Configuration perception and face memory, and face context effects in developmental prosopagnosia

Elisabeth Huis in ‘t Veld a, Jan Van den Stock b & Beatrice de Gelder a b c

a Cognitive and Affective Neuroscience Laboratory, Tilburg University, Tilburg, The Netherlands
b Brain and Emotion Laboratory Leuven (BELL), Division of Psychiatry, Department of Neurosciences, KU Leuven, Leuven, Belgium
c Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands

Published online: 29 Nov 2012.

To cite this article: Elisabeth Huis in ‘t Veld, Jan Van den Stock & Beatrice de Gelder (2012) Configuration perception and face memory, and face context effects in developmental prosopagnosia, Cognitive Neuropsychology, 29:5-6, 464-481, DOI: 10.1080/02643294.2012.732051

To link to this article: http://dx.doi.org/10.1080/02643294.2012.732051

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the “Content”) contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or
Configuration perception and face memory, and face context effects in developmental prosopagnosia

Elisabeth Huis in ’t Veld¹, Jan Van den Stock², and Beatrice de Gelder¹,²,³
¹Cognitive and Affective Neuroscience Laboratory, Tilburg University, Tilburg, The Netherlands
²Brain and Emotion Laboratory Leuven (BELL), Division of Psychiatry, Department of Neurosciences, KU Leuven, Leuven, Belgium
³Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands

This study addresses two central and controversial issues in developmental prosopagnosia (DP), configuration- versus feature-based face processing and the influence of affective information from either facial or bodily expressions on face recognition. A sample of 10 DPs and 10 controls were tested with a previously developed face and object recognition and memory battery (Facial Expressive Action Stimulus Test, FEAST), a task measuring the influence of emotional faces and bodies on face identity matching (Face–Body Compound task), and an emotionally expressive face memory task (Emotional Face Memory task, FaMe-E). We show that DPs were impaired in upright, but not inverted, face matching but they performed at the level of controls on part-to-whole matching. Second, DPs showed impaired memory for both neutral and emotional faces and scored within the normal range on the Face–Body Compound task. Third, configural perception but not feature-based processing was significantly associated with memory performance. Taken together the results indicate that DPs have a deficit in configural processing at the perception stage that may underlie the memory impairment.

Keywords: Developmental prosopagnosia; Configural processing; Inversion effect; Face memory; Emotion.

The face provides us with a wealth of information about a person (Bruce & Young, 1986), first and foremost the identity, but also other major facial attributes like the facial expression. Information from these different channels is normally processed automatically and effortlessly. There are notorious exceptions to this ability, and maybe the most striking one, a deficit in recognizing a person by the face, is called prosopagnosia. In extreme cases, people with prosopagnosia cannot recognize the face of their own spouse or children. The face specificity of this person recognition deficit is underscored by the fact that identity can still be gleaned from other features such as the individual’s voice, gait, or clothing.

Prosopagnosia was initially identified as a face identity recognition deficit resulting from brain damage in adulthood (acquired prosopagnosia), and quite a few cases have been reported over the last hundred years (Farah, 1990). With a few
exceptions, almost all reports concern single cases (Damasio, Damasio, & Van Hoesen, 1982; Landis, Cummings, Christen, Bogen, & Imhof, 1986; Levine & Calvano, 1989; Meadows, 1974; Sergent & Signoret, 1992; Wada & Yamamoto, 2001). In contrast with prosopagnosia caused by acquired brain damage or congenital brain abnormalities, there is increasing evidence for face recognition disorders without clear evidence of brain damage (Barton, Cherkasova, & O’Connor, 2001; Behrmann, Avidan, Marotta, & Kimchi, 2005; Bentin, Deouell, & Soroker, 1999; de Gelder & Rouw, 2000a; de Haan & Campbell, 1991; Duchaine, Yovel, Butterworth, & Nakayama, 2006; Hasson, Avidan, Deouell, Bentin, & Malach, 2003; McConachie & Helen, 1976; Nunn, Postma, & Pearson, 2001; Palermo, Rivolta, Wilson, & Jeffery, 2011; Rivolta, Palermo, Schmalzl, & Coltheart, 2012; Stollhoff, Jost, Elze, & Kennerknecht, 2010; Van den Stock, Van de Riet, Righart, & de Gelder, 2008). The term developmental prosopagnosia (DP) was coined in order to stress that this disorder is most likely a result of a failure to acquire normal face recognition skills in the course of otherwise normal cognitive development and for reasons still very poorly understood.

In recent years, brain imaging has been a powerful research tool for face perception researchers, but functional magnetic resonance imaging (fMRI) investigations have not yet yielded a clear picture on how the areas and networks normally related to face processing function in people with developmental prosopagnosia. Some studies find abnormal face-specific activations (Dinkelacker et al., 2011; Furl, Garrido, Dolan, Driver, & Duchaine, 2011; Hadjikhani & de Gelder, 2002) while others report normal activity (Avidan, Hasson, Malach, & Behrmann, 2005; Hasson et al., 2003; Marotta, Genovese, & Behrmann, 2001) or normal activity within the putative face recognition network but abnormal activation in the extended networks (Avidan & Behrmann, 2009). In addition, other possible neurological explanations have been suggested. For example, diminished cortical grey matter volume (Dinkelacker et al., 2011; Garrido et al., 2009), disrupted connectivity (Thomas et al., 2009), or cerebellar hypoplasia (Van den Stock, Vandenbulcke, Zhu, Hadjikhani, & de Gelder, 2012) may be held accountable. Furthermore, research on hereditary disorders and (neuro)genetics may give rise to further explanations on how this developmental process may go astray (Grueter et al., 2007; Kennerknecht, Kischka, Stemper, Elze, & Stollhoff, 2011). Anomalous development may have multiple phenotypes, depending on the onset of the pathology in the acquisition process.

