, p < 0.001$. The interaction reflects the fact that the effect of the Voice was larger at the ambiguous levels of the face continuum. These results closely replicate those of **de Gelder and Vroomen (2000)**. Of most interest in the present study is the comparison between groups. There was a marginally significant main effect of Group, $F(1,26)=3.73, p=0.065$, because normal controls were more biased to respond ‘sad’ than schizophrenics. More interestingly, there was a significant interaction between Face and Group, $F(10,240)=2.85, p < 0.002$, and a significant interaction between Voice and Group, $F(2,48)=5.72, p < 0.006$. The latter indicated that the effect of the voice (i.e., the average difference between the identification functions of the ‘happy’ and ‘sad’ voice) was about four times bigger in normal controls (0.21) than in schizophrenics (0.05). Normal con-

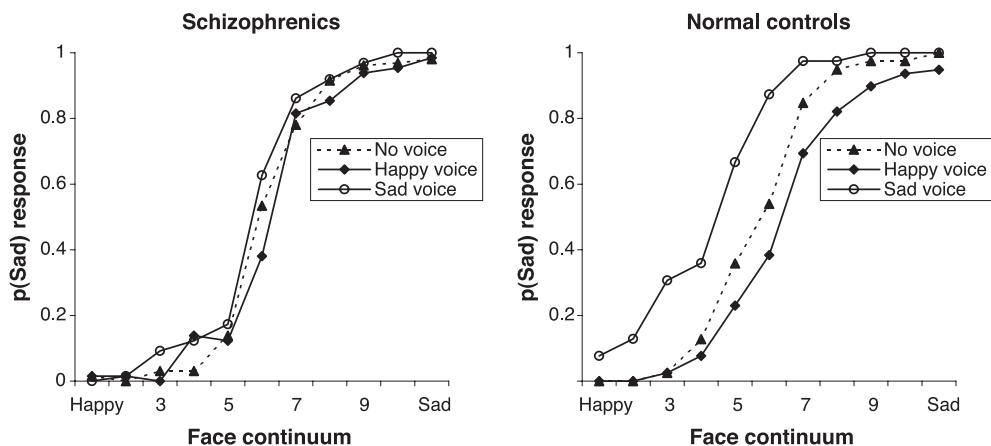


Fig. 1. Face categorization: The proportion of ‘sad’ responses as a function of the face continuum when combined with a happy, sad, or no voice, separately for the schizophrenic (left panel) and normal group (right panel). Face categorization of the schizophrenics was less affected by the voice than normal controls.

trols were thus more influenced by the emotion in the voice than schizophrenics.

2.3. Discussion

The goal of this experiment was to investigate the effect of the emotional content of a spoken sentence on the categorisation of facial expressions. The results obtained from the normal control group replicate our previous study with college students (de Gelder and Vroomen, 2000). They indicate that in this normal group, crossmodal bias of voice on emotion in the face is reduced, corresponding to *hypo-integration*. Yet, the data clearly indicate that like the normal controls, the schizophrenic subjects performed the task as instructed, by attending to the face. They were categorising the faces and not randomly responding sometimes to the voice only and sometimes to the face only. If the responses were solely based on the voice expressions we would have obtained two flat response curves corresponding to the voice expression only, as we reported previously for a patient who was unable to recognise facial expressions (de Gelder et al., 2000). It is also worth stressing that we do not observe in the schizophrenic group evidence for an impaired ability to focus attention only on the aspects of the stimulus situation that are relevant for the task. Such a failure of selective attention would show itself as a *larger* effect of the task-irrelevant voice expression.

Our next question was whether a similarly reduced cross-modal bias effect also obtains in a situation where the target is the voice and the distractor is the face.

3. Experiment 2. The effect of a face on voice categorization

Experiment 2 used a similar situation as in Experiment 1, but this time, we studied the effect of a face on the perception of expression in the voice, using a voice continuum obtained by morphing between a naturalistic token of a sentence expressing fear and one expressing happiness.

3.1. Method

3.1.1. Participants

The same group was tested as before.

