Clinical Neurophysiology 119 (2008) 2004–2010

www.elsevier.com/locate/clinph

Atypical processing of fearful face–voice pairs in Pervasive Developmental Disorder: An ERP study

Maurice J.C.M. Magnée[a,b,*], Beatrice de Gelder[b,d], Herman van Engeland[a], Chantal Kemner[a,c]

[a] Department of Child and Adolescent Psychiatry, Rudolf Magnus Institute of Neuroscience, University Medical Center, B01.201, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands
[b] Laboratory of Cognitive and Affective Neuroscience, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands
[c] Section Biological Developmental Psychology, Faculty of Psychology, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands
[d] Martinos Center for Biomedical Imaging, Massachusetts General Hospital and Harvard Medical School, Charlestown, MA 02129, USA

Accepted 11 May 2008
Available online 19 June 2008

Abstract

Objective: An important premise for successful social-affective communication is rapid perception of visual and auditory emotional cues, as well as their multisensory integration (MSI). We investigated to what extent a deficit in recognition of emotions in individuals with Pervasive Developmental Disorder (PDD) may have its roots in abnormal MSI of emotional cues provided by the sight of a facial expression and an emotional tone of voice.
Methods: In twelve high-functioning, adult PDD individuals and thirteen age- and IQ-matched controls, (1) the processing of fearful faces was compared with that of happy faces; (2) MSI was assessed by characterizing the interaction effects of crossmodal presentation, using EEG.
Results: Increased P1 and N170 amplitudes were seen in response to fearful faces compared with happy faces in both groups. However, PDD individuals differed from healthy controls in MSI of fearful information from visual and auditory cues.
Conclusions: Both groups show a similar pattern as concerns the early components of visual emotion processing, but there are anomalies in processing of fearful face–voice combinations in the PDD group.
Significance: Because of the importance of rapid MSI for social competence, MSI anomalies in PDD may be linked to the observed deficits in their emotional behavior.

Keywords: Autism; Facial expression; Multisensory processing; Voice prosody; EEG

1. Introduction

Pervasive Developmental Disorder (PDD) refers to a group of neurodevelopmental disorders characterized by impairments in social–emotional behavior and communication, co-occurring with repetitive, stereotyped behaviors (APA, 1994). A poor insight into the emotions of others seems to play a central role in the social interaction problems in this group, and it has been suggested that this is related to poor recognition of facial expressions (Bachevalier and Loveland, 2006). Recent evidence points towards a specific deficit in the recognition of threatening information (e.g., Ashwin et al., 2007; Humphreys et al., 2007). However, an alternative explanation is that these symptoms arise from an earlier impairment in the processing of (fearful) stimuli, more related to perception, rather than a more cognitive recognition deficit. Very few studies on emotion processing in PDD have focused on this question, but

* Corresponding author. Address: Department of Child and Adolescent Psychiatry, University Medical Center Utrecht, B01.201, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands. Tel.: +31 88 7556026; fax: +31 88 7555444.
E-mail address: M.J.C.M.Magnee@umcutrecht.nl (M. J.C.M. Magnée).

1388-2457/$34.00
doi:10.1016/j.clinph.2008.05.005

evidence for an abnormal perceptual component in autism is increasing (Behrmann et al., 2006; Dakin and Frith, 2005).

An important aspect of processing of emotional stimuli in every day life is that signals arrive from different modalities at the same time. For example, visual-emotional signals (facial expressions) are usually accompanied by specific auditory signals (vocal expressions). There is increasing evidence that multisensory integration (MSI) of these stimuli is of great importance for fast and accurate emotion recognition. This is shown in a study that found improved judgment accuracy and speed for bimodal vs. unimodal recognition of emotions (de Gelder and Vroomen, 2000). Therefore, the question can be raised to what extent social-affective impairments in PDD originate from impairments in perceptual integration of affective information.

In one of the few studies investigating MSI of emotions in PDD, Hobson (1986) found behavioral impairments in the autistic children's ability to match photographed facial expressions with emotional gestures and voices, although in a replication of this experiment (Prior et al., 1990), no impairments were found. In a PET study, Hall and colleagues found a different pattern of cerebral blood flow in individuals with PDD than controls during the processing of facial expressions accompanied by prosodic information. However, there is little information what stage of emotion processing is affected in PDD, and how this is related to MSI. One way to reliably assess the different temporal phases of the processing of emotional stimuli is by using Event Related Potentials (ERPs).

