CORTEX 176 (2024) 37–52

Available online at www.sciencedirect.com

ScienceDirect

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

Research Report

Perceptual heterogeneity in developmental prosopagnosia is continuous, not categorical

Joseph DeGutis [a,b,*], Leah Kirsch [a,b], Travis C. Evans [a,c], Regan Fry [a], Daniel J. Lee [c,e], Maruti Mishra [a,b,d] and Alison Campbell [a,c]

[a] Boston Attention and Learning Laboratory, VA Boston Healthcare System, Boston, MA, USA
[b] Department of Psychiatry, Harvard Medical School, Boston, MA, USA
[c] Department of Psychiatry, Boston University Chobanian & Avedisian School of Medicine, Boston, MA, USA
[d] California State University, Bakersfield, CA, USA
[e] National Center for PTSD, VA Boston Healthcare System, Boston, MA, USA

ARTICLE INFO

Article history:
Received 16 September 2023
Reviewed 26 November 2023
Revised 23 January 2024
Accepted 21 March 2024
Action editor Holger Wiese
Published online 25 April 2024

Keywords:
Developmental prosopagnosia
Face perception
Holistic processing
Subgroups

ABSTRACT

Developmental prosopagnosia (DP) is associated with considerable perceptual heterogeneity, though the nature of this heterogeneity and whether there are discrete subgroups versus continuous deficits remains unclear. Bennetts et al. (2022) recently found that holistic versus featural processing deficits distinguished discrete DP subgroups, but their sample was relatively small ($N = 37$), and subgroups were defined using a single task. To characterize perceptual heterogeneity in DPs more comprehensively, we administered a broad face perception battery to a large sample of 109 DPs and 134 controls, including validated measures of face matching (Cambridge Face Perception Test – CFPT, Computerized Benton Facial Recognition Test, Same/Different Face Matching Task), holistic processing (Part-Whole Task), and feature processing (Georges Task and Part-Whole part trials). When examining face matching measures, DPs exhibited a similar distribution of performance as controls, though shifted towards impairment by an average of 1.4 SD. We next applied Bennetts (2022) hierarchical clustering approach and *k*-means clustering to the CFPT upright, inverted, and inversion index measures, similarly finding one group of DPs with poorer inverted face performance and another with a decreased face inversion effect (holistic processing). However, these subgroup differences failed to generalize to other measures of feature and holistic processing beyond the CFPT. We finally ran hierarchical and *k*-means cluster analyses on our larger battery of face matching, feature, and holistic processing measures. Results clearly showed subgroups with generally better versus worse performance across all measures, with the distinction between groups being somewhat arbitrary. Together, these findings support a continuous account of DP perceptual heterogeneity, with performance differing primarily across all aspects of face perception.

Published by Elsevier Ltd.

* *Corresponding author*. VA Boston Healthcare System (182JP), 150 S. Huntington Avenue, Boston, MA 02130, USA.
E-mail address: degutis@wjh.harvard.edu (J. DeGutis).
https://doi.org/10.1016/j.cortex.2024.03.011
0010-9452/Published by Elsevier Ltd.

1. Introduction

Developmental prosopagnosia (DP) is a severe lifelong impairment in the ability to recognize familiar faces as well as learn new faces, despite otherwise normal cognitive and visual functioning (Duchaine & Nakayama, 2006). DP affects up to 3% of the population (when considering mild and major forms, DeGutis et al., 2023), is associated with significantly compromised social functioning, and can limit employment opportunities (Murray et al., 2018; Yardley et al., 2008). Face perception deficits in DPs have been frequently observed (e.g., Barton et al., 2003; Bate et al., 2019; Behrmann et al., 2005; Biotti et al., 2019; Dalrymple et al., 2014; Duchaine et al., 2007; Le Grand et al., 2006; Mishra et al., 2021; Yovel & Duchaine, 2006) and are thought to be key mechanisms underlying DP face recognition impairments (see Biotti et al., 2019). However, studies have also demonstrated that DPs may perform within the normal range on some face perception tasks (e.g., composite face task, Biotti et al., 2017; Susilo et al., 2010) or that deficits may only occur in a subset of DPs[1] (e.g., Le Grand et al., 2006; adult DPs, Dalrymple et al., 2014; Ulrich et al., 2017). The goal of the current study was to administer a comprehensive face perception battery to a large sample of DPs ($N = 109$) to better characterize DP face perception deficits and examine whether perceptual deficits occur in discrete subgroups of DPs or rather vary on a continuum.

DPs consistently demonstrate face perception deficits when using face matching tasks, i.e., determining whether two face images are the same or different individuals, particularly when images vary in lighting, viewpoint, emotion, or a combination of these variations (e.g., Computerized Benton Facial Recognition Test – BFRT-c, Rossion & Michel, 2018; Oxford Face Matching Test, Stantic et al., 2022). However, the processes that underlie DPs' face matching deficits are currently debated. DPs have shown to have reduced, though not abolished, holistic face processing, the simultaneous integration of part and spacing information into a single perceptual representation (Rossion, 2008). Holistic processing is likely not a unitary construct and can vary depending on how it is measured (e.g., facilitation in part-whole paradigm versus interference in the composite task, for a review see Richler et al., 2012; Rezlescu et al., 2017 found little-to-no correlation between part-whole and composite effects in controls). Reflecting this variance, holistic processing deficits in DPs have been consistently observed using some measures, including the face inversion effect (e.g., Klargaard et al., 2018, Bennetts et al., 2022, though see DeGutis, Chatterjee, et al., 2012 for an exception), part-whole effect (e.g., Chapman et al., 2018; DeGutis et al., 2012), and gaze-contingent masking (Verfaillie et al., 2014), but have been less consistently shown using the composite face paradigm (Biotti et al., 2017; Susilo et al., 2010; though see Avidan et al., 2011; Palermo et al., 2011). Further, measures of holistic processing may differ in their specificity. For example, face inversion effects have been shown for face parts (for a review, see McKone & Yovel, 2009) and objects (Gerlach et al., 2023), and in some cases inverted faces have even shown to engage holistic processing (Murphy et al., 2020; Richler et al., 2011). In contrast, the part-whole effect has been shown to be more specific to upright faces (for a review, see Tanaka & Simonyi, 2016).

Beyond holistic processing, studies have demonstrated facial feature processing deficits in DPs (e.g., Le Grand et al., 2006; Yovel & Duchaine, 2006), particularly with the eye region (e.g., Berger et al., 2022; Fisher et al., 2016). This is relevant in that several critical features for determining face identity are in the eye region (eye shape, eyebrows, eye color, Abudarham et al., 2019). DPs have shown to be less sensitive to changes in the eye region (Berger et al., 2022; Fisher et al., 2016) and, compared to controls, may not use information around the eye region as much during face recognition judgments (Tardif et al., 2019). Notably, Berger et al. (2022) showed that, in a combined group of DPs and controls, holistic processing (measured by the part-whole effect) and eye discrimination ability each predicted unique variance in performance on a face matching task. Together, this suggests that holistic processing and eye processing are important aspects of DP face perception deficits. However, it remains unclear how deficits in holistic and featural/eye processing are related to one another and how they may combine to produce face perception impairments in DPs.

In addition to an incomplete understanding of the mechanisms underlying DP face perception deficits, it is currently debated whether some DPs show distinct patterns of perceptual deficits, i.e., subtypes or subgroups, or if deficits lie on a continuous spectrum (Barton & Corrow, 2016). Identifying subtypes or important continuous dimensions of perceptual ability could both improve the mechanistic understanding of DPs and lead to more targeted and individualized interventions (e.g., Corrow et al., 2019; DeGutis et al., 2014). Regarding overall face perception abilities in DP, studies have supported the continuum model, with DPs showing a similar distribution of scores as controls, but shifted towards impairment (Biotti et al., 2019; Mishra et al., 2021). Contrasting these findings, a recent study by Bennetts et al. (2022) found that DPs did not differ in overall face perception abilities, but rather discrete DP subgroups were differentiated based on holistic versus featural processing abilities. Bennetts et al. (2022) performed hierarchical and *k*-means cluster analyses of 37 DPs using the Cambridge Face Perception Test (CFPT, Duchaine et al., 2007) upright, inverted, and face inversion index scores. Across both analyses, two DP subgroups emerged with similar upright face perception performance, though one subgroup showed holistic processing deficits (reduced face inversion index, $n = 21$) while the other showed featural deficits (reduced inverted face accuracy, $n = 16$). Though this potentially suggests two separable deficiencies to produce face perception impairments in DPs, further research is necessary to confirm these findings. In particular, the CFPT inversion index was highly correlated with CFPT inverted scores (–.71), higher than reported reliabilities for CFPT inverted scores (e.g., ~.50, Rezlescu et al., 2017), which raises questions about the validity of including both measures. Other limitations are that the face inversion effect is not a very specific measure of holistic processing (e.g., Gerlach &

[1] It should be noted that these are studies of adult DPs. Results from Dalrymple et al. (2014) suggest that, in contrast to adult DPs that may or may not have face perception deficits, all child DPs they tested had face perception deficits.