Whatever the neurological underpinnings of prosopagnosia, a major focus in the psychological literature to date is whether there is a deficit in configural perception and whether this is associated with or compensated for by more than average skill at feature processing. These debates are complicated by the fact that in the various studies, notions like configuration processing and feature processing are very general but also very vague. They acquire their meaning only in reference to the specific tasks used to measure them in each different study.

Configural processing generally refers to the ability of apprehending the whole configuration of the face in a single sweep. Damage to the brain areas involved in normal face recognition results in a loss of this processing routine. The test of configuration ability that still occupies central place in the assessment of intact face perception is the inversion effect. Following the initial observation (Yin, 1969) that recognition performance for inverted objects dropped excessively for faces, more so than for any other object category with a canonical orientation (de Gelder, Bachoud-Levi, & Degos, 1998), it was found that some prosopagnosic patients were more than normally sensitive to inversion and that their inversion sensitivity went in the opposite direction to that of controls (Barton, Zhao, & Keenan, 2003; Busigny & Rossion, 2010; de Gelder & Rouw, 2000b; Farah, Wilson, Drain, & Tanaka, 1995). This phenomenon was variously labelled inverted face inversion effect (Farah et al., 1995), inversion superiority (de Gelder et al., 1998), and the “paradoxical inversion” effect by de Gelder.
and collaborators (de Gelder & Rouw, 2000b). However, the occurrence of a paradoxical inversion effect went against the then dominant notion that loss of configuration processing and its replacement by feature processing is at the core of acquired prosopagnosia (Levine & Calvanio, 1989; Sergent & Signoret, 1992). If the ability to process the configuration would simply have been wiped out by the brain lesion, stimuli that normally trigger configuration-processing routines (e.g., upright faces) and stimuli that do not depend crucially on orientation-sensitive processes (like inverted faces and a host of other non-orientation-specific objects) would be treated similarly and recognized equally well or equally poorly. However, when detailed results began to show that upright and inverted faces are not processed similarly, it became difficult to conclude that the core of the prosopagnosic deficit is a loss of configuration perception and its replacement by feature processing. To understand this pattern of a conflict between processing routines, we developed the notion that faces are processed by two different routes, one that we called the face detection system, the other the face recognition system that contains both whole-based and part-based processes (de Gelder & Rouw, 2001). That view has since been confirmed in other studies (e.g., Busigny & Rossion, 2010; Rossion, Dricot, Goebel, & Busigny, 2011).

Another important issue in prosopagnosia research is how to establish whether an individual with poor face recognition skills specifically suffers from prosopagnosia. Several different and widely varying tests and tasks have been developed to date, such as the Cambridge Face Memory Test (Duchaine & Nakayama, 2006), the Benton Facial Recognition Test (Benton, Sivan, Hamsher, Varney, & Spreen, 1983), the Warrington Recognition Memory for Faces (Warrington, 1984), and various tasks using famous faces. In our opinion, there are several important aspects that must be taken into account when developing a test for prosopagnosia. First, granting that configuration is a crucial aspect of normal face perception, a proper assessment of face recognition difficulties requires comparable task settings and cognitive task requirements using not only faces, but also objects selected to be the best comparable object category given the specific focus of the task. For example, if one measures configuration processing by using the inversion effect, the resulting score is a difference score expressing the relative loss of recognition for faces compared to objects. Secondly, the same comparative requirement for face and non-face materials and tasks must apply when measuring feature processing. From a configuration-gestalt perspective, a feature is defined not by its abstract identity, but by its role in the context of the configuration and spacing is a defining property of a feature. Thirdly, there should be separate tasks with and without a memory component.

These considerations led us over time to develop a novel face test (de Gelder & Bertelson, 2009) that combines the relative assessment of face inversion versus object inversion with a direct assessment of feature-based recognition and its selectivity for faces. Furthermore, besides the inversion effect, configuration- versus feature-based processing can also be investigated more directly by part-to-whole matching tasks. To assess these processes, the test battery was extended to include a set of face and object perception materials for the assessment of both memory and configural and feature-based processing and designs (de Gelder, Frissen, Barton, & Hadjikhani, 2003), which led to the development of the Facial Expressive Action Stimulus Test (FEAST; de Gelder, Van den Stock, & Huis in ’t Veld, 2012).

Lastly, even though DPs are probably not impaired in their ability to recognize emotion (Duchaine, Parker, & Nakayama, 2003; Humphreys, Avidan, & Behrmann, 2007; Palermo, Willis, et al., 2011), little is known about the role that facial and bodily expressions play in the facial recognition and memory abilities in DP. There is increasing evidence that there is no complete dissociation between processes related to perceiving facial identity and facial expression (see Calder & Young, 2005, for a review). For example, the amygdala, orbitofrontal cortex, parahippocampal cortex, and...
lateral temporal regions may play an important role during the retrieval of faces previously seen with an emotional expression (Satterthwaite et al., 2009; Sergerie, Lepage, & Armony, 2005, 2006; Sterpenich et al., 2006). Therefore, it makes sense to expect that the presence of an emotional expression influences facial identity processes in DP.

Against this background, the aim of the current study is three-fold. First, we present data on the FEAST from a relatively large and diverse group of people suffering from longstanding face recognition deficits that to the best of our knowledge are not related to any known neurological condition. Secondly, we wanted to assess the effects of emotional context, such as facial and bodily expression, on face recognition and memory since there is accumulating evidence indicating that face recognition is sensitive to contextual influences such as facial and bodily expressions (de Gelder et al., 2006; de Gelder & Van den Stock, 2011b). Finally, we wanted to explore how face recognition mechanisms at the perceptual stage are associated with subsequent performance at the memory stage.