3.1.2. Auditory materials

The same materials were used as described in de Gelder and Vroomen (2000, Experiment 3). In short, the prosody of a Dutch sentence with semantically neutral content ('Zijn vriendin kwam met het vliegtuig', meaning 'His girlfriend came by plane') was modified so as to create a continuum between 'fear' and 'happy'. Changing simultaneously the duration, pitch range and pitch register of the utterances created the continuum. In order to change the pitch in equal steps, the original pitch contour was replaced by a minimal sequence of straight-line approximations, while the perceptual identity remained close to the original one. A program computed the various pitch movements by superimposing them on a declination line. Then, only two free parameters need to be set: the excursion size of the pitch movements in semitones and the place in the pitch register. For the 'happy' endpoint of the continuum, the excursion size was set at 10 semitones and the end-frequency at 150 Hz. For each next stimulus in the continuum, the excursion size was decreased with one semitone and the end-frequency was increased with 12 Hz. Thus, the seventh stimulus at the 'fear' endpoint had an excursion size of four semitones and an end-frequency of 222 Hz. The duration of the thus created utterances was then linearly compressed. The duration of the utterance at the 'happy' endpoint was left at 100% (i.e., 1.58 s) and the duration of each next stimulus in the continuum was decreased with 2% so that duration at the 'fear' endpoint was 88% (i.e., 1.39 s). All pitch and time manipulations were based on direct waveform manipulations so that the tokens sounded natural. These auditory stimuli were played directly from hard disk and presented at a comfortable listening level over loudspeakers.

3.1.3. Visual materials

The visual stimuli consisted of two facial expressions ('happy' and 'fear') of the male actor who had recorded the original sentences. The two black-and-white pictures were positioned in a 23 × 16-cm frame against a dark background. A face was presented 300 ms after the onset of the voice and stayed on the screen for the duration of the utterance. Participants sat at a distance of approximately 60 cm from a PC screen on which the pictures were presented.

3.1.4. Design and procedure

The experiment consisted of 70 experimental trials (5 repetitions of the 2 face \times 7 voice combinations). All stimuli were presented in five pseudo-randomly ordered blocks. Before testing started, subjects were given a short practice session. Subjects were instructed to decide whether the voice expressed fear (left button) or happiness (right button). They were told to base their judgement on the voice only and to ignore the face.

3.2. Results

Fig. 2 presents the proportion of ‘happy’ responses as a function of the auditory continuum, separately for each group. As is apparent from this figure, schizophrenics were, in comparison with the normal controls, more variable in labelling the voice tokens. Moreover, the cross-modal effect of the face on the voice was larger in schizophrenics than in normal controls. An ANOVA on the proportion of happy responses was performed with Group (normal or schizophrenic) as between-subjects factor and Face (happy, fear) and Voice as within-subjects factors. As expected, the average proportion of ‘happy’ responses increased as the Voice moved from the ‘fear’ to the ‘happy’ end of the continuum, $F(6,144)=60.85, p<0.001$. The effect of the Face was significant, $F(1,24)=31.87, p<0.001$, as was the interaction between Face and Voice, $F(6,144)=3.85, p<0.001$. The interaction reflects the fact that the effect of the face was larger at the

ambiguous levels of the voice continuum. There was a significant interaction between Face and Group, $F(1,24)=12.34, p<0.001$, indicating that the impact of the face (as measured as the average difference between the identification functions of the voice when combined with a happy versus a fear face) was smaller in normal controls (i.e., 0.04) than in schizophrenics (0.20). There was also a significant interaction between Voice and Group, $F(6,144)=8.34, p<0.001$, as reflected in the less steep and thus more variable identification functions of the schizophrenics.

3.3. Discussion

In Experiment 2, we find again that schizophrenics present an abnormal pattern of multisensory integration. This time, the anomalous pattern consists, however, in an excessive rather than a reduced bias from the to be ignored input channel. In contrast with the pattern of hypo-integration obtained in the previous experiment, we refer to this pattern as *hyper-integration*. This reversal of the previous kind of anomaly is interesting for a number of reasons. One is that it directly speaks to one possible explanation of the reduced bias observed in Experiment 1, an absence of integration processes between the input to the two channels. The excessive integration that is observed here, rules out that the two sensory systems providing emotional information do not connect at all in schizophrenics. If that was the case, we would have