Mogensen, 2023), the sample size may be inadequate for cluster analyses, and effects observed may be due to idiosyncrasies of the CFPT task (for a discussion about issues with relying on a single face perception task, see Bobak et al., 2023). To address these issues, we sought to replicate Bennetts et al. in a larger sample of DPs and further test whether there are DP subgroups based on holistic versus featural deficits across a battery of tasks.

In the current study, we sought to characterize DP perceptual heterogeneity more comprehensively by administering a broad face perception battery to a sample of 109 DPs. We assessed general face perception ability, i.e., face matching, using validated measures including the CFPT, BFRT-c (Rossion & Michel, 2018), and Same/Different Face Matching Task (SDFMT, Mishra et al., 2021; Berger et al., 2022). We assessed holistic face processing using the CFPT inversion effect index (replication of Bennetts et al., 2022) as well as the part-whole task, which is a more specific measure of holistic processing (Gerlach & Mogensen, 2023). Finally, similar to previous studies, we assessed facial feature processing by administering the Georges task (Berger et al., 2022; Pancaroglu et al., 2016) and using the eye and mouth part trials in the Part-Whole task (Fry et al., 2023). We first attempted to replicate Bennetts et al.'s HCA and *k*-means approach with CFPT measures in our larger sample. We next tested for DP subgroups using our entire battery of face perception measures and performed HCA with more widely used clustering parameters (Ferreira & Hitchcock, 2009; Abu-Jamous et al., 2015) as well as *k*-means cluster analysis.

2. Methods

2.1. Participants

A total of 109 developmental prosopagnosics (DPs) completed the study (44 in lab, 65 online). Participants were recruited via our database of DPs who previously had participated in laboratory studies, references from other research laboratories (e.g., Dr. Brad Duchaine, Dartmouth College, www.faceblind.org), advertising on the Massachusetts Bay Transportation Authority subway system, and from our study listed on clinicaltrials.gov. To be eligible for the study, all participants first underwent a brief phone screen to ensure they reported symptoms consistent with developmental prosopagnosia. Exclusionary criteria included prosopagnosic symptoms after a brain injury, a history of any significant neurological disorders (e.g., Alzheimer's Disease, multiple sclerosis), lifetime moderate or severe traumatic brain injury (TBI) or mild TBI within the last 6 months, musculoskeletal or sensory impairments that would impact test performance, lack of English proficiency, psychiatric disorders including schizophrenia or bipolar disorder, or current dependence on alcohol or other substances. Our inclusion/exclusion criteria were established prior to data analysis and were similar to our recent studies (Fry et al., 2023; Stumps et al., 2020).

All DP participants reported severe lifelong face recognition deficits and were classified into mild and major DPs based on their CFMT (using Duchaine & Nakayama, 2006 to calculate z-scores) and Famous Faces Memory Test (FFMT, using the current control sample to calculate z-scores) performance and DSM-5 neurocognitive disorder criteria (z-score ≤ -2 on CFMT and FFMT for major and z-score ≤ -1 on CFMT and FFMT for mild prosopagnosia, see DeGutis et al., 2023).[2] The vast majority of participants (95/109) were major DPs and 14 were mild DPs. Removing these participants had no appreciable effects on the key results below (see Supplementary Materials). All DPs had normal or corrected-to-normal vision, scoring within the normal threshold on the Leuven Perceptual Organization Screening Test (L-POST, $\leq$3 subtests below the 10th percentile; Torfs et al., 2014). We also administered the 20-item Prosopagnosia Index (PI20; Shah et al., 2015) to assess self-reported face recognition difficulties in daily life and the Autism Quotient questionnaire (Baron-Cohen et al., 2001), a 50-item self-report questionnaire designed to identify symptoms of autism spectrum disorder in adults. Because our recent work showed remarkably similar face perception performance in DPs with higher versus lower autism traits (Fry et al., 2023), we did not use autism traits as an exclusionary criterion.

A total of 134 control participants between the ages of 20 and 69 years old (M = 39.29, SD = 12.22, 100 females) also took part in the study (25 in lab, 109 online). Subjects were recruited from the community primarily via flyers and the Harvard Decision Science lab as well as through Prolific.com, an online subject recruitment platform. All participants had normal or corrected-to-normal vision, scoring within the normal range on the L-POST, reported having never experienced difficulties with face recognition, and scored 45 or above on the CFMT. There were 3 additional control participants that were recruited but did not complete all the tests in the battery and were thus excluded from analysis, as well as 5 control participants who were removed for evidence of reduced effort, having scored at or near chance performance on one or more of our face perception tests. All DPs were also screened to ensure they were giving full effort (e.g., not speeding through the tasks and pressing the same response key repeatedly). DPs with evidence of giving full effort but who had poor performance on the face tasks were not removed, since poor performance could be due to their prosopagnosia.

All participants gave informed consent according to the Declaration of Helsinki and in compliance with the Institutional Review Boards of the VA Boston Healthcare System and Harvard University. They were then tested at the VA Boston Medical Center in Jamaica Plain, MA, Harvard Decision Science Lab at Harvard University in Cambridge, MA, or online.

2.2. Procedure

The experiments were implemented in either PsychoPy v1.85.4 and JavaScript (for CFMT) and run on a laptop computer (34.5 × 19.5 cm display, 1920 × 1080 pixels, 60 Hz) or were

[2] This more inclusive approach to diagnosing DP is consistent with Burns et al., 2022; Lowes et al., 2024, who show that DPs diagnosed based on self-reported face recognition difficulties often show reduced accuracy and reaction times (as well as balanced integration scores in Lowes et al., 2024) on face recognition tests compared to controls. Our approach differs from theirs in that we require both self-report face recognition difficulties as well as evidence of objective deficits on two validated face recognition measures to regard a participant as having DP.

run online using equivalent paradigms on the www.testable.org platform. Participants were asked to sit 60 cm from the computer screen and instructed to indicate their responses using either a keyboard or a computer mouse. The study included 5 face perception tasks (see Fig. 1). Written instructions were provided. The order of the computerized tests was fixed across participants: a) CFPT, b) BFRT-c, c) Part-Whole Task, d) SDFMT, and e) Georges Task. This was done to reduce order-related individual differences (Ruiz et al., 2019) and to detect training-related changes more sensitively for the DPs who went on to perform cognitive training.

2.2.1. *Computerized Benton Facial Recognition Test*

The computerized Benton Facial Recognition Test (BFRT-c, Fig. 1A) is a face matching task that emphasizes speed as well as accuracy (Rossion & Michel, 2018). The test uses grayscale photographs of unfamiliar faces (3 cm × 3.5 cm) presented with little visible hair and all external information cropped out. On each trial, a target face is presented at the top of the screen with six faces below in two rows. The test consists of two sections, with the first section (6 trials) asking participants to select one out of the six faces that matches the target face (the rest of the faces have a small change in size or contrast from the target face). This is typically very easy, even for DPs. The second, critical section (16 trials, see Fig. 1A) asks participants to pick three out of the six options that match the target face. Participants are instructed, "You will see one face at the top of the screen that you will have to match to three faces presented below. Click on the 3 matching faces. Try to respond as quickly and accurately as possible." In this section, the six faces have either all lighting changes or all viewpoint changes (see Fig. 1A for viewpoint changes). The stimuli are displayed until the participant completes their responses. Scoring was calculated by receiving either a 1 or a 0 for the first 6 trials and a score of 3, 2, 1, or 0 on the remaining trials based on how many faces were correctly identified as the target. For the BFRT-c accuracy, we used the total score out of 54 items. Completion time was examined for outliers but was not used

Fig. 1 – Face perception battery. ***Note.*** **Representative stimuli from the face perception battery: A) Computerized Benton Face Recognition Test (BFRT-c), B) Cambridge Face Perception Test (CFPT), C) Same/Different Face Matching Test (SDFMT), D) Part-Whole Task, and E) Georges Task. Correct answers: A) top-middle, bottom-left, and bottom-middle faces, B) faces displayed in correct order from most to least similar from left to right, C) same identity, D) whole: left face / part: right eyes, and E) top face is different, eye horizontal spacing change.**

as a dependent measure because it was recently found not to be indicative of face recognition ability or differentiate DPs and controls above and beyond accuracy (DeGutis et al., 2022).

