CORTEX 108 (2018) 112–126

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research Report

Holistic face perception is impaired in developmental prosopagnosia

John Towler [a], Katie Fisher [b] and Martin Eimer [b,*]

[a] Department of Psychology, College of Human and Health Sciences, Swansea University, Swansea, UK
[b] Department of Psychological Sciences, Birkbeck, University of London, UK

ARTICLE INFO

Article history:
Received 11 December 2017
Reviewed 30 January 2018
Revised 21 March 2018
Accepted 27 July 2018
Action editor Stefan Schweinberger
Published online 6 August 2018

Keywords:
Face perception
Face recognition
Developmental prosopagnosia
Holistic face processing
N250r component

ABSTRACT

Individuals with developmental prosopagnosia (DP) have severe difficulties recognising familiar faces. A current debate is whether these face recognition impairments derive from problems with face perception and in particular whether individuals with DP cannot utilize holistic representations of individual faces. To assess this hypothesis, we recorded event-related potentials (ERPs) during a sequential face identity matching task where successively presented pairs of upright faces were either identical or differed with respect to their internal features, their external features, or both. Participants with DP and age-matched controls reported on each trial whether the face pair was identical or different. To track the activation of cortical visual face memory representations, we measured N250r components over posterior face-selective regions. N250r components to full face repetitions were strongly attenuated for DPs as compared to control participants, indicating impaired face identity matching processes in DP. In the Control group, the N250r to full face repetitions was superadditive (i.e., larger than the sum of the two N250r components to partial repetitions of external or internal features). This demonstrates that holistic face representations were involved in identity matching processes. In the DP group, N250r components to full and partial identity repetitions were strictly additive, indicating that the identity matching of external and internal features operated in an entirely part-based fashion, without any involvement of holistic representations. In line with this conclusion, DPs also made a disproportionate number of errors on partial repetition trials, where they often failed to report a change of internal facial features. This suggests an atypical strategy for encoding external features as cues to identity in DP. These results provide direct electrophysiological and behavioural evidence for qualitative differences in the representation of face identity in the occipital-temporal face processing system in developmental prosopagnosia.

* *Corresponding author.*
E-mail address: m.eimer@bbk.ac.uk (M. Eimer).
https://doi.org/10.1016/j.cortex.2018.07.019

1. Introduction

Developmental prosopagnosia (DP) is a neurodevelopmental disorder characterised by a severe and specific difficulty recognising the faces of familiar people in daily life (for recent reviews, see: Towler, Fisher, & Eimer, 2017; Susilo & Duchaine, 2013; Towler & Eimer, 2012). Individuals with DPs do not appear to have suffered from brain injury, have normal intelligence, and intactlvisual and social cognitive abilities. Instead, these individuals appear to have specifically failed to develop the normal cognitive and neural mechanisms that allow for the rapid and effective recognition of individual faces. All individuals with DP have trouble recognising familiar faces, but the mechanisms that are responsible for this impairment are not yet fully understood. One fundamental question is whether DP is the result of some form of visual-perceptual face processing deficit, or whether it exclusively reflects later memory-related or associative impairments (e.g. Bate & Tree, 2017). Some investigations found deficits in perceptual face matching tasks for DPs (e.g. Duchaine, Yovel, & Nakayama, 2007; White, Rivolta, Burton, Al-Janabi, & Palermo, 2017; Yovel & Duchaine, 2006), while others failed to find such impairments (e.g. Le Grand et al., 2006; Ulrich et al., 2017). However, because different aspects of perceptual face processing and their possible impairment in DP have so far not been studied systematically, it remains unclear if and to what degree specific deficits in face perception contribute to the face recognition problems experienced by individuals with DP. The goal of the present study was to investigate whether the holistic perceptual processing of faces is selectively impaired in DP.

Holistic face processing refers to the ability to simultaneously apprehend the whole face in a single glance. This ability involves the integration of the internal facial features (such as the eyes, nose, and mouth), along with external facial features such as the hair and overall shape of the head, into a single visual representation that can be used for fast and effective face recognition. Perhaps the most compelling demonstration of holistic face processing comes from the composite face task (Young, Hellawell, & Hay, 1987) where participants have to match the identity of the top half of face pairs while ignoring the task-irrelevant bottom halves. Performance is impaired when the bottom halves depict different individuals, and this interference is abolished or strongly reduced when the top and bottom face halves are spatially misaligned (thus breaking the canonical face configuration) or when faces are inverted. This composite face illusion (CFI) provides direct evidence for the holistic processing of upright faces (for a review, see Rossion, 2013). Studies with DP using this task have produced inconsistent results, with some reporting a reduced CFI for DPs (e.g. Avidan, Tanzer, & Behrmann, 2011), while others find no difference in the size of the CFI between DPs and control participants (e.g., Biotti, Wu, Yang, Jiahui, Duchaine, Cook, 2017). Clearer evidence for deficits in holistic face processing deficits for DPs comes from part-whole face matching tasks. In this task, participants encode the identity of a whole face, are tested with either a whole face or face parts, and have to decide whether the whole face is the same as the sample face, or whether the face part is the same or different to the part presented in the sample face (e.g., Tanaka & Farah, 1993). Performance is generally better when the test image is a whole face. Importantly, this whole-face advantage is abolished by face inversion or by the spatial scrambling of face parts, suggesting that it reflects benefits produced by holistic face processing. DPs show a normal whole-face-advantage for the mouth, but no such effect for the eyes (DeGutis, Cohan, Mercado, Wilmer, & Nakayama, 2012). This result suggests a deficit of holistic face processing in DP that is specific to the eye region. Individuals with DP may be impaired in integrating the eye region within the context of the rest of the face. A third line of evidence for atypical holistic face processing in DP comes from observations that the effects of face inversion on performance in face identity matching tasks are often absent or reduced in individuals with DP (e.g. Duchaine et al., 2007). Because face inversion effects are often seen as the hallmark of holistic face processing, their absence may be interpreted as indication that this type of processing is impaired or absent in DP.

As the current evidence for deficits of holistic face processing in individuals with DP from composite and part-whole face matching tasks is limited and inconclusive, we employed a new face matching task that was designed to reveal such impairments with both behavioural and electrophysiological measures, and was first used in a previous study with participants with typical face recognition ability (Towler & Eimer, 2016). In this task, participants' attention is directed to the entire face and a holistic style of face processing is required for successful task performance. On each trial, a pair of face images is presented sequentially, and participants are instructed to detect repetitions and changes of these faces. Critically, repetitions or changes in the internal features of (the eyes, nose, and mouth) and external features (hair, ears, and head outline) of these face pairs are orthogonally varied. On half of all trials, internal and external features are both identical or both different (full repetition and full change trials). On the other half, there is a change in the internal features while the external features are repeated or vice versa (external or internal feature repetition trials). Participants' task is to report whether the two faces were identical or whether there was a change between them (either a partial change of external or internal features, or a full change). Because response selection cannot be based exclusively on repetition or changes of internal or external features alone, this task encourages participants to form holistic face representations that integrate across both types of features.

Our previous study (Towler & Eimer, 2016) provided behavioural and electrophysiological evidence for holistic face processing in participants with unimpaired face recognition ability. In different parts of this experiment, the face images were either presented in their normal upright orientation or upside-down. Face inversion increased the percentage of incorrect responses, specifically for trials with a partial change in either external or internal facial features, which were more likely to be reported as full face repetitions. This suggests that holistic face representations were

available only for upright faces, which allows for the efficient detection of partial changes, but not for inverted faces. More direct evidence for this conclusion was provided by event-related brain potentials (ERPs) that were recorded during task performance. Here, we focused on the N250r component that is elicited during sequential face matching tasks as an enhanced negativity in response to face identity repetitions as compared to identity changes, and typically emerges around 220 ms after the onset of a repeated versus changed face image. N250r components have been consistently observed over bilateral posterior occipito-temporal electrode locations, are accompanied by a fronto-central positivity, and are assumed to reflect a match between a face representation in working memory and an on-line perceptual representations of a particular face (e.g. Fisher, Towler, & Eimer, 2016; Schweinberger, 2011; Schweinberger, Huddy, & Burton, 2004; Towler, Kelly, & Eimer, 2016; Zimmermann & Eimer, 2013). Although the N250r is usually largest in response to repetitions of two physically identical face images (e.g., Schweinberger, Pickering, Jentzsch, Burton, & Kaufmann, 2002), this component is also reliably present for repetitions of different images of the same face (Bindemann, Burton, Leuthold, & Schweinberger, 2008; Zimmermann & Eimer, 2013, 2014), demonstrating that the underlying face identity matching processes are at least partially image-independent.

In our previous study (Towler & Eimer, 2016), N250r components to full face repetitions versus full face changes were delayed and reduced in amplitude when faces were inverted (see also Jacques & Rossion, 2009; Schweinberger et al., 2004), demonstrating that face inversion impairs both the speed and precision of face identity matching processes. An N250r was not only elicited by full face repetitions, but also on trials where only the internal or external features were repeated, although these N250r components to partial repetitions were smaller than the N250r in response to full face repetitions. Critically, we employed these N250r components observed on full and partial repetition trials to investigate the part-based versus holistic nature of the representations involved in these face matching processes. If external and internal features were represented in an entirely independent part-based fashion, the sum of the two N250r components to internal and external feature repetitions should be equal in size to the N250r component triggered on full repetition trials. In contrast, if these features were represented in an integrated holistic fashion, the N250r to full repetitions should be larger than the sum of the two N250r components to partial feature repetitions (i.e., super-additive). In blocks with inverted faces, N250r components to full face repetitions were entirely additive, that is, identical to the sum of the N250rs to internal and external repetitions. This indicates that the internal and external features of inverted faces were registered and represented separately and in parallel in a part-based fashion. A qualitatively different pattern was found for upright faces. Here, the N250r to full face repetitions was larger than the sum of the N250r components elicited on the two types of partial face repetition trials. This superadditivity of N250r components for full repetitions of upright faces provides strong evidence for the holistic nature of the underlying face representations. It suggests that the matching of upright faces involved representations that integrate across the internal and external facial features. Because the N250r tends to be larger for identical image repetitions (Schweinberger et al., 2002), the N250r superadditivity for upright faces could in principle be due to the fact that identical images were shown on full repetition trials but not on partial repetition trials. However, this was also the case for blocks with inverted faces, where no superadditive N250r was observed, demonstrating that this effect is not linked to image-based factors.