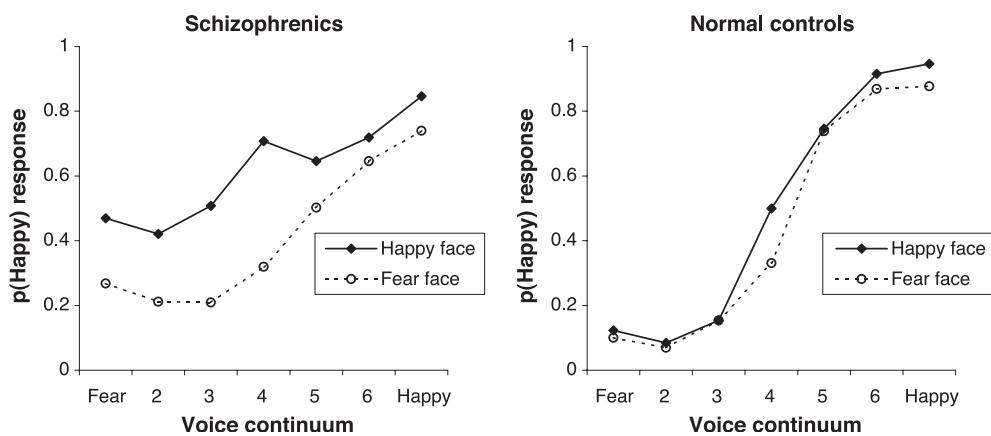


Fig. 2. Voice categorization: The proportion of ‘happy’ responses as a function of the voice continuum when combined with a happy or sad face, separately for the schizophrenic (left panel) and normal group (right panel). Voice categorization of the schizophrenics was more affected by the voice than normal controls.

observed no difference between the two voice conditions as the participants' responses would only have reflected the content of the voice expressions. How can this be explained? Task difficulty appears as a likely candidate as indicated by the finding that voice categorisation of the schizophrenics is very noisy even at the two endpoints of the continuum where the auditory tokens belong unambiguously to one or the other category. Apparently, schizophrenics found auditory tokens more difficult to recognize than visual tokens. Models of intersensory integration predict that with bi-sensory input, subjects' responses to a difficult or ambiguous target stimulus are more influenced by the secondary stimulus, even if this is irrelevant for performing the task (Massaro, 1998). Since schizophrenics have more difficulty categorizing the auditory tokens, as is apparent in the light of the noisy categorisations at the two endpoints, they are more affected by the secondary visual stimulus. A further, more interesting issue is then why this auditory task is so difficult for the schizophrenic group. Is it because of the nature of the auditory materials or because of the specific task instructions to ignore the visual input? Previous studies have provided evidence for impairments in perception of emotional prosody in schizophrenics (Ross et al., 2001). A different explanation, but obviously not one that is incompatible, is related to the present task instructions. Participants were asked to respond by categorising the voice expressions and pay no attention to the face expressions. The focussed attention required here might be difficult to achieve for schizophrenics and more so when the target materials are difficult.

4. General discussion

Results of Experiment 1 clearly showed a reduced impact of the voice on face categorization, while in Experiment 2, this pattern was reversed. These results replicate findings from a pilot study of voice-to-face integration (de Gelder et al., 1997). A comparison with the findings from our previous study (de Gelder et al., 2003) allows us to exclude a generalized audiovisual integration deficit that would affect any situation in which an auditory and a visual input are to be combined. Rather, the most straightforward explanation seems to be that schizophrenics have a relative

difficulty recognizing emotions in the voice, which explains the observed pattern in the bimodal situations. A separate but related issue is that the contrast between hypo- and hyper-integration is the result of a more general factor like stronger visual than auditory dominance in audiovisual perception.

As indicated by our previous findings, attention is by itself an unlikely candidate for explaining intersensory integration (Vroomen et al., 2001). As findings from our own and other laboratories clearly indicate, attention is not the critical element in intersensory integration (see de Gelder and Bertelson, 2003 for review). But this does not mean that attention cannot influence performance indirectly and play a role in multiple stimulation setting requiring a focussed attention task; and its role can be more important as the kind of integration that does not require attention fails. Therefore, at this stage, we do not want to exclude the possibility that attention problems including attention disengagement problems might partly explain the present anomalous results of the schizophrenic group. Further research is needed to disentangle the role of on-line integration and that of post-perceptual influences due to attentional factors. The observed pattern might also be partly related to a phenomenon of *auditory hypoemotionality*, similar to the pattern of visual hypoemotionality reported by Bauer (1982). Alternatively, there could be a higher dominance of visual (or just facial) inputs over auditory (or just vocal) ones in this group. Related to this, the pattern of dominance might also depend on the specific emotions being tested.