Method

Participants

Developmental prosopagnosia (DP) group. Participants were recruited between 2008 and 2012 via an announcement on our website (http://www.tilburguniversity.edu/gezichtsblindheid). Participants applied for participation through email or by telephone. An initial invitation letter was sent out in which the procedure and purpose of the study were explained. An appointment at Tilburg University was made when the participants consented to participate. Participants signed for informed consent. Participation was rewarded with reimbursement of travel costs and payment of €10 an hour. The total testing time was about 3 hours, divided in two or three sessions, according to the preference of the participant. Afterwards, the participants received a short written overview of the purpose of each test and their scores. The study was approved by an ethical committee. The sample used in this study consisted of 10 DPs who met the following inclusion criteria: complaints of longstanding difficulties with face recognition, normal or corrected-to-normal vision, and normal basic visual functions as assessed by the Birmingham Object Recognition Battery (line length, size, orientation, gap, minimal feature match, foreshortened view, and object decision; Riddoch & Humphreys, 1992). A history of psychiatric or neurological problems was an exclusion criterion. This resulted in the inclusion of 9 women and 1 man between the ages of 22 and 65 years ($M = 43.3, SD = 15.4$). All DPs reported problems with face recognition since childhood, such as recognizing friends and family. The DPs also indicated that they had problems with watching movies due to not being able to recognize the characters. Also, they complained about friction in personal relationships that were caused by failing to recognize familiar people and reported that they had been described as arrogant or aloof due to these problems with recognition. All DPs reported that they tried to recognize others by actively focusing on nonfacial identity features, such as attributes, haircuts, the voice, mannerisms, or body posture. As a first exploration and in order to establish continuity with the literature, we assessed face recognition using the Benton Facial Recognition Test (BFRT; Benton et al., 1983). The DPs scored significantly lower on the BFRT than did the controls, $t(18) = 3.38, p < .01$.

Control group. The control group was recruited among the acquaintances of the lab members. The control group consisted of 4 women and 6 men between the ages of 21 and 59 years ($M = 36.4, SD = 13.0$) with matched backgrounds and education levels. Participation was voluntarily, and the controls were not given a monetary reward. The control group did not differ from the DPs with regard to age, $t(18) = -1.08$.

Basic test battery (FEAST)

The FEAST (Facial Expressive Action Stimulus Test; de Gelder, Van den Stock & Huis in ’t Veld, 2012) consists of a number of subtests designed to provide a full picture of different
aspects of face recognition ability. The subtests have been extensively described and validated on the occasion of prosopagnosia case reports. The FEAST consists of the following face and object perception and memory tests.

Neutral Face Memory task (FaMe-N). The first subtest of the FEAST is a face memory test, designed along the lines of the Warrington face memory test (Warrington, 1984), consisting of an encoding and a recognition phase. In the encoding phase, the participants passively viewed 50 greyscale Caucasian faces (25 male) with a neutral facial expression, taken from our own database. Every actor was photographed in front view, with direct gaze, and the stimulus included the original haircut of the actor. Participants were instructed to study each face carefully and were told that their memory for the faces would be tested afterwards. Each face was presented for 3,000 ms with an intertrial interval of 1,000 ms.

During the recognition phase following immediately afterwards, the subjects were presented with 50 trials, each consisting of two faces. The task was to indicate which of the two faces in the pair had been seen previously. Each trial consisted of two stimuli: one target stimulus (exactly the same picture as that seen in the encoding phase) and one distractor stimulus (a new face). The stimuli pairs were matched for gender, facial attributes, and hairstyle. The facial expression of both target and distractor was always neutral. A trial lasted until a response was given. (See Figure 1.)

Faces and objects matching test. The faces and objects matching test (de Gelder et al., 1998; de Gelder & Bertelson, 2009) was used to assess configural perception and the inversion effect for faces and objects. The test consisted of a 2 (category: faces and shoes) × 2 (orientation: upright and inverted) factorial design. The materials consisted of greyscale photographs of shoes (8 unique shoes) and faces (4 male, 4 female; neutral facial expression). Each face and each shoe were photographed once in front view and once in three-quarter profile view. A trial consisted of three pictures: one frontal view picture on top and two three-quarter profile view pictures underneath. One of the two bottom pictures was of the same identity as the one on top (target), and the other was a distractor. The target and distractor pictures of the faces were matched for gender, facial attributes, and hairstyle. Each trial was presented for 750 ms after which participants were instructed to indicate by a button press which of the two bottom pictures represented the same exemplar as the one on top. Following the response, a black screen with a fixation cross was shown for a variable duration (1,000–2,000 ms). The experiment consisted of four blocks (one block per condition). In each block, 16 stimuli were presented four times in a pseudorandomized order, adding up to a total of 64 trials per block, and each block was preceded by 4 practice trials, during which the participants received feedback about their response. (See Figure 2.)

Face and house part-to-whole matching test. This test is used to assess feature-based processing.
The test consisted of a 2 (category: faces and houses) × 2 (orientation: upright and inverted) factorial design. Materials consisted of greyscale pictures of eight faces (four male; neutral facial expression, photographed in front view and with direct gaze) and eight houses. From each face, part stimuli were constructed by extracting the rectangle containing the eyes and the rectangle containing the mouth. House-part stimuli were created using a similar procedure, but the parts consisted of the door or a window. The procedure consisted of presenting a picture of a target whole face or house on top and two part stimuli underneath. The target part stimulus was taken from the whole face or house displayed on top; the distractor was taken from another stimulus. Participants were instructed to indicate by a button press which of the two bottom part stimuli depicted the same exemplar as that contained in the one on top. Stimulus presentation time was 750 ms. Following the response, a blank screen was presented for 1,000 ms. The experiment was divided into two blocks per condition. Each block comprised 32 trials (two randomized presentations of 16 stimuli). Within blocks, the presentation of the two possible parts (eyes or mouth, window or door) was randomized in order to prevent participants paying attention only to one specific feature. The order of the blocks was counterbalanced. Each block was preceded by 4 practice trials, during which the participants received feedback about their response. (See Figure 3.)

**Experimental stimuli and design**

**Experiment 1: Emotional Face Memory task (FaMe-E).** This task was designed by adapting the FaMe-N task by using stimuli containing emotional instead of neutral faces. Images were taken from the NimStim database (Tottenham et al., 2009) and our own database. As in the FaMe-N, the actors (28 female, 20 male, Caucasian) were photographed in front view with direct eye gaze, and the pictures contained the original hairstyle. The individuals in
the stimuli express fear (16 trials), sadness (16 trials), or happiness (16 trials). There was no overlap in identities with the FaMe-N. The encoding phase consisted of passively viewing each face for 3,000 ms each with a 1,000-ms interstimulus interval, and the participants were instructed to encode the identity of each face.

The procedure of the recognition phase was the same as the procedure of the neutral faces memory task. Participants were presented with 48 face-pairs and were instructed to indicate which individual in the pair was seen during the encoding phase. The target stimulus was the exact same stimulus as that seen in the encoding phase, and the distractor stimulus was matched on emotion and also as much as possible on hairstyle and facial features, but not in all cases on gender. (See Figure 4.)

**Experiment 2: Face–Body Compound (FBC) matching task.** Pictures of facial expressions were taken from the Averaged Karolinska Directed Emotional Faces (AKDEF; Lundqvist & Litton, 1998) and from our own database. In a pilot study, the faces were randomly presented one by one on a screen, and participants (N = 20) were instructed to categorize the emotion expressed in the face in a seven-alternative forced-choice paradigm (anger, disgust, fear, happiness, neutral, surprise, or sadness). On the basis of this pilot study, we selected 80 fearful (40 female) and 80 neutral (40 female) facial expressions, all recognized...
correctly by at least 75% of the participants. All the faces were photographed in front view and with direct gaze. Stimuli of whole body expressions were taken from our own Bodily Expressive Action Stimulus Test (BEAST) database and were selected on the basis of a similar pilot study (de Gelder & Van den Stock, 2011a). The selected stimuli displayed fearful body postures and neutral body postures. An instrumental action (pouring water in a glass) was used as neutral (not fearful) body postures, because, like the fearful expressions, instrumental actions elicit movement and action representation, and we wanted to control for these variables. Forty fearful (20 female) and 40 instrumental (20 female) body expressions were selected. We created face–body compounds by carefully resizing and combining both the facial and bodily expressions. A total of 80 compound stimuli were created following a 2 (face: fearful and neutral) \(\times\) 2 (body: fearful and neutral) factorial procedure, resulting in 20 stimuli (10 male) per condition. Face and body were always of the same gender, but in only half of the compound stimuli did the face and body express the same emotion. A trial consisted of one compound stimulus presented on top and two faces presented left and right underneath. The target stimulus was the same as the face of the compound stimulus, and the other was a distractor matched on emotional expression as well as main visual features, such as hair colour and gender. Participants were instructed to indicate which of the two bottom faces matched the one of the compound stimulus. The stimuli were presented for 750 ms, and the interstimulus interval was 2,000 ms. The experiment started with two practice trials, during which the subject received feedback. (See Figure 5.)

**Results**

Accuracies were calculated as the total proportion of correct responses both for the total score of each test and for each condition separately. Average response times from stimulus onset were calculated over the correct responses only. See Table 1 for an overview of the accuracy and reaction times (RTs) of both controls and DPs on the FEAST and the experimental tasks, and z scores per DP calculated with the mean and standard deviations of the control group.

**Basic test battery (FEAST)**

**Neutral Face Memory task (FaMe-N).** The accuracy of the DPs was significantly lower than that of controls, \(t(18) = 3.56, p < .01\), but there were no differences in RTs, \(t(18) = -1.03\). (See Figure 6.)

**Faces and objects matching test.** Repeated measures analysis of variance (ANOVA) with category (faces and shoes) and orientation (upright and inverted) as within-subject factors and group (DP and control) as between-subjects factor was carried out on the accuracy and response time data. For accuracy there was an interaction effect of Group \(\times\) Category, \(F(1, 18) = 17.68, p < .001\), and Group \(\times\) Orientation, \(F(1, 18) = 5.92, p < .05\). Post hoc tests show that the Group \(\times\) Category interaction was a result of a higher accuracy of controls on (upright and inverted) faces than of DPs, while controls and DPs scored similarly on the (upright and inverted) shoes condition. A similar pattern explains the Group \(\times\) Orientation interaction; controls have higher accuracy ratings in the upright (faces and shoes) condition than DPs; however, controls and DPs scored similarly in the inverted (faces and shoes) condition.

The same repeated measures ANOVA on the RT data resulted in a Group \(\times\) Orientation, \(F(1, 18) = 25.40, p < .001\), interaction. DPs are significantly slower than controls, and this lag is more pronounced when the stimuli are presented upright. Additionally, to explicitly test the face inversion effect for faces and shoes, \(t\) tests comparing the inversion effect (calculated by subtracting results on the inverted condition from those on the upright condition) were performed. DPs showed face inversion superiority (\(M = -2.0, SD = 3.74\)), and controls showed the normal face inversion effect (\(M = 1.9, SD = 1.72\)), \(t(18) = 2.99, p < .05\). However, no significant differences between DPs and controls were found for shoe inversion
Table 1. Means of accuracy and reaction times as a function of group