An early ERP response that shows sensitivity to faces is the P1 component (Itier and Taylor, 2002). The P1 is an occipital positive potential around 120 ms that is enhanced by selective attention and presumably generated in the extra-striate cortex. The N170 component occurs later in time and reflects a distinct stage of processing. The N170 is a negative deflection around 170 ms at bilateral (although slightly right-lateralized) occipital-temporal sites. It is most commonly associated with higher level visual processing of faces such as structural encoding of facial configuration (Bentin et al., 1996). Several studies now show that P1 as well as N170 amplitudes are influenced by facial expression, especially fear (Batty and Taylor, 2003; Halgren et al., 2000; Pourtois et al., 2005; Stekelenburg and de Gelder, 2004). These results confirm the hypothesis that emotional expression can modulate face processing even before structural encoding of the face is completed (Streit et al., 2003). So far, some studies in PDD have looked at the N170 as an index of facial configuration processing (e.g., O'Connor et al., 2005), but no studies have looked at these early modulations by facial expressions.

In addition, several studies with typically developing individuals now provide evidence of an early perceptual integration of face with voice information in the processing of affect (de Gelder et al., 1999; Pourtois et al., 2000). Furthermore, recent studies indicate that ERP peaks that are usually thought to reflect early perceptual processing in sensory-specific cortices can be affected by crossmodal interactions. For instance, the frontal–central P2 component, which is known to reflect activity from auditory cortical areas, has been shown to be sensitive to the congruency between emotions conveyed through the face and the voice (Pourtois et al., 2002). Additionally, an fMRI study investigating MSI of emotional stimuli showed enhanced neural activity in the fusiform gyrus to fearful faces only in congruency with fearful voices (Dolan et al., 2001). These results provide evidence of crossmodal effects on sensory-specific brain areas.

Our experimental paradigm consisted of the measurement of ERPs while participants were presented pairs of emotional faces and voices. Each trial started with the presentation of a happy or a fearful face, which was after a 900-ms delay followed by a voice uttering a short sentence in either a happy or fearful tone of voice. This design allowed us to: (1) compare the processing of fearful faces with that of happy ones; (2) assess MSI by characterizing the interaction effects of crossmodal presentation. In the first place, finding increased P1 and N170 amplitudes in response to the presentation of fearful faces compared to happy faces, would indicate enhanced visual processing of the fearful event. Second, we examined electrophysiological responses to emotionally congruent and incongruent face–voice pairs at frontal–central (reflecting activity from auditory cortical areas) and occipital-temporal (reflecting activity from visual cortical areas) sites. In line with previous research, we hypothesized that in controls the largest audiovisual (AV) modulation of sensory-specific activity would be seen in response to congruent fearful face–voice pairs, on both visual and auditory cortical ERP components. A result indicating that this effect is not obtained in individuals with PDD, would provide evidence for a disintegrated processing of visual and auditory emotions in this group.

2. Methods

2.1. Participants

Twelve high functioning, medication free, adult males with PDD (average age 21.5, SD 4.0; one left-handed) and thirteen healthy adult males (average age 23.0, SD 2.9; one left-handed) participated in the study. All individuals were administered the Wechsler Adult Intelligence Scale, Dutch edition (WAIS-III-NL). Mean age and total IQ scores were similar for individuals with PDD (IQ 122.4, SD 9.2) and individuals from the control group (IQ 127, SD 14.4). All individuals were free of recent substance abuse, seizure disorders, neurological diseases, head trauma or mental retardation. Additionally, all individuals in the control group were free of neurological and psychiatric history, and familial history of psychiatric disorders. All diagnoses of PDD (either Autistic Disorder or Asperger