2.2.2. *Cambridge Face Perception Test*

The CFPT (Duchaine et al., 2007, Fig. 1B) is a computerized face sorting task. On each trial, participants arrange six front-view face images (3 cm × 4 cm) according to the similarity with the 3/4th profile view of the target face (left-being most similar, right-most dissimilar). Participants are given one minute to complete each trial. The six faces are generated by morphing a varying proportion of the identity of the target face with six new individual faces. Eight sorts were created, each with upright and inverted face trials that were intermixed in the block. Following Rezlescu et al. (2017), we calculated the correct score as (100 − % total errors) as our dependent measure. Chance level is 35.6%. The CFPT inversion index was calculated as (upright − inverted)/(upright + inverted), where a positive inversion index indicates better performance for upright than inverted faces, i.e., greater holistic processing (see Bennetts et al., 2022).

2.2.3. *Same/Different Face Matching Task*

This task (see Fig. 1C) has been previously reported in Mishra et al. (2021) and Berger et al. (2022) and has shown to robustly correlate with other face perception assessments (e.g., SDFMT *vs* BFRT-c in controls r = .59, Mishra et al., 2021). Face images were from the multi-PIE database (Gross et al., 2010) and were neutral in expression. They were converted to grayscale and cropped to remove external features such as hair or clothing. Individual foil faces were selected to be matched to each individual target so they had similar verbal descriptions (e.g., female, 20s, thin eyebrows, dark eyes). In this task, participants were presented with two face images side-by-side (4.5 × 6 cm each) for 3 sec and had to press 1 or 0 to indicate whether the faces were the same (50% of trials) or different identities (50% trials), respectively. There was a 1 sec intertrial interval. There were six different trial types: 1) Same identity from front view (face images were taken on different days), 2) different identity from front view, 3) same identity with lighting change (fully lit *vs* lit from the side), 4) different identity with lighting change, 5) same identity with viewpoint change (front view *vs* 3/4 view), and 6) different identity with viewpoint change. There were 30 trials per trial type which were randomly intermixed for a total of 180 trials.

2.2.4. *Part-Whole Task*

We used the version of the Part-Whole task from Tanaka et al. (2004) that has been previously used in individual differences (Rezlescu et al., 2017) and DP studies (DeGutis et al., 2012), see Fig. 1D. Target faces were created using the outline of one Caucasian male face. By inserting a combination of six different pairs of eyes, noses, and mouths, six unique target faces were created. For whole trials, foil faces were created by switching one of the three facial features (eyes, nose, or mouth) with that of a different target face. For part trials, foil stimuli were an isolated facial feature (eyes, nose, or mouth) from another target face. Each trial began with a central fixation display presented for 500 msec. Next, one of the six target faces was centrally presented for 1000 msec, and participants attempted to encode this face. Next, a scrambled face mask was displayed for 500 msec. During the subsequent test period, participants were presented with a pair of probe images side by side, either whole faces (whole trials) or isolated features (part trials). One image matched the target, and the other image was a foil. Stimuli remained on the screen until participants indicated with a button press which probe stimulus matched the target face (participants responded 1 for left image, 2 for right image). For whole trials, participants choose between the whole target face and a whole foil face, which was the same as the target face except that one of the features (eyes, nose, or mouth) was replaced with a foil feature. For part trials, participants choose between a face part from the target face (eyes, nose, or mouth) and the same facial feature from a foil face. On a given trial, participants are given no indication on which feature they would be tested, nor did they know whether isolated features or whole faces would be shown during the test period. There were 72 trials (36 parts trials and 36 whole trials), 24 for each feature type. We calculated the holistic advantage by first regressing the part trial 'control condition' from the whole trial 'condition of interest' (using the regression equation in the control sample, e.g., see DeGutis et al., 2013), and then applied this equation to calculate residuals for DPs and controls.

2.2.5. *Georges Task*

The Georges task has been used in previous studies as a measure of the ability to discriminate feature identities and configurations (e.g., Malcolm et al., 2004; Pancaroglu et al., 2016, see Fig. 1E). In the Georges task, participants are presented with frontal views of three same-identity unfamiliar faces in a triangular arrangement for 2 sec, with the lower two faces slightly offset horizontally. Two faces were identical and one had a single manipulation. These faces were manipulated in six ways involving three categories of change: internal feature position, feature size, and external contour. Each category of change had one manipulation in the upper face and one in the lower face. To manipulate feature position, either the eyes were edited to be closer together or the mouth was edited to be higher on the face. To manipulate feature shape, the size of the eyes or mouth was increased. To manipulate external contour, the hairline was elevated or the chin was narrowed. The different trial types were randomly intermixed so participants did not know where to expect changes. Participants indicated which of the three faces differed from the other two using the left, right, and up arrow keys. There were 108 trials, with six different face identities.

2.3. *Analyses*

2.3.1. *Reliability*

We computed Cronbach's α and Guttman's λ_2 for each of our measures of interest separately for DPs and controls (see Supplementary Table S3). For the composite eye and mouth measures, we used the Wang and Stanley (1970) composite reliability formula. For the part-whole residuals, we used the Malgady and Colon-Malgady (1991) residual reliability

formula. For the CFPT inversion index, we calculated the Spearman-Brown corrected split-half reliability of the inversion index for all the odd trials versus all the even trials.

2.3.2. Cluster analyses

First, we aimed to directly replicate the hierarchical cluster analysis (HCA) performed by Bennetts et al. (2022) by using the same cluster analysis parameters (single linkage and correlation as the distance measure) and measures, including CFPT upright, CFPT inverted, and CFPT inversion index. Second, we conducted an additional cluster analysis that utilized more widely accepted methods (Ward's linkage and Euclidean distance, Ferreira & Hitchcock, 2009; Abu-Jamous et al., 2015) and a more comprehensive battery of face processing measures, including CFPT upright, BFRT-c, SDFMT, part-whole holistic advantage, eye composite (average of Part-Whole eyes and Georges eyes trials), and mouth composite (average of part-whole eyes and Georges eyes trials). To examine the consistency of these results, we followed up both approaches by using *k*-means cluster analyses with the number of clusters indicated by HCA. By conducting the cluster analyses in this manner, we were able to gauge the degree to which results were dependent on specific tasks and data analytic approaches. Prior to conducting both sets of cluster analyses, we screened for multivariate outliers to ensure that multivariate distributional assumptions were met. Based on a Maholonobis distance criterion, we removed any multivariate outliers that exhibited significant departure from multivariate normality ($p < .001$). One DP was removed from the Bennetts HCA replication analysis with this algorithm. We also reran all these analyses on the subgroup of 95 major DPs, individuals with z-scores ≤ -2 on both the CFMT and FFMT (see Supplementary Materials).

We conducted HCA using R software with accompanying libraries (R Core Team, 2013; http://www.R-project.org/). The specific code can be found at https://osf.io/d5kbq/. Briefly, HCA is performed in a data-driven, iterative fashion that primarily aims to identify clusters of data that exhibit maximal within-cluster similarity and maximal between-cluster differences. In HCA, each participant is initially assigned to a unique cluster. Next, cluster formation is performed by combining clusters that exhibit the minimal amount of multivariate distance from each other. Throughout this process, clusters are combined until all participants are represented within a single cluster that contains all data points. By performing cluster formation in this manner, HCA identifies multiple possible cluster solutions that range from a minimum number of 2 clusters ($k = 2$) to a maximum number of clusters based on the number of participants ($k = n - 1$). Specifically, we used the nbClust library in R to evaluate potential cluster solutions based on 28 different criteria (e.g., silhouette width, Charrad et al., 2014). We evaluated a range of cluster solutions ranging from $k = 2$ to $k = 6$ based on a conservative heuristic guided by sample size (2^k; Formann, 1984).