Our earlier results (Towler & Eimer, 2016) suggest that N250r components and in particular the additivity or super-additivity of N250r components to full versus partial face repetitions can dissociate part-based and holistic face identity matching processes in participants with intact face processing. In the present study, we employed the same task procedures and analysis logic to investigate in a large sample of 14 individuals with DP to find out whether face identity matching is impaired in DP, and critically, whether such impairments are associated with a selective deficit in the holistic processing of faces. DPs and age-matched control participants had to report the presence or absence of a change between two successively presented face images. Performance and ERPs were measured separately for trials with full repetitions and full changes, as well as for trials with internal or external feature repetitions. Because our previous study found evidence for holistic face processing for upright but not for inverted faces, faces were always presented in their standard upright orientation in the current experiment.

Face identity matching processes could either operate in a qualitatively different fashion in DPs and neurotypical individuals, or there may only be quantitative differences. For example, DPs may be slower or less efficient in matching facial identities, but the underlying mechanisms may still be the same. If this was the case, DPs should show slower reaction times, increased error rates, and smaller and delayed N250r components relative to control participants (see Fisher, Towler, & Eimer, 2017, for evidence that N250r amplitudes are attenuated in individuals with DP). However, these performance impairments should be equally present on all types of trials, and N250r components for DP participants should still show some evidence of holistic face processing (i.e., superadditivity for full versus partial identity repetitions). In contrast, if the difficulty with face identity matching in DP specifically stems from an underlying deficit in constructing holistic representations of face images, performance should be disproportionately impaired on partial face repetitions trials as compared to full face repetitions or changes in the DP group. In addition, and critically, N250r components for full face repetitions in this group should be strictly additive (i.e., identical to the summed N250r components triggered by partial internal and external repetitions), similar to what has been observed for inverted faces in neurotypical participants (Towler & Eimer, 2016). A third possibility is that DPs have no impairments in the image-based face identity matching task employed here. In this case, there should be no differences in face matching performance and N250r components between the DP and Control groups, including fully super-additive N250r components for full versus partial face repetitions for participants with DP.

2. Materials and methods

2.1. Participants

Twenty-eight paid volunteers were tested. Fourteen individuals (eight female, mean age: 31.46 years, ages ranged from 21 to 46 years) were tested as a control group, and fourteen individuals with developmental prosopagnosia were also tested (nine female, mean age: 31.92 years, ages ranged from 19 to 49 years). Each DP participant was individually aged-matched to one control participant within an age range of ±4 years. All participants had normal or corrected-to normal vision, and gave written and verbal informed consent prior to testing.

DP participants were recruited through two research websites (http://www.faceblind.org; http://www.prosopagnosia.bbk.ac.uk). All reported difficulties with face recognition since childhood. We confirmed the presence of their face recognition impairments with a battery of behavioural tests. Long-term face memory was assessed with the Famous Faces Test (FFT, Duchaine & Nakayama, 2005), which required participants to identify 60 people who are famous in popular culture, e.g. actors, musicians, politicians. Memory for new unfamiliar faces was assessed with the Cambridge Face Memory Test (CFMT). Participants were required to memorize faces of six target individuals shown from different viewpoints which they then had to identify among other similar distractor faces in a test array (see Duchaine & Nakayama, 2006, for a detailed description). The Old–New Face Recognition Test (ONT, Duchaine & Nakayama, 2005) also tested face learning. DP participants had to memorize 10 faces, and to distinguish these learned faces from 30 novel faces by making an old/new judgement for each item. The Cambridge Face Perception Test (CFPT, Duchaine et al., 2007) assessed the ability of DPs to perceptually process faces in the absence of memory demands. Participants were shown a target face presented together with six-front view morphed test faces that resembled the target face to varying degrees. These test faces had to be rearranged in order of their degree of similarity to a target face. DPs completed this task when the target and test faces were upright, and when they were inverted. Individual z-scores for each of the 14 DP participants for these four behavioural tests are shown in Table 1. Because impaired face recognition is the defining feature of DP, the criterion employed to classify a particular individual as DP was that they were impaired (below −2 z-scores of the mean) in at least two of the three face recognition tests (FFT, CFMT, ONT). All DPs scored below −4 z-scores of the mean in the FFT and below −2 z-scores in the ONT, and all except one below −2 z-scores in the CFMT. Performance was more variable in the CFPT, where the majority of DPs performed within the normal range. All control participants reported that they were confident in their face recognition abilities.

Table 1 – Z-values for 16 DP participants in the Cambridge Face Memory Test (CFMT), the Cambridge Face Perception Test (CFPT) for upright and inverted faces, Famous Faces Test (FFT), and the Old–New Test (ONT).

	CFMT	CFPT Upright	CFPT Inverted	FFT	ONT
DM	−3.78	−.92	−.06	−4.25	−7.13
CM	−4.29	−3.1	−2.89	−7.72	−14.34
TW	−2.52	−1.74	.79	−9.46	−3.61
SK	−1.25	−.78	−0.2	−5.21	−3.36
KT	−2.52	−.92	−0.2	−5.98	−1.54
KS	−2.9	−.92	−1.05	−8.49	−9.03
DD	−2.77	.17	−.77	−5.21	−3.36
LR	−2.39	−.38	−.63	−6.56	−4.9
MF	−2.14	−2.29	0.5	−5.96	−10.35
ZS	−2.14	−.92	−.35	−6.95	−2.04
PH	−3.02	−3.24	−1.48	−8.49	−5.52
MM	−2.26	−1.60	−.20	−5.79	−3.93
MC	−3.02	−1.19	1.07	−4.83	−4.90
DB	−4.03	−.10	.50	−6.76	−5.41

2.2. Stimuli and procedure

Participants were seated in a dimly lit, sound-attenuated and electrically shielded chamber. The face stimuli were created using 10 images of male Caucasian faces obtained from the PUT Face Database (Kasiński, Florek, & Schmidt, 2008) and the University of Stirling Online Database. All images were converted to greyscale, and were edited using Adobe Photoshop to homogenise large differences in overall luminance, and skin tone and hair. Distinguishing characteristics (e.g., piercings or blemishes) were removed from the images. The internal features of each of the ten faces were paired with the external features of each of the other ten faces to create a total of 100 face stimuli (ten original faces, 90 newly created composite faces; see Fig. 1 for examples). All stimuli were presented on a CRT monitor against a near black background (.4 cd/m^2) at a viewing distance of 100 cm. Stimulus presentation, timing and response recording were controlled by E-Prime (Psychology Software Tools, Pittsburgh, PA). On each trial, two faces were presented in rapid succession. The first face (S1) was presented for 400 ms and the second face (S2) was presented for 200 ms. These two face images were separated by a 200 ms interstimulus interval. The intertrial interval was 1500 ms. S1 stimuli occupied a visual angle of 5.8° × 8°. S2 stimuli were 10% larger than S1 stimuli in order to avoid pixel-wise matching, in particular on trials where S1 and S2 images were otherwise identical. All face images were presented in their standard upright orientation. The average luminance of all face stimuli was 21 cd/m^2.

There were four types of S1-S2 sequences that occurred in random order and with equal probability. On full repetition trials, the S1 and S2 face images were identical (except that S2 images were larger). On full change trials, the S1 and S2 faces differed both in terms of their external and internal features. On external feature repetition trials, the internal features of the two faces differed but their external features were identical. On internal feature repetition trials, the external features of the S1 and S2 faces differed, but their internal features were identical. Participants were instructed to encode both the internal and external features of the first face and to decide whether both were repeated in the second face, or whether there was a change between the two faces. They signalled a full face repetition by pressing one response key and the presence of a change by pressing another response key. A "change" response was required when there was a