Syndrome) were based on DSM-IV criteria and were made by a child psychiatrist. Additionally, all patients were administered the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 1989) by a trained rater, and their parents were informants on the Autism Diagnostic Interview Revised (ADI-R; Lord et al., 1994). Mean ADOS scores were 4.1 (SD 1.8) for the subscale 'Communication' (cut-off score 2), and 8.4 (SD 3.6) for 'Social behavior' (cut-off score 4). Mean ADI scores were 19.2 (SD 3.6) for 'Social behavior' (cut-off score 10), 15.6 (SD 4.9) for 'Communication' (cut-off score 8), 5.6 (SD 3.3) for 'Repetitive behaviors' (cut-off score 3), and 2.7 (SD 1.1) for 'Age of onset' (cut-off score 1). All participants had normal or corrected to normal vision. They were all paid for their participation. Written informed consent was obtained for each participant before the session, according to the Declaration of Helsinki (2000). Approval of the medical ethics committee of the University Medical Center Utrecht was obtained prior to the study.

2.2. Stimuli and procedure

Stimuli consisted of AV stimulus pairs with either a congruent or an incongruent affective content. The visual component consisted of one of six happy and six fearful faces (equally matched between male and female pictures) taken from the Ekman series (Ekman and Friesen, 1976). The auditory component consisted of spoken sentence fragments with a neutral content, which were pronounced in either a happy or fearful tone of voice (the Dutch sentence fragment 'met het vliegtuig' meaning 'by plane'). Auditory stimuli were tested to eight volunteers who did not participate in the experiment to ensure that stimuli were perceived as belonging to the appropriate category. Each visual stimulus was combined with a different spoken fragment, resulting in 12 congruent and 12 incongruent stimulus pairs. Each face–voice stimulus pair was derived from actors of the same sex and the same actors participated in the happy and fearful stimulus combinations. Size of the face pictures was 19 cm height by 13 cm width, which at the mean viewing distance of 80 cm corresponds to a visual angle of $13.5° \times 9.2°$. Luminance was 38 cd/m^2 on a 2.5 cd/m^2 background. The mean level for sound was 60 dB(a), delivered over one loudspeaker placed directly below the screen. Mean levels for sound and luminance were equal across both happy and fearful stimuli.

A trial started with the presentation of the face. After 900 ms, the auditory stimulus was presented, whereas the face remained on screen until the end of the voice fragment. This delay was introduced to be able to analyze the visual as well as the AV ERP responses separately.

The six resulting stimulus categories were as follows: visual happy (H), visual fear (F), congruent AV happy (Hh), congruent AV fear (Ff), incongruent visual fear-auditory happy (Fh) and incongruent visual happy-auditory fear (Hf). Participants were comfortably seated in a chair in a soundproof experimental chamber. They were instructed to judge the sex of each stimulus pair, by pushing one of two designated buttons on a response box. To avoid any response-related components in the ongoing ERP signal, they were instructed not to respond until after the visual stimulus was withdrawn. Intertrial interval was chosen randomly between 1000 and 1500 ms, immediately after the participant's response. During this interval a central fixation cross was presented on screen. Stimuli within a total of eight blocks of 24 AV trials were presented randomly (96 H and 96 F trials; 48 trials for each AV stimulus combination).

2.3. Recordings

EEG's were recorded from 48 locations using standard Ag/AgCl pin-type active electrodes (BIOSEMI) mounted in an elastic cap, referenced to an additional active electrode (Common Mode Sense) during recording. EEG signals were band-pass filtered (0.1–30 Hz) at a sample rate of 512 Hz and offline referenced to an average reference. Horizontal and vertical EOG's were measured for offline correction.

The raw data were segmented into epochs for visual and AV categories separately. All categories consisted of 900-ms epochs, including a 100-ms prestimulus baseline. After EOG correction, epochs with amplitudes exceeding ± 100 μV at any channel were automatically rejected. Lowest allowed activity was 3 μV/200 ms, and the maximal allowed voltage step per sampling point was 50 μV. ERPs were averaged separately for all six stimulus categories.