Recent studies of stimulus integration and related theoretical models have begun to highlight impairments in how context influences stimulus processing. Abnormal changes in context sensitivity could be related to NMDA hypoactivity in schizophrenia, a hypothesis explored and defended in Phillips and Silverstein (in press). Some recent results appear particularly important for the present discussion like impaired sensory gating in auditory context (Jentsch and Roth, 1999), abnormal pattern on Stroop test (Barch et al., 1999) or impaired ability to alter perceptual organisation of ambiguous stimuli based on current context (Silverstein et al., 1996).

In conclusion, we have observed two patterns of anomalous audiovisual integration of emotional information, one corresponding to reduced integration, the other to exaggerated integration. Further research is

needed to explore whether the perceptual system has a deficit in automatic processes such that what are normally smooth automatic perceptual processes not requiring attention, spill over into the realm of post-perceptual decisions. This in turn opens the door for attentional factors to play a role. We are currently testing a new group of schizophrenic patients in order to explore anomalous attentional effects by means of a dual-task methodology (Vroomen et al., 2001).

Acknowledgements

Research was partly supported by Janssen Pharmaceutica. We are grateful to L. Annen for the assistance with data collection and to all subjects for their participation.

References

- Barch, D.M., Carter, C.S., Hachten, P.C., Usher, M., Cohen, J.D., 1999. The 'benefits' of distractability: mechanisms underlying increased stroop effects in schizophrenia. *Schizophrenia Bulletin* 25, 749–762.
- Bauer, R.M., 1982. Visual hypoemotionality as a symptom of visual-limbic disconnection in ma. *Archives of Neurology* 39 (11), 702–708.
- Bertelson, P., de Gelder, B., in press. The psychology of multi-sensory perception. In Spence, C. and Driver, J. (Eds.), *Cross-Modal Space and Cross-Modal Attention*. Oxford: Oxford University Press.
- Borod, J.C., Martin, C.C., Alpert, M., Brozgold, A., Welkowitz, J., 1993. Perception of facial emotion in schizophrenia and right brain-damaged patients. *Journal of Nervous and Mental Disorders* 181, 494–502.
- Borod, J.C., Cicero, B.A., Obler, L.K., Welkowitz, J., Erhan, H.M., Santschi, C., Grunwald, I.S., Agosti, R.M., Whalen, J.R., 1998. Right hemisphere emotional perception: evidence across multiple channels. *Neuropsychology* 12, 446–458.
- de Gelder, B., Bertelson, P., 2003. Multisensory integration, levels of processing and ecological validity. *Trends in Cognitive Science* 7 (10), 460–467.
- de Gelder, B., Vroomen, J., 2000. The perception of emotions by ear and by eye. *Cognition and Emotion* 14, 289–311.
- de Gelder, B., Vroomen, J., Teunisse, J.P., 1995. Perceiving emotions by ear and by eye. *Bulletin of the Psychonomic Society* 30, 50 (abstract).
- de Gelder, B., Parnas, J., Bovet, P., Vroomen, J., Popelier, T., Innocenti, G.M., 1997. Impaired integration of audition and vision in schizophrenics. *Experimental Brain Research* 117, 23 (abstract).
- de Gelder, B., Böcker, K., Tuomainen, J., Hensen, M., Vroomen, J., 1999. The combined perception of emotion from voice and face: early interaction revealed by electric brain responses. *Neuroscience Letters* 260, 133–136.
- de Gelder, B., Pourtois, G., Vroomen, J., Bachoud-Levi, A.C., 2000. Covert processing of faces in prosopagnosia is restricted to facial expressions: evidence from cross-modal bias. *Brain and Cognition* 44, 425–444.
- de Gelder, B., Vroomen, J., Annen, L., Masthof, E., Hodamont, P., 2003. Audio-visual integration in schizophrenia. *Schizophrenia Research* 59 (2–3), 211–218.