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>z</td>
<td>z</td>
<td>z</td>
<td>z</td>
<td>z</td>
<td>z</td>
<td>z</td>
<td>z</td>
</tr>
<tr>
<td>BFRT</td>
<td>Accuracy (%)</td>
<td></td>
<td>46</td>
<td>3</td>
<td>42**</td>
<td>2</td>
<td>-2.19</td>
<td>0.09</td>
<td>-0.69</td>
<td>-2.19</td>
<td>-0.99</td>
<td>-2.19</td>
<td>-1.59</td>
<td>-1.29</td>
</tr>
<tr>
<td></td>
<td>RT (ms)</td>
<td></td>
<td>2,757</td>
<td>904</td>
<td>3,238</td>
<td>1,176</td>
<td>-1.62</td>
<td>1.39</td>
<td>-0.93</td>
<td>2.62</td>
<td>1.06</td>
<td>-0.03</td>
<td>-0.25</td>
<td>1.88</td>
</tr>
<tr>
<td>Face matching</td>
<td>Accuracy (%)</td>
<td>Upr</td>
<td>96</td>
<td>4</td>
<td>82**</td>
<td>7</td>
<td>-2.25</td>
<td>-5.21</td>
<td>-5.21</td>
<td>-6.06</td>
<td>-5.64</td>
<td>-1.82</td>
<td>-1.40</td>
<td>-3.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inv</td>
<td>93</td>
<td>5</td>
<td>85**</td>
<td>7</td>
<td>0.18</td>
<td>-1.62</td>
<td>-2.22</td>
<td>-2.51</td>
<td>-0.72</td>
<td>-1.62</td>
<td>-0.42</td>
<td>-2.22</td>
</tr>
<tr>
<td>Shoe matching</td>
<td>Accuracy (%)</td>
<td>Upr</td>
<td>91</td>
<td>6</td>
<td>89</td>
<td>5</td>
<td>-0.06</td>
<td>0.77</td>
<td>-1.99</td>
<td>-1.16</td>
<td>0.77</td>
<td>-0.33</td>
<td>-0.61</td>
<td>-0.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inv</td>
<td>90</td>
<td>8</td>
<td>91</td>
<td>7</td>
<td>0.64</td>
<td>0.44</td>
<td>-1.37</td>
<td>1.05</td>
<td>1.25</td>
<td>0.44</td>
<td>0.04</td>
<td>-1.57</td>
</tr>
<tr>
<td>Face matching</td>
<td>RT (ms)</td>
<td>Upr</td>
<td>1,138</td>
<td>153</td>
<td>1,457</td>
<td>375</td>
<td>1.54</td>
<td>3.08</td>
<td>-0.19</td>
<td>3.01</td>
<td>7.47</td>
<td>0.52</td>
<td>0.73</td>
<td>-0.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inv</td>
<td>1,111</td>
<td>125</td>
<td>1,316</td>
<td>267</td>
<td>0.89</td>
<td>2.07</td>
<td>0.20</td>
<td>2.48</td>
<td>6.04</td>
<td>0.14</td>
<td>0.39</td>
<td>-0.99</td>
</tr>
<tr>
<td>Shoe matching</td>
<td>RT (ms)</td>
<td>Upr</td>
<td>1,100</td>
<td>129</td>
<td>1,309</td>
<td>211</td>
<td>1.32</td>
<td>-0.17</td>
<td>1.72</td>
<td>0.44</td>
<td>1.51</td>
<td>0.41</td>
<td>0.98</td>
<td>2.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inv</td>
<td>1,135</td>
<td>175</td>
<td>1,187</td>
<td>204</td>
<td>0.89</td>
<td>-0.32</td>
<td>-0.44</td>
<td>0.21</td>
<td>-0.28</td>
<td>-0.51</td>
<td>-0.09</td>
<td>-0.24</td>
</tr>
<tr>
<td>Face-PM</td>
<td>Accuracy (%)</td>
<td>Upr</td>
<td>78</td>
<td>14</td>
<td>71</td>
<td>15</td>
<td>1.53</td>
<td>-1.51</td>
<td>0.12</td>
<td>0.12</td>
<td>-0.21</td>
<td>-0.64</td>
<td>-0.86</td>
<td>-2.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inv</td>
<td>68</td>
<td>13</td>
<td>67</td>
<td>12</td>
<td>2.40</td>
<td>-0.40</td>
<td>-0.75</td>
<td>0.07</td>
<td>-0.51</td>
<td>-0.40</td>
<td>-0.40</td>
<td>-0.63</td>
</tr>
<tr>
<td>House-PM</td>
<td>Accuracy (%)</td>
<td>Upr</td>
<td>83</td>
<td>13</td>
<td>78</td>
<td>13</td>
<td>1.14</td>
<td>-1.42</td>
<td>-0.61</td>
<td>-2.01</td>
<td>0.09</td>
<td>0.68</td>
<td>-0.02</td>
<td>-0.61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inv</td>
<td>85</td>
<td>15</td>
<td>83</td>
<td>10</td>
<td>0.95</td>
<td>-1.08</td>
<td>-0.76</td>
<td>-0.12</td>
<td>0.42</td>
<td>0.31</td>
<td>0.10</td>
<td>-0.65</td>
</tr>
<tr>
<td>Face-PM</td>
<td>RT (ms)</td>
<td>Upr</td>
<td>1,359</td>
<td>221</td>
<td>1,647</td>
<td>553</td>
<td>6.82</td>
<td>1.53</td>
<td>-0.60</td>
<td>-0.01</td>
<td>3.00</td>
<td>-0.28</td>
<td>-1.13</td>
<td>-0.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inv</td>
<td>1,411</td>
<td>257</td>
<td>1,614</td>
<td>578</td>
<td>6.30</td>
<td>0.52</td>
<td>-0.66</td>
<td>0.22</td>
<td>1.05</td>
<td>-0.74</td>
<td>0.06</td>
<td>-1.49</td>
</tr>
<tr>
<td>House-PM</td>
<td>RT (ms)</td>
<td>Upr</td>
<td>1,318</td>
<td>251</td>
<td>1,444</td>
<td>324</td>
<td>3.33</td>
<td>0.69</td>
<td>-1.04</td>
<td>0.85</td>
<td>0.40</td>
<td>-0.27</td>
<td>-0.16</td>
<td>-0.59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inv</td>
<td>1,286</td>
<td>214</td>
<td>1,425</td>
<td>296</td>
<td>4.01</td>
<td>0.49</td>
<td>0.29</td>
<td>1.04</td>
<td>0.20</td>
<td>-0.22</td>
<td>-0.32</td>
<td>-0.59</td>
</tr>
</tbody>
</table>

Note: Accuracy: percentage correct. FEAST = Facial Expressive Action Stimulus Test; BFRT = Benton Facial Recognition Test; FaMe-N = Neutral Face Memory task; FaMe-E = Emotional Face Memory task; FBC = Face–Body Compound task; RT = reaction time; PM = part matching; Upr = upright orientation; Inv = inverted orientation; DPs = developmental prosopagnosia group; L.F. to M.R. = z scores of the individual prosopagnosics.

*p < .05. **p < .01.
scores (DP: $M = -1.7$, $SD = 3.56$; controls: $M = 0.40$, $SD = 4.06$), $t(18) = 1.23, p = .24$.

**Face and house part-to-whole matching test.** Repeated measures ANOVA with category (faces and houses) and orientation (upright and inverted) as within-subject factors and group (DP and control) as between-subjects factor was carried out on the accuracy and response time data. This revealed for the accuracy data a Category $\times$ Orientation interaction effect, $F(1, 18) = 6.82, p < .05$. Accuracy scores on house-part matching were on average higher than the accuracy scores on face-part matching. However, accuracy on house-part matching was the same regardless of orientation. In contrast, accuracy decreased on inverted face-part matching compared to upright face-part matching. No significant differences were found between controls or DPs.

Repeated measures ANOVA on the RTs resulted in a main effect of category, $F(1, 18) = 8.86, p < .001$; RTs were higher for face-part matching than for house-part matching. In addition, $t$ tests comparing the inversion effects were performed, but yielded no significant results.

---

**Figure 4.** (a) Example of a trial displaying a sad stimulus in the encoding phase of the Emotional Face Memory task (FaMe-E) with an exemplar of a happy and sad stimulus. (b) Example of a possible trial in the recognition phase.

**Figure 5.** Example of a trial of each of the four conditions of the Face–Body Compound (FBC) matching task. (A) Face fearful, body fearful, (B) face neutral, body fearful, (C) face fearful, body neutral, (D) face neutral, body neutral.
Results of experimental test battery

Experiment 1: Emotional Face Memory task (FaMe-E)

Mean accuracies (proportion correct responses) and response times were calculated for every condition. The results are shown in Figure 6. Repeated measures ANOVA with emotion (fear, happy, sad) as within-subject factor and group (DP and control) as between-subjects factor was carried out on the accuracy (proportion correct responses) and response time data of both tasks (see Figure 6). This revealed for the accuracy data a main effect of group, \( F(1, 18) = 15.57, p = .001 \); controls had higher accuracy ratings than DPs. Also, a main effect of emotion was found, \( F(3, 16) = 17.47, p < .001 \); Bonferroni corrected pairwise comparisons showed that accuracy was higher in the neutral condition than in the fear \((p < .01)\) and happy \((p < .01)\) conditions, and accuracy ratings were lower in the fear condition than in the sad condition \((p < .001)\) for both groups.

The same repeated measures ANOVA on the reaction time data resulted in a significant Group \( \times \) Emotion interaction effect, \( F(3, 16) = 4.11, p < .05 \). Controls were faster than DPs in the neutral, happy, and sad conditions, but slower in the fear condition.

Comparing results on the FaMe-N and the FaMe-E. Repeated measures ANOVA on the combined results of the FaMe-N and the FaMe-E with emotion (neutral, fear, happy, sad) as within-subject factor and group (DP and control) as between-subjects factor was carried out on the accuracy (proportion correct responses) and response time data of both tasks (see Figure 6). This revealed for the accuracy data a main effect of group, \( F(1, 18) = 15.57, p = .001 \); controls had higher accuracy ratings than DPs. Also, a main effect of emotion was found, \( F(3, 16) = 17.47, p < .001 \); Bonferroni corrected pairwise comparisons showed that accuracy was higher in the neutral condition than in the fear \((p < .01)\) and happy \((p < .01)\) conditions, and accuracy ratings were lower in the fear condition than in the sad condition \((p < .001)\) for both groups.

Experiment 2: Face–Body Compound (FBC) matching task

Mean accuracies (proportion correct responses) and response times from stimulus onset for the correct trials only were calculated for every condition. The results are shown in Figure 7. A repeated measures ANOVA with facial expression (fearful and neutral) and bodily expression (fearful and neutral) as within-subject factors and group (DP and control) as between-subjects factor was carried out on the accuracy and response time data. This revealed for the accuracy data a main effect of group, \( F(1, 17) = 5.58, p < .05 \); controls had higher accuracy scores than DPs overall. Also, a main effect of facial expression, \( F(1, 17) = 11.46, p < .05 \), was
found. Accuracy was lower on conditions with a fearful body, compared to a neutral body. Further explorations comparing the scores of controls and DPs on each condition separately using $t$ tests show that controls had higher accuracy scores than DPs only when both the body and the face were neutral, $t(17) = 2.66, p < .05$.

For the RTs, a Group $\times$ Facial Expression, $F(1, 17) = 7.66, p = .01$, and a trend for a Facial Expression $\times$ Bodily Expression, $F(1, 17) = 4.22, p = .056$, interaction effect were found. The Group $\times$ Facial Expression interaction is a result of higher RTs when the facial expression was fearful than when it was neutral for the DPs, while in contrast, the RTs were the same for both conditions for the controls. The Facial Expression $\times$ Bodily Expression effect was caused by higher RTs when both the face and bodily expression express fear.

**Relations between results on the face battery subtests and experimental tasks**
The different subtests and experiments measure different aspects of face recognition. The relation between them is not yet clearly understood, and calculating the relations between performances in the two groups provides useful insights in this question.

**Effect of configural and feature-based processing on memory task performance.** In order to evaluate whether mechanisms measured at the perception stage (i.e., configural processing and feature-based processing) are predictive of memory performance, total scores on upright and inverted whole face and whole shoe matching (as measured by the faces and objects matching task) and total scores on upright and inverted face and house part-to-whole matching (as measured with the face and house part-to-whole matching task) were entered simultaneously in a regression analysis, to assess which of these predictors is significantly related to the total accuracy score on the FaMe-N. The same procedure was followed to assess the strength of association with the FaMe-E separately (see Table 2).

We observed that only the ability to match upright whole faces significantly and positively predicted face memory, and this was the case for neutral faces as well as for faces with an emotional expression.

**Relationship between the inversion effect and face memory**
To further explore the relationship between the (paradoxical) inversion effect and memory performance, a score for the strength of the inversion effect for each stimulus category in both the faces and objects matching task and the face and house part-to-whole matching task was calculated by subtracting the total accuracy on the inverted condition from the upright condition. Entering these predictors simultaneously in a linear regression model to predict the total accuracy scores on the FaMe-N and the FaMe-E separately revealed that the strength of the face inversion effect...
significantly predicted accuracy scores on the FaMe-E (see Table 3).

Discussion

The aim of the current study was to investigate face-processing deficits of a relatively large group of DPs and to focus specifically on the relation between configural and feature-based processes tested in separate tasks of faces and objects recognition. In addition, we considered the role of these two processes for face memory of unfamiliar faces. We also took into account the possible role of affective information for face memory and studied neutral faces and facial expressions separately. Finally, we investigated the role of realistic facial and bodily expressions contexts on face recognition. In support of the notion of face specificity, we established that the DP group showed impaired ability on matching upright faces, but not objects. Most importantly, the controls showed the expected inversion effect for faces, but the DPs did not. These findings are in line with other studies showing that DPs are impaired in configural processing specifically for faces (Avidan, Tanzer, & Behrmann, 2011; Behrmann et al., 2005; de Gelder & Rouw, 2000a; Duchaine et al., 2006; Duchaine, Yovel, & Nakayama, 2007; Farah et al., 1995;

Table 2. Regression coefficients of the total accuracy scores on the task conditions for configural and feature-based processing on the total accuracy scores of the Face Memory–Neutral and the Face Memory–Emotional task

<table>
<thead>
<tr>
<th>Task</th>
<th>Orientation</th>
<th>Face Memory–Neutral</th>
<th></th>
<th>Face Memory–Emotional</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>SE B</td>
<td>(^\hat{a})</td>
<td>B</td>
</tr>
<tr>
<td>Face matching</td>
<td>Upr</td>
<td>0.651</td>
<td>0.270</td>
<td>0.937(^*)</td>
<td>0.823</td>
</tr>
<tr>
<td></td>
<td>Inv</td>
<td>−0.406</td>
<td>0.379</td>
<td>−0.468</td>
<td>−0.849</td>
</tr>
<tr>
<td>Shoe matching</td>
<td>Upr</td>
<td>0.616</td>
<td>0.373</td>
<td>0.514</td>
<td>0.532</td>
</tr>
<tr>
<td></td>
<td>Inv</td>
<td>−0.235</td>
<td>0.306</td>
<td>−0.270</td>
<td>−0.401</td>
</tr>
<tr>
<td>Face–PM</td>
<td>Upr</td>
<td>−0.052</td>
<td>0.157</td>
<td>−0.119</td>
<td>0.370</td>
</tr>
<tr>
<td></td>
<td>Inv</td>
<td>0.193</td>
<td>0.109</td>
<td>0.380</td>
<td>0.003</td>
</tr>
<tr>
<td>House–PM</td>
<td>Upr</td>
<td>−0.189</td>
<td>0.124</td>
<td>−0.496</td>
<td>−0.104</td>
</tr>
<tr>
<td></td>
<td>Inv</td>
<td>0.067</td>
<td>0.198</td>
<td>0.129</td>
<td>−0.269</td>
</tr>
</tbody>
</table>

\[ R^2 = .58; F(8, 19) = 1.92 \]

Note: PM = part matching; Upr = upright orientation; Inv = inverted orientation.
\(^*p < .05.\)

Table 3. Regression coefficients of the inversion scores on the tasks for configural and feature-based processing on the total scores of the Face Memory–Neutral and the Face Memory–Emotional task

<table>
<thead>
<tr>
<th>Task</th>
<th>Difference score</th>
<th>B</th>
<th>SE B</th>
<th>(^\hat{a})</th>
<th>B</th>
<th>SE B</th>
<th>(^\hat{a})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face matching</td>
<td>Upr – Inv</td>
<td>0.519</td>
<td>0.285</td>
<td>0.442</td>
<td>0.794</td>
<td>0.307</td>
<td>0.576(^*)</td>
</tr>
<tr>
<td>Shoe matching</td>
<td>Upr – Inv</td>
<td>0.192</td>
<td>0.289</td>
<td>0.182</td>
<td>0.196</td>
<td>0.311</td>
<td>0.159</td>
</tr>
<tr>
<td>Face–PM</td>
<td>Upr – Inv</td>
<td>−0.091</td>
<td>0.096</td>
<td>−0.213</td>
<td>0.074</td>
<td>0.104</td>
<td>0.148</td>
</tr>
<tr>
<td>House–PM</td>
<td>Upr – Inv</td>
<td>−0.082</td>
<td>0.127</td>
<td>−0.164</td>
<td>−0.042</td>
<td>0.137</td>
<td>−0.072</td>
</tr>
</tbody>
</table>

\[ R^2 = .30, F(4, 19) = 1.622 \]

Note: PM = part matching; Upr = upright orientation; Inv = inverted orientation.
\(^*p < .05.\)
Palermo, Willis, et al., 2011). Furthermore, six DPs even had higher accuracy ratings on inverted faces than upright faces, a surprising pattern known as the paradoxical inversion superiority effect. In contrast, it is worth noting that DPs are equally able as controls to match parts in the context of a whole stimulus, whether a whole face or a whole house, and this indicates a normal feature-processing ability. This latter result taken together with the inversion effect result indicates that DP is not simply a matter of a loss of configuration perception combined with an intact processing of features (de Gelder & Rouw, 2001). If that were the case, DPs would be able to apply their normal feature perception skills to match whole faces, and their performance would be the same whether the stimulus is upright or inverted. Note that there is no difference in feature-processing ability between stimulus categories in the sense that DPs are not better than controls at feature matching, which might have been evidence for a compensation strategy for impaired configuration matching. Based on these results, we conclude that the DP group shows evidence of a specific deficit on configuration-sensitive face tasks. However, this deficit is not to be viewed as a complete loss or insensitivity to the face configuration.

Our next question concerns the role of affective information in the face and the context of the body. Interestingly, when a forced-choice face-matching task is conducted, similar to that in the faces and objects matching test but with additional expressive faces and bodies, the differences in accuracy between controls and DPs are clearly less pronounced. Controls still score better at this task overall, but this is mainly due to a better performance of controls on the specific condition in which both the facial and bodily expression of the stimuli are neutral. This thus confirms the previous result of a deficit with neutral faces. But when the face or body expresses a fearful emotion, DPs perform at the level of controls, and this is even the case when the task is one of face identity matching. These findings are in line with previous reports showing that emotional information reduces the face perception impairments in both acquired and developmental prosopagnosia (de Gelder et al., 2003; Van den Stock et al., 2008). The results also indicate that controls and DPs are equally influenced in their ability to match identity when a fearful body is present. This is in line with the results from previous studies that find normal body processing in DPs (Duchaine et al., 2006; Van den Stock et al., 2008), but see Moro et al., 2012, for a report of body agnosia in a case of acquired prosopagnosia.

Finally, our assessment of memory for faces shows that DPs are significantly impaired when they are asked to remember neutral faces compared to controls, and this is in line with the results from previous studies (Duchaine, Germine, & Nakayama, 2007; Duchaine & Nakayama, 2006; Righart & de Gelder, 2007; Stollhoff, Jost, Elze, & Kennerknecht, 2011; Van den Stock et al., 2008). The DPs also score significantly worse than the controls on the emotional memory test. Furthermore, memory for faces is impaired more strongly when the face expresses emotion, especially fear. Previous studies assessing the effect of emotion on memory for faces have yielded inconsistent results. Johansson, Mecklinger, and Treese (2004) found that the discrimination between previously seen or new faces was unaffected by emotional expression and argued that valence differentially affected processes underlying recognition by familiarity and recollection. Other studies find increased accuracy for retrieving faces seen with a fearful expression (Righi et al., 2012; Sergerie et al., 2005). The pattern of responses on the RTs, however, is interesting: Emotional expression does not influence RTs for the DPs. This contrasts with the performance of controls, who are slower than DPs when the face expresses fear, but faster when the expression is sad or happy. Additionally, DPs but not controls are overall faster in the emotional face memory task than in the neutral task. A possible explanation can be found in a neuroimaging study (Van den Stock et al., 2008) showing that fusiform face area (FFA) activation in DPs was lower for neutral faces, but comparable to that in controls for emotional faces. We suggested that...
this increase of activation in response to emotion may result from a boost of the emotion-processing system but not specifically the face-processing system. The results of the current experiments seem to partly support this claim, as emotional expression decreased response times for DPs. However this boost of emotional expression is not associated with an increased recognition of facial identity in a memory task. In short, even though DPs are most likely not impaired in emotion recognition (Duchaine et al., 2003), the role of emotional expression in face recognition and the underlying neurological networks in developmental (and acquired) prosopagnosia is not yet clear (Humphreys et al., 2007; Peelen, Lucas, Mayer, & Vuilleumier, 2009), and the relative autonomy between expression and identity processes may be much less established in DP than in normal controls.

Our final question concerns possible relations between the different tasks and the abilities they measure. Our results indicate a positive relationship between the (in)ability to remember faces and the ability to match upright faces as measured with the face-matching task. Only accuracy scores on the upright faces condition in the faces and shoes task significantly predicted scores on the neutral and emotional memory tasks, and the strength of the (paradoxical) face inversion effect predicted performance on the emotional face memory task. We tentatively conclude that impaired configural processing may play an important role in consolidation of a face in memory and/or memory retrieval. These results are in line with accumulating evidence that the ability to process faces configurally is positively related to face recognition ability (Wang, Li, Fang, Tian, & Liu, 2012), also in DPs (Richler, Cheung, & Gauthier, 2011), as measured with the composite face task, and in “super-recognizers” (Russell, Duchaine, & Nakayama, 2009). More research is needed to further explore these results with additional measures of holistic (face) processing, such as the composite face task or a global/local task (Navon, 1977). Also, the precise processes and underlying neurological correlates of this effect need further exploration.

In conclusion, our study shows that the face-processing impairment in DP is specifically related to configural processes but does not affect feature processing in either a positive or a negative direction, and that identity recognition deficit is reduced when the face conveys emotional information. Furthermore, configural processing at the perception stage is predictive of face recognition at the memory stage. Lastly, the comparative approach at the basis of FEAST makes it a useful tool in prosopagnosia research and clinical practice.

REFERENCES

CONFIGURATION PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

Furl, N., Garrido, L., Dolan, R. J., Driver, J., & Duchaine, B. (2011). Fusiform gyrus face selectivity...