2.4. Data analysis

Electrode and time window selection for P1 and N170 was based on previous studies (Batty and Taylor, 2003); P1 was measured as mean of the peak latencies and amplitudes over occipital sites (O1/O2) and occipital-temporal sites (PO3/PO4, PO7/PO8) in the time window 80–140 ms. Mean N170 was measured over bilateral occipital-temporal sites (P5/P6, P7/P8, PO7/PO8) as maximal negative peaks in the time window 130–210 ms. Mean amplitudes and latencies of auditory evoked potentials were measured relative to the maximum positivity at frontal–central electrodes (F3, Fz, F4, FC1, FC2) in the 160–240 ms interval (P2 component). Furthermore, characteristics of the auditory response were measured at occipital-temporal sites (P5/P6, P7/P8, PO7/PO8), as maximal negative peaks in the time window 160–240 ms (for electrode positions see Fig. 1). Electrode selection for this peak was based on previous research yielding these sites as the ERP correlates of the fusiform gyrus (Itier and Taylor, 2002). At these sites clear negative peaks were observed around 200 ms; therefore this potential is referred to as the auditory N2 component.

Separate multivariate analyses for repeated measures were performed to analyze the effects of visual and AV stimuli. We first tested whether P1 and N170 amplitudes

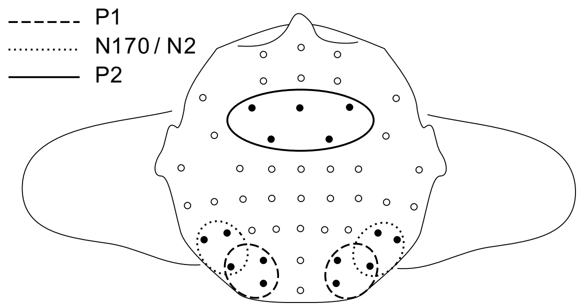

Fig. 1. Scalp locations on which visual P1, visual N170, auditory P2, and auditory N2 were measured.

and latencies differed in response to happy and fearful faces, and whether there were differences in this respect between groups. Both analyses consisted of one between-subjects factor Group (PDD vs. control group), and the within-subjects factors Hemisphere (left vs. right) and Emotion (happy vs. fear).

Second, we wanted to determine interaction effects between processing of facial expressions and emotional tone of voice. For this purpose, we tested whether auditory P2 amplitudes and latencies differed for the four AV stimulus categories. Analyses consisted of the between-subjects factor Group (PDD vs. control group), and the within-subjects factors Emotion of the auditory stimulus (happy vs. fear), and Congruency (congruent vs. incongruent). Furthermore, to study whether interaction effects were also seen on (bilateral) visual cortical areas, additional analyses were done with a different set of more occipitally located electrodes (N2 component). Analyses were arranged in similar within-subjects factors and one new factor, namely, Hemisphere (left vs. right). Possible differences between groups related to AV interaction effects would be demonstrated as a specific interaction between the factors Emotion of the auditory stimulus, Congruency and Group. In case this planned interaction was found, we continued with determining how the response to an emotional tone of voice was modulated by the preceded processing of a particular emotional expression, and how this differed between both groups. Therefore, we compared congruent with incongruent AV interactions for both auditory conditions (happy and fear) separately, using multivariate analyses for repeated measures. To test the effect for the fearful auditory condition, we compared the Ff with the Hf condition; for the happy auditory condition we compared the Hh with the Fh condition.

3. Results

3.1. Visually evoked potentials

P1 amplitude values of one PDD individual deviated more than three standard deviations from the group mean. This individual was excluded from further P1 analysis. There was a main effect of Emotion on P1 amplitude, $F(1,22) = 7.20$, $p < 0.05$, reflecting increased amplitudes for fearful faces (mean ± SE; 8.9 μV ± 0.8) compared to happy faces (8.2 μV ± 0.8). This effect was observed similarly across both groups, and indeed there was no significant interaction between Emotion and Group. The factor Emotion had no effect on P1 latencies, (mean latency 122.2 ms ± 1.9).

Second, the influence of facial expressions on the N170 component was tested. A main effect was observed for the factor Emotion, $F(1,23) = 8.02$, $p < 0.01$, indicating larger negative amplitudes for fearful faces (−0.83 μV ± 0.6) compared to happy faces (−0.46 μV ± 0.5). This difference was significant across both groups (see Table 1), with no interactions. Emotion had no significant effect on N170 latencies (mean latency 175.2 ms ± 3.8).

3.2. Evoked potentials related to AV interaction

The factor Emotion of the auditory stimulus had a significant effect on the (auditory) P2 amplitudes, $F(1,23) = 7.79$, $p < 0.01$, meaning that both auditory fear conditions (Ff and Hf) showed increased amplitudes (3.21 μV ± 0.3) compared to both auditory happy (Hh and Fh) conditions (2.58 μV ± 0.3). There was no interaction between Emotion and Congruency, and no differences were found between groups. There were no significant effects of Emotion of the auditory stimulus on P2 latencies.