- Dolan, R., Morris, J., de Gelder, B., 2001. Crossmodal binding of fear in voice and face. *Proceedings of the National Academy of Science* 98, 10006–10010.
- Ekman, P., Friessen, W.V., 1976. Measuring facial movement. *Journal of Environmental Psychology and Non-Verbal Behavior* 1, 56–75.
- Feinberg, T.E., Rifking, A., Schaffer, C., Walker, E., 1986. Facial discrimination and emotional recognition in schizophrenic and affective disorders. *Archives of General Psychiatry* 43, 276–279.
- Heinrichs, D.W., Buchanan, R.W., 1988. Significance and meaning of neurological signs in schizophrenia. *American Journal of Psychiatry* 145, 11–18.
- Ismail, B., Cantor-Graae, E., McNeil, T.F., 1998. Neurological abnormalities in schizophrenic patients and their siblings. *American Journal of Psychiatry* 155, 84–89.
- Jentsch, J.D., Roth, R.H., 1999. The neuropsychopharmacology of phencyclidine: from NMDA receptor hypofunction to the dopamine hypothesis of schizophrenia. *Neuropsychopharmacology* 20, 201–225.
- Kohler, C.G., Bilker, W., Hagendoorn, M., Gur, E.E., Gur, R.C., 2000. Emotion recognition deficit in schizophrenia: association with symptomatology and cognition. *Biological Psychiatry* 48, 127–136.
- Mandal, M.K., Pandey, R., Prasad, A.B., 1998. Facial expressions of emotions and schizophrenia: a review. *Schizophrenia Bulletin* 24, 399–412.
- Massaro, D.W., 1998. *Perceiving Talking Faces: From Speech Perception to a Behavioural Principle*. MIT Press, Cambridge.
- Massaro, D.W., Egan, P.B., 1996. Perceiving affect from the voice and the face. *Psychonomic Bulletin and Review* 3, 215–221.
- Mueser, K.T., Doonan, R., Penn, D.L., Blanchard, J.J., Bellack, A.S., Nishith, P., DeLeon, J., 1996. Emotion recognition and social competence in chronic schizophrenia. *Journal of Abnormal Psychology* 105, 271–275.
- Phillips, W.A., Silverstein, S.M., 2003. Convergence of biological and psychological perspectives on cognitive coordination in schizophrenia. *Brain and Behavioral Sciences* 26, 65–82.
- Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B., Crommelink, M., 2000. The time course of intermodal binding between seeing and hearing affective information. *NeuroReport* 11, 1329–1333.
- Pourtois, G., Debatissé, D., Despland, P.A., de Gelder, B., 2002. Facial expressions modulate the time course of long latency auditory brain potentials. *Cognitive Brain Research* 14, 99–105.
- Ross, D.E., Buchanan, R.W., Medoff, D., Lahti, A.C., Thaker, G.K., 1998. Association between eye tracking disorder in schizophrenia

- and poor sensory integration. *American Journal of Psychiatry* 155, 1352–1357.
- Ross, E.D., Orbelo, D.M., Cartwright, J., Hansel, S., Burgard, M., Testa, J.A., Buck, R., 2001. Affective-prosodic deficits in schizophrenia: comparison to patients with brain damage and relation to schizophrenic symptoms. *Journal of Neurosurgery and Psychiatry* 70, 597–604.
- Silverstein, S.M., Matteson, S., Knight, R.A., 1996. Reduced top-down influence in auditory perceptual organization in schizophrenia. *Journal of Abnormal Psychology* 105, 663–667.
- Streit, M., Wölwer, W., Gaebel, W., 1997. Facial-affect recognition and visual scanning behaviour in the course of schizophrenia. *Schizophrenia Research* 24, 311–317.
- Vroomen, J., Driver, J., de Gelder, B., 2001. Is cross-modal integration of emotional expressions independent of attentional resources? *Cognitive, Affective, and Behavioral Neuroscience* 1, 382–387.
- Wölwer, W., Streit, M., Polzer, U., Gaebel, W., 1996. Facial affect recognition in the course of schizophrenia. *European Archives of Psychiatry and Clinical Neurosciences* 246, 165–170.