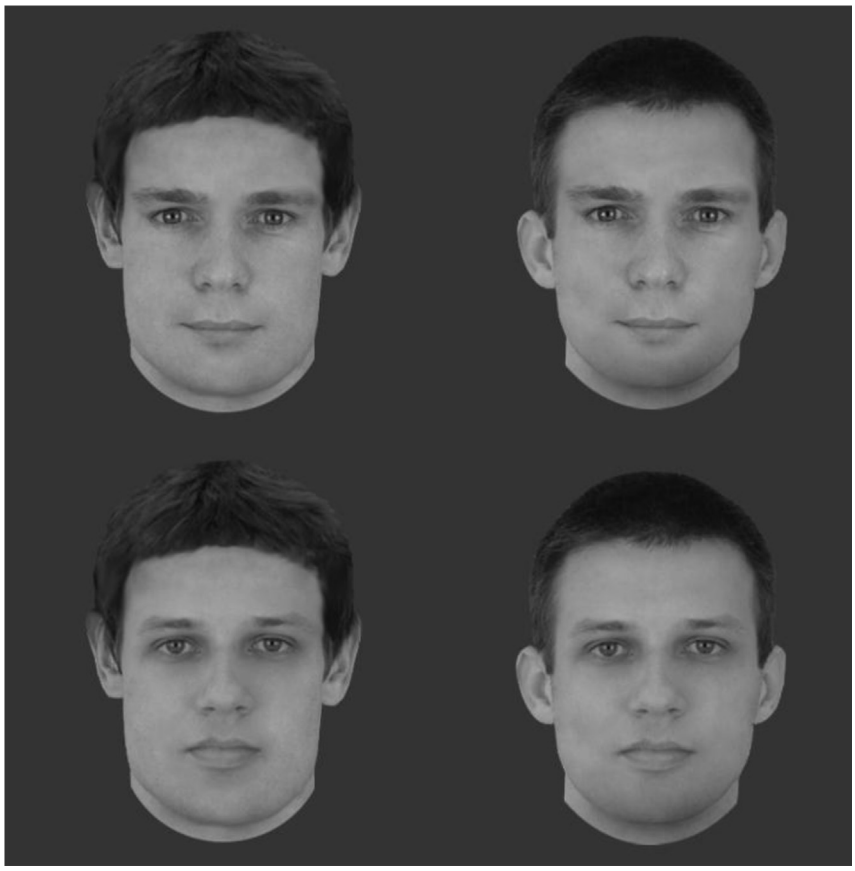

Fig. 1 – Examples of different face images shown in the current experiment. The faces in the top versus bottom row have the same external features, but different internal features. The faces on the left versus right side have the same internal features, but different external features.

change in the external features, the internal features, or both. Responses were executed with the index and middle finger of one hand. Response hand was counterbalanced across participants.

The experiment contained 12 blocks, with 50 trials per block, resulting in 600 trials in total. There were 150 trials for each of the four different types of S1-S2 sequences. Individual face images were presented in a pseudorandom order in which each face (of the 100 faces) was presented once as the S1 face within a two-block period (100 trials). All participants were given a training block of 50 trials before starting the first experimental block. After each block, they received on-screen feedback about their average accuracy and response times (RTs) in this block.

2.3. *EEG recording and data analysis*

EEG was DC-recorded with a BrainAmps DC amplifier (upper cut-off frequency 40 Hz, 500 Hz sampling rate) and Ag-AgCI electrodes mounted on an elastic cap from 27 scalp sites (Fpz, F7, F3, Fz, F4, F8, FC5, FC6, T7, C3, Cz, C4, T8, CP5, CP6, P7, P9, P3, Pz, P4, P8, P10, PO9, PO7, PO8, PO10 and Oz, according to the extended international 10–20 system). Bipolar horizontal electrooculogram (HEOG) was recorded from the outer canthi of both eyes. An electrode placed on the left earlobe served as reference for online recording, and EEG was re-referenced off-line to the common average of all scalp electrodes. Electrode impedances were kept below 5 kΩ. No additional off-line filters were applied. ERPs in response to the S2 face on each trial were computed on the basis of EEG epochs obtained between 50 ms before to 500 ms after S2 onset, relative to a 100 ms baseline from 50 ms before to 50 ms after S2 onset. This non-standard baseline was chosen to minimize the presence of ERP components triggered in response to the S1 face in the baseline period (see Towler & Eimer, 2016, for identical procedures). Epochs with activity exceeding ±30 μV in the HEOG channel (reflecting horizontal eye movements) or ±60 μV at Fpz (indicating eye blinks or vertical eye movements) were excluded from all analyses, as were epochs with voltages exceeding ±80 μV at any other electrode. Trials with incorrect responses were also excluded from the EEG analysis.

Following artifact rejection, EEG epochs were averaged to compute ERP waveforms for the four trial types (full repetition, full change, internal feature repetition, external feature repetition). ERP mean amplitudes in the N250r time window (230–280 ms after S2 onset) were measured at four lateral posterior electrode sites over the left hemisphere (P7, PO7, P9, and PO9), and at the corresponding electrodes over the right hemisphere (P8, PO8, P10, and PO10). Statistical analyses of N250r amplitudes were conducted with mixed-design ANOVAs for the within-participants factors external feature repetition (repetition *vs* change), internal feature repetition

(repetition *vs* change), hemisphere (left *vs* right), and electrode site (four lateral posterior electrode positions), and the between-participants factor group (developmental prosopagnosics *vs* control participants). Additional analyses were conducted separately for developmental prosopagnosics and control participants. The factors internal feature repetition, external feature repetition, and group were employed for the analyses of behavioural performance.

To evaluate the reliability of N250r components at the level of individual participants, a non-parametric bootstrap procedure (Di Nocera & Ferlazzo, 2000) was employed. This procedure assesses the reliability of ERP amplitude differences between two conditions by resampling and averaging two sets of trials that are drawn randomly (with replacement) from the combined dataset, and computing differences between the two resulting ERPs. This procedure was repeated 10,000 times in the current study, resulting in a distribution of difference values with a mean value of zero, as both sample pairs were drawn from the same dataset. Based on this distribution, the reliability of an empirically observed ERP difference between conditions was determined for individual participants. If the probability of obtaining the observed difference by chance is below 5%, it is accepted as statistically significant (see Dalrymple et al., 2011; Towler, Gosling, Duchaine, & Eimer, 2012; Towler, Gosling, Duchaine, & Eimer, 2016; Fisher et al., 2016, 2017, for previous applications of this procedure in ERP studies of prosopagnosia). In the present experiment, this bootstrap procedure was based on EEG mean amplitudes obtained between 230 and 280 ms after S2 onset on full repetition and full change trials (collapsed the eight lateral posterior electrodes over the left and right hemisphere).

3. Results

3.1. Behavioural performance

Error rates. Fig. 2 shows error rates (bars) and response times (RTs; lines) for the Control Group (grey) and for the DP Group (black). Individuals with DP performed generally worse on the face matching task than control participants, with an overall error rate of 13% (SE = 1.61) as compared to 3% (SE = .68) for the Control group. This was reflected my a main effect of group, F (1,26) = 35.2, $p < .001$, $\eta_p^2 = .575$. As shown in Fig. 2, the performance of DPs was particularly impaired on external feature repetition trials where errors (i.e., an incorrect "repetition" response when internal features changed) were much more frequent than on all other types of trials (29% as compared to less than 11%). Control participants showed no such selective performance deficit for external feature repetition trials. This was reflected in interactions between group and internal feature repetition, F (1,26) = 6.43, $p < .02$, $\eta_p^2 = .198$, group and external feature repetitions, F (1,26) = 11.24, $p < .002$, $\eta_p^2 = .302$, as well as, most importantly, a highly significant three way interaction (group x internal feature repetitions x external feature repetition: F (1,26) = 14.44, $p < .001$, $\eta_p^2 = .357$). Additional ANOVAs were performed for each group separately. For the DP group, a significant main effect of external feature repetition, F (1,13) = 14.67, $p = .002$, $\eta_p^2 = .53$, and a non-significant trend for internal feature repetition, F (1,13) = 3.40, $p = .088$, $\eta_p^2 = .207$, were accompanied by a significant interaction between these two factors F (1,13) = 21.19, $p < .001$, $\eta_p^2 = .62$. For the Control group, there was a main effect of internal feature repetition, F (1,13) = 8.97, $p = .01$, $\eta_p^2 = .408$, with more errors on trials where internal features were repeated. The effect of external feature repetition was not significant [F (1,13) = 3.317, $p = .092$, $\eta_p^2 = .203$], and there was no interaction between these two factors for control participants, F < 1.2.

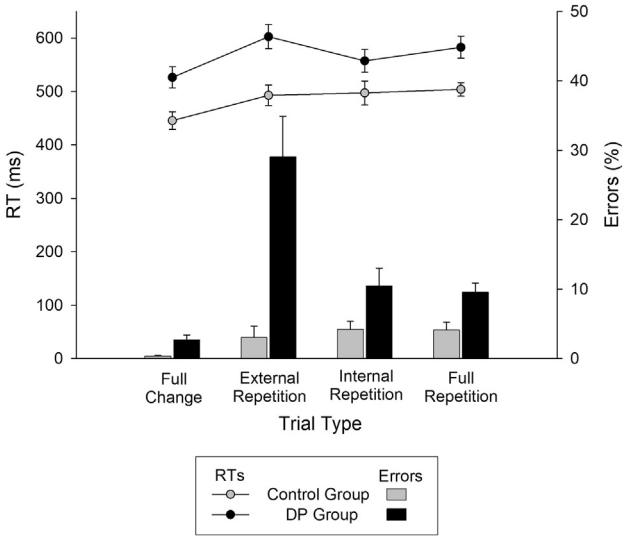

Fig. 2 – Response times (RTs, line graphs) and error rates (bar graph) for the four different trial types (Full Change, External Repetition, Internal Repetition, Full Repetition), shown separately for the Control group and the DP group. Error bars represent one standard error of the mean.

To assess whether the face identity matching problems shown by DPs specifically on trials where external features were repeated but internal features changed would improve as a result of learning, error rates for the DP group were calculated separately for the first half (blocks 1–6) and second half (blocks 7–12) of the experiment. There was indeed a considerable improvement for external repetition/internal change trials in the second experimental half, with incorrect "repetition" responses on 24% of all trials as compared to 34% in the first half, t (13) = 2.38, $p < .05$. A smaller practice-related improvement was also present on internal repetition/external change trials (8% *vs* 13%), but this difference only approached significance, t (13) = 1.84, $p = .088$. No such reduction of error rates in the second experimental half were observed for full change trials (3% in both halves) and full repetition trials (10% *vs* 9%; both t < 1). In spite of their improved performance on external repetition/internal change trials, DPs continued to show disproportionate deficits on these trials also in the second experimental half. This was confirmed by ANOVAs conducted separately for the first and second half of the experiment, which showed significant interactions between internal and external feature repetition for both halves (first half: F (1,13) = 22.881, $p < .001$, $\eta_p^2 = .638$, second half: F (1,13) = 11.383, $p = .005$, $\eta_p^2 = .467$). Correlational analyses revealed that the DP participants with the highest error rates

on external repetition/internal change trials in the first half improved the most in the second half of the experiment, r = .842, $p < .001$. A similar pattern was also found for internal repetition/external change trials, r = .568, $p < .05$.

Response times. Responses in the DP group were approximately 80 ms slower in than in the control group (567 ms [SE = 19.88] *vs* 485 ms [SE = 16.20]; main effect of group: F (1,26) = 11.11, $p < .003$, $\eta_p^2 = .299$). There was an interaction between internal feature repetition and group, F (1,26) = 5.09, $\eta_p^2 = .164$, but no further significant main effects or interactions, all F < 3.1. Additional ANOVAs were conducted for each group separately. For DP participants, there was no effect of internal feature repetition, F < 1, but the effect of external feature repetition was significant, F (1,13) = 22.23, $p < .001$, $\eta_p^2 = .631$, with slower RTs when external features were repeated than when they changed. In line with the accuracy results, there was an interaction between these two factors for the DP group, F (1,13) = 44.28, $p < .001$, $\eta_p^2 = .773$, as RTs were disproportionately delayed on external feature repetition trials (see Fig. 2). For the Control group, main effects of internal feature repetitions, F (1,13) = 14.38, $p < .005$, $\eta_p^2 = .525$, and external feature repetition, F (1,13) = 10.73, $p = .006$, $\eta_p^2 = .452$ were present, and the interaction between these factors approached significance, F (1,13) = 3.45, $p = .086$, $\eta_p^2 = .21$. This pattern of effects is primarily due to the fact that RTs in the Control group were faster on full change trials (446 ms) relative to the other three types of trials (range: 493–504 ms).

3.2. ERP results

3.2.1. N250r components

Fig. 3 (top panels) shows ERPs elicited at the four lateral posterior electrodes over the left and right hemisphere in response to S2 faces on the four different types of trials, separately for control and DP participants. Only trials with correct responses contributed to these ERP waveforms. An N250r to full repetitions versus full changes was present in both groups, but appeared to be smaller in the DP group. N250r components to internal or external feature repetitions were clearly smaller than the N250r to full repetitions. Fig. 3 (bottom panel) shows the scalp topography of N250r components on full repetition trials, separately for the Control and DP groups, based on difference amplitudes obtained in the N250r measurement interval (230–280 ms post-stimulus) after subtracting ERPs on full change trials from ERPs on full repetition trials. The scalp distribution of N250r components is similar for both groups, but N250r amplitudes are clearly attenuated in the DP group.

An overall ANOVA with Group as a between-subject factor revealed main effects of internal feature repetition, F (1,26) = 45.99, $p < .001$, $\eta_p^2 = .639$, and external feature repetition, F (1,26) = 69.05, $p < .001$, $\eta_p^2 = .726$, confirming the presence of N250r components to both internal and external feature repetitions versus changes across all participants tested. Importantly, both factors significantly interacted with group, (internal feature repetition: F (1,26) = 14.69, $p < .001$, $\eta_p^2 = .361$; external feature repetition: F (1,26) = 4.66, $p = .04$, $\eta_p^2 = .152$, as N250r components were smaller in the DP group for both types of repetitions. As expected, there was a significant interaction between internal and external feature repetition, F (1,26) = 16.56, $p < .001$, $\eta_p^2 = .39$, indicative of the presence of superadditive N250r components for full as compared to partial face repetitions. Critically, there was also a three-way interaction between internal feature repetition, external feature repetition, and group, F (1,26) = 9.24, $p = .005$, $\eta_p^2 = .262$, suggesting that the superadditivity of N250r components might differ between DPs and control participants. To investigate this, and to assess the presence of reliable N250r components in both groups, separate additional ANOVAs were run for the DP and Control groups.

For the Control Group, there were main effects of both internal feature repetition, F (1,13) = 54.27, $p < .001$, $\eta_p^2 = .807$, and external feature repetition, F (1,13) = 46.30, $p < .001$, $\eta_p^2 = .781$, and, crucially, a highly significant interaction between these two factors, F (1,13) = 49.12, $p < .001$, $\eta_p^2 = .791$. These results confirm the presence of superadditive N250r components for full as compared to partial face repetitions in the Control group. This is illustrated in Fig. 4 (top panel), shows an ERP difference wave for the N250r to full face repetitions (obtained by subtracting full change from full repetition trials) and a difference wave illustrating the sum of the contributions of internal and external repetitions to the N250r. This difference wave was obtained by first subtracting ERPs on full change trials from ERPs on internal and external repetition trials, respectively, and then adding the two resulting difference waves. As is obvious from Fig. 4, the N250r to full repetitions in the Control group is substantially larger (i.e, superadditive) than the sum of the N250r components to internal and external feature repetitions. For the DP group, reliable main effects were also found for both internal feature repetition, F (1,13) = 7.43, $p = .017$, $\eta_p^2 = .364$, and external feature repetition, F (1,13) = 23.17, $p < .001$, $\eta_p^2 = .641$], indicating that both types of repetitions triggered N250r components. Critically, and in marked contrast to the Control group, there was no evidence for an interaction between these two factors, F < 1, which strongly suggests that no superadditive N250r component to full as compared to partial repetitions was elicited for participants with DP. This is shown in Fig. 4 (bottom panel), which shows the N250r to full face repetitions and the sum of the N250r components to internal and internal feature repetitions in the DP group. In contrast to control participants, there is little evidence for a superadditive N250r to full face repetitions for participants with DP.

However, Fig. 4 suggests that some residual N250r superadditivity might emerge at a later point in time than the DP Group. To test this possibility, we conducted additional analyses within a later time window (280–330 ms after S2 onset), separately for both groups. There was no evidence for an interaction between internal and external feature repletion for the DP group, F < 1.3. For the Control group, this interaction now only approached significance, F (1,13) = 4.01, $p = .067$, $\eta_p^2 = .236$]. In other words, there was no indication for any N250r superadditivity during a later post-stimulus period for DPs, and additional evidence that this superadditivity is a transient phenomenon that dissipates around 300 ms after S2 onset for control participants. Fig. 4 also suggests that while a superadditive N250r emerges from control participants from about 230 ms post-stimulus, the initial phase of the N250r shows no superadditivity in either group (see also Towler & Eimer, 2016, for analogous results). This was confirmed by