To replicate the approach of Bennetts et al. (2022), we utilized CFPT upright, CFPT inverted, and CFPT inversion index scores, and used Pearson correlation as the multivariate distance metric to compare participants/clusters. Likewise, we similarly used the single-linkage method for cluster formation. Finally, we selected the cluster solution supported by the highest number of evaluative metrics described above. For the exploratory cluster analysis with our broader perceptual battery, we utilized several parameters that differ from those employed in Bennetts et al. (2022). First, rather than restricting our dataset to the CFPT upright, inverted, and inversion index measures (3 input measures total), we included additional face processing measures that included the BFRT-c, CFPT upright, SDFMT, part-whole holistic advantage, eye composite, and mouth composite (6 input measures total). Studies have found that the part-whole effect is a more specific holistic processing measure than the face inversion effect (Gerlach & Mogensen, 2023), so we focused on the part-whole holistic advantage. We then z-transformed these scores relative to control means and standard deviations. Second, we utilized squared Euclidean distance rather than Pearson correlation to compute multivariate distance between participants/clusters, which permits utilization of more reliable cluster-linkage methods. Finally, we used Ward's minimum variance linkage rather than single-linkage clustering, which is more robust against violations of multivariate distributional assumptions (Ward, 1963).

For all the hierarchical cluster analyses, we also aimed to determine the generalizability of cluster solutions across cluster analytic techniques. To this end, we used the *k*-means algorithm to extract the same number of clusters that were identified with HCA approaches and compared the degree to which DPs were similarly assigned to clusters using each approach. In contrast to HCA, *k*-means categorizes data such that relatively equal proportions of participants are assigned to each cluster. Further, *k*-means uses a more basic centroid algorithm to determine cluster assignment. To determine the degree of generalization across HCA and *k*-means approaches, we calculated inter-rater reliability using Cohen's Kappa (κ). Consistent with recent guidelines (McHugh, 2012), we interpreted inter-rater reliability of κ values $< .40$ as minimal, κ values between .40 and .59 as weak, κ values between .60 and .79 as moderate, and κ values $> .80$ as excellent.

2.3.3. Using the DSM-5 approach to determine prevalence of major and mild perceptual impairments in DPs

Considering our large DP sample, we sought to determine how many DPs have mild and major face perception deficits and how many are unimpaired perceptually. We used the DSM-5 approach to determining mild and major neurocognitive impairment (Sachdev et al., 2014). Based on poor diagnostic reliability in using any single measure (Holdnack et al., 2017), the DSM-5 recommends using at least two objective validated cognitive measures within a domain showing impairment (z-score ≤ -2 for major and ≤ -1 for mild neurocognitive disorder). For our objective measures, we used the previously validated CFPT, BFRT-c, and SDFMT face matching tasks (Mishra et al., 2021), with z-scores calculated from our control group. We also examined how our control sample performed across our battery of matching tasks. We first re-calculated

control z-scores on each task using a leave-one-subject-out (LOSO) approach (Esterman et al., 2010). In this method, a single subject is iteratively left out of the first stage group analysis, in this case calculating means and standard deviations to derive z-scores, allowing the rest of the group to serve as an independent sample on which to compare each individual.

2.4. Sample size justification

Our sample size was guided by several factors. First, we wanted to have a substantially larger sample than Bennetts et al. (2022), who performed HCA and *k*-means using a sample of 37 DP participants. Also, guidelines suggest that the minimum cluster analysis sample size should be 10 times the number of variables (Qiu & Joe, 2009). Considering that we had six variables when using our broader face perception battery, this recommends a minimum sample size of 60 DPs. Finally, when running simulated data sets through cluster analyses, Dalmaijer et al. (2022) found that 20–30 participants per subgroup were necessary to detect reliable subgroup differences. Considering that we wanted sensitivity to detect up to 3–4 potential subgroups, we recruited a sample size of 109 DPs.

2.5. Stimuli, analysis code, and data availability and preregistration

Cluster analysis code and de-identified data is available at https://osf.io/d5kbq/. Regarding the face stimuli, because sufficient authorization was not obtained to publicly share images of individuals' faces, we are unable to make the face stimuli publicly available. However, face stimuli can be obtained by emailing the corresponding author on the original publication for each task. No part of the study procedures or analyses were pre-registered prior to the research being conducted. We report how we determined our sample size, all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study.

3. Results

3.1. Participants and diagnostic test performance

Our sample consisted of 109 DPs (77 females, 3 nonbinary) and 134 controls (100 females) with a similar mean age of 41.11 (SD = 14.85) and 39.29 (SD = 12.22), respectively (see Table 1). According to the DSM-5 criteria of cognitive impairment ($z \leq -1$ on two or more tests for mild, ≤ -2 for major), based on diagnostic face recognition measures our DP sample included 95 major DPs and 14 mild DPs.[3] As expected, the entire DP group performed substantially worse than controls on the CFMT and FFMT (Table 1). The control group CFMT M and SD were similar to the original Duchaine and Nakayama (2006) study and there were only three controls with scores of 70 or above, suggesting that controls with very high face recognition abilities were not driving DP versus control group differences.

3.2. Developmental prosopagnosics versus controls face perception performance and reliability

Before examining perceptual heterogeneity in DPs, we confirmed that the DP group had worse face perception abilities compared to controls. We briefly summarize the findings, though more details can be found in the Supplementary Material and Table S1. Similar to previous studies (e.g., Mishra et al., 2021), DPs performed significantly worse at face matching than controls on the CFPT, BFRT-c, and SDFMT (Cohen's $d = 1.2, 1.6, 1.2$, respectively; all p's $< .001$, see Fig. 1). DPs also showed evidence of reduced holistic processing compared to controls on the CFPT task (reduced inversion effect, $p < .001$) and Part-Whole task (reduced part-whole advantage, $p < .001$, see Supplementary Table S1 for full results). Finally, DPs displayed evidence of reduced feature processing compared to controls in the Georges task (particularly the eye spacing, eye size, and mouth spacing trials) and the Part-Whole part trials (particularly the eye part trials). To assess DPs' feature processing abilities more reliably, similar to Berger et al. (2022), we averaged Part-Whole part trials (e.g., eye part trials) and conditions from the Georges task (e.g., eye size and eye spacing trials). We did this separately for the eye region and mouth region since some DPs have shown to have selective deficits with the eye region (Berger et al., 2022).

For the face perception measures of interest, we separately calculated reliability/internal consistency for DPs and controls (see Supplementary Table S2). Consistent with previous studies (Berger et al., 2022; Bobak et al., 2023; Rezlescu et al., 2017), we found that only a fraction of our measures showed Guttman's $\lambda_2 > .70$, including the SDFMT in DPs and controls, CFPT in DPs, and eye composite in controls. Most of the other measures were in the .59–.70 range, while lower reliabilities were found when examining difference scores, residuals, and inversion index scores (.33–.54).

3.3. Distributional analyses of face matching tests

With DPs demonstrating clear face perception deficits similar to previous studies, we sought to test the subgroup versus shifted distribution models of DP face perception performance using our three validated face matching measures. As can be seen in Fig. 2, DPs had consistently lower accuracy on face matching tasks (M difference = 1.4 SD), though the distribution of DP perceptual performance was generally similar to controls across the measures. To test the similarity of the DP and control distributions, we first shifted the DP data for each task by adding the difference between the DP and control means to each DP score and then performed two-sample Kolmogorov–Smirnov tests on the distributions. Results indicated no significant differences between DP and control distributions across all the face matching measures (BFRT-c $z = 1.00$, $p = .268$; CFPT $z = .90$, $p = .400$; SDFMT $z = .82$, $p = .518$). These results support the shifted distribution model of DP perceptual performance rather than DPs having more

[3] We conducted all analyses excluding the 14 mild DPs and the results were similar with a few exceptions, see Supplementary Materials.

Table 1 – Developmental prosopagnosic versus control group demographics and face recognition assessments.

	N	Gender	Age	CFMT	FFMT	PI20
DP	109	77:29:3	41.11 ± 14.85	39.17 ± 5.15	.32 ± .17	84.60 ± 8.21
Control	134	100:34	39.29 ± 12.22	61.57 ± 7.48	.80 ± .12	37.00 ± 8.04
p-value	–	.989	.306	<.001	<.001	<.001
Cohen's *d*	–	–	.14	3.42	3.31	5.86

Note. Mean ± Standard Deviation, Gender = Female:Male:Non-binary, CFMT = Cambridge Face Memory Test, FFMT = Famous Faces Memory Test, PI20 = 20-Item Prosopagnosia Index.

distinct subgroups (e.g., such as a bimodal distribution would indicate).

3.4. *Cluster analyses*

Though DPs' face matching performance across the BFRT-c, CFPT, and SDFMT showed shifted distributions compared to controls (and showed shifted distributions for other feature and holistic processing measures, see Supplementary Fig. S1), this does not rule out the possibility of discrete DP subgroups. As suggested by Bennetts et al. (2022), DPs may exhibit more complex differences across face processing measures such as exhibiting similar face matching performance, but differential holistic versus featural processing. To test this, we took two approaches. First, we directly replicated the Bennetts et al. (2022) DP hierarchical cluster analysis (HCA), using the same CFPT upright, inverted, and inversion index measures and Pearson correlation as a distance measure. We also assessed how the subgroups identified using this approach performed on other measures and additionally performed *k*-means cluster analysis of the CFPT measures. Our second approach to test for DP subgroups involved performing HCA using our broader set of face processing measures including several measures of face matching and feature processing and a more specific measure of holistic processing (part-whole effect). In this HCA, we implemented more standard clustering techniques, using a Euclidean multivariate distance metric in conjunction with Ward's linkage algorithms (Ferreira & Hitchcock, 2009; Abu-Jamous et al., 2015). Finally, to examine the consistency of the subgroups, we applied *k*-means cluster analysis to this broad set of face processing measures.

3.5. *Hierarchical cluster analysis using CFPT measures (Bennetts 2022 replication)*

Identical to Bennetts et al. (2022), in our DP sample we used DP-relative z-scores of the CFPT upright, CFPT inverted, and CFPT inversion index, selected Pearson correlation as the distance measure, and applied a single-linkage HCA. We identified an optimal cluster solution of six clusters (favored by 9/28 metrics), which slightly outperformed both a five-cluster solution and a two-cluster solution (favored by 7/28 metrics), as well as all

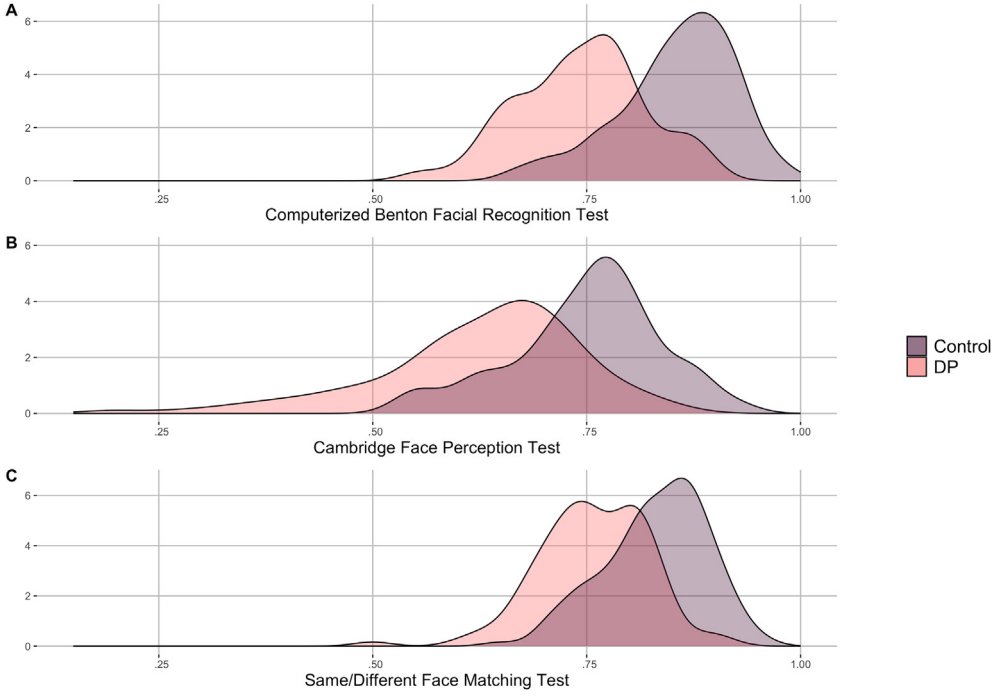

Fig. 2 – Developmental prosopagnosic and control group face matching task performance. *Note.* Each x-axis indicates task accuracy and y-axis is normalized density. DP and control distributions are displayed with the same area under the curve for the sake of comparison.

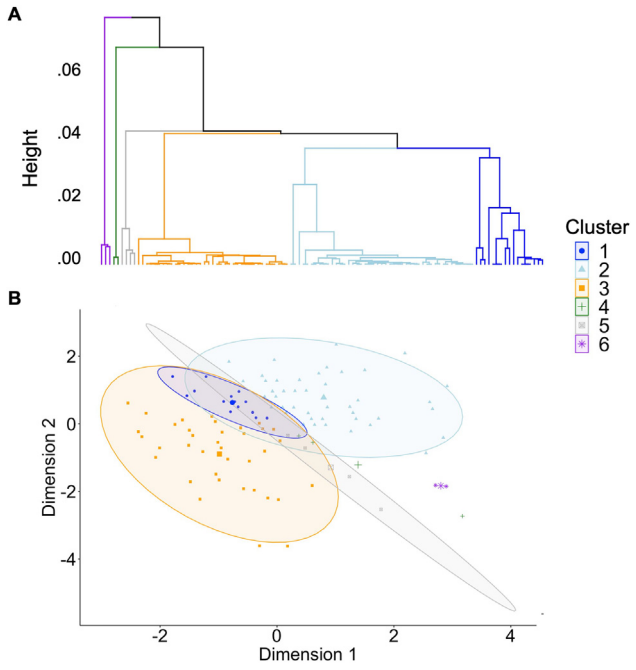

Fig. 3 – Bennetts replication of DP hierarchical cluster analysis using CFPT measures. *Note.* A) Dendrogram showing the 6-cluster solution using the Bennetts et al. (2022) hierarchical cluster analysis approach using the Cambridge Face Perception Test upright, inverted, and inversion index. B) Scatterplot of 2-dimensional visualization of the 6 clusters, with dimensions being scaled scores from the first two extracted components of a principal components analysis performed on the CFPT measures. Larger markers denote the center of each cluster distribution.

other potential cluster solutions (each favored by ≤ 4 metrics). Within the six-cluster solution, most participants were assigned to Cluster 1 (n = 17), Cluster 2 (n = 37), or Cluster 3 (n = 45). In contrast, only three participants were assigned to Cluster 4, only four participants into Cluster 5, and only two participants were assigned to Cluster 6.[4] Given the extremely small number of participants assigned to Clusters 4, 5, and 6, we focused on results for Clusters 1, 2, and 3 (see Fig. 3).

As can be seen in Fig. 4A, the CFPT results of the larger clusters showed a very similar pattern to the results of Bennetts et al. (2022), with Cluster 2 and Cluster 3 exhibiting differential featural versus holistic processing performance. Specifically, Clusters 2 (n = 37) and 3 (n = 45) performed comparably on the upright CFPT ($M_{Cluster2}$ = −1.21, $SD_{Cluster2}$ = 1.14; $M_{Cluster3}$ = −1.52, $SD_{Cluster3}$ = 1.14; p > .05), while Cluster 2 performed better on the inverted CFPT trials ($M_{Cluster2}$ = .84, $SD_{Cluster2}$ = .48; $M_{Cluster3}$ = −.69, $SD_{Cluster3}$ =.55; p < .001) and Cluster 3 showed a larger CFPT inversion index ($M_{Cluster2}$ = −1.67, $SD_{Cluster2}$ = .68; $M_{Cluster3}$ = −.21, $SD_{Cluster3}$ = .48; p < .001). These results suggest that DPs assigned to Cluster 2 exhibited greater holistic processing deficits, whereas DPs assigned to Cluster 3 exhibited greater featural processing deficits. A notable difference from the Bennetts et al. solution is that we found a third, smaller cluster in our solution (Cluster 1, n = 17). This cluster performed much better than the larger two clusters on the CFPT upright trials while also performing well on the inverted trials and only having a mildly impaired inversion index, indicative of overall milder perceptual deficits than the other two clusters.

To determine the generalizability of this pattern, we examined how these CFPT-derived DP clusters performed on face matching, featural, and holistic processing measures beyond the CFPT. In contrast to the CFPT results, these clusters showed generally more similar performance across the extended battery of measures, with a few exceptions. For face matching tasks, Cluster 3 scored numerically worse than Cluster 2 on both the BFRT-c (M = −1.88, SD = 1.02; M = −1.46, SD = 1.14, respectively) and SDFMT (M = −1.61, SD = 1.21; M = −.87, SD = .99, respectively), despite scoring similarly on the CFPT. Notably, Cluster 3, which performed worse on CFPT inverted trials suggesting featural processing deficits, did not score worse on featural processing when examining the eye and mouth composite scores. Likewise, Cluster 2, which had a reduced CFPT face-inversion effect, did not display significantly worse holistic processing than Cluster 3 on the Part-Whole holistic advantage (Cluster 2 M = −.70, Cluster 3 M = −.74, p > .05). Cluster 1 also had a different pattern of performance than for the CFPT, showing deficient performance across all measures except for relatively normal holistic processing on the Part-Whole task.

3.6. *K-means cluster analysis using CFPT measures (Bennetts 2022 replication)*

To examine the consistency of the Bennetts et al. (2022) approach across analytic techniques, we next performed k-means cluster analysis of the CFPT measures (see Supplementary Materials) and examined the agreement of DP assignment between k-means and HCA methods. We removed the participants placed into Clusters 4, 5, and 6 before selecting a 3-cluster k-means solution. This k-means solution was very similar to the HCA solution above, identifying one less impaired cluster and two clusters with differential deficits on the CFPT inverted trials and inversion index. In the expanded battery, the three clusters similarly showed generally better versus worse performance (see Supplementary Fig. S2). The inter-rater reliability between the HCA and k-means approaches indicated a modest agreement between cluster assignments (κ = .53, p < .001).

In sum, the similar patterns observed in the HCA and k-means analysis of CFPT measures partially replicate the clusters reported by Bennetts et al. (2022), but importantly these clusters did not generalize to perceptual performance on tasks beyond the CFPT. Differences in holistic and featural processing on the CFPT were not found when examining other valid measures of holistic and featural processing. Together, this suggests that subgroups identified by the Bennetts approach are tenuous at best and may capture differences

[4] The smaller number of participants in Clusters 4, 5, and 6 may be attributable to the fact that single-linkage clustering methods are highly sensitive to divergent data points that may produce small clusters within a dataset (Almeida et al., 2007).

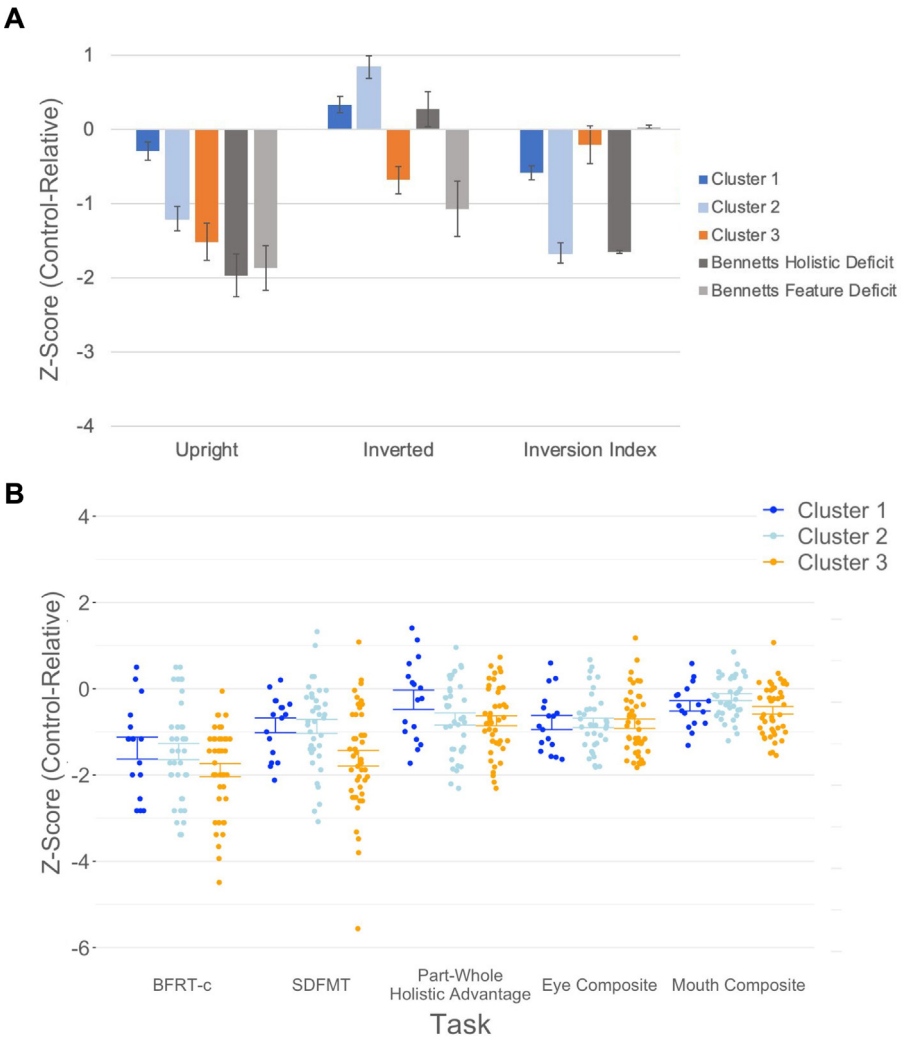

Fig. 4 – Bennetts replication of CFPT-derived DP hierarchical cluster analysis: Cluster performance on (A) CFPT and (B) Extended Perception Battery. *Note.* A) Average DP Cambridge Face Perception Test (CFPT) performance from our three large clusters (Clusters 1, 2, and 3 see Fig. 3) compared to Bennetts et al. (2022) clusters (plotted via reported means, standard deviations, and sample size). B) Clusters 1, 2, and 3 expanded face perception battery performance. Upright = CFPT upright trial accuracy, Inverted = CFPT inverted trial accuracy, Inversion Index = [CFPT upright – CFPT inverted]/[CFPT upright + inverted], BFRT-c = Computerized Benton Face Recognition Test, SDFMT = Same/Different Face Matching Task, Eye and Mouth Composites = Part-Whole and Georges eye and mouth trials. Error bars represent standard error.

specific to the CFPT upright, inverted, and inversion index measures, rather than more general DP ability differences.

3.7. *Hierarchical cluster analysis using multiple measures of face matching, holistic, and feature processing*

In a further attempt to identify subgroups of DPs, we performed cluster analyses using our broader battery of face processing measures, including face matching (CFPT, BFRT-c, SDFMT), a more specific measure of holistic processing (Part-Whole holistic advantage), and measures of eye and mouth feature processing ability (composite of Part-Whole part trials and Georges task). We also used Ward's method for clustering linkage and Euclidean distance as a distance metric (Ferreira & Hitchcock, 2009; Abu-Jamous et al., 2015). Using this approach, we observed an optimal cluster solution of two clusters (favored by 12/28 metrics, see Fig. 5), outperforming a three-cluster solution (favored by 9/28 metrics) and all other potential solutions (favored by ≤ 4/28 metrics).

As can be seen in Fig. 6A, Cluster 1 (n = 91) was characterized by better performance across all CFPT measures, while Cluster 2 (n = 18) was characterized by worse performance across all measures. This contrasts Bennetts et al. (2022) results and our Bennetts replication results above using CFPT measures (Fig. 4A), where clusters of DPs had similar CFPT upright but differential CFPT inverted and inverted index scores. We also found that Cluster 1 consistently outperformed Cluster 2 on our other face matching

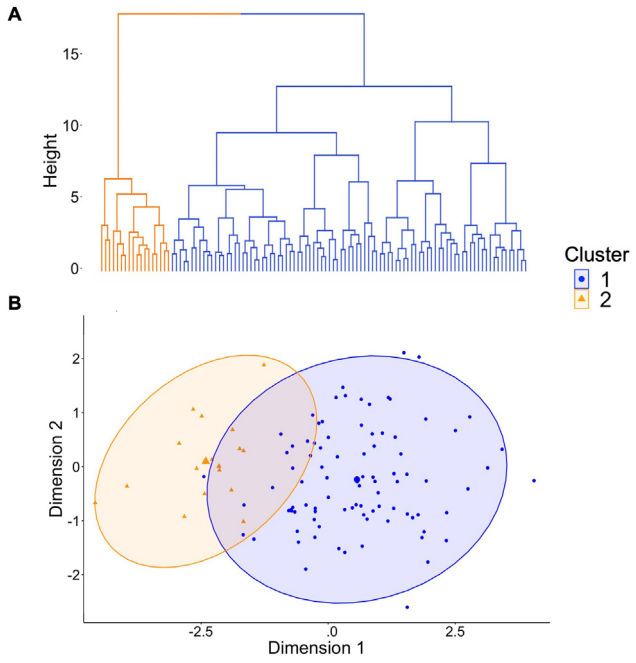

Fig. 5 – DP hierarchical cluster analysis using an extended battery of face perception measures. *Note.* A) Dendrogram showing the 2-cluster solution when using hierarchical cluster analysis (Ward's Method) and the Cambridge Face Perception Test upright, Computerized Benton Facial Recognition Test, Same/Different Face Matching Task, holistic processing (Part-Whole Task), and feature processing (Georges Task and Part-Whole part trials). B) Scatterplot of 2-dimensional visualization of the 2 clusters, with dimensions being scaled scores from the first two extracted components of a principal components analysis performed on the entire perceptual battery. Larger markers denote the center of each cluster distribution.

(BFRT-c, SDFMT), holistic processing (Part-Whole holistic advantage), and feature processing (eye and mouth composites) measures, see Fig. 6B.

3.8. *K-means cluster analysis using multiple measures of face matching, holistic, and feature processing*

We next performed *k*-means cluster analysis on the larger perceptual battery and examined the agreement of DP assignment between *k*-means and HCA methods. Similar to the HCA cluster analyses above, the *k*-means solution identified two clusters with slightly below average versus very below average perceptual performance across the entire battery (see Supplementary Fig. S3). The main difference was that the number of DPs assigned to each cluster was more evenly distributed using *k*-means than HCA (*k*-means: $n = 56$ *vs* $n = 53$; HCA: $n = 92$ *vs* $n = 17$). Because of this, we observed only a modest degree of inter-rater reliability between the *k*-means and HCA cluster solutions ($\kappa = .35$, $p < .001$). In short, both *k*-means and HCA clustering approaches similarly grouped DPs into generally better versus worse perceptual performance, though *k*-means and HCA clustering approaches did somewhat differ in which DPs were assigned to better or worse performance clusters.

3.9. *Summary of cluster analyses*

First, we were able to replicate the Bennetts HCA and *k*-means CFPT-derived clustering approaches for the CFPT measures, finding clusters with differential holistic versus feature processing. However, this pattern failed to generalize to other holistic and featural measures, suggesting that the clusters identified were highly task-dependent (see Fig. 4B). Further, the results of HCA and *k*-means clustering using our expanded battery of face perception measures demonstrated that perceptual heterogeneity in DPs is more of a global unidimensional continuum rather than DPs having differential featural versus holistic processing deficits. When performing these cluster analyses using our expanded task battery in a subset of 95 DPs with major face recognition impairment (see Supplementary Materials), the two-cluster solution replicated this global unidimensional perceptual continuum, while the preferred three-cluster solution showed subgroups with differential Part-Whole holistic processing versus eye processing. However, this pattern did not generalize to other holistic (CFPT inversion index) or feature processing measures (CFPT inverted or mouth composite) and notably, global perceptual ability better accounted for variance in DP performance than holistic versus feature processing. Finally, the different clusters chosen by HCA and *k*-means approaches for the expanded battery suggests that assignment of individual DPs varies somewhat depending on the clustering approach. Additionally, the highly overlapping distributions of individual DPs across these clusters (see Figs. 5B and 6B and Supplementary Figs. S2 and S3) indicate fuzzy boundaries between clusters, which is consistent with more dimensional differences in face processing across DPs rather than categorical.

3.10. *What is the prevalence of major and mild face perception impairments in DPs?*

Our final analysis was to use our large sample of DPs to quantify the prevalence of major and mild face perception impairments. Because both the shifted distribution and cluster analyses suggest that DP face perception deficits lie on a continuum, we used the rigorous psychometric approach the DSM-5 recommends (z-score ≤ -2 on 2 or more tests for major and z-score ≤ -1 on 2 or more tests for mild) to identify mild and major cognitive deficits in populations with continuous deficits (e.g., mild cognitive impairment, dementia, Sachdev et al., 2014). We applied these criteria to our three most reliable, validated measures of face perception (BFRT-c, CFPT, SDFMT). This identified 27 DPs (24.8%) with major face perception impairments, 45 DPs (41.3%) with mild impairments, and 37 unimpaired DPs (33.9%). Interestingly, even many of the 37 unimpaired DPs showed below average performance on multiple assessments (e.g., BFRT-c and eye composite, see Supplementary Materials and Fig. S4). For comparison, when performing leave-one-subject-out (see Methods) z-score calculations with controls, we found that no

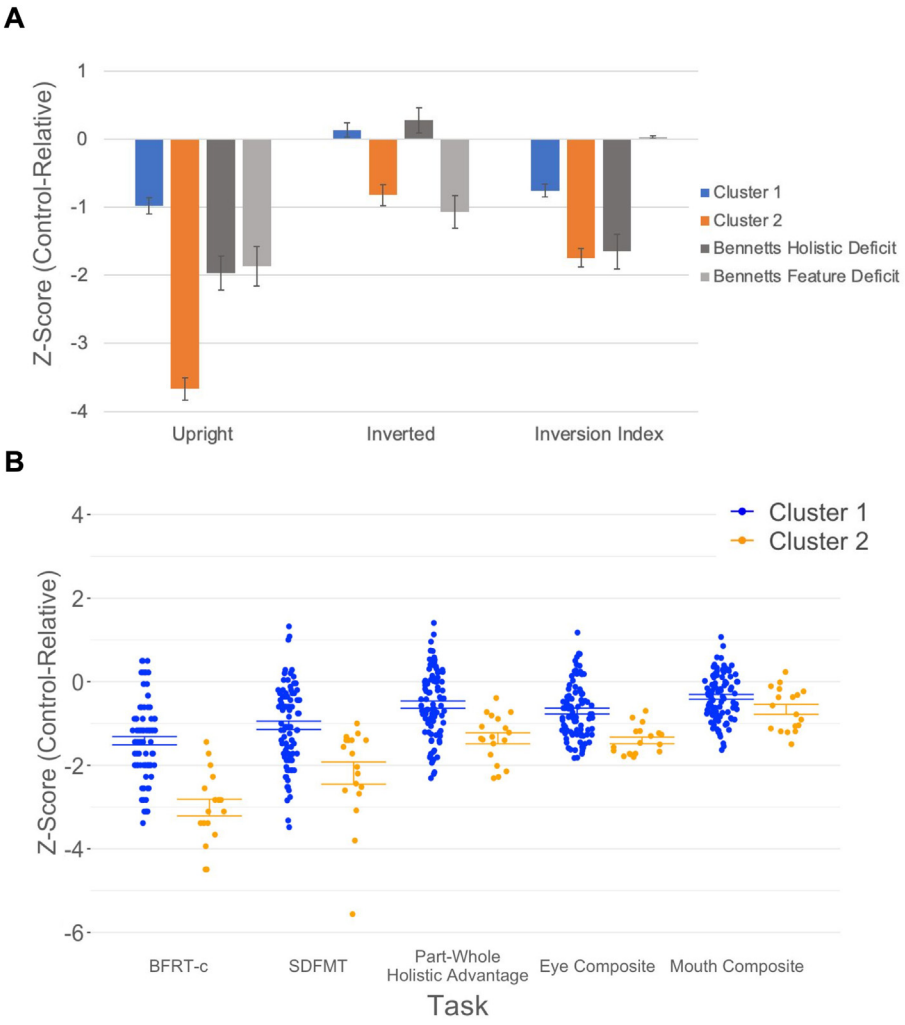

Fig. 6 – DP hierarchical cluster analysis using an extended battery of face perception measures: Cluster performance on (A) CFPT measures and (B) Larger Perception Battery. ***Note.*** **A) Average DP Cambridge Face Perception Test (CFPT) performance from our two observed clusters compared to Bennetts et al. (2022) clusters (plotted via reported means, standard deviations, and sample size). B) DP expanded face perception battery performance by cluster. Upright = CFPT upright trial accuracy, Inverted = CFPT inverted trial accuracy, Inversion Index = [CFPT upright − CFPT inverted]/[CFPT upright + inverted], BFRT-c = Computerized Benton Face Recognition Test, SDFMT = Same/Different Face Matching Task, Eye and Mouth Composites = Part-Whole and Georges eye and mouth trials. Error bars represent standard error of the mean.**

controls were majorly impaired, 12 (9.0%) had mild impairments, and 122 controls (91.0%) were unimpaired.

4. Discussion

The present study in a large group of DPs clearly demonstrates that DP face perception performance falls on a unidimensional continuum (i.e., generally better *vs* worse overall face perception) shifted towards impairment rather than having discrete face perception subgroups. This was reflected both in distributional analyses of validated face matching tasks (CFPT, BFRT-c, SDFMT) and across multiple cluster analyses, either using CFPT measures or our broader battery of face matching, featural, and holistic measures. We also quantified the prevalence of face perception impairments using DSM-5 guidelines, estimating that 24.8% of DPs have major face perception deficits, while 41.3% have mild deficits, and 33.9% are perceptually unimpaired. These results provide both mechanistic insights into DP as well as have important clinical and treatment implications.

First, we found consistent evidence to support the continuous/shifted distribution model of DP face perception deficits, rather than the discrete subgroups model (see Barton & Corrow, 2016). This was clear when examining our face matching tests, the CFPT, BFRT-c, and SDFMT, which demonstrated that the distribution of DP performance was similar to controls and shifted towards impairment by an average of 1.4 SD across tasks. We also showed this pattern of graded DP impairments on holistic and featural processing as

well (see Supplementary Fig. S1). Further, although cluster analyses identified putative subgroups of DPs, these subgroups largely exhibited unidimensional patterns of perceptual deficits and were consistently overlapping, arguing against the clusters representing qualitatively distinct subgroups. These findings align with previous DP studies demonstrating graded perceptual deficits in face matching ability (Biotti et al., 2019; Mishra et al., 2021), holistic face processing (DeGutis, Cohan, et al., 2012), and feature processing abilities (Berger et al., 2022). The current findings extend these studies by using a larger sample of DPs and demonstrating results across a broader battery of face perception tasks. Our observation of continuous face perception deficits in DPs is consistent with other developmental disorders having continuous dimensions rather than qualitatively distinct subgroups, including aspects of autism (e.g., repetitive behaviors, Zheng et al., 2019) and attention deficit hyperactivity disorder (e.g., inattention, hyperactivity/impulsivity, Marcus & Barry, 2011). In fact, a recent meta-analysis of 187 studies of psychopathological disorders (including developmental disorders) found that roughly 5 times more disorders showed evidence that individual differences in underlying factors were continuous rather than categorical (Haslam et al., 2020). Together, this suggests our finding that DP perceptual deficits exist on a continuum is consistent with symptom profiles of many other developmental and psychological disorders.

The current findings support that DPs most strongly differ on a single general perceptual dimension, rather than primarily exhibiting differential featural versus holistic processing deficits, as suggested by Bennetts et al. (2022). When using the same HCA and *k*-means cluster analysis of CFPT measures as Bennetts et al. (2022), we similarly observed DP subgroups with differential featural versus holistic processing deficits on the CFPT. However, when we compared these resulting subgroups on other feature and holistic (Part-Whole) processing tasks, we failed to find any performance differences. This suggests that the clusters found using the Bennetts approach are not generalizable but rather are specific to the CFPT measures. Further, we found a similar pattern of measure-specific Part-Whole holistic advantage versus featural processing differences (not generalizing to the mouth composite, CFPT inverted, or CFPT inversion index) for the cluster analysis of the extended battery in the 95 major DPs (see Supplementary Materials). Together, this suggests that DPs can exhibit differential feature versus holistic processing abilities but these may be relatively specific to how these abilities are measured. It also suggests that differential feature versus holistic processing is secondary to general perceptual abilities. This is consistent with studies showing that eye/feature processing (Fisher et al., 2016; Tardif et al., 2019) and holistic processing abilities (DeGutis, Cohan, et al., 2012; Klargaard et al., 2018) are unique and important contributors to DPs' overall face perception abilities (see Berger et al., 2022; along with other factors such as preferential fixation location, Peterson et al., 2019, see below), but on their own do not differentiate DPs as well as overall face perception abilities. This can be seen in Figs. 4B and 6B as well as Supplementary Figs. S11B and S12B, where DPs span a substantially greater range of performance on face matching measures than on holistic and featural processing measures.

Our DSM-5-based prevalence estimates suggest that DP face perception impairments are common but in most cases do not fully explain the severity of face recognition deficits. Several previous studies have quantified DP face perception deficits using single measures (e.g., Biotti et al., 2019; Murray et al., 2022; Stantic et al., 2022). However, single measures often show limited reliability and may only modestly correlate with other face perception measures (see Bobak et al., 2023). The current study is the first to use multiple validated face perception measures in a large sample and apply DSM-5 criteria to better quantify the prevalence of face perception impairments. We found that 24.8% of DPs had major face perception impairments, while 41.3% had mild impairments, and 33.9% were unimpaired, i.e., 2 out of every 3 DPs had some degree of face perception impairment. This could suggest the need to use more sensitive face perception tests (e.g., Oxford Face Matching Test, Stantic et al., 2022) and incorporate more sophisticated methods (e.g., confidence ratings to generate face matching ROC curves, Fitousi, 2023) to better quantify the extent of DPs' face perception deficits. It also suggests that face perception impairment alone may not explain the full extent of face recognition deficits in DPs and other deficient processes are likely involved. One potential extra-perceptual mechanism is face recollection memory, the all-or-none retrieval of qualitative, contextual, or semantic information associated with the face (Yonelinas, 2002). Using an old-new unfamiliar face recognition task with confidence ratings, Stumps et al. (2020) found that deficient face recollection explained variance in DP versus control group membership above and beyond face perception deficits. It could be that deficits in face perception and memory interact and even mild face perception deficits alongside face memory deficits could combine to produce DPs' severe face recognition difficulties. Further characterizing memory mechanisms in DPs and the interaction between perception and memory deficits would be a fruitful future direction.

The current findings have important clinical and treatment implications for DP. First, they suggest that individual differences in DPs may be best captured by more general face perception tests, such as the CFPT and BFRT-c. They also suggest that most DPs have some face perception deficit (even perceptually 'unimpaired' DPs were worse than controls, see Supplementary Fig. S4) and may benefit from training perceptual mechanisms (e.g., face similarity training, Corrow et al., 2019; holistic processing training, DeGutis et al., 2014). The current results also emphasize the need to develop additional interventions to target extra-perceptual deficits. For example, repetition lag training, which focuses on improving face recollection, has recently shown promising benefits in improving face memory in DPs, especially those with better face perception abilities (Kirsch et al., 2023).

Though the results of the present study are compelling, they have limitations. Despite recruiting a large sample of DPs and using multiple clustering approaches, cluster analyses are more stable with larger samples, and we could have employed additional clustering techniques (e.g., latent profile analysis).

That being said, when we combined DPs and controls to increase the sample size and reran HCA and *k*-means analyses as well as latent profile analysis (see Supplementary Materials, Fig. S14), we found very similar results of unidimensional grouping. Another limitation is that we did not characterize all aspects of face perception (e.g., eye movements and preferential fixation location, Peterson et al., 2019), and there may be other behavioral or neural markers that would identify discrete DP subgroups. Large DP fMRI studies have failed to show any evidence of discrete DP perