Neuropsychologia 124 (2019) 285–298

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Is developmental prosopagnosia best characterised as an apperceptive or mnemonic condition?

Federica Biotti[a,*], Katie L.H. Gray[b], Richard Cook[c]

[a] Department of Psychology, City, University of London, London, UK
[b] School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK
[c] Department of Psychological Sciences, Birkbeck, University of London, London, UK

ARTICLE INFO

Keywords:
Developmental prosopagnosia
Face memory
Apperceptive deficit
Face recognition
Object recognition

ABSTRACT

Traditionally, developmental prosopagnosia (DP) has been thought of as an apperceptive condition that hinders individuals' ability to encode face structure. However, several authors have recently raised the possibility that many DPs may be able to form accurate percepts, but be unable to maintain those percepts over time. The present study sought to distinguish these possibilities. In our first experiment 16 DPs and 22 typical controls completed a delayed match-to-sample task with face and car stimuli, with a retention interval of 1-second (low demand) or 6-seconds (high demand). As expected, the participants with DP were worse than the controls at face matching, and were disproportionately impaired at matching faces relative to cars. However, the relative degree of impairment seen in the DPs did not interact with retention interval; they exhibited similar levels of impairment when matching faces with 1- and 6-second delays. Next, we compared the performance of 72 DPs and 54 typical controls on the Cambridge Face Perception Test (CFPT), a task that measures face perception ability in a way that minimises the memory demands. As expected, we found that the DPs were impaired at the group level. This difference was not attributable to a few individuals with an apperceptive profile; rather we found evidence that the distribution of CFPT scores seen in the DP sample was shifted relative to that of typical controls. Some heterogeneity is likely in any neurodevelopmental population, and DP is no different. Generally, however, these findings suggest that selective STFM impairment may be relatively uncommon in this population. Instead, deficits of perceptual encoding may play a larger role in DP than currently acknowledged.

1. Introduction

Developmental prosopagnosia[1] (DP) is a neurodevelopmental condition associated with difficulties recognising familiar faces and distinguishing unfamiliar faces, that occurs in people with normal intelligence and typical visual acuity, and in the absence of manifest brain injury (Behrmann and Avidan, 2005; Duchaine and Nakayama, 2006b; Susilo and Duchaine, 2013). Historically, the condition was thought to be rare (McConachie, 1976), but current estimates suggest that 2% of the general population may experience face recognition difficulties severe enough to disrupt their daily lives (Kennerknecht et al., 2006, 2008). The fact that DP often runs in families suggests the condition has a genetic component (Duchaine et al., 2007; Johnen et al., 2014; Schmalzl et al., 2008), a finding that accords with the broader view that face recognition ability is a heritable trait (Shakeshaft and Plomin, 2015; Wilmer et al., 2010; Zhu et al., 2010). At the neural level, studies suggest that DP is associated with reduced structural (Gomez et al., 2015; Song et al., 2015; Thomas et al., 2009) and functional (Avidan and Behrmann, 2009; Lohse et al., 2016; Rosenthal et al., 2017) connectivity within the occipito-temporal face processing network. Due to their characteristic deficits, DPs often rely on non-facial cues like voice, hairstyle, and walking gait to recognise familiar others (Cook and Biotti, 2016; Shah et al., 2015b).

1.1. Apperceptive characterisation

Traditionally DP has been thought of as an apperceptive form of prosopagnosia (De Renzi et al., 1991); a condition with a perceptual origin that hinders individuals' ability to encode the structure of faces (Behrmann and Avidan, 2005; Duchaine and Nakayama, 2006b; Susilo

* Corresponding author.
E-mail address: Federica.Biotti@city.ac.uk (F. Biotti).

[1] We use the term *developmental prosopagnosia* instead of *congenital prosopagnosia* to indicate the possibility that in some cases the disorder may appear during development and not necessarily from birth.

https://doi.org/10.1016/j.neuropsychologia.2018.11.014
Received 24 April 2018; Received in revised form 2 November 2018; Accepted 26 November 2018
Available online 28 November 2018
0028-3932/

and Duchaine, 2013). Consistent with this view, many DPs exhibit difficulties distinguishing unfamiliar faces presented simultaneously (Avidan, Tanzer, and Behrmann, 2011; Biotti and Cook, 2016; Biotti et al., 2017a; Duchaine et al., 2007; Shah et al., 2015a; White et al., 2017) or sequentially, either side of sub-second interval (Duchaine et al., 2006; Fisher et al., 2017; Le Grand et al., 2006; Yovel and Duchaine, 2006). In addition to problems matching or recognising facial identities, many DPs appear to have problems recognising facial emotion (Biotti and Cook, 2016; Burns et al., 2017; Duchaine et al., 2006), facial age (Ariel and Sadeh, 1996), and facial gender (Ariel and Sadeh, 1996; Esins et al., 2016). Moreover, electrophysiological markers thought to index early face encoding (e.g. the N170 ERP component) are often atypical in cases of DP (Fisher et al., 2016; Towler et al., 2017, 2012, 2016). This profile of deficits is consistent with a locus of impairment early in the face processing stream, before the processing of identity and other facial attributes bifurcates (Bruce and Young, 1986; De Renzi et al., 1991; Haxby et al., 2000).

According to one influential apperceptive account, a failure to process faces holistically – whereby facial features are integrated into a non-decomposable whole (Farah et al., 1998; McKone and Yovel, 2009; Piepers and Robbins, 2013) – may underlie the face recognition difficulties seen in DP (Avidan et al., 2011; DeGutis et al., 2012, 2014b; Liu and Behrmann, 2014; Palermo et al., 2011). Consistent with this view, individuals with DP are thought to be less sensitive to facial orientation (Duchaine et al., 2006; Shah et al., 2015a; Tree and Wilkie, 2010), and sometimes have problems distinguishing faces using feature configurations (Le Grand et al., 2006; Yovel and Duchaine, 2006). It has also been argued that some DPs show reduced susceptibility to visual illusions thought to index holistic face processing, including the part-whole (DeGutis et al., 2012) and composite face effects (Avidan et al., 2011; Liu and Behrmann, 2014; Palermo et al., 2011). Where observed, aberrant processing of configurations may extend to non-face stimuli (Avidan et al., 2011).

1.2. A deficit of perceptual encoding or perceptual maintenance?

The case for an apperceptive characterisation of DP is not as strong as it first appears. Several findings suggest that some DPs may encode face structure typically; for example, some individuals with DP exhibit broadly typical discrimination of unfamiliar faces presented simultaneously (Bowles et al., 2009; Dalrymple et al., 2014; McKone et al., 2011; Ulrich et al., 2017), and apparently normal recognition of facial emotion (Dobel et al., 2007; Humphreys et al., 2007; Lee et al., 2010; Palermo et al., 2011), facial age and facial gender (Chatterjee and Nakayama, 2013; DeGutis et al., 2014a). Many DPs also exhibit typical susceptibility to visual illusions thought to arise from the holistic encoding of facial structure, in particular the composite face effect (Biotti et al., 2017b; Le Grand et al., 2006; Susilo et al., 2010; Ulrich et al., 2017). Notably, Biotti et al. (2017b) recently described two group studies – using independent samples of 16 and 24 DPs – neither of which found evidence of reduced composite effects. These behavioural results indicate that early structural encoding may be intact in many cases of DP.

Rather than characterise DP as an apperceptive condition, several authors have raised the possibility that in many cases, DP may be caused by impaired short-term face memory (STFM); that many DPs may be able to form accurate percepts, but be unable to maintain those percepts over time (Dalrymple et al., 2014; Dalrymple and Palermo, 2016; Jackson et al., 2017; Stollhoff et al., 2011; Ulrich et al., 2017). A similar possibility has been suggested in the literature on autism spectrum disorder (ASD), where a systematic review concluded that a delay of a few seconds between the presentation of the target and test faces disproportionately impairs matching or recognition performance in this population (Weigelt et al., 2012). While the suggestion that faces may benefit from domain-specific memory processing is relatively new, the implied dissociation between perceptual processes responsible for face encoding, and memory processes responsible for maintaining face representations, is consistent with evidence that face memory follows a different developmental trajectory relative to perceptual memory for other objects (Weigelt et al., 2013).

Consistent with the possibility that DP may be caused by aberrant STFM, many cases of DP have been described (Bowles et al., 2009; Dalrymple et al., 2014; McKone et al., 2011; Ulrich et al., 2017) who exhibit impaired performance on diagnostic tests with a memory component such as the Cambridge Face Memory Test (CFMT; Duchaine and Nakayama, 2006a), but perform within the typical range on tests with a minimal memory component such as the Cambridge Face Perception Test (CFPT; Duchaine et al., 2007). When DPs are required to retain faces in memory for brief periods, functional magnetic resonance imaging (fMRI) reveals wider activation in prefrontal regions implicated in working memory, relative to typically developed (TD) controls (Avidan et al., 2005), suggesting that percept retention may be effortful. Similarly, where observed, neural differences in DP are sometimes more pronounced in anterior (extended) regions of the face processing network, than in posterior (core) areas thought to be responsible for early structural encoding (Avidan et al., 2014).

1.3. Present study

In the present study we investigated whether DP is best characterised as i) a disorder of STFM, where these individuals initially form accurate perceptual descriptions of faces, but struggle to maintain these representations over time; or ii) as an apperceptive condition, where face recognition difficulties arise from poor encoding of face structure. In our first experiment, we compared the face-matching ability of 16 DPs following 1- and 6-second retention intervals. Contrary to the predictions of the impaired STFM hypothesis, we find that DPs show similar levels of matching impairment relative to controls at short and long retention intervals. Next, we examined the performance of a large sample of DPs (N = 72) on the CFPT, a task that measures face perception in a way that minimises participants' memory load and is therefore thought to index structural encoding ability. Consistent with an apperceptive characterisation, we find that DPs perform poorly on the CFPT at the group-level and show signs of a shifted distribution.

2. Do face matching deficits seen in DP increase as a function of retention interval?

It has been proposed that many DPs exhibit intact encoding of face structure, but experience deficits of STFM (Dalrymple et al., 2014; Dalrymple and Palermo, 2016; Jackson et al., 2017; Stollhoff et al., 2011; Ulrich et al., 2017). Experiments that directly test whether the face matching and face recognition deficits seen in DP are sensitive to memory load are therefore particularly important. For this reason, we sought to revisit a finding described by Shah et al. (2015a). This previous study utilised a delayed match-to-sample task whereby participants were required to identify a target stimulus from a test display of four items (target plus three lures). Memory demands were manipulated by varying the delay between the presentation of the target and the test array. This approach is useful as it allows systematic manipulation of the memory component of the task, but ensures the perceptual demands – associated with the encoding of target and test items – are held constant (Shah et al., 2015a). If DP is associated with impaired STFM, disproportionate impairment should be seen after longer retention intervals, relative to shorter retention intervals. Contrary to this prediction, however, Shah and colleagues found that their DP sample (N = 15) exhibited comparable deficits at short (2-second) and long (8-second) intervals.

In the original study described by Shah et al. (2015a) the same images were used to present items in the study and test phases. Consequently, targets were always seen from the same frontal viewpoint. In the present study, we examined observers' ability to match items viewed from the same frontal perspective (constant-viewpoint matching), and across a viewpoint disparity of 45° (different-viewpoint matching). While constant- and different-viewpoint matching appear similar, they may differ substantially in their perceptual and mnemonic demands. First, observers sometimes match unfamiliar faces using superficial pictorial cues (Hancock et al., 2000; Megreya and Burton, 2006). Because rotation introduces substantial disparity between target and test images, different-viewpoint matching is less susceptible to this strategy than constant-viewpoint matching (Longmore et al., 2008). Instead, different-viewpoint matching is thought to tax observers' ability to form and maintain a view-invariant structural description (Bruce and Young, 1986; Marr and Nishihara, 1978). Second, a particular type of short-term memory – visual working memory (Baddeley, 1992, 1993, 2010) – has been hypothesised that supports the rotation and manipulation of percepts. While constant- and different-viewpoint face matching both tap some short-term memory processes, different-viewpoint matching places greater demands on visual working memory. In light of their different mnemonic demands, these two tasks may behave differently as a function of retention interval, and be differentially affected in DP.