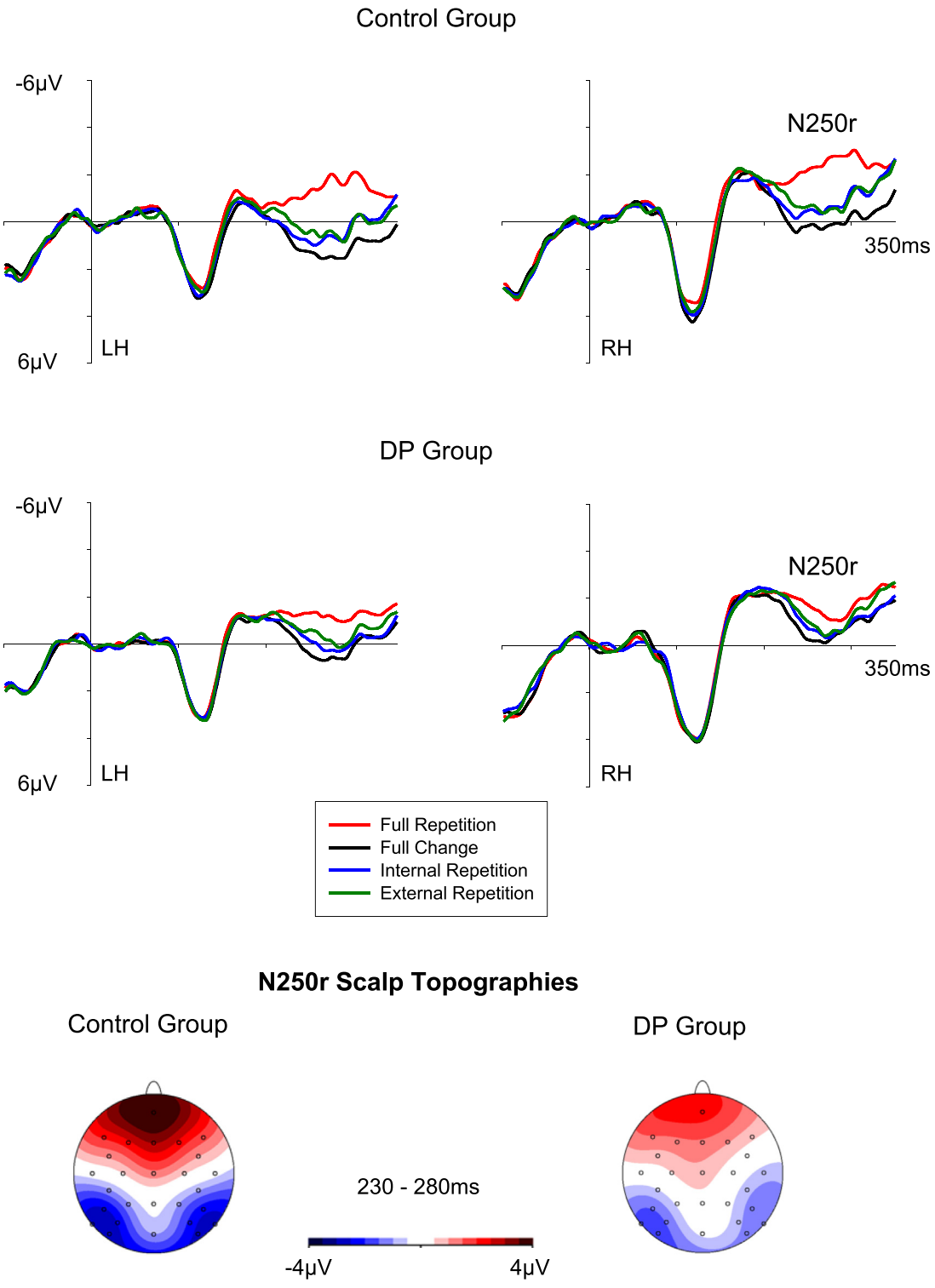

Fig. 3 – Top panels: Grand averaged event-related brain potentials (ERPs) measured in the 350 ms interval after the onset of the S2 face at lateral posterior electrodes over the left hemisphere (LH) and right hemisphere (RH). ERPs elicited on the four different types of trials are shown separately for the Control and DP groups. Bottom panel: Topographical maps showing the scalp distribution of N250r components to full face repetitions in the Control group (left) and DP group (right). Maps were generated by subtracting ERP mean amplitudes measured in the 230–280 ms post-stimulus time window on full change trials from ERPs on full repetition trials.

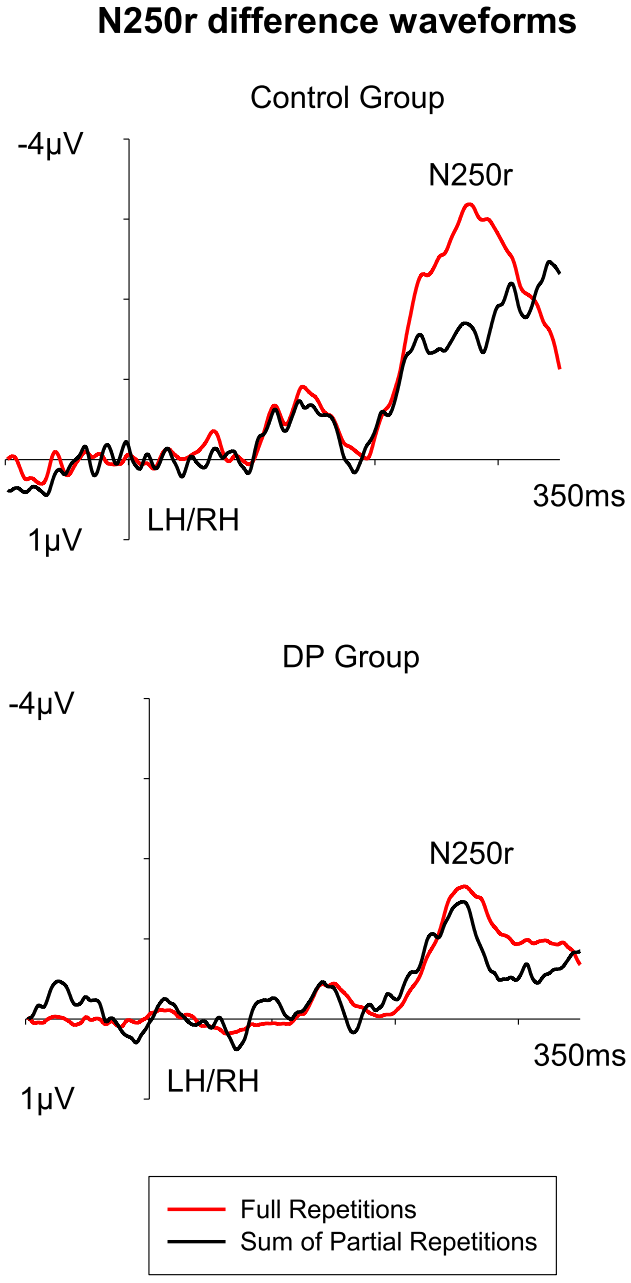

Fig. 4 – N250r difference waveforms during the 350 ms interval after S2 onset for the Control and DP groups (top and bottom panels). All difference waves are collapsed across hemisphere and lateral posterior electrode sites. N250r difference waves for full face repetitions were obtained by subtracting ERPs on full change from full repetition trials. The difference waves representing the sum of partial face repetitions were computed by subtracting full change trials from internal repetition trials, and full change trials from external repetition trials, and summing the resulting two N250r difference waveforms. The superadditivity of N250r components was assessed from 230 to 280 ms after S2 onset and in an additional later time window (280–330 ms post-stimulus).

additional analyses that were conducted for an earlier time window (200–230 ms after S2 onset). In the analysis including both groups, there were significant effects of both internal feature repetitions, F (1,26) = 5.81, p = .023, η_p^2 = .183, and external feature repetition, F (1,26) = 21.12, p < .001, eta = .448, but no interaction between these factors, F < 1. This interaction was also absent for either group when ERPs for DP and Control participants were analysed separately.

Because the analyses of error rates in the DP group revealed evidence for improved performance in the second half of the experiment (see above), we conducted an exploratory analysis of N250r components measured for participants with DP in the first versus second experimental half for all four different types of trials. This analysis, which was based on the data of those 12 of the originally tested 14 DPs who had sufficiently large number of epochs remaining after artefact rejection, found no evidence for any systematic differences of N250r components between experimental halves. There was no interaction between experimental half and either internal or external feature repetition, and no three-way interaction between these factors (all F < 1).

3.2.2. *Individual N250r amplitudes and correlation analyses*

Because DP is a heterogeneous condition, it is important to assess atypical face processing in DP not just on the basis of differences between DP and Control groups, but also at the level of individual participants. Fig. 5 (top panel) shows N250r mean amplitudes for full repetition versus full change trials (collapsed across hemispheres) for each individual participant with DP (black bars) and each control participant (grey bars). The bottom panel of Fig. 5 shows the size of the superadditive N250r for each individual participant (computed as the difference between N250r amplitudes on full change versus full repetition trials and the sum of the two N250r amplitudes obtained for internal repetition versus full change and external repetition versus full change trials). In both plots, participants are ordered from left to right as a function of the size of individual N250r components. For both N250r amplitudes and N250r superadditivity, control participants tended to cluster on the left, and DPs on the right, reflecting the overall attenuation of N250r components and N250r superadditivity in the DP group. There was however some overlap between the two groups, with some DPs showing N250r effects in the normal range, and some control participants with small N250r effects. For N250r amplitudes, the presence of reliable effects at the level of individual participants was determined with a non-parametric bootstrap analysis (Di Nocera & Ferlazzo, 2000), as indicated in Fig. 5 (top panel) by asterisks. At the descriptive level, eleven of the fourteen control participants tested showed a reliable N250r on full repetition trials. In contrast, only seven of the fourteen DPs had a significant negative-going N250r component. In addition, one DP participant (DM) had a significant positive N250r component, reflecting atypical larger N250r amplitudes on identity change than on identity repetition trials. For N250r superadditivity (Fig. 5, bottom panel), no such bootstrap-based statistical analyses could be conducted because this measure was not based on simple amplitude differences between task conditions, but was computed in a more complex way (as the difference between the N250r on full repetition trials and the

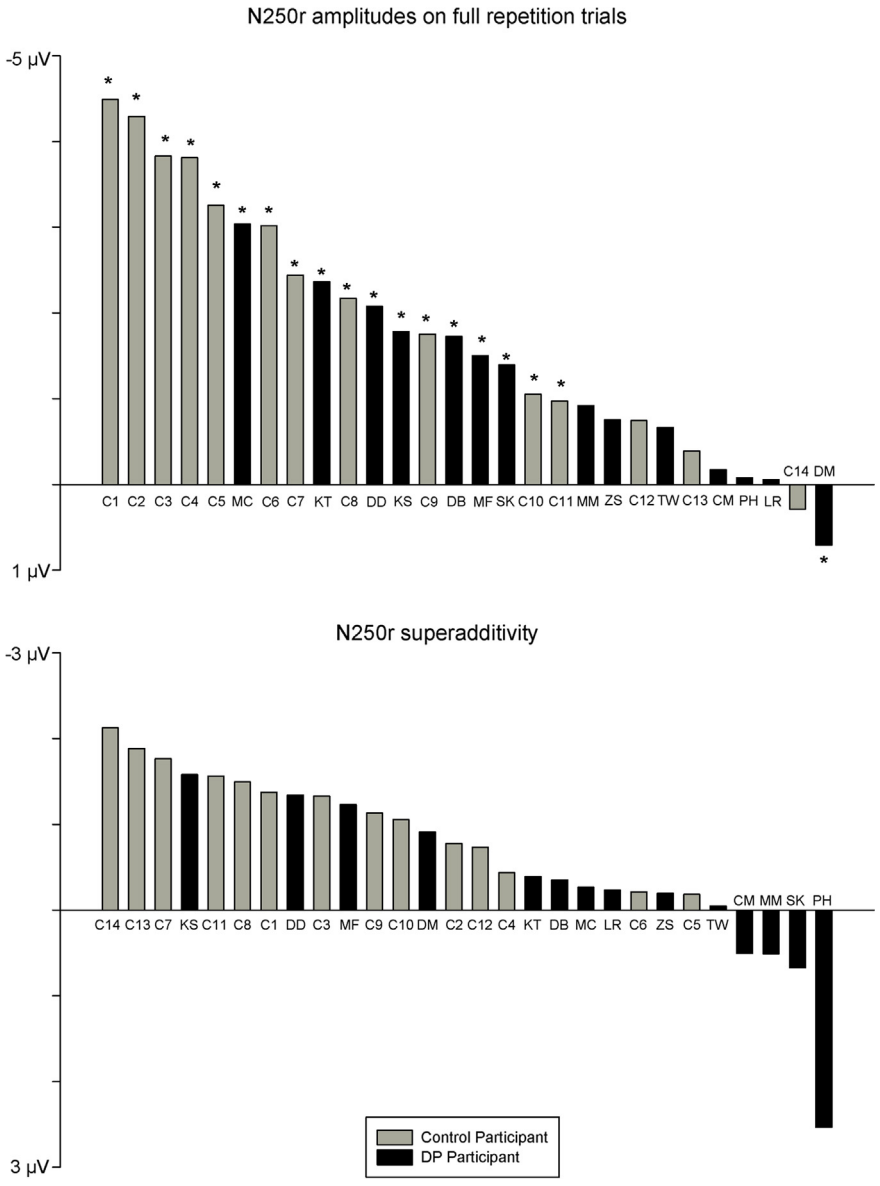

Fig. 5 – Top panel: N250r amplitudes for individual participants with DP (black bars) and control participants (grey bars). These amplitude values were calculated by subtracting ERP mean amplitudes measured 230–280 after S2 onset on full change trials from mean amplitudes on full repetition trials, and averaging across all eight lateral posterior electrodes over the left and right hemispheres. Individual DP participants are labelled with their initials, corresponding to Table 1. Asterisks indicate reliable N250r components, as determined by bootstrap analyses. Bottom panel: N250r superadditivity effects for individual DPs and control participants. Amplitude values were computed by subtracting N250r amplitudes on full repetition versus full change trials from the sum of the two N250r amplitudes obtained for internal repetition versus full change trials and external repetition versus full change trials, respectively.

summed N250r components for internal and external feature repetition trials, see above).

To assess whether differences in N250r amplitudes or N250r superadditivity between individual DPs are linked to individual differences in face processing ability, we correlated these two N250r-based measures with the performance of each DP in the standardised behavioural tests. N250r amplitudes and N250r superadditivity were not correlated with CFMT performance or any other face recognition test score. However, these two markers were linked to individual performance on the CFPT. Generic N250r amplitudes were correlated with CFPT performance with inverted faces, $r = .46$, $p < .05$, with larger N250r amplitudes predicting better performance. In contrast, N250r superadditivity was correlated with upright-face CFPT performance, with stronger N250r superadditivity effects associated with better performance, $r = -.59$, $p < .02$, but not with CFPT scores with inverted faces, $p > .1$. Finally, we tested whether the performance of individual DPs in the current face matching task, and specifically their notable impairment on trials with external feature

repetitions (see Fig. 2), was associated with N250r components or scores in standardised tests. Higher error rates on external feature repetition trials predicted poorer performance in the CFMT, $r = -.48$, $p < .05$, and were associated with larger N250r amplitudes for full face repetitions, $r = -.55$, $p < .05$. Because the sample size of DP participants was relatively small ($N = 14$), these correlations were no longer significant when using non-parametric Spearman's Rho (all 95% confidence intervals overlapped with $r = 0$).

4. Discussion

The current study utilised event-related brain potentials and measured performance in a face identity matching task to evaluate the locus of perceptual impairments in developmental prosopagnosia. Two upright face images were presented sequentially, and full face repetitions, full face changes, and partial face repetitions of either the internal or external features occurred with equal probability. Participants' task was to report the presence or absence of any change across the two faces, in order to emphasize internal and external features equally, and to discourage any part-based encoding and matching strategies. A group of 14 DPs and a group of 14 age-matched control participants were tested. To track visual face memory matching processes, and in particular the involvement of holistic and part-based face representations during face matching, N250r components were measured to full and partial repetitions of internal and external facial features.

The results obtained for control participants were consistent with the observations from our previous study in blocks with upright faces (Towler & Eimer, 2016). Accuracy was close to ceiling, and there was little interference from partial face repetitions, suggesting the involvement of holistic face representations. RTs were slightly delayed on internal and external repetition trials as compared to full change trials in the Control group. This indicates that partial repetitions of internal or external features were registered (in line with the involvement of part-based face representations, see below), but could be rapidly rejected when there was a change in the other set of features. As expected, N250r components were present not only for full face repetitions, but also for both partial repetitions of the internal and external features. Importantly, and in line with our previous observations (Towler & Eimer, 2016), the N250r response for full face repetitions was larger than the sum of the N250r components obtained on partial repetition trials. This super-additivity of the N250r component demonstrates that holistic face representations were activated during the face identity matching process, and that individual faces were visually represented as more than the sum of their parts. The presence of N250r components to partial repetitions of external or internal features demonstrates that additional part-based face representations were also active during this face matching task.

A different pattern of results was obtained for individuals with developmental prosopagnosia. As expected, DPs generally performed worse in the face identity matching task than control participants. However, these impairments did not affect all trial types equally. Disproportionate performance costs were observed on partial repetition trials, and specifically on trials where the external facial features were repeated and internal features changed. On nearly one third of these trials, participants with DP reported incorrectly that the entire face had been repeated. Even when an internal feature change was reported correctly on these trials, RTs were slower than on all other types of trials. These novel findings suggest that individuals with DP have an atypical bias towards encoding facial features that are contained in the external parts of the face (e.g., hair and head outline) during face matching tasks, and place less emphasis on the representation of internal features. The selective impairment observed for DPs on external repetition/internal change trials improved in the course of the experiment, with a reduction of error rates on these trials from 34% to 24%. This shows that DPs are capable of learning to perform face identity matching tasks more effectively, and provides support for the view that DPs are slow face learners. Face processing deficits in DP are not static and unchanging, but are malleable and can be improved through training (e.g., DeGutis, Cohen, & Nakayama, 2014). Participants who performed poorly on these trials during the first half showed the largest improvement in the second half. This suggests that even the DPs with the most pronounced deficits were able to learn to perform face identity matching tasks more effectively, possibly by basing their same/different responses more equally on both external and internal face parts. In spite of these learning-related improvements, a disproportionate deficit on external repetition/internal change trials was still evident for DPs in the second half of the experiment, suggesting that atypical spatial biases remained present even after extended practice. It should also be noted that the moderate improvements of face matching performance for DPs in the second part of the experiment were not matched by corresponding electrophysiological differences at the level of N250r components between experimental halves. This suggests that these behavioural learning effects were not primarily a result of improved perceptual face matching reflected by the N250r, but were mainly generated at later stages associated with explicit identity judgments and response selection processes.

The analysis of N250r components to full face repetitions in the DP group confirmed previous observations (Fisher et al., 2017) that DPs have significant but strongly reduced N250r components as compared to age-matched control participants without face processing impairments. Individuals with DP showed reliable N250r components not only in response to full face repetitions versus changes, but also to partial repetitions of both internal and external features. However, and critically, the sum of the two N250r components on external and internal feature repetition trials was equal in size to the N250r component observed for full face repetition trials. This dissociation between a superadditive N250r for the Control group and a strictly additive N250r for the DP group provides novel evidence for qualitative differences in perceptual face processing between DPs and individuals with typical face processing abilities. The lack of any N250r superadditivity in the DP group strongly suggests that face identity matching processes were based on additive contributions from matching processes that operated independently for internal and external facial features, without an involvement of holistic

face representations. Thus, the performance impairments observed for individuals with DP in the current face identity matching task may reflect a deficit in forming holistic face representations that encompass both the internal and external facial features.

It is striking that the pattern of error rates and N250r components observed for DPs in the current study is similar to the pattern observed for unimpaired control participants in our previous study in blocks with inverted faces (Towler & Eimer, 2016). In these blocks, the superadditivity of N250r components that was observed with upright faces was abolished, and error rates increased specifically on partial repetition trials, indicating that face inversion prevented the formation of holistic face representations. These effects of face inversion on face identity matching in neurotypical participants may therefore provide a useful model for understanding the nature of the impairments of face representation and matching mechanisms in individuals with developmental prosopagnosia. Their face recognition deficits may be at least partially caused by difficulties in activating, maintaining, and utilizing holistic face representations. Without such holistic representations, face matching and face recognition have to be based exclusively on part-based representations of individual facial features, and thus cannot benefit from any identity-related signals produced by holistic face processing. In addition to the apparent absence of holistic face representations, the pattern of performance observed for individuals with DP (see above) strongly suggests that they have a specific bias towards encoding external facial features. This is different from the way in which unimpaired control participants process inverted faces.

In our previous study (Towler & Eimer, 2016), face inversion caused a small but reliable increase in errors and response times for both internal and external feature repetitions, suggesting that these two types of features are equally represented, and thus produce similar costs on partial repetition trials. The fact that DPs showed disproportionate performance costs on external feature repetition trials in the present study, suggests that they may have particular difficulties encoding and maintaining internal facial features, and thus prioritise external features during face identity matching processes. In addition to a general deficit in holistic face processing, the face recognition deficits in DP may also be due to an atypical bias towards external features. It is possible that a lifelong and chronic impairment of holistic face processing causes a compensatory strategy of attending to external features, which may even further impair face recognition ability. A strong bias towards prioritizing external facial features in DPs could also be reflected by a tendency to preferentially fixate external instead of internal facial features. Because eye position was only monitored with horizontal EOG but not with additional eye tracking procedures, the existence of such an eye movement bias towards external features in DPs could not be assessed in the present study (but see Schwarzer et al., 2007, for initial evidence that DPs fixate more on external features during face recognition tasks).

To complement the analyses at the group level, we also quantified N250r components and N250r superadditivity separately for each individual participant with DP. Confirming previous observations (Fisher et al., 2017), we found substantial differences between individual DPs in the size of their N250r amplitudes to full face repetitions versus changes, with some DPs showing reliable N250r components in the normal range (Fig. 5, top panel). A similar pattern was also found for the superadditivity of N250r components for full versus partial feature repetitions (Fig. 5, bottom panel). These findings underline the importance of not only focusing on group differences between DPs and control participants, but also documenting differences at the individual level. A correlation analyses suggested a link between the superadditivity of the N250r component and the performance of individual DPs on an independent standardised test of face perception (CFPT), specifically for upright faces. Better CFPT performance was associated with a stronger tendency towards a superadditive N250r to full versus partial face repetitions, which could reflect a link between the ability to represent faces holistically and the ability to perceptually process and match upright faces. Larger N250r components on full face repetition trials were linked to better performance on the CFPT with inverted faces, but also to higher error rates on external feature repetition trials. This pattern suggests that where N250r components are observed in DP individuals, they primarily reflect the strength of part-based face representations, with an additional bias towards representing external features, rather than the activation of holistic representations. A stronger reliance on face parts may facilitate the perceptual matching of inverted faces, but can produce costs in situations where repetitions/changes of internal and external features are incongruent. However, these conclusions have to remain tentative at present. Because the sample size of DPs was relatively small, these correlations were no longer reliable when non-parametric analyses were used.

In contrast to a previous study (Fisher et al., 2017), N250r amplitudes on full face repetition trials were not correlated with CFMT performance in the DP group (see also Wirth, Fisher, Towler, & Eimer, 2015, for a link between N250r components and CFMT scores in unimpaired participants). This difference may simply reflect variability between different relatively small samples of DPs. Alternatively, it could be linked to differences in task demands. In these previous studies, face matching tasks were used where the sample and test faces showed different images of the same individuals, and therefore required individuals to form image-invariant face representations. In the present study, two identical face images were presented on full repetition trials, and either internal or external features remained unchanged on partial repetition trials, so that image-based matching processes were sufficient to discriminate repetitions and changes. Previous research has shown that individuals with DP have disproportionate difficulties in face identity matching when the sample and test images visually vary more substantially (for example, the same person posing different facial expressions as compared to a similar expression; e.g. Fisher et al., 2017; White et al., 2017). This suggests that face recognition deficits in DP arise at the stage of forming image-invariant face identity representations. Although DPs sometimes have difficulties on face matching tasks with identical images, this is not always the case (e.g. Le Grand et al., 2006; Ulrich et al., 2017). Because no image-invariant identity matching was required in the present study, the atypical N250r components

observed for DPs might reflect deficits at an earlier stage of forming image-dependent face representations – in particular for internal facial features. The absence of correlations between N250r components and CFMT scores suggest that such image-specific face perception impairments may not be systematically related to face memory ability that is measured with the CFMT.

What do the current findings imply for the cognitive and neural mechanisms that are responsible for impaired face recognition in DP, and in particular for the role of holistic processing? Previous behavioural studies using different face matching paradigms (composite face and part-whole tasks) have shown that individuals with DP can have problems with holistic face processing, but that this is not always the case (Avidan et al., 2011; Biotti et al., 2017; De Gutis et al., 2012). Demonstrations that DPs may have normal holistic face processing abilities have contributed to the debate about the extent to which holistic face perception is related to face recognition ability in the general population (DeGutis, Wilmer, Mercado, & Cohan, 2013; Richler, Cheung, & Gauthier, 2011). A recent study suggests that different kinds of task measure different aspects of holistic face perception that may be differentially dissociable from one another and from face recognition ability itself (Rezlescu, Susilo, Wilmer, & Caramazza, 2017). Previous studies investigating the neural basis of face perception and recognition in DP have also been inconclusive with respect to the possibility of impaired holistic face processing. Face-selective neural responses in the occipital and temporal lobes are generally found in individuals with DP, as revealed by both fMRI and ERP measures (e.g. Avidan et al., 2014; Furl, Garrido, Dolan, Driver, & Duchaine, 2011; Towler et al., 2012; Towler et al., 2016), but these could primarily reflect responses to face parts. Previous ERP studies of the face-sensitive N170 component have found atypical cortical responses to face inversion, and scrambling the locations of internal facial features in DP – two manipulations known to abolish holistic face processing (Towler, Parketny, & Eimer, 2016; Towler et al., 2012). This is in line with the current evidence for impaired holistic face processing in DP. Selective impairments in processing the contrast polarity of the eye region in DPs have also demonstrated for the N170 component (Fisher et al., 2016). These observations suggest that while specialised face processing is present in individuals with DP, the face representations that are generated during face perception are atypical in at least two respects. On the one hand, these representations appear to be predominantly part-based rather than holistic. On the other hand, they may be relatively insensitive to identity-related information provided by internal facial features, particularly to contrast-related signals from the eye region.

A recent study has provided a possible mechanistic explanation for holistic face processing deficits in DP (Witthoft et al., 2016). In this study, population receptive field sizes were estimated from functional MRI data for various visual regions in the ventral visual pathway. Individuals with DP had smaller population receptive fields in face-selective and some additional object-selective cortical regions than individuals without face processing impairments. Reduced receptive field sizes may prevent face-selective brain regions from receiving the combined inputs from lower-level visual feature analysers areas across the entire visual field. A prerequisite for generating holistic face representations is that visual information is pooled from a wide area of visual space that covers the spatial extent of the whole face. If individuals with DP have face-selective brain regions with small receptive fields, this requirement may not be met, resulting in face representations that are strongly part-based. We suggest that the super-additive N250r components observed when control participants match upright faces (in both the present study and our previous study, Towler & Eimer, 2016) reflects the activation of identity-sensitive visual neurons with large receptive fields which encompass all internal and external facial features simultaneously. Purely additive N250r components triggered by inverted faces may represent the independent activation of visual neurons with smaller receptive fields that encompass only a subset of these features (e.g., the hairline, eyes, or mouth, which are the most prominent external and internal features in the face images used here). The presence of reliable N250r components on partial repetition trials and fully additive N250r components on full repetition trials in the DP group is therefore in line with the hypothesis that the receptive field sizes of face-selective neurons tuned to face identity are spatially restricted in developmental prosopagnosia.

Overall, the current study has revealed two interrelated aspects in which face identity matching processes differ between DPs and control participants. First, and most importantly, while holistic face representations are clearly involved when control participants match two successively presented face images, this process appears to be much more part-based in individuals with DP, and holistic face representations may not be activated at all. In addition, DPs have a strong tendency to prioritise the encoding of external over internal facial features. If individuals with DP apply this atypical strategy of relying on face parts and specifically on external features when encoding or remembering the identity of individual faces, this will inevitably result in impaired face recognition. External facial features are suboptimal cues for reliably indicating an individual's identity, because they are more variable than internal features, and can change from moment to moment, unlike the relatively invariant identity cues provided by the internal facial features. Individuals with DP often rely on hairstyle to recognise familiar people, and experience severe recognition difficulties when individuals change their hairstyle. This atypical bias for encoding the external facial features for face recognition may be a combined result of an impaired sensitivity to internal facial features and a deficit in constructing holistic perceptual representations of upright faces. In short, we propose that the inability to construct holistic visual representations of the entire face in a single glance is a critical perceptual factor in DP. Face recognition processes are impaired in individuals with DP because they are based primarily on part-based representations. This core deficit is exacerbated by the fact that these part-based representations are strongly biased towards external rather than internal facial features. Understanding the contributions of atypical holistic face processing to the face recognition impairments in DP is not only theoretically important, but also has implications for the remediation of these impairments. For example, a recent training study showed that holistic face training can improve

face recognition ability in a group of individuals with DP (DeGutis et al., 2014). The current results suggest that such training programmes might additionally benefit from taking into account the fact that DPs tend to have a bias to use the external facial features during face identity matching and face recognition.

Competing financial interests

All authors confirm that no competing financial interests apply.

Acknowledgement

This work was supported by a grant (ES/K002457/1) from the Economic and Social Research Council (ESRC), UK.

REFERENCES

Avidan, G., Tanzer, M., & Behrmann, M. (2011). Impaired holistic processing in congenital prosopagnosia. *Neuropsychologia, 49*, 2541–2552. http://10.1016/j.neuropsychologia.2011.05.002.

Avidan, G., Tanzer, M., Hadj-Bouziane, F., Liu, N., Ungerleider, L. G., & Behrmann, M. (2014). Selective dissociation between core and extended regions of the face processing network in congenital prosopagnosia. *Cerebral Cortex, 24*, 1565–1578. https://doi.org/10.1093/cercor/bht007.

Bate, S., & Tree, J. J. (2017). The definition and diagnosis of developmental prosopagnosia. *The Quarterly Journal of Experimental Psychology, 70*, 193–200. https://doi.org/10.1080/17470218.2016.1195414.

Bindemann, M., Burton, A. M., Leuthold, H., & Schweinberger, S. R. (2008). Brain potential correlates of face recognition: Geometric distortions and the N250r brain response to stimulus repetitions. *Psychophysiology, 45*, 535–544. https://doi.org/10.1111/j.1469-8986.2008.00663.x.

Biotti, F., Wu, E., Yang, H., Jiahui, G., Duchaine, B., & Cook, R. (2017). Normal composite face effects in developmental prosopagnosia. *Cortex, 95*, 63–76. http://10.1016/j.cortex.2017.07.018.

Dalrymple, K. A., Oruç, I., Duchaine, B., Pancaroglu, R., Fox, C. J., Iaria, G., et al. (2011). The anatomic basis of the right face-selective N170 IN acquired prosopagnosia: A combined ERP/fMRI study. *Neuropsychologia, 49*(9), 2553–2563. https://doi.org/10.1016/j.neuropsychologia.2011.05.003.

DeGutis, J., Cohan, S., Mercado, R. J., Wilmer, J., & Nakayama, K. (2012). Holistic processing of the mouth but not the eyes in developmental prosopagnosia. *Cognitive Neuropsychology, 29*, 419–446. http://10.1080/02643294.2012.754745.

DeGutis, J., Cohen, S., & Nakayama, K. (2014). Holistic face training enhances face processing in developmental prosopagnosia. *Brain, 137*(6), 1781–1798. https://doi.org/10.1093/brain/awu062.

DeGutis, J., Wilmer, J., Mercado, R. J., & Cohan, S. (2013). Using regression to measure holistic face processing reveals a strong link with face recognition ability. *Cognition, 126*, 87–100. http://10.1016/j.cognition.2012.09.004.

Di Nocera, F., & Ferlazzo, F. (2000). Resampling approach to statistical inference: Bootstrapping from event-related potentials data. *Behavior Research Methods Instruments and Computers, 32*, 111–119. https://doi.org/10.3758/BF03200793.

Duchaine, B., & Nakayama, K. (2005). Dissociations of face and object recognition in developmental prosopagnosia. *Journal of Cognitive Neuroscience, 17*, 249–261. http://doi.org/10.1162/0898929053124857.

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia, 44*, 576–585. https://doi.org/10.1016/j.neuropsychologia.2005.07.001.

Duchaine, B., Yovel, G., & Nakayama, K. (2007). No global processing deficit in the Navon task in 14 developmental prosopagnosics. *Social Cognitive and Affective Neuroscience, 2*, 104–113 [electronic Resource] http://10.1093/scan/nsm003.

Fisher, K., Towler, J., & Eimer, M. (2016). Reduced sensitivity to contrast signals from the eye region in developmental prosopagnosia. *Cortex, 81*, 64–78. https://doi.org/10.1016/j.cortex.2016.04.005.

Fisher, K., Towler, J., & Eimer, M. (2017). Face identity matching is selectively impaired in developmental prosopagnosia. *Cortex, 89*, 11–27. http://10.1016/j.cortex.2017.01.003.

Furl, N., Garrido, L., Dolan, R. J., Driver, J., & Duchaine, B. (2011). Fusiform gyrus face selectivity relates to individual differences in facial recognition ability. *Journal of Cognitive Neuroscience, 23*, 1723–1740. https://doi.org/10.1162/jocn.2010.21545.

Jacques, C., & Rossion, B. (2009). The initial representation of individual faces in the right occipito-temporal cortex is holistic: Electrophysiological evidence from the composite face illusion. *Journal of Vision, 9*, 8-8 https://doi.org/10.1167/9.6.8.

Kasinski, A., Florek, A., & Schmidt, A. (2008). The PUT face database. *Image Processing and Communications, 13*, 59–64.

Le Grand, R., Cooper, P. A., Mondloch, C. J., Lewis, T. L., Sagiv, N., de Gelder, B., et al. (2006). What aspects of face processing are impaired in developmental prosopagnosia? *Brain and Cognition, 61*, 139–158. https://doi.org/10.1016/j.bandc.2005.11.005.

Rezlescu, C., Susilo, T., Wilmer, J. B., & Caramazza, A. (2017). The inversion, part-whole, and composite effects reflect distinct perceptual mechanisms with varied relationships to face recognition. *Journal of Experimental Psychology Human Perception and Performance, 43*, 1961–1973.

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011). Holistic processing predicts face recognition. *Psychological Science, 22*, 464–471. http://doi.org/10.1177/0956797611401753.

Rossion, B. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition, 21*, 139–253. https://doi.org/10.1080/13506285.2013.772929.

Schwarzer, G., Huber, S., Grüter, M., Grüter, T., Groß, C., Hipfel, M., et al. (2007). *Psychological Research, 71*, 583. https://doi.org/10.1007/s00426-006-0068-0.

Schweinberger, S. R. (2011). Neurophysiological correlates of face recognition. In A. Calder, G. Rhodes, M. Johnson, & J. Haxby (Eds.), *The Oxford handbook of face perception* (pp. 345–366). New York: Oxford University Press.

Schweinberger, S. R., Huddy, V., & Burton, A. M. (2004). N250r: A face-selective brain response to stimulus repetitions. *NeuroReport, 15*, 1501–1505. https://doi.org/10.1097/01.wnr.0000131675.00319.42.

Schweinberger, S. R., Pickering, E. C., Jentzsch, I., Burton, A. M., & Kaufmann, J. M. (2002). Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Cognitive Brain Research, 14*, 398–409. https://doi.org/10.1016/S0926-6410(02)00142-8.

Susilo, T., & Duchaine, B. (2013). Advances in developmental prosopagnosia research. *Current Opinion in Neurobiology, 23*, 423–429. https://doi.org/10.1016/j.conb.2012.12.011.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology Section A, 46*, 225–245. https://doi.org/10.1080/14640749308401045.

Towler, J., & Eimer, M. (2012). Electrophysiological studies of face processing in developmental prosopagnosia: Neuropsychological and neurodevelopmental perspectives. *Cognitive Neuropsychology, 29*, 503–529. https://doi.org/10.1080/02643294.2012.716757.

Towler, J., & Eimer, M. (2016). Electrophysiological evidence for parts and wholes in visual face memory. *Cortex, 83*, 246–258. https://doi.org/10.1016/j.cortex.2016.07.022.

Towler, J., Fisher, K., & Eimer, M. (2017). The cognitive and neural basis of developmental prosopagnosia. *The Quarterly Journal of Experimental Psychology, 70*, 316–344. https://doi.org/10.1080/17470218.2016.1165263.

Towler, J., Gosling, A., Duchaine, B., & Eimer, M. (2012). The face sensitive N170 component in developmental prosopagnosia. *Neuropsychologia, 50*, 3588–3599. https://doi.org/10.1016/j.neuropsychologia.2012.10.017.

Towler, J., Gosling, A., Duchaine, B., & Eimer, M. (2016). Normal perception of Mooney faces in developmental prosopagnosia: Evidence from the N170 component and rapid neural adaptation. *Journal of Neuropsychology, 10*, 15–32. https://doi.org/10.1111/jnp.12054.

Towler, J., Kelly, M., & Eimer, M. (2016). The focus of spatial attention determines the number and precision of face representations in working memory. *Cerebral Cortex, 26*, 2530–2540. https://doi.org/10.1093/cercor/bhv083.

Towler, J., Parketny, J., & Eimer, M. (2016). Perceptual face processing in developmental prosopagnosia is not sensitive to the canonical location of face parts. *Cortex, 74*, 53–66. https://doi.org/10.1016/j.cortex.2015.10.018.

Ulrich, P. I. N., Wilkinson, D. T., Ferguson, H. J., Smith, L. J., Bindemann, M., Johnston, R. A., et al. (2017). Perceptual and memorial contributions to developmental prosopagnosia. *The Quarterly Journal of Experimental Psychology, 70*, 298–315. https://doi.org/10.1080/17470218.2016.1177101.

White, D., Rivolta, D., Burton, A. M., Al-Janabi, S., & Palermo, R. (2017). Face matching impairment in developmental prosopagnosia. *The Quarterly Journal of Experimental Psychology, 70*, 287–297. https://doi.org/10.1080/17470218.2016.1173076.

Wirth, B. E., Fisher, K., Towler, J., & Eimer, M. (2015). Facial misidentifications arise from the erroneous activation of visual face memory. *Neuropsychologia, 77*, 387–399. https://doi.org/10.1016/j.neuropsychologia.2015.09.021.

Witthoft, N., Poltoratski, S., Nguyen, M., Golarai, G., Liberman, A., LaRocque, K. F., et al. (2016). Reduced spatial integration in the ventral visual cortex underlies face recognition deficits in developmental prosopagnosia. *BioRxiv*, 051102. https://doi.org/10.1101/051102.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*, 747–759.

Yovel, G., & Duchaine, B. (2006). Specialized face perception mechanisms extract both part and spacing information: Evidence from developmental prosopagnosia. *Journal of Cognitive Neuroscience, 18*(4), 580–593. https://doi.org/10.1162/jocn.2006.18.4.580.

Zimmermann, F. G. S., & Eimer, M. (2013). Face learning and the emergence of view-independent face recognition: An event-related brain potential study. *Neuropsychologia, 51*, 1320–1329. https://doi.org/10.1016/j.neuropsychologia.2013.03.028.

Zimmermann, F. G., & Eimer, M. (2014). The activation of visual memory for facial identity is task-dependent: Evidence from human electrophysiology. *Cortex, 54*, 124–134. https://doi.org/10.1016/j.cortex.2014.02.008.