Regarding the (visual) N2 amplitudes, there was a significant effect of the main factor Emotion of the auditory stimulus, $F(1,23) = 6.78$, $p < 0.05$, with largest N2 amplitudes in the auditory fearful condition (−2.40 μV ± 0.2), compared to auditory happy (−1.90 μV ± 0.3), and no significant group interaction. Differences between groups related to AV interaction effects resulted in the predicted interaction between the factors Emotion of the auditory stimulus, Congruency and Group, $F(1,23) = 4.11$, $p < 0.1$. Next, we determined how the response to an emotional tone of voice was modulated by the concurrent processing of a particular emotional expression, and how this differed between both groups. For the auditory fearful condition, we found a significant Congruency × Group interaction effect, $F(1,23) = 14.27$, $p < 0.01$. Congruent fearful AV conditions led in the control group to a significant increase in N2 amplitude (Ff −3.02 μV ± 0.3 vs. Hf −2.19 μV ±0.4), $F(1,12) = 5.83$, $p < 0.05$, but to a significant decrease among PDD individuals (Ff −1.70 μV ± 0.3 vs. Hh −2.70 μV ± 0.4), $F(1,11) = 8.67$, $p < 0.05$. Auditory happy conditions across both groups did not significantly differ between congruent and incongruent conditions (see Fig. 2). No significant hemispheric differences were found, and no significant effects were found on N2 latencies (mean latency 201.1 ms ± 1.1).

4. Discussion

The goal of the present study was to examine ERPs to emotional faces and voices in high-functioning individuals

Table 1
MANOVA summary for visual (P1 and N170) and auditory (P2 and N2) ERP peaks

		F	*p*	Results
Visual P1	Emotion	7.2	*	
	Emotion * Group	1.17	ns	Both groups fearful > happy
Visual N170	Emotion	8.01	**	
	Emotion * Group	0.01	ns	Both groups fearful > happy
Auditory P2	Emotion	7.79	**	
	Emotion * Group	0.01	ns	
	Emotion * Congruency	0.35	ns	Both groups fearful > happy
Auditory N2	Emotion	6.77	*	
	Emotion * Group	1.67	ns	
	Congruency	0.42	ns	
	Congruency * Group	5.35	*	
	Emotion * Congruency * Group	4.11	0.054	Group differences in AV congruency
→ Ff vs. Hf	Congruency	0.12	ns	
	Congruency * Group	14.3	**	Group differences in auditory fearful condition
→ Ff vs. Hf control group only	Congruency	5.83	*	Ff > Hf in control group
→ Ff vs. Hf PDD group only	Congruency	8.67	*	Ff < Hf in PDD group
→ Hh vs. Fh	Congruency	0.36	ns	
	Congruency * Group	0.11	ns	No Group differences in auditory happy condition

N2 Group interactions were post hoc analyzed for both auditory conditions separately. Fearful auditory conditions were analyzed for both Groups separately. *$p < 0.05$. **$p < 0.01$.

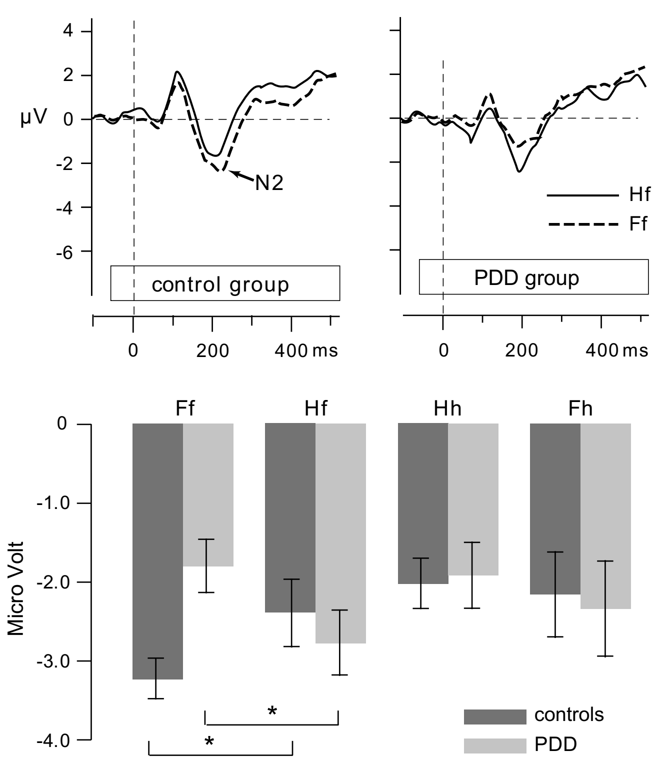

Fig. 2. The top panel depicts mean N2 peaks to congruent Fearful face–fearful voice (Ff) and incongruent Happy face–fearful voice (Hf) conditions for the control group (left) and the PDD group (right). The bottom panel shows mean N2 amplitudes (±SE) for both groups to all AV face–voice pairs: Fearful face–fearful voice (Ff), Happy face–fearful voice (Hf), Happy face–happy voice (Hh), Fearful face–happy voice (Fh). Asterisks indicate significant *p* values ($p < 0.05$).

with PDD and healthy controls. Specifically, we were interested in early perceptual responses to happy and fearful facial expressions and in MSI of affective information. Our results showed equivalent increases in P1 and N170 amplitudes at occipital-temporal sites in response to fearful faces as compared to happy faces in both groups. We can conclude that these rapid perceptual responses to facial expressions are intact in PDD individuals, and seem therefore not related to the displayed problems in emotional interaction. Furthermore, the results clearly indicated enhancement of the auditory P2 amplitude in response to fearful voices as compared to happy voices, an effect which was present in both the control group and the patient group. This P2 component is known to reflect activity from auditory cortical areas. Increased activity in auditory cortex has been found in response to negative emotional voices relative to neutral voices, which might suggest a similar mechanism as described for the fusiform cortical areas in the visual domain (Grandjean et al., 2005).

Second, we wanted to determine interaction effects between the processing of facial expressions and emotional tone of voice. For this purpose, we measured the effects of crossmodal presentation on ERP peaks that are thought to reflect perceptual processes in sensory-specific cortices. In healthy controls N2 amplitudes at occipital-temporal sites were enhanced to presentation of fearful voices compared to happy ones, but only when presented in the context of a fearful facial expression. On the contrary, individuals with PDD showed diminished N2 activity when fearful voices were processed in the context of fearful facial expressions. These results suggest that individuals with PDD differ from healthy controls in how they integrate threatening information from visual and auditory cues.

In view of this result, an important question is what this modulation of activity in visual cortex to fearful faces and voices reflects. Evidence from single-cell recordings in monkeys has shown enhanced firing of specific neurons to emotional (particularly fearful) compared to neutral images (Sugase et al., 1999). Recent brain imaging studies investigating the effects of emotions on visual processing suggest that this modulation of visual cortical activation is provided by functional coupling of the amygdala to the fusiform cortex, resulting in enhanced processing of the emotional event (Amaral et al., 2003). Interestingly, modulation of the fusiform response has been related to the integration of AV emotional information as well. Using fMRI, Dolan et al. (2001) and colleagues found increased amygdala and fusiform activation in response to fearful faces only when these were emotionally congruent with fearful voices. As in the present study, other emotional face–voice interactions showed no such effect. The authors argued that a functional consequence of this interaction between fearful information from faces and voices is enhanced attention and processing of the emotion-eliciting stimulus. One might hypothesize that the AV fear-related modulation of ERP activity in occipital-temporal areas we observed in the present study may be the electrophysiological correlate of these fMRI findings.

A number of research groups have proposed that PDD may be characterized by generalized abnormalities of neural connectivity (Brock et al., 2002; Just et al., 2004). We suggest that the present study provides evidence against a wide application of this model. Specifically, finding intact modulation of activity in visual cortical areas in response to fearful facial expressions argues for intact structural connections of these areas with the amygdala. On the other hand, the aberrant modulation of visual cortical activity to fearful AV stimuli might involve other brain areas, such as the STS, which appears to play an important role in MSI at later processing stages (Baylis et al., 1987). It is believed to be involved in cortical integration of both sensory and limbic information, and is furthermore associated with social perceptual skills (Allison et al., 2000). Based on the present findings it could be argued that individuals with PDD have impaired functional connectivity between the fusiform gyrus and STS.

However, besides interpreting the present data in the light of possible impairments in MSI, one should not overlook the possibility that the observed results in the patient group are due to other causes. Attention, for instance, can already modulate MSI at the earliest stages of the process, finding the strongest interactions when attention is divided between both unisensory objects (Talsma et al., 2006). In the present study attention was equally needed in all conditions through the use of a concurrent gender recognition task, although we cannot exclude that the gender decision was based on paying attention to the visual stimulus only. Therefore, differences in this respect between the two groups may have confounded the results. Future research has to be undertaken to disentangle attentional deficits from true integration abilities associated with the disorder. Further, it is important to recognize that in the present study all participants were young adults with high IQ. Interestingly, Putzar et al. (2007) argued that full development of MSI depends on adequate sensory input during the first months of life. Therefore, prospective developmental studies are needed to establish how early sensory processing is related to MSI in PDD. Also, from the present study it is not clear whether the results can be generalized to individuals with PDD suffering from intellectual disabilities. Our group sizes were relatively small, which may have reduced statistical power of our effects. However, we took great effort to put together a homogeneous group of individuals without mental retardation, as a result of which we can conclude that the found effects can be attributed to PDD only.

Taken together, the present results argue for intact processing of fearful facial expressions at the perceptual level, as indexed by ERP peaks P1 and N170, in adults with PDD. Furthermore, this study raises the matter of crossmodal processing of emotional faces and voices, as the ability to integrate multiple sources of perceptual input adds to the extent in which an individual efficiently processes environmental cues. Our findings indicate a modulation of visual cortical activity during the perception of fearful faces, only when they are accompanied by fearful voices and this effect is present differently in the control group than in the patients. A difference in this integration effect in individuals with PDD raises important implications concerning how impairments in emotional reciprocity might arise in this group. Given the developmental origin of the syndrome, one might even speculate that such a dysfunctional modulation early in life contributes to the development of abnormal social perception in this syndrome. Therefore, the present study argues in favor of the view that the observed problems in social-affective abilities in PDD are not entirely social in origin, but may have an early perceptual component. This might be of potential importance for exploring treatment options in individuals with PDD. Treatment is momentarily primarily based on social features of PDD, while perceptual resources, such as adjustments in stimulus environment, may also proove to be effective. However, there are alternative explanations for the findings in the present study, and there is certainly need for further exploration of the determinants of MSI in PDD.

Acknowledgement

The work described was supported by an Innovational Research Incentives Grant (VIDI-scheme, 402-01-094) of the Netherlands Organization for Scientific Research (NWO) to Chantal Kemner.

References

Allison T, Puce A, McCarthy G. Social perception from visual cues: role of the STS region. Trends Cogn Sci 2000;4:267–78.

Amaral DG, Behniea H, Kelly JL. Topographic organization of projections from the amygdala to the visual cortex in the macaque monkey. Neuroscience 2003;118:1099–120.

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4th ed. Washington, DC: American Psychiatric Association; 1994.

Ashwin C, Baron-Cohen S, Wheelwright S, O'Riordan M, Bullmore ET. Differential activation of the amygdala and the 'social brain' during fearful face-processing in Asperger Syndrome. Neuropsychologia 2007;45:2–14.

Bachevalier J, Loveland K. The orbitofrontal-amygdala circuit and self-regulation of social–emotional behavior in autism. Neurosci Biobehav Rev 2006;45:2–14.

Batty M, Taylor MJ. Early processing of the six basic facial emotional expressions. Cogn Brain Res 2003;17:613–20.

Baylis GC, Rolls ET, Leonard CM. Functional subdivisions of the temporal lobe neocortex. J Neurosci 1987;7:330–42.

Behrmann M, Thomas C, Humphreys K. Seeing it differently: visual processing in autism. Trends Cogn Sci 2006;10:258–64.

Bentin S, Allison T, Puce A, Perez E, McCarthy G. Electrophysiological studies of face perception in humans. J Cogn Neurosci 1996;8:551–65.

Brock J, Brown CC, Boucher J, Rippon G. The temporal binding deficit hypothesis of autism. Dev Psychopathol 2002;14:209–24.

Dakin S, Frith U. Vagaries of visual perception in autism. Neuron 2005;48:497–507.

de Gelder B, Böcker KB, Tuomainen J, Hensen M, Vroomen J. The combined perception of emotion from voice and face: early interaction revealed by human electric brain responses. Neurosci Lett 1999;260:133–6.

de Gelder B, Vroomen J. The perception of emotions by ear and eye. Cogn Emotion 2000;14:289–311.

Dolan RJ, Morris JS, de Gelder B. Crossmodal binding of fear in voice and face. Proc Natl Acad Sci USA 2001;98:10006–10.

Ekman P, Friesen WV. Pictures of facial affect. Palo-Alto: Consulting Psychologists Press; 1976.

Grandjean D, Sander D, Pourtois G, Schwartz S, Seghier ML, Scherer KR, et al. The voices of wrath: brain responses to angry prosody in meaningless speech. Nat Neurosci 2005;8:145–6.

Halgren E, Raij T, Marinkovic K, Jousmäki V, Hari R. Cognitive response profile of the human fusiform face area as determined by MEG. Cereb Cortex 2000;10:69–81.

Hall GBC, Szechtman H, Nahmias C. Enhanced salience and emotion recognition in autism: a PET study. Am J Psychiatry 2003;160:1439–41.

Hobson RP. The autistic child's appraisal of expressions of emotion. J Child Psychol Psychiatry 1986;27:321–42.

Humphreys K, Minshew N, Leonard GL, Behrmann M. A fine-grained analysis of facial expression processing in high-functioning adults with autism. Neuropsychologia 2007;45:685–95.

Itier RJ, Taylor MJ. Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: a repetition study using ERPs. Neuroimage 2002;15:353–72.

Just MA, Cherkassky VL, Keller TA, Minshew NJ. Cortical activation and synchronization during sentence comprehension in high-functioning autism: evidence of underconnectivity. Brain 2004;127:1811–21.

Lord C, Rutter M, Goode S, Heemsbergen J, Jordan H, Mawhood L, et al. Autism Diagnostic Observation Schedule: a standard observation of communicative and social behavior. J Autism Dev Disord 1989;19:185–212.

Lord C, Rutter M, Le Couteur A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. J Autism Dev Disord 1994;24:659–85.

O'Connor K, Hamm JP, Kirk IJ. The neurophysiological correlates of face processing in adults and children with Asperger's syndrome. Brain Cogn 2005;59:82–95.

Pourtois G, Dan ES, Grandjean D, Sander D, Vuilleumier P. Enhanced extrastriate visual response to bandpass spatial frequency filtered fearful faces: time course and topographic Evoked-Potential Mapping. Hum Brain Mapp 2005;26:65–79.

Pourtois G, Debatisse D, Despland PA, de Gelder B. Facial expressions modulate the time course of long latency auditory brain potentials. Cogn Brain Res 2002;14:99–105.

Pourtois G, de Gelder B, Vroomen J, Rossion B, Crommelinck M. The time-course of intermodal binding between seeing and hearing affective information. Neuroreport 2000;11:1329–33.

Prior M, Dahlstrom B, Squires TL. Autistic children's knowledge of thinking and feeling states in other people. J Child Psychol Psychiatry 1990;31:587–601.

Putzar L, Goerendt I, Lange K, Rösler F, Röder B. Early visual deprivation impairs multisensory interactions in humans. Nat Neurosci 2007;10:1243–5.

Stekelenburg JJ, de Gelder B. The neural correlates of perceiving human bodies: an ERP study on the body-inversion effect. Neuroreport 2004;15:777–80.

Streit M, Dammers J, Simsek-Kraues S, Brinkmeyer J, Wölwer W, Ioannides A. Time course of regional brain activations during facial emotion recognition in humans. Neurosci Lett 2003;342:101–4.

Sugase Y, Yamane S, Ueno S, Kawano K. Global and fine information coded by single neurons in the temporal visual cortex. Nature 1999;400:869–73.

Talsma D, Doty TJ, Woldorff M. Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? Cereb Cortex 2006;17:679–90.