On half the trials, we used a retention interval of 1-second (low demand); on half the trials, we used a retention interval of 6-seconds (high demand). The short interval used in this study (1-second) is shorter than that employed previously (2-seconds; Shah et al., 2015a), thereby reducing further the memory demands in the low demand condition. We recognise, however, that the retention of percepts for 1-second still represents a memory demand. Crucially, our aim in the short interval condition was to minimise, not to eliminate, the memory demands of the matching task.[2] Participants' face matching ability was compared to that seen with cars to determine if deficits, where observed, were face-specific, or whether they extended to a non-face object category.

2.1. Methods

2.1.1. Participants

Sixteen adults with DP (6 males; M_{age} = 41.50 years, SD_{age} = 12.58 years) and 22 TD controls (9 males; M_{age} = 38.23 years, SD_{age} = 13.39 years) completed the experiment. None of the DPs were included in the sample described by Shah et al. (2015a). Neither participant age [$t(36) = 0.763$, $p = .451$] nor proportion of males [$X^2(1) = 0.045$, $p = .551$] differed significantly between the two groups. Ethical approval was granted by the local ethics committee. The study was conducted in line with the ethical guidelines provided by the 6th (2008) Declaration of Helsinki. All participants provided informed consent and were fully debriefed after the experimental procedure (i.e., the aims and rationale of the study were explained).

2.1.2. Diagnostic testing

DP participants were recruited through www.troublewithfaces.org and reported lifelong face recognition difficulties in the absence of brain injury or psychiatric disorder (e.g., ASD, schizophrenia). Diagnostic decisions were based primarily on participants' scores on the Twenty-Item Prosopagnosia Index (PI20; Gray et al., 2017; Shah et al., 2015b) and the CFMT (Duchaine and Nakayama, 2006a). Diagnostic information for each DP is provided in Table 1. The development of standardised diagnostic criteria for DP still appears some way off (Barton and Corrow, 2016; Dalrymple and Palermo, 2016; Shah et al., 2015b). However, the use of convergent self-report evidence and scores on objective, computer-based tasks may be a particularly effective approach to the identification and classification of DP; for example, less than 1.5% of the general population score below 65% on the CFMT *and* more than 65 on the PI20 (see Gray et al., 2017).

As expected, the TD controls (M_{CFMT} = 85.1%, SD_{CFMT} = 10.2%; M_{PI20} = 39.0, SD_{PI20} = 9.0) differed significantly from the DPs (M_{CFMT} = 55.9%, SD_{CFMT} = 7.9%; M_{PI20} = 79.9, SD_{PI20} = 8.1) in their CFMT [$t(36) = 9.574$, $p < .001$] and PI20 [$t(36) = 14.390$, $p < .001$] scores. In addition to the CFMT and the PI20, all participants also completed the CFPT to measure their face encoding ability, and the Cambridge Car Memory Test (CCMT; Dennett et al., 2011) to measure their non-face object recognition ability. The TD controls (M = 26.27, SD = 9.61) made fewer errors than the DPs (M = 50.38, SD = 17.33) on the upright condition of the CFPT [$t(36) = 5.483$, $p < .001$]. The TD controls (M = 73.9%, SD = 12.8%) and the DPs (M = 63.5%, SD = 8.4%) also differed significantly in terms of their performance on the CCMT [$t(35) = 2.837$, $p = .008$]. All participants were also screened for colour blindness using Ishihara's Tests for Colour-Blindness (Ishihara, 1993).

2.1.3. Stimuli and procedure

Each category (faces, cars) comprised 50 exemplars. Both categories were further organised into 5 subsets of 10 exemplars based on approximate similarity. Cars were sorted into subsets based on their size and class (e.g. Saloons/Sedans/SUVs). Faces were sorted based on aspect-ratio, pigmentation, and eye-brow colour. Each exemplar was depicted twice: once in frontal view, once in 3/4 view. When viewed at 57 cm, the faces subtended 5° of visual angle vertically; the cars subtended 3° vertically. Face stimuli (male Caucasian faces) were created using FaceGen Modeller Version 3.3 (Singular Inversions Inc.). Car stimuli were generated through www.3dtuning.com.

The structure of the delayed matching task is shown in Fig. 1. Each trial started with a fixation point (750 ms) on a blank screen. A single target stimulus was then presented centrally for 400 ms. Targets were always shown in frontal view. A given facial identity or car model could appear as a target only once in each viewing condition. In all other respects, the choice of target was randomly determined by the experimental program. The offset of the target was followed by a retention interval during which a mask image was presented. The mask was constructed by recombining regions cropped from other target images from the same category. An array of four test items followed the retention interval. The array comprised the target and three lures selected at random from the same within-category subset. On half of the trials, test stimuli were presented in frontal view (here, the target and test stimuli were shown from the same viewpoint). On the remaining trials, test stimuli were presented in 3/4 view (here, the target and test stimuli were shown from different viewpoints). Test arrays were visible until a keypress response was registered. Participants were asked to respond with speed and accuracy.

The factorial combination of Stimulus Type (faces, cars), Retention Interval (short, long), and Viewpoint (frontal, 3/4) yielded eight types of trial, which were randomly interleaved. There were 20 trials of each type, yielding 160 trials in total. Given the large number of face and car stimuli required by the procedure it was necessary to recycle stimuli from each pool of 50 items. Some stimulus items therefore appeared multiple times across the procedure, either as targets or lures. Six practice trials preceded the experiment. No feedback was provided during the procedure. The task lasted approximately 45 minutes and included three short breaks. The task was programmed in MATLAB using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997).

[2] The key strength of this paradigm is that it allows the manipulation of memory demands in a way that leaves the perceptual demands of the task unaltered. Having a no interval condition (i.e., where the target is presented alongside the array of 4 test items) would have violated this logic. Although presenting the 5 faces simultaneously would further reduce the memory demands, this reduction would be confounded with an increase in perceptual and attentional demands.

Table 1
Scores of each developmental prosopagnosic (DP) on the 20 Item Prosopagnosia Index (PI20), the Cambridge Face Memory Test (CFMT), and The Cambridge Face Perception Test (CFPT). The z-scores provided for the CFPT are based on performance in the upright condition.

Participant	Age	PI20	CFMT %	CFPT Upright (Errors)	CFPT inverted (Errors)	PI20 z-scores	CFMT z-scores	CFPT z- scores
F1	21	59	63.89	30	64	− 2.31	− 2.36	− 0.06
F2	48	85	63.89	60	70	− 5.17	− 2.36	− 3.27
F3	22	80	40.28	66	98	− 4.62	− 5.01	− 3.91
F4	66	79	61.11	40	70	− 4.51	− 2.68	− 1.13
F5	48	78	58.33	26	64	− 4.40	− 2.99	0.36
F6	25	70	44.44	74	68	− 3.52	− 4.54	− 4.77
F7	38	73	59.72	48	64	− 3.85	− 2.83	− 1.99
F8	40	85	59.72	34	75	− 5.17	− 2.83	− 0.49
F9	53	84	52.78	62	58	− 5.06	− 3.61	− 3.49
F10	38	90	50.00	66	80	− 5.72	− 3.92	− 3.91
M1	30	85	59.72	48	84	− 5.17	− 2.83	− 1.99
M2	47	77	68.06	42	72	− 4.29	− 1.90	− 1.35
M3	52	94	55.56	88	64	− 6.17	− 3.30	− 6.27
M4	55	80	50.00	44	84	− 4.62	− 3.92	− 1.56
M5	41	80	61.11	44	70	− 4.62	− 2.68	− 1.56
M6	40	79	45.83	34	76	− 4.51	− 4.39	− 0.49
DP mean	41.5	79.87	55.90	50.37	72.56			
DP SD	12.6	8.13	7.89	17.33	0.09			
TD mean	39.2	37.96	84.98	29.41	63.37			
TD SD	13.4	9.09	8.92	9.35	15.74			

Note. The prosopagnosics' scores on the diagnostic procedures were compared with the group of 54 controls described in Experiment 2 (23 males). All but one of the DPs scored at least two standard deviations below the comparison average on the PI20 and the CFMT. The case for including this individual (M2) in our DP sample was bolstered by his poor score (< 3 SDs below the mean) on a UK variant of the Famous Face Recognition Task.

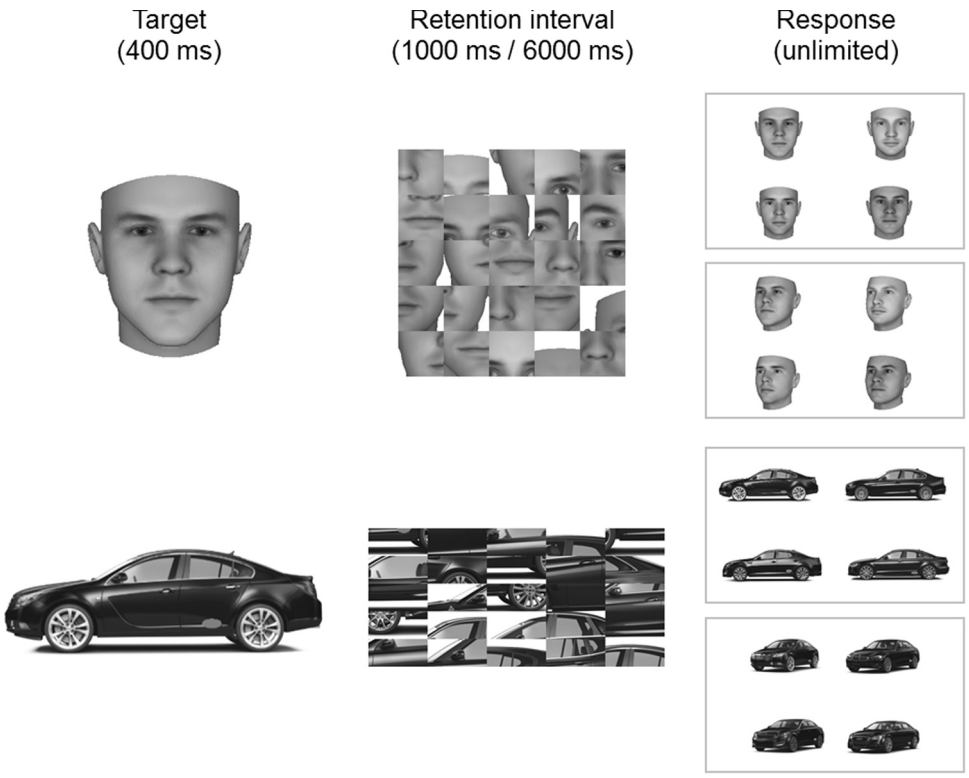

Fig. 1. Illustration of the stimuli and procedure employed in our delayed matching task.

2.2. Results and discussion

2.2.1. Group analyses

Matching accuracy (Fig. 2a) was analysed using ANOVA with Stimulus Type (faces, cars), Retention Interval (short, long), and Viewpoint (frontal, 3/4) as within-subjects factors, and Group (DP, TD) as a between-subjects factor. The analysis revealed significant main effects of Viewpoint [$F(1,36) = 52.44$, $p < .001$, $\eta_p^2 = .59$] and Retention Interval [$F(1,36) = 48.41$, $p < .001$, $\eta_p^2 = .57$], whereby a change of viewpoint and a longer retention interval were associated with poorer matching accuracy, respectively. However, there was no main effect of Stimulus Type [$F(1,36) = 3.16$, $p = .084$, $\eta_p^2 = .081$], nor did we see a Retention Interval × Viewpoint interaction [$F(1,36) = 0.001$, $p = .982$, $\eta_p^2 < .001$]. We observed a significant main effect of Group [$F(1,36) = 10.35$, $p = .003$, $\eta_p^2 = .22$], but this was qualified by a significant Group × Stimulus Type interaction [$F(1,36) = 6.11$, $p = .018$, $\eta_p^2 = .145$]. When matching accuracy for cars and faces was analysed in separate ANOVAs, we observed a significant effect of Group

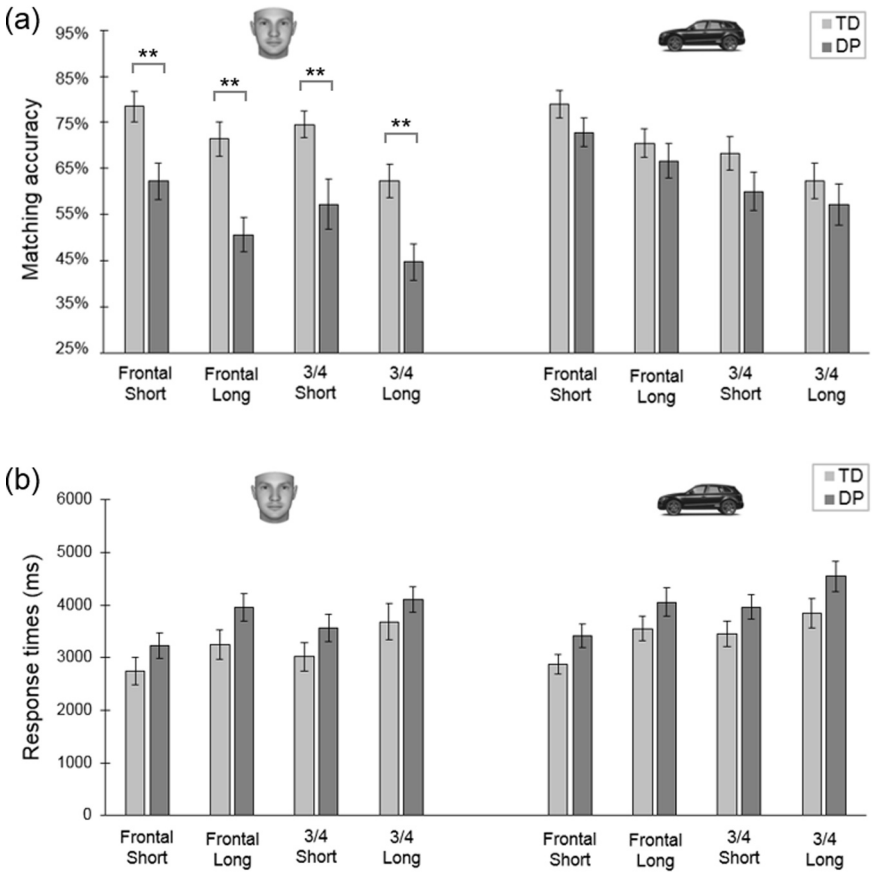

Fig. 2. Mean (a) accuracy and (b) response times for the two groups on the delayed-matching task. Performance is broken down by Viewpoint (frontal, 3/4) and Retention Interval (short, long). Simple contrasts were non-significant unless otherwise indicated. ***$p < .001$; **$p < .01$; *$p < .05$. Error bars denote ± 1SEM.

for face trials [$F(1,36) = 15.072$, $p < .001$, $\eta_p^2 = .295$], but not for car trials [$F(1,36) = 1.869$, $p = .180$, $\eta_p^2 = .049$]. Crucially, no further interactions with Group were seen on the face (all $Fs < 0.45$, $ps > 0.50$) or car trials (all $Fs < 0.60$, $ps > 0.45$).

To evaluate the effects of the two within-subjects manipulations we computed measures expressing each observer's viewpoint effect (same-viewpoint accuracy – different-viewpoint accuracy) and their retention interval effect (short-interval accuracy – long-interval accuracy). The retention interval effects of the TDs ($M = 9.7\%$, $SD = 12.1\%$) and the DPs ($M = 12.0\%$, $SD = 8.7\%$) did not differ [$t(36) = 0.689$, $p = .495$] and all DPs exhibited retention interval effects within 2 SDs of the typical mean. Similarly, the viewpoint effects of the TDs ($M = 6.5\%$, $SD = 8.5\%$) and DPs ($M = 5.5\%$, $SD = 11.1\%$) did not differ [$t(36) = 0.319$, $p = .752$] and all DPs exhibited viewpoint effects within 2 *SDs* of the typical mean.

We also analysed participants' response times (Fig. 2b) using ANOVA with Stimulus Type (faces, cars), Retention Interval (short, long), and Viewpoint (frontal, 3/4) as within-subjects factors, and Group (DP, TD) as a between-subjects factor. The analysis revealed main effects of Stimulus Type [$F(1,36) = 5.33$, $p = .027$, $\eta_p^2 = .129$], Viewpoint [$F(1,36) = 52.48$, $p < .001$, $\eta_p^2 = .593$], and Retention Interval [$F(1,36) = 98.94$, $p < .001$, $\eta_p^2 = .733$]. Overall, participants responded faster on face trials than on car trials, were faster when identifying frontal views of targets than 3/4 views, and were faster following short retention intervals than long retention intervals. The analysis revealed no main effect of Group [$F(1,36) = 2.65$, $p = .112$, $\eta_p^2 = .069$], nor a Group × Stimulus Type interaction [$F(1,36) = 0.01$, $p = .931$, $\eta_p^2 < .001$]. No further interactions with Group were seen on the face (all $Fs < 0.75$, $ps > 0.39$) or car trials (all $Fs < 0.90$, $ps > 0.35$). When analysed in separate ANOVAs, the response times of the DPs and the TD controls did not differ significantly on either face [$F(1,36) = 2.012$, $p = .165$, $\eta_p^2 = .053$] or car trials [$F(1,36) = 2.845$, $p = .100$, $\eta_p^2 = .073$].

In both the accuracy and response time analyses, Group failed to interact significantly with either Retention Interval or Viewpoint. In order to evaluate the strength of evidence provided by these null results, we subjected these interaction effects to Bayesian analysis in JASP-Team (2018) with default prior width. Analysis of the Group × Retention Interval interaction seen in the accuracy data indicated that the observed results were 2.64 times more likely to occur under the null model, than under an alternative. The observed Group × Viewpoint interaction was 3.02 times more likely to occur under the null model, than under an alternative. Analysis of the Group × Retention Interval interaction seen in the response time data indicated that the observed results were 3.05 times more likely to occur under the null model, than under an alternative. The observed Group × Viewpoint interaction was 2.90 times more likely to occur under the null model, than under an alternative.

2.2.2. Correlational analyses

The group analyses described above reveal comparable deficits at short and long retention intervals, replicating the findings of Shah et al. (2015a). The insensitivity of the DP deficit to retention interval suggests that poor perceptual encoding – not aberrant STFM – may be responsible for the face recognition problems seen in this population. If this view is correct, performance in our matching task should correlate with participants' scores on the CFPT – a measure of face encoding ability. Consistent with this prediction, overall face matching accuracy (i.e., collapsing across viewpoint and interval conditions) correlated closely with performance on the CFPT, [$r = -0.743$, $p < .001$, $CI_{95\%}$: -0.593 to -0.857] (Fig. 3a). Highly significant correlations were seen between CFPT scores and face matching accuracy in all conditions

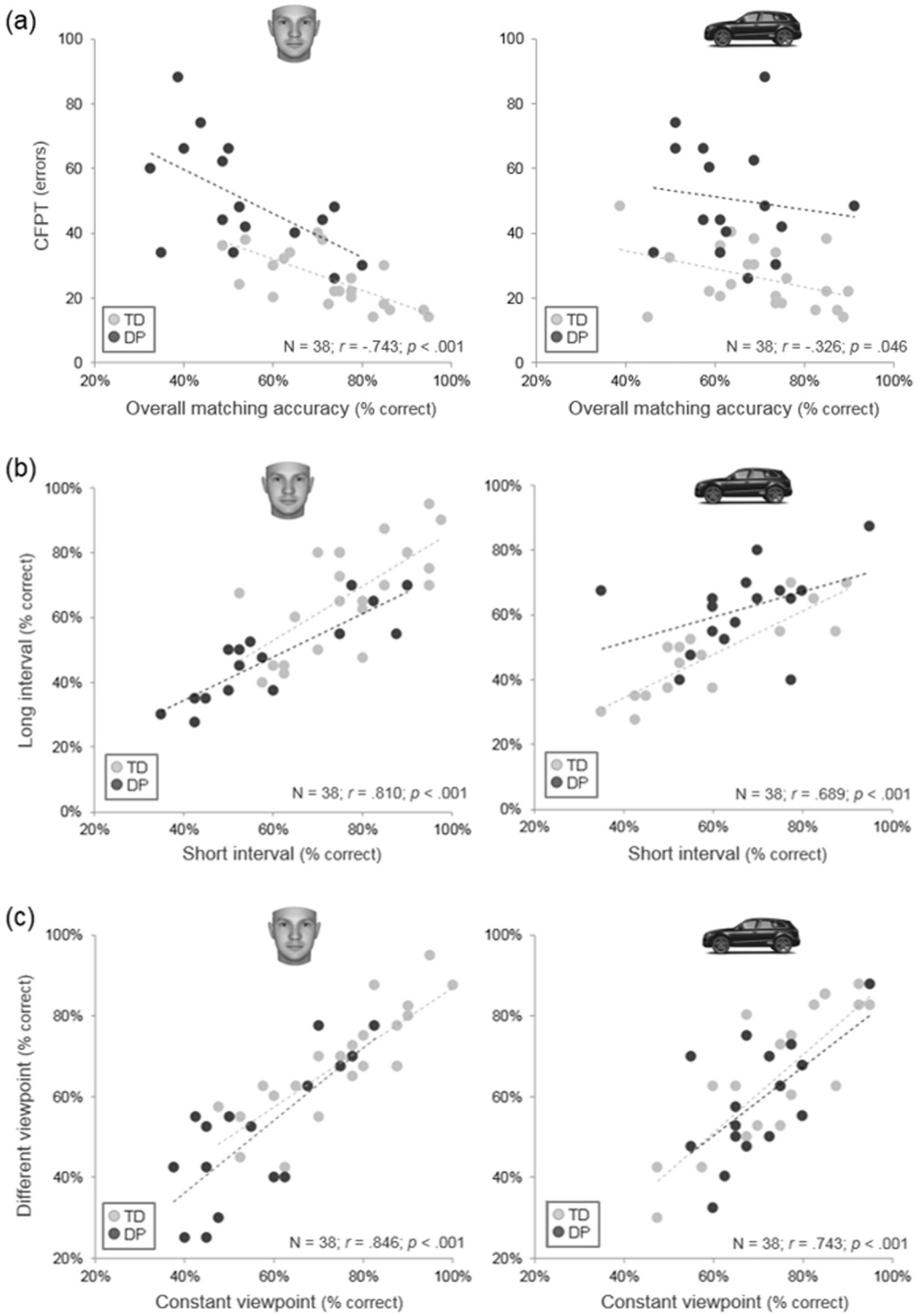

Fig. 3. (a) Scatterplots depicting the relationship between observers' CFPT scores and their face (left) and car (right) matching ability. (b) The relationship between long- and short-interval matching accuracy for faces (left) and cars (right). (c) The relationship between constant- and different-viewpoint matching for faces (left) and cars (right). In each case, the line of best-fit was modelled separately for the TD and DP groups.

(Table 2). In the combined sample, CFPT scores also correlated with overall car matching accuracy [r = -0.326, p = .046, $CI_{95\%}$: -0.025 to -0.621], however this correlation was significantly weaker than that seen between the CFPT and face matching [z = 2.59, p < .001]. As expected, observers' matching accuracy at short intervals correlated closely with their performance at longer intervals for both faces [r = 0.810, p < .001, $CI_{95\%}$: 0.669–0.900] and cars [r = 0.689, p < .001, $CI_{95\%}$: 0.438–0.879] (Fig. 3b).

The group analyses also indicate that, relative to controls, DPs showed similar levels of impairment in the constant- and different-viewpoint conditions. This finding suggests that observers may be using the same perceptual strategy to achieve both types of matching. Consistent with this possibility, we found that observers' different-viewpoint face matching ability correlated closely with their constant-viewpoint face matching ability [r = 0.846, p < .001, $CI_{95\%}$: 0.763–0.906]. A similar relationship was seen for cars [r = 0.743,

Table 2
Correlations seen between participants' scores on the $CFPT_{upright}$ and their accuracy and response time (RT) performance in each of the matching conditions. 95% confidence intervals are shown in brackets. ***$p < .001$; **$p < .01$; *$p < .05$.

	Same view Short interval	Same view Long interval	Different view Short interval	Different view Long interval
Face accuracy	− 0.696*** (−0.515: −0.819)	− 0.703*** (−0.566: −0.825)	− 0.613*** (−0.380: −0.771)	− 0.621*** (−0.420: −0.794)
Face RTs	0.347* (0.059: 0.624)	0.411** (0.171: 0.656)	0.352* (0.108: 0.598)	0.273 (0.011: 0.524)
Car accuracy	− 0.337* (−0.024: −0.614)	− 0.286 (0.039: −0.585)	− 0.334* (−0.030: −0.618)	− 0.143 (0.137: −0.428)
Car RTs	0.280 (−0.080: 0.563)	0.222 (−0.125: 0.544)	0.226 (−0.139: 0.536)	0.173 (−0.218: 0.537)

$p < .001$, $CI_{95\%}$: 0.522–0.868] (Fig. 3c). Some correlation was also seen between same-viewpoint face matching and same-viewpoint car matching [$r = 0.376$, $p = .02$, $CI_{95\%}$: 0.044–0.651], and between different-viewpoint face matching and different-viewpoint car matching [$r = 0.387$, $p = .016$, $CI_{95\%}$: 0.112–0.599]. However, both between-class correlations were significantly lower than the within-class correlations seen for faces ($z = 3.54$, $p < .001$; $z = 3.49$, $p < .001$) and cars ($z = 2.35$, $p = .019$; $z = 2.30$, $p = .021$). Having collapsed across viewing angle and retention interval, a correlation was seen between observers' face and car matching in the combined sample [$r = 0.437$, $p = .006$, $CI_{95\%}$: 0.154–0.683].

The correlations seen between observers' short- and long-interval matching accuracy (faces: $r = 0.810$; cars: $r = 0.689$), and between their constant- and different-viewpoint matching accuracy (faces: $r = 0.846$; cars: $r = 0.743$) indicate that the task – in particular, the face conditions – has good reliability. Reassuringly, matching accuracy for faces and cars also correlated with our other measures of face and car processing (Table 3). In particular, strong correlations were observed in the combined sample between participants' face matching accuracy and their scores on the CFMT [$r = 0.671$, $p < .001$, $CI_{95\%}$: 0.449–0.810], but not the CCMT [$r = 0.246$, $p = .140$, $CI_{95\%}$: -0.140-0.552]. Conversely, car matching accuracy correlated with scores on the CCMT [$r = 0.546$, $p < .001$, $CI_{95\%}$: 0.343–0.738], but not the CFMT [$r = 0.286$, $p = .082$, $CI_{95\%}$: -0.036-0.585].

3. Can DPs sort simultaneously presented faces by resemblance?

Previous studies have found that small samples of DPs make more errors on the CFPT than groups of matched TD controls (e.g., Shah et al., 2015a). As has been noted elsewhere, however, individual DPs often fail to exhibit significant impairment at the single-case level – i.e., they score within 2 *SDs* of mean typical performance on this task (e.g., Bowles et al., 2009; Ulrich et al., 2017). In the past, such observations have been cited as evidence that many DPs encode faces typically, and that their face recognition difficulties therefore arise from a deficit of STFM (Dalrymple et al., 2014; Dalrymple and Palermo, 2016; Jackson et al., 2017; Stollhoff et al., 2011; Ulrich et al., 2017). According to this view, group differences in CFPT performance, where observed, are driven by a subset of DPs with apperceptive impairments who produce outlying error scores.

Where observed, ambiguous z-scores (e.g., ~ -1) do not constitute significant single-case evidence of impairment on the CFPT. Nor, however, do they constitute strong evidence for normal perceptual encoding. In principle, a sample of DPs could all score within 2 *SDs* of the typical mean on a task, but exhibit a highly significant group difference.

We therefore sought to consider a second possibility – that apperceptive deficits are widespread in the DP population, but that the CFPT does not always reveal clear evidence of impairment. The use of the CFPT in the diagnosis of DP is discouraged due to its relatively poor psychometric qualities (Bowles et al., 2009). Given that the CFPT yields relatively noisy estimates of perceptual ability, the distribution of CFPT scores produced by DPs and controls might be expected to overlap to some degree. Moreover, the simultaneous sorting task employed by the CFPT may also render it susceptible to compensatory strategies; for example, the side-by-side presentation of the to-be-sorted faces, and the opportunity to study each trial display for a minute, may help DPs detect trivial details that help them achieve the correct solution.

It is difficult to distinguish these rival views by examining the scores from single cases of DP. However, these accounts make different predictions about the distributions of CFPT scores that should be seen in DP samples. According to the apperceptive subset view, the distribution of CFPT scores produced by TDs and DPs should differ only in terms of the lower tail of their distributions; i.e. the DP distribution should be identical to that of controls, with the exception of some outlying individuals at the lower tail who make a disproportionate number of errors. According to the shifted distribution view, however, evidence of impairment should be seen in both the upper and lower tail of the DP

Table 3
Correlations between participants' scores on the Cambridge Face Memory Test (CFMT), the upright condition of the Cambridge Face Perception Test (CFPT), the 20-item Prosopagnosia Index (PI20), the Cambridge Car Memory test (CCMT), and face and car matching performance. Accuracy and response time (RT) measures have been collapsed across viewing conditions. 95% confidence intervals are shown in brackets. ***$p < .001$; **$p < .01$; *$p < .05$.

	Faces		Cars	
	Accuracy	*RTs*	*Accuracy*	*RTs*
CFMT	0.671*** (0.464: 0.810)	− 0.192 (0.070: −0.518)	0.286 (−0.012: 0.591)	− 0.120 (0.226: −0.462)
CFPT	− 0.743*** (−0.584: .−864)	0.359* (0.103: 0.601)	− 0.326* (−0.016: −0.604)	0.234 (−0.155: 594)
PI20	− 0.628*** (−0.423: −0.777)	0.345* (0.073: 0.683)	− 0.283 (0.004: −0.589)	0.384* (0.059: 0.671)
CCMT	0.246 (−0.148: 0.559)	− 0.319 (−0.052: −0.518)	0.546*** (0.336: 0.734)	− 0.199 (0.099: −0.448)

Table 4
Diagnostic information for the DP and TD samples employed in Experiment 2.

	PI20				CFMT (%)			
	M	SD	Min	Max	M	SD	Min	Max
Typical controls (N = 54)	37.96	9.09	20	56	84.98	8.92	65.28	100.00
Prosopagnosics (N = 72)	80.94	7.45	59	97	56.00	7.83	34.72	68.02

distribution – not only should the worst DPs make more errors than the worst controls, but the best DPs should be unable to achieve scores comparable with the best controls. We sought to test these rival predictions by examining the distribution of CFPT scores produced by a large sample of DPs and controls.

3.1. Methods

3.1.1. Participants

In total, we considered data from 126 adults, 72 with DP (30 males; M_{age} = 42.34 years, SD_{age} = 11.77 years) and 54 typically developed (TD) controls (23 males; M_{age} = 39.20 years, SD_{age} = 13.36 years). These groups include the 16 DPs and 22 TD controls from the first experiment. Summary statistics for both groups are provided in Table 4 and detailed diagnostic information for each DP is provided as supplementary material. As expected, the groups differed significantly in their PI20 [$t(124) = 29.156$, $p < .001$] and CFMT scores [$t(124) = 19.357$, $p < .001$]. Neither participant age [$t(124) = 1.400$, $p = .164$] nor proportion of males [$X^2(1) = 0.01$, $p = .920$] differed significantly between the two groups.

3.1.2. Stimuli and procedure

The CFPT assesses face perception ability in such a way as to minimise the memory demand on participants. Trials present a target face and a series of six faces that resemble the target to varying degrees (Fig. 4a). Participants have 60 s to sort the six faces in order of target-face similarity. Eight trials present the target and test faces upright, eight present the faces inverted. Trials are scored by calculating deviations from the correct order. Participants were given the option of completing the CFPT using a trackpad or mouse – whichever they found easier to use. All participants were tested individually at the *troublewithfaces.org* lab, under tightly controlled conditions, in return for a small honorarium.

3.2. Results and discussion

Participants' scores on the CFPT were analysed using ANOVA with Orientation (upright, inverted) as a within-subjects factor, and Group (DP, TD) as a between-subjects factor (Fig. 4b). The analysis revealed main effects of Orientation [$F(1,124) = 370.862$, $p < .001$, $\eta_p^2 = .749$] and Group [$F(1,124) = 10.650$, $p < .001$, $\eta_p^2 = .079$] with more errors seen when faces were inverted and less precise sorting exhibited by the DP group. However, it also yielded a significant Orientation × Group interaction [$F(1,124) = 251.784$, $p < .001$, $\eta_p^2 = .670$]. The DPs ($M = 50.64$, $SD = 15.35$) made disproportionately more errors than the controls ($M = 29.41$, $SD = 9.35$) on the upright trials of the CFPT [$t(124) = 9.601$, $p < .001$]. However, the DPs ($M = 69.86$, $SD = 13.11$) also made more errors than controls ($M = 63.37$, $SD = 15.74$) on the inverted trials [$t(124) = 2.522$, $p = .013$].

While the scores of the TD observers were more sensitive to the orientation manipulation (upright vs. inverted presentation), this may simply reflect the fact that the DPs are closer to floor performance in the upright condition (also see Klargaard et al., 2018). In addition to the group difference (DPs < TDs) seen for the inverted trials of the CFPT, we found evidence of correlation between observers' scores on the upright and inverted trials [$r = 0.370$, $p < .001$, $CI_{95\%}$: 0.225–0.521] (Fig. 4c). When considered separately, this correlation was seen in the TD group [$r = 0.530$, $p < .001$, $CI_{95\%}$: 0.332–0.695] but not in the DPs [$r = 0.207$, $p = .081$, $CI_{95\%}$: -0.001-0.432]. These findings accord with the view that the visual processing of upright and inverted faces may differ quantitatively (Gold et al., 2012; Murphy and Cook, 2017; Sekuler et al., 2004; Susilo et al., 2013), not qualitatively (McKone and Yovel, 2009; Rossion, 2008).

Next, we ranked the TD (N = 54) and DP (N = 72) samples based on individuals' performance on the upright trials of the CFPT and split each distribution into thirds: best performing TDs (N = 18, M_{age} = 37.83) and DPs (N = 24, M_{age} = 43.08), intermediate TDs (N = 18, M_{age} = 37.94) and DPs (N = 24, M_{age} = 41.38), and poorest performing TDs (N = 18, M_{age} = 41.83) and DPs (N = 24, M_{age} = 42.58). Strikingly, the TD controls outperformed the DPs at each level of their respective distributions: best performers [$t(40) = 11.304$, $p < .001$], intermediate performers [$t(40) = 15.596$, $p < .001$], poorest performers [$t(40) = 13.051$, $p < .001$] (Fig. 4d). This pattern argues against the view that group differences on the CFPT reflect the presence of a few individual DPs with an apperceptive deficit. Instead, these results favour the view that the entire distribution of CFPT scores produced by the DPs is shifted relative to that of TD controls.

To illustrate how apperceptive impairment in DP might produce a shifted distribution of CFPT scores similar to that observed, we have shown the effects of inflating each typical observer's CFPT error score by 80% (Fig. 5a). This inflation coefficient is akin to the application of a hypothetical apperceptive deficit that increases the number of sorting errors made. As can be seen, this simple model provides a reasonable approximation of the distribution of scores seen in the DP sample. To be clear, we are not claiming that DP always impairs perceptual encoding of faces by 80%; rather, we present this demonstration as a proof-of-principle. We merely seek to illustrate that an apperceptive deficit might plausibly produce the distribution of CFPT scores seen in our DP sample.

In the present study, decisions to classify people as DP were based principally on individuals' PI20 and CFMT scores. We note, however, that observers' CFPT scores correlated with their CFMT scores (Fig. 5b). This relationship was seen in the combined sample [$r = -0.665$, $p < .001$, $CI_{95\%}$: -0.569 to -0.740], and independently in the TD [$r = -0.371$, $p = .006$, 95% $CI_{95\%}$: -0.081 to -0.626] and DP groups [$r = -0.299$, $p = .011$, $CI_{95\%}$: -0.113 to -0.470]. Contrary to the prevailing view that the CFMT is a test of 'face memory', this finding suggests that the individual differences revealed by the CFMT may be strongly influenced by individuals' ability to encode face structure. In other words, we may be able to predict with a fair degree of accuracy whether an individual's CFMT score will fall in the DP range, using estimates of their perceptual encoding ability such as their CFPT score.

4. General discussion

In the present study we considered whether DP is best characterised as i) a disorder of STFM, where individuals initially form accurate perceptual descriptions of faces, but struggle to maintain these representations over time; or ii) as an apperceptive condition, where face recognition difficulties arise from poor encoding of face structure. In our first experiment, participants (16 DPs and 22 TD controls) completed a delayed match-to-sample task for faces and cars, with a

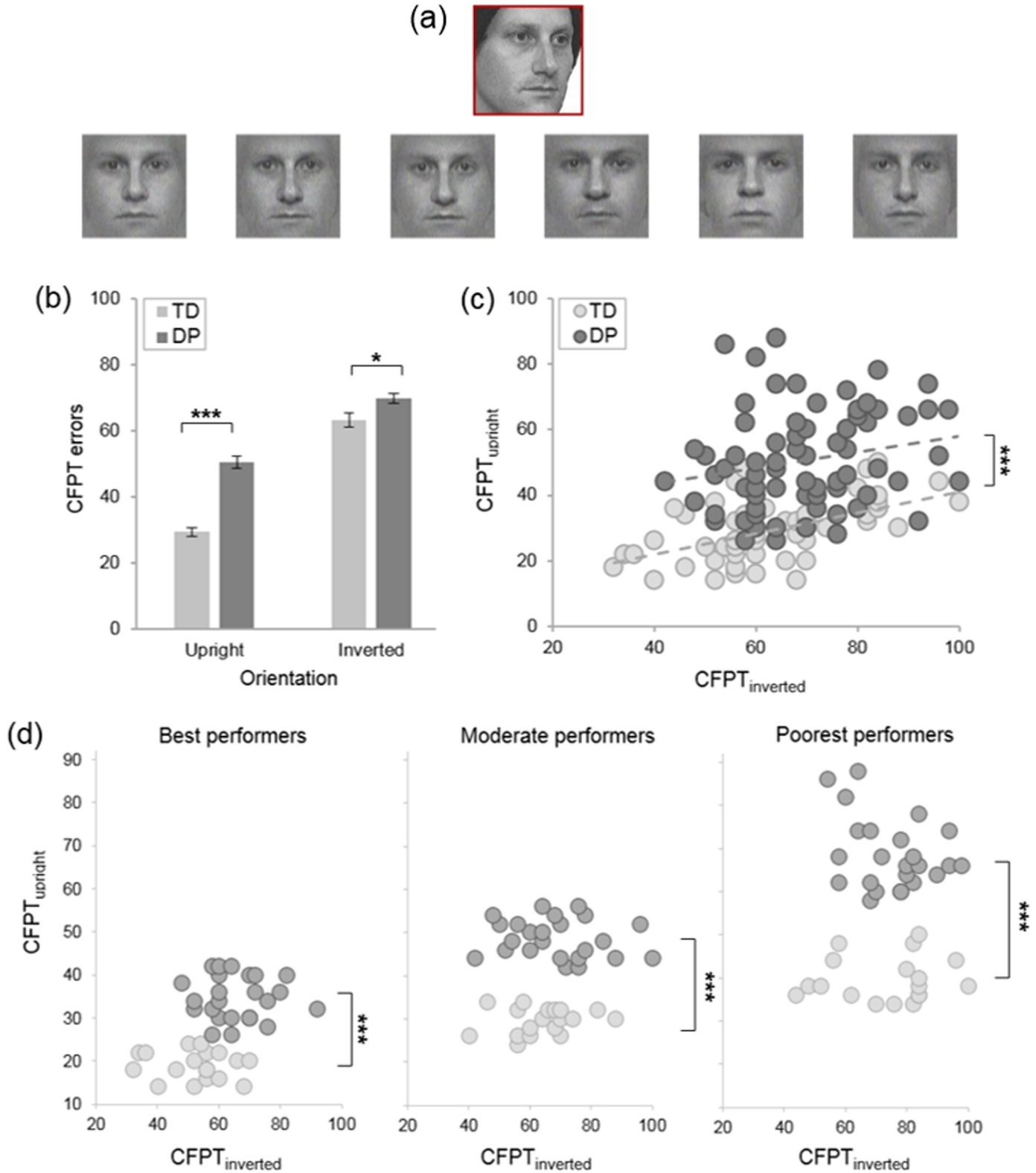

Fig. 4. (a) Each trial of the CFPT presents simultaneously a target face and a series of six faces that resemble the target to varying degrees. Participants have 60 s to sort the six items in order of target-face similarity. (b) Mean performance of the TD (N = 54) and DP (N = 72) groups in the upright and inverted conditions of the CFPT. (c) Each participant's performance on the upright trials plotted against their inverted performance. (d) Analysis of the best, moderate, and worst performers from the sample indicated that the entire distribution of DP scores was shifted relative to the distribution of TD scores. ***$p < .001$; **$p < .01$; *$p < .05$. Error bars denote ± 1SEM.

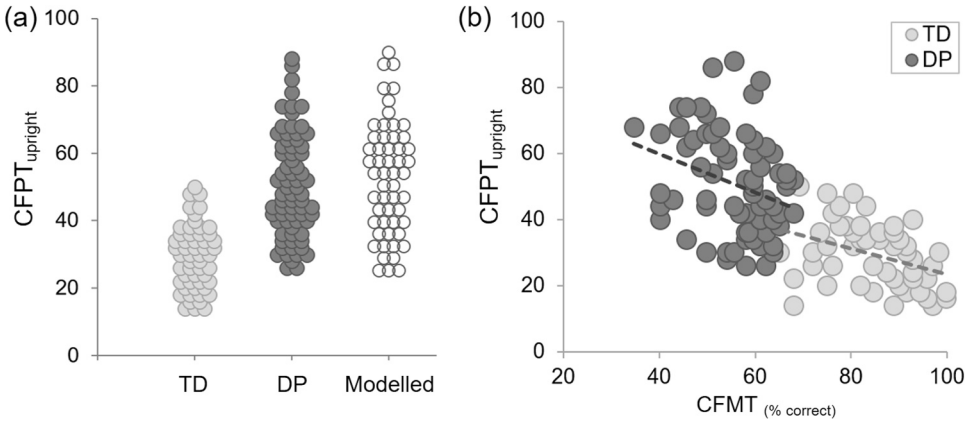

Fig. 5. (a) To illustrate the shifted distribution account, we modelled the effects of inflating each typical observer's error score by 80%, akin to the application of a hypothetical apperceptive deficit. (b) Scatterplot showing the relationship between the participants' CFPT and CFMT scores. The line of best-fit is modelled separately for the TD and DP groups.

retention interval of 1-second (low demand) or 6-seconds (high demand). As expected, participants with DP were worse than TD controls at face matching. Interestingly, however, the relative degree of impairment exhibited by the DPs did not interact with retention interval. Next, we analysed the performance of 72 DPs and 54 TD controls on the CFPT. We found that the DPs were clearly impaired at the group level, and showed signs of a shifted distribution.

4.1. Evidence for an apperceptive characterisation

In our first experiment, we found that the face matching deficits seen in DP were insensitive to retention interval; i.e., that very similar levels of impairment were seen at the short and long intervals. To date, only one other study has used a delayed match-to-sample task to explore the perceptual and mnemonic contributions to DP (Shah et al., 2015a). In this study, the authors found that 15 DPs exhibited comparable face matching deficits at short (2-second and longer (8-seconds intervals. We replicated this result in a sample of 16 different DPs using a short interval condition of 1-second. In addition, the present results show that DPs exhibit similar impairments at short and long retention intervals when a 45° viewpoint disparity exists between the target and test items. This finding excludes the possibility that DPs have a particular problem retaining percepts in a way that supports rotation and manipulation (working memory; Baddeley, 1992, 1993, 2010).

The view that the face matching deficits seen in DP are relatively insensitive to memory load is also suggested by a finding recently described by Jackson et al. (2017; Experiment 1). Rather than vary retention interval, the authors manipulated memory load by increasing the number of target faces observers had to memorise (one, two, three, or four). Participants were asked whether a single test image presented a second later was one of the targets. As expected, the authors found that matching accuracy decreased as a function of the number of target faces held in memory (a main effect of Memory Load), and that relative to controls, DPs performed poorly in all conditions (a main effect of Group). Crucially, however, the relative impairment of the DPs did not increase with memory load.[3] The insensitivity of the DPs' deficits to the memory load manipulation mirrors the findings of the present study. Once again, this result suggests that the matching deficits observed have a perceptual origin; for example, the DPs in the experiment described by Jackson and colleagues (2017) may have had problems forming a perceptual description of the test face, and thus exhibited poor matching at all levels of the memory load manipulation.

Given the apparent insensitivity of their deficits to the memory demands of face matching tasks (Experiment 1; see also Shah et al., 2015a; Jackson et al., 2017), we re-examined DPs performance on the CFPT. The fact that many DPs exhibit clear impairment on the CFMT (a matching task with substantial perceptual and memory components), but show only mild difficulties on the CFPT (a sorting task that measures face perception in a way that minimises participants' memory load) forms a key line of evidence for the mnemonic account of DP (Bowles et al., 2009; Dalrymple et al., 2014; McKone et al., 2011; Ulrich et al., 2017). Having analysed the CFPT performance of 72 DPs, however, we found clear evidence of impairment at the group level. This difference was not attributable to a few individuals with an apperceptive profile; rather we found evidence that the distribution of CFPT scores seen in the DP sample was shifted relative to that of typical controls. Not only did the worst DPs make more errors than the worst controls, but the best DPs were unable to achieve scores comparable with the best controls. Indeed, only three of the 72 DPs made fewer errors than the mean of the control group.

In sum, several group studies have now failed to find effects of memory load on face matching deficits in DP, including the present study (N = 16 DPs), Shah et al. (2015a) (N = 15 DPs), and Jackson et al., (N = 10 DPs). In contrast, our analysis of the CFPT scores produced by 72 DPs, suggests that perceptual encoding problems may be widespread within this population, and not limited to a small apperceptive subgroup. Some heterogeneity is likely in any neurodevelopmental population, and DP is no different. Generally, however, these findings suggest that selective STFM impairment may be relatively uncommon in this population. Instead, deficits of perceptual encoding may play a larger role in DP than currently acknowledged.

4.2. Putative dissociations between performance on the CFMT and CFPT

Many DP samples include individual DPs who show marked impairment on the CFMT (e.g., < 2*SDs* below the TD mean) but who exhibit only marginal impairment on the CFPT. Our sample is no different (see Supplementary material). For example, 21% of the DPs described met the criteria for a putatively classical dissociation and 4% a strong dissociation between their CFMT and CFPT performance (Crawford and Garthwaite, 2007). As described above, the fact that the CFMT (a task with both perceptual and memory components) is more likely to reveal clear deficits at the single-case level than the CFPT (a task that assesses face perception with minimal memory demands) has led many to speculate that DP may often be caused by aberrant STFM, and not impaired perceptual encoding (Bowles et al., 2009; Dalrymple et al.,

[3] In the second experiment described by Jackson et al. (2017), trials presented four faces sequentially for 500 ms each, followed by a maintenance interval of one second. Participants were asked whether a single test image presented a second later was one of the targets. In their analysis, the authors examined how observers' discrimination varied as a function of the position of the target in the sequence (first, second, third, fourth). Although the DPs performed relatively poorly in all categories, their serial-position functions closely resembled those of the controls; for example, both the DPs and controls were more accurate when matching recently presented targets.

2014; McKone et al., 2011; Ulrich et al., 2017). Where observed, however, we recommend authors treat these apparent dissociations with caution.

First, artefactual dissociations between CFMT and CFPT performance will arise from the fact that one measure plays a key role in the diagnosis of DP, while the other does not. A clear deficit on the CFMT (e.g., < 2 *SDs* below the TD mean) is widely seen as necessary for a DP diagnosis (e.g., Barton and Corrow, 2016; Dalrymple and Palermo, 2016). Where individuals fail to reach this criterion, they are often excluded from DP research. In contrast, CFPT scores are free to vary; where reported, they are provided only as an indication of whether a DP is apperceptive or mnemonic (e.g., Biotti and Cook, 2016; Biotti et al., 2017b). Some DPs will inevitably under-score on the CFMT and over-score on the CFPT (i.e., their scores on these measures under- and over-estimate their true ability, respectively) as a result of the measurement error in these instruments. Because individuals can receive a diagnosis of DP without clear impairment on the CFPT, individuals with this profile are free to participate in DP research. Conversely, other DPs will over-score on the CFMT and under-score on the CFPT. Individuals with this profile are less likely to meet the criteria for inclusion in DP samples and are at risk of being excluded from research. In light of this prevailing bias, it is unsurprising that the DP literature includes many individuals who exhibit a clear deficit on the CFMT but not on the CFPT. The practice of preselecting individuals based on extremely poor CFMT scores (< 2 *SDs* below the mean), and then reporting single-case analyses that purport to show that an individual's CFMT deficit exceeds their CFPT deficit is akin to 'double dipping' (Kriegeskorte et al., 2009).

Second, the CFPT may simply be less likely to detect significant impairments at the single-case level than the CFMT. The CFMT is thought to have good internal reliability (e.g., α's of ~.85; Bowles, et al., 2017). In contrast, the fact that the CFPT has fewer trials, the means by which test faces are sorted, and the way performance is scored, may compromise its psychometric properties (α's of ~.74; Bowles, et al., 2017). While poor reliability may not systematically suppress the scores of DPs in particular, the noisy performance of everyone makes it less likely that any given DP will fall 2 *SDs* outside the typical range. The format of the CFPT may also render it more susceptible to compensatory strategies, than the CFMT; for example, the side-by-side presentation of the to-be-sorted faces, and the opportunity to study each trial display for a minute, may help DPs detect trivial details that help them achieve the correct solution. Due to its relatively poor reliability, authors have been discouraged from using the CFPT in the diagnosis of DP (Bowles et al., 2009). If the CFPT lacks the reliability necessary to diagnose someone as DP or not DP, we should be cautious about making apperceptive vs. mnemonic classifications on this basis.

Third, it is not clear whether meaningful theoretical inferences can be drawn from differential impairments on the CFPT and the CFMT. Critically, the CFPT and CFMT differ not only in terms of their respective memory components, but also in their fundamental *perceptual* demands. The CFMT and CFPT present different facial identities under different viewing conditions; for example, the CFMT, but not the CFPT, includes trials where targets are obscured by high-spatial frequency noise. Moreover, the CFPT requires individuals to sort six faces, presented side-by-side, based on their resemblance to a target face, whereas the CFMT requires individuals to identify a recently encountered individual from a line-up of three different facial identities. These two tasks – sorting morphed faces by resemblance (the CFPT), and facial individuation and identification (the CFMT) – may depend on different types of cue, and tax different types of perceptual process (e.g., White et al., 2017). It is therefore misleading to equate the CFMT and CFPT to perceptual and mnemonic conditions in a controlled experimental manipulation. Although the CFMT and the CFPT differ in their respective memory demands, the differential memory load is confounded with numerous perceptual differences.

4.3. The broader case against apperceptive accounts

Individuals with an apperceptive face processing deficit would be expected to exhibit aberrant perception and recognition of facial expression (Biotti and Cook, 2016, 2017; De Renzi et al., 1991; Duchaine et al., 2003). Studies describing (seemingly) typical recognition of facial emotion in DP (e.g., Humphreys et al., 2007; Ulrich et al., 2017) therefore appear to challenge the view that the majority of individuals with DP exhibit some degree of apperceptive impairment. We note, however, that sensitive psychophysical tasks – *and appropriate analyses* – may be required to detect expression recognition difficulties arising from impoverished structural description. Having employed expression morphing and the estimation of psychometric functions, Biotti and Cook (2016) found that subtle expression recognition deficits were relatively common in a sample of 17 DPs (see also Burns et al., 2017). In contrast, tasks that simply require participants to label prototypical expressions ('basic emotions') may be prone to ceiling effects and lack the sensitivity necessary to detect subtle deficits (for related discussion, see Ipser and Cook, 2015). The development of sensitive, reliable emotion recognition tasks will help to establish the prevalence of apperceptive deficits in the DP population.

Recent evidence suggests that most individuals with DP show typical susceptibility to the composite face effect (Biotti et al., 2017b; Esins et al., 2016; Le Grand et al., 2006; Susilo et al., 2010; Ulrich et al., 2017), a visual illusion thought to index holistic face processing (Murphy et al., 2017; Rossion, 2013). While these results suggest that holistic face processing may be intact in DP, they by no means exclude all apperceptive accounts of the condition. For example, DPs may have an apperceptive problem that affects local feature descriptions. Consistent with this possibility, many DPs struggle to make judgements about local regions shown in isolation (Biotti and Cook, 2016; Duchaine et al., 2006; Liu and Behrmann, 2014). We also note recent evidence from aperture viewing paradigms suggesting that the ability to process local regions may be a key determinant of face recognition performance (Murphy and Cook, 2017).

4.4. Insensitivity of face matching deficits to viewpoint disparity

Different-viewpoint matching is thought to be a better test of face perception ability than constant-viewpoint matching (e.g., Duchaine and Nakayama, 2006a). To match unfamiliar faces across different viewpoints, observers must infer the 3D structure of a target face from an ambiguous 2D image depicting a single view. This represents a substantial computational challenge (Todd, 2004). In the absence of an image-change, constant-viewpoint matching can in principle be accomplished using superficial pictorial cues (Hancock et al., 2000; Megreya and Burton, 2006). One might therefore expect DPs to show greater impairment, relative to controls, when matching across different viewpoints. The fact that our DPs exhibited similar deficits when matching faces shown from the same viewing angle, and from different viewing angles (Experiment 1), is therefore striking. Rather than dissociation between constant-viewpoint and different-viewpoint face-matching, our results suggest association: our participants appear to have used a similar process in both conditions. This is further suggested by the fact that participants' constant-viewpoint matching ability was closely related to their different-viewpoint matching ability.

One possibility is that observers accomplished both types of face matching through superficial pictorial cues, and that DPs experience difficulties using this image matching strategy. This seems unlikely for two reasons. First, image matching is by definition a domain-general process (Hancock et al., 2000; Megreya and Burton, 2006). Crucially, however, our DPs were unimpaired at car matching in our first experiment. Similarly, the DPs tested by Shah et al. (2015a) showed typical matching of chairs, butterflies, and hands. These convergent findings argue against a simple image matching deficit. Second, face

matching accuracy – but not car matching accuracy – correlated with the face-recognition problems encountered by observers outside the lab, as measured by the PI20 (e.g., mistaking familiar people for strangers, failing to recognise people in the absence of vocal cues, problems recognising people wearing hats or different hairstyles). These difficulties seem unlikely to reflect aberrant processing of trivial pictorial cues. Instead, this correlation underscores the fact that the processes measured by our matching task have meaningful consequences for the day-to-day social interactions of our participants.

Instead, we favour the view that observers use 'face-centred' (Bruce and Young, 1986; Marr and Nishihara, 1978) structural descriptions to achieve both constant-viewpoint and different-viewpoint face matching.[4] We speculate that i) these structural descriptions augmented the matching performance of typical observers in both the constant-viewpoint and different-viewpoint matching conditions; and ii) the DPs were outperformed in all viewing conditions because they were hampered by imprecise structural descriptions. There is little doubt that seeing to-be-learned individuals in different poses, with different expressions, from different viewing angles (so-called exemplar variation) aids face learning (e.g., Ipser et al., 2016; Murphy et al., 2015). The suggestion that observers form face-centred descriptions of unfamiliar faces from a single 2D image may therefore seem counter-intuitive. Consider, however, that computer programs have been described that do precisely this; i.e., extrapolate a morphable, posable 3D model of a human face from a single image of a novel face, using the covariation present in a set of training images (e.g., FaceGen Modeller). Once derived, these morphable posable models can be used to estimate how the target face will appear from different viewing angles (e.g., Jones et al., 2017). In a similar way, the human visual system may use the statistical regularities present in the faces it has encountered in the past to estimate the likely 3D structure of novel faces.

4.5. Is DP associated with a face-specific or domain-general deficit?

It remains unclear whether the deficit seen in DP is face-specific or indicative of a domain-general impairment (Gerlach et al., 2016; Geskin and Behrmann, 2017). On the one hand, we observed a significant group difference on the CCMT and a correlation ($r = 0.437$) between face and car matching accuracy. We also found that the DP group made more errors than the typical controls when sorting inverted faces, regarded by some as a measure of domain-general perceptual ability (e.g., Rossion, 2008, 2013). On the other hand, our DPs were unimpaired in the car matching condition of Experiment 1, and other authors, for example, Shah et al. (2015a; N = 15 DPs) and Esins et al. (2016a; N = 16 DPs), have found that DPs' performance on the CCMT is comparable with matched controls.

Further research is needed to elucidate the nature of the object recognition difficulties seen in DP (Gerlach et al., 2016; Geskin and Behrmann, 2017). However, evidence of idiosyncratic, inconsistent deficits accords well with the independent disorders hypothesis - the view that forms of developmental agnosia affecting faces and objects are best thought of as independent neurodevelopmental conditions (Gray and Cook, 2018). This account predicts the existence of 'pure' cases of DP and developmental object agnosia (DOA), individuals who experience impaired face recognition but typical object recognition (Duchaine et al., 2006), and vice versa (Germine et al., 2011). However, the independent disorders hypothesis also predicts that the incidence of DOA will be higher in DP than in the wider population due to common genetic or environmental risk factors. For example, susceptibility to aberrant structural development of occipito-temporal cortex (e.g., reduced density and coherence of white matter tracts or atypical neural migration) may be a common risk factor for DP and DOA (see also Susilo and Duchaine, 2013).

4.6. Limitations and future research

As noted above, several group studies have now failed to find an effect of memory load on face matching deficits in DP, including the present study (N = 16 DPs), Shah et al., 2015a (N = 15 DPs), and Jackson et al., 2017 (N = 10 DPs). Together, these results suggest that cases of DP arising solely from a short-term memory impairment may be less common than currently believed. However, the foregoing studies investigated mnemonic effects with retention intervals of 8-seconds (Shah et al., 2015a) or less (present study; Jackson et al., 2017). These results do not exclude the possibility that some DPs have a mnemonic deficit that impairs longer term face memory (e.g., Stollhoff et al., 2011) or face learning (e.g., Ipser et al., 2016; Murphy et al., 2015). It is important that future empirical work explore potential deficits in these domains. In order to facilitate this work, we encourage proponents to articulate more clearly the parameters of mnemonic accounts of DP; for example, what types of memory process are thought to be impaired, whether perceptual encoding is preserved entirely, and what constitutes 'short-term' and 'long-term' face memory.

Traditionally, researchers have sought to evidence mnemonic cases of DP by reporting deficits on face recognition tasks with substantial memory demands, but (relatively) intact performance on perceptual tasks with minimal memory demands (e.g. Ulrich et al., 2017). This approach relies on *negative evidence* – a failure to detect significant perceptual impairment. Instead, we encourage future research to seek *positive evidence* of memory deficits. As we have described, DPs with a selective deficit of STFM should show greater impairment on a face matching task with a high memory demand (e.g., a long retention interval), than when performing the same task under conditions of low memory demand (e.g., a short retention interval). By keeping the perceptual demands in these two conditions identical, it should be possible to exclude apperceptive interpretations, where deficits increase as a function of memory demands. Moreover, if the DP population includes a subgroup who exhibit selective problems retaining faces in memory, but who are able to encode faces without any impairment whatsoever, it should be possible to find individuals who achieve *good* levels of performance on sensitive measures of face encoding, including psychophysical tests of emotion, age, and gender classification ability. Positive evidence for a memory deficit, together with evidence of strong perceptual encoding ability, would represent a compelling demonstration of a mnemonic case of DP.

4.7. Conclusion

There has been considerable speculation that a subgroup exists within the DP population, comprised of individuals who can form accurate face percepts, but are unable to maintain those percepts over time. While this mnemonic hypothesis is an interesting idea, it currently lacks convincing empirical support. In particular, several group studies have now failed to find effects of short-term memory load on the face matching ability of DPs. In contrast, our analysis of a large sample of DPs indicates that deficits of perceptual encoding are widespread in this population. We would stop short of saying that *all* cases of DP are apperceptive: neurodevelopmental populations are rarely this homogenous. However, our results suggest that cases of DP may typically have an apperceptive origin.

[4] We use the term face-centred rather than view-invariant to reflect the fact that these representations do not exhibit perfect view-invariance. We note, however, that observers' matching performance – in the present study and elsewhere – typically far exceeds chance even when pairs of unfamiliar faces are presented with large viewpoint disparities. Given the highly complex 3D shape of the human face, and the fact 3D structure must be recovered from a highly ambiguous 2D image, this is a remarkable achievement of the human visual system.

Acknowledgements

FB was supported by a doctoral studentship awarded by City, University of London. RC is supported by a Starting Grant awarded by the European Research Council (ERC-2016-StG-715824).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.neuropsychologia.2018.11.014.

References

Ariel, R., Sadeh, M., 1996. Congenital visual agnosia and prosopagnosia in a child: a case report. Cortex 32 (2), 221–240.
Avidan, G., Behrmann, M., 2009. Functional MRI reveals compromised neural integrity of the face processing network in congenital prosopagnosia. Curr. Biol. 19 (13), 1146–1150.
Avidan, G., Hasson, U., Malach, R., Behrmann, M., 2005. Detailed exploration of face-related processing in congenital prosopagnosia: 2. functional neuroimaging findings. J. Cogn. Neurosci. 17 (7), 1150–1167.
Avidan, G., Tanzer, M., Behrmann, M., 2011. Impaired holistic processing in congenital prosopagnosia. Neuropsychologia 49 (9), 2541–2552.
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. Cereb. Cortex 24 (6), 1565–1578.
Baddeley, A.D., 1992. Working memory. Science 255 (5044), 556–559.
Baddeley, A.D., 1993. Verbal and visual subsystems of working memory. Curr. Biol. 3 (8), 563–565.
Baddeley, A.D., 2010. Working memory. Curr. Biol. 20 (4), R136–R140.
Barton, J.J.S., Corrow, S.L., 2016. The problem of being bad at faces. Neuropsychologia 89, 119–124.
Behrmann, M., Avidan, G., 2005. Congenital prosopagnosia: face-blind from birth. Trends Cogn. Sci. 9 (4), 180–187.
Biotti, F., Cook, R., 2016. Impaired perception of facial emotion in developmental prosopagnosia. Cortex 81, 126–136.
Biotti, F., Cook, R., 2017. Impaired perception of facial emotion in developmental prosopagnosia: a reply to Van den Stock's commentary. Cortex 101, 298–299.
Biotti, F., Gray, K.L.H., Cook, R., 2017a. Impaired body perception in developmental prosopagnosia. Cortex 93, 41–49.
Biotti, F., Wu, E., Yang, H., Jiahui, G., Duchaine, B., Cook, R., 2017b. Normal composite face effects in developmental prosopagnosia. Cortex 95, 63–76.
Bowles, D.C., McKone, E., Dawel, A., Duchaine, B., Palermo, R., Schmalzl, L., Yovel, G., 2009. Diagnosing prosopagnosia: effects of ageing, sex, and participant-stimulus ethnic match on the Cambridge Face Memory Test and Cambridge Face Perception Test. Cogn. Neuropsychol. 26 (5), 423–455.
Brainard, D.H., 1997. The psychophysics toolbox. Spat. Vis. 10 (4), 433–436.
Bruce, V., Young, A.W., 1986. Understanding face recognition. Br. J. Psychol. 77, 305–327.
Burns, E.J., Martin, J., Chan, A.H.D., Xu, H., 2017. Impaired processing of facial happiness, with or without awareness, in developmental prosopagnosia. Neuropsychologia 102, 217–228.
Chatterjee, G., Nakayama, K., 2013. Normal facial age and gender perception in developmental prosopagnosia. Cogn. Neuropsychol. 29 (5–6), 482–502.
Cook, R., Biotti, F., 2016. Developmental prosopagnosia. Curr. Biol. 26 (8), R312–R313.
Crawford, J.R., Garthwaite, P.H., 2007. Comparison of a single case to a control or normative sample in neuropsychology: development of a Bayesian approach. Cogn. Neuropsychol. 24, 343–372.
Dalrymple, K.A., Garrido, L., Duchaine, B., 2014. Dissociation between face perception and face memory in adults, but not children, with developmental prosopagnosia. Dev. Cogn. Neurosci. 10, 10–20.
Dalrymple, K.A., Palermo, R., 2016. Guidelines for studying developmental prosopagnosia in adults and children. Wiley Interdiscip. Rev.: Cogn. Sci. 7 (1), 73–87.
De Renzi, E., Faglioni, P., Grossi, D., Nichelli, P., 1991. Apperceptive and associative forms of prosopagnosia. Cortex 27 (2), 213–221.
DeGutis, J., Chatterjee, G., Mercado, R.J., Nakayama, K., 2014a. Face gender recognition in developmental prosopagnosia: evidence for holistic processing and use of configural information. Vis. Cogn. 20 (10), 1242–1253.
DeGutis, J., Cohan, S., Mercado, R.J., Wilmer, J., Nakayama, K., 2012. Holistic processing of the mouth but not the eyes in developmental prosopagnosia. Cogn. Neuropsychol. 29 (5–6), 419–446.
DeGutis, J., Cohan, S., Nakayama, K., 2014b. Holistic face training enhances face processing in developmental prosopagnosia. Brain 137 (Pt 6), 1781–1798.
Dennett, H.W., McKone, E., Tavashmi, R., Hall, A., Pidcock, M., Edwards, M., Duchaine, B., 2011. The Cambridge Car memory test: a task matched in format to the Cambridge face memory test, with norms, reliability, sex differences, dissociations from face memory, and expertise effects. Behav. Res. Methods 44 (2), 587–605.
Dobel, C., Bölte, J., Aicher, M., Schweinberger, S.R., 2007. Prosopagnosia without apparent cause: overview and diagnosis of six cases. Cortex 43 (6), 718–733.
Duchaine, B., Germine, L., Nakayama, K., 2007. Family resemblance: ten family members with prosopagnosia and within-class object agnosia. Cogn. Neuropsychol. 24 (4), 419–430.
Duchaine, B., Nakayama, K., 2006a. 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 (4), 576–585.
Duchaine, B., Nakayama, K., 2006b. Developmental prosopagnosia: a window to content-specific face processing. Curr. Opin. Neurobiol. 16, 166–173.
Duchaine, B., Parker, H., Nakayama, K., 2003. Normal recognition of emotion in a prosopagnosic. Perception 32 (7), 827–838.
Duchaine, B., Yovel, G., Butterworth, E., Nakayama, K., 2006. Prosopagnosia as an impairment to face-specific mechanisms: elimination of the alternative hypotheses in a developmental case. Cogn. Neuropsychol. 23 (5), 714–747.
Esins, J., Schultz, J., Stemper, C., Kennerknecht, I., Bulthoff, I., 2016. Face perception and test reliabilities in congenital prosopagnosia in seven tests. i-Percept. 7 (1), 1–37.
Farah, M.J., Wilson, K.D., Drain, M., Tanaka, J.N., 1998. What is "special" about face perception? Psychol. Rev. 105 (3), 482–498.
Fisher, K., Towler, J., Eimer, M., 2016. Reduced sensitivity to contrast signals from the eye region in developmental prosopagnosia. Cortex 81, 64–78.
Fisher, K., Towler, J., Eimer, M., 2017. Face identity matching is selectively impaired in developmental prosopagnosia. Cortex 89, 11–27.
Gerlach, C., Klargaard, S.K., Starrfelt, R., 2016. On the relation between face and object recognition in developmental prosopagnosia: no dissociation but a systematic association. PLoS One 11 (10), e0165561.
Germine, L., Cashdollar, N., Düzel, E., Duchaine, B., 2011. A new selective developmental deficit: impaired object recognition with normal face recognition. Cortex 47, 598–607.
Geskin, J., Behrmann, M., 2017. Congenital prosopagnosia without object agnosia? A literature review. Cogn. Neuropsychol. 22, 1–51.
Gold, J.M., Mundy, P.J., Tjan, B.S., 2012. The perception of a face is no more than the sum of its parts. Psychol. Sci. 23 (4), 427–434.
Gomez, J., Pestilli, F., Witthoft, N., Golarai, G., Liberman, A., Poltoratski, S., Grill-Spector, K., 2015. Functionally defined white matter reveals segregated pathways in human ventral temporal cortex associated with category-specific processing. Neuron 85 (1), 216–227.
Gray, K.L.H., Bird, G., Cook, R., 2017. Robust associations between the 20-item prosopagnosia index and the Cambridge Face Memory Test in the general population. R. Soc. Open Sci. 4 (3), 160923.
Gray, K.L.H., Cook, R., 2018. Should developmental prosopagnosia, developmental body agnosia, and developmental object agnosia be considered independent neurodevelopmental conditions? Cogn. Neuropsychol. 35 (1–2), 59–62.
Hancock, P.J., Bruce, V.V., Burton, A.M., 2000. Recognition of unfamiliar faces. Trends Cogn. Sci. 4 (9), 330–337.
Haxby, J.V., Hoffman, E.A., Gobbini, M.I., 2000. The distributed human neural system for face perception. Trends Cogn. Sci. 4, 223–233.
Humphreys, K., Avidan, G., Behrmann, M., 2007. A detailed investigation of facial expression processing in congenital prosopagnosia as compared to acquired prosopagnosia. Exp. Brain Res. 176 (2), 356–373.
Ipser, A., Cook, R., 2015. Inducing a concurrent motor load reduces categorization precision for facial expressions. J. Exp. Psychol.: Hum. Percept. Perform. 42 (5), 706–718.
Ipser, A., Ring, M., Murphy, J., Gaigg, S.B., Cook, R., 2016. Similar exemplar pooling processes underlie the learning of facial identity and handwriting style: evidence from typical observers and individuals with Autism. Neuropsychologia 85, 169–176.
Ishihara, S., 1993. Ishihara's Tests for Colour-Blindness. Kanehara, Tokyo, Japan.
Jackson, M.C., Counter, P., Tree, J.J., 2017. Face working memory deficits in developmental prosopagnosia: tests of encoding limits and updating processes. Neuropsychologia 106, 60–70.
JASP-Team, 2018. JASP (Version 0.9) [Computer Software]. Amsterdam, The Netherlands.
Johnen, A., Schmukle, S.C., Hüttenbrink, J., Kischka, C., Kennerknecht, I., Dobel, C., 2014. A family at risk: congenital prosopagnosia, poor face recognition and visuoperceptual deficits within one family. Neuropsychologia 58, 52–63.
Jones, S.P., Dwyer, D.M., Lewis, M.B., 2017. The utility of multiple synthesized views in the recognition of unfamiliar faces. Q. J. Exp. Psychol. 70 (5), 906–918.
Kennerknecht, I., Grüter, T., Welling, B., Wentzek, S., Horst, J., Edwards, S., Grüter, M., 2006. First report of prevalence of non-syndromic hereditary prosopagnosia (HPA). Am. J. Med. Genet. 140A (15), 1617–1622.
Kennerknecht, I., Ho, N.Y., Wong, V.C.N., 2008. Prevalence of heriditary prosopagonsia (HPA) in Hong Kong Chinese population. Am. J. Med. Genet. 146A (22), 2863–2870.
Klargaard, S.K., Starrfelt, R., Gerlach, C., 2018. Inversion effects for faces and objects in developmental prosopagnosia: a case series analysis. Neuropsychologia 113, 52–60.
Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S., Baker, C.I., 2009. Circular analysis in systems neuroscience: the dangers of double dipping. Nat. Neurosci. 12 (5), 535–540.
Le Grand, R., Cooper, P.A., Mondloch, C.J., Lewis, T.L., Sagiv, N., de Gelder, B., Maurer, D., 2006. What aspects of face processing are impaired in developmental prosopagnosia? Brain Cogn. 61 (2), 139–158.
Lee, Y., Duchaine, B., Wilson, H.R., Nakayama, K., 2010. Three cases of developmental prosopagnosia from one family: detailed neuropsychological and psychophysical investigation of face processing. Cortex 46 (8), 949–964.
Liu, T.T., Behrmann, M., 2014. Impaired holistic processing of left-right composite faces in congenital prosopagnosia. Front. Hum. Neurosci. 8, 750.
Lohse, M., Garrido, L., Driver, J., Dolan, R.J., Duchaine, B.C., Furl, N., 2016. Effective connectivity from early visual cortex to posterior occipitotemporal face areas supports face selectivity and predicts developmental prosopagnosia. J. Neurosci. 36 (13), 3821–3828.
Longmore, C.A., Liu, C.H., Young, A.W., 2008. Learning faces from photographs. J. Exp. Psychol.: Hum. Percept. Perform. 34 (1), 77–100.
Marr, D., Nishihara, H.K., 1978. Representation and recognition of the spatial

organization of three-dimensional shapes. Proc. R. Soc. B: Biol. Sci. 200 (1140), 269–294.
McConachie, H.R., 1976. Developmental prosopagnosia. A single case report. Cortex 12 (1), 76–82.
McKone, E., Hall, A., Pidcock, M., Palermo, R., Wilkinson, R.B., Rivolta, D., O'Connor, K.B., 2011. Face ethnicity and measurement reliability affect face recognition performance in developmental prosopagnosia: evidence from the Cambridge Face Memory Test-Australian. Cogn. Neuropsychol. 28 (2), 109–146.
McKone, E., Yovel, G., 2009. Why does picture-plane inversion sometimes dissociate the perception of features and spacing in faces, and sometimes not? Toward a new theory of holistic processing. Psychon. Bull. Rev. 16 (5), 778–797.
Megreya, A.M., Burton, A.M., 2006. Unfamiliar faces are not faces: evidence from a matching task. Mem. Cogn. 34 (4), 865–876.
Murphy, J., Cook, R., 2017. Revealing the mechanisms of human face perception using dynamic apertures. Cognition 169, 25–35.
Murphy, J., Gray, K.L.H., Cook, R., 2017. The composite face illusion. Psychon. Bull. Rev. 24 (2), 245–261.
Murphy, J., Ipser, A., Gaigg, S.B., Cook, R., 2015. Exemplar variance supports robust learning of facial identity. J. Exp. Psychol.: Hum. Percept. Perform. 41 (3), 577–581.
Palermo, R., Willis, M.L., Rivolta, D., McKone, E., Wilson, C.E., Calder, A.J., 2011. Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. Neuropsychologia 49 (5), 1226–1235.
Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat. Vis. 10 (4), 437–442.
Piepers, D.W., Robbins, R.A., 2013. A review and clarification of the terms "holistic," "configural," and "relational" in the face perception literature. Front. Psychol. 3 (559), 1–11.
Rosenthal, G., Tanzer, M., Simony, E., Hasson, U., Behrmann, M., Avidan, G., 2017. Altered topology of neural circuits in congenital prosopagnosia. Elife 6.
Rossion, B., 2008. Picture-plane inversion leads to qualitative changes of face perception. Acta Psychol. 128 (2), 274–289.
Rossion, B., 2013. The composite face illusion: a whole window into our understanding of holistic face perception. Vis. Cogn. 21 (2), 139–253.
Schmalzl, L., Palermo, R., Coltheart, M., 2008. Cognitive heterogeneity in genetically based prosopagnosia: a family study. J. Neuropsychol. 2 (Pt 1), 99–117.
Sekuler, A.B., Gaspar, C.M., Gold, J.M., Bennett, P.J., 2004. Inversion leads to quantitative, not qualitative, changes in face processing. Curr. Biol. 14 (5), 391–396.
Shah, P., Gaule, A., Gaigg, S.B., Bird, G., Cook, R., 2015a. Probing short-term face memory in developmental prosopagnosia. Cortex 64, 115–122.
Shah, P., Gaule, A., Sowden, S., Bird, G., Cook, R., 2015b. The 20-item prosopagnosia index (PI20): a self-report instrument for identifying developmental prosopagnosia. R. Soc. Open Sci. 2 (6), 140343.
Shakeshaft, N.G., Plomin, R., 2015. Genetic specificity of face recognition. Proc. Natl. Acad. Sci. USA 112 (41), 12887–12892.
Song, S., Garrido, L., Nagy, Z., Mohammadi, S., Steel, A., Driver, J., Furl, N., 2015. Local but not long-range microstructural differences of the ventral temporal cortex in developmental prosopagnosia. Neuropsychologia 78, 195–206.
Stollhoff, R., Jost, J., Elze, T., Kennerknecht, I., 2011. Deficits in long-term recognition memory reveal dissociated subtypes in congenital prosopagnosia. PLoS One 6 (1), e15702.
Susilo, T., Duchaine, B., 2013. Advances in developmental prosopagnosia research. Curr. Opin. Neurobiol. 23, 423–429.
Susilo, T., McKone, E., Dennett, H., Darke, H., Palermo, R., Hall, A., Rhodes, G., 2010. Face recognition impairments despite normal holistic processing and face space coding: evidence from a case of developmental prosopagnosia. Cogn. Neuropsychol. 27 (8), 636–664.
Susilo, T., Rezlescu, C., Duchaine, B., 2013. The composite effect for inverted faces is reliable at large sample sizes and requires the basic face configuration. J. Vision. 13 (13), 1–9.
Thomas, C., Avidan, G., Humphreys, K., Jung, K.J., Gao, F., Behrmann, M., 2009. Reduced structural connectivity in ventral visual cortex in congential prosopagnosia. Nat. Neurosci. 12, 29–31.
Todd, J.T., 2004. The visual perception of 3D shape. Trends Cogn. Sci. 8, 115–121.
Towler, J., Fisher, K., Eimer, M., 2017. The cognitive and neural basis of developmental prosopagnosia. Q. J. Exp. Psychol. 70 (2), 316–344.
Towler, J., Gosling, A., Duchaine, B., Eimer, M., 2012. The face-sensitive N170 component in developmental prosopagnosia. Neuropsychologia 50 (14), 3588–3599.
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.
Tree, J.J., Wilkie, J., 2010. Face and object imagery in congenital prosopagnosia: a case series. Cortex 46 (9), 1189–1198.
Ulrich, P.I., Wilkinson, D.T., Ferguson, H.J., Smith, L.J., Bindemann, M., Johnston, R.A., Schmalzl, L., 2017. Perceptual and memorial contributions to developmental prosopagnosia. Q. J. Exp. Psychol. 70 (2), 298–315.
Weigelt, S., Koldewyn, K., Dilks, D.D., Balas, B., McKone, E., Kanwisher, N., 2013. Domain-specific development of face memory but not face perception. Dev. Sci. 17 (1), 47–58.
Weigelt, S., Koldewyn, K., Kanwisher, N., 2012. Face identity recognition in autism spectrum disorders: a review of behavioral studies. Neurosci. Biobehav. Rev. 36 (3), 1060–1084.
White, D., Rivolta, D., Burton, A.M., Al-Janabi, S., Palermo, R., 2017. Face matching impairment in developmental prosopagnosia. Q. J. Exp. Psychol. 70 (2), 287–297.
Wilmer, J.B., Germine, L., Chabris, C.F., Chatterjee, G., Williams, M., Loken, E., Duchaine, B., 2010. Human face recognition ability is specific and highly heritable. Proc. Natl. Acad. Sci. USA 107 (11), 5238–5241.
Yovel, G., Duchaine, B., 2006. Specialized face perception mechanisms extract both part and spacing information: evidence from developmental prosopagnosia. J. Cogn. Neurosci. 18 (4), 580–593.
Zhu, Q., Song, Y., Hu, S., Li, X., Tian, M., Zhen, Z., Liu, J., 2010. Heritability of the specific cognitive ability of face perception. Curr. Biol. 20 (2), 137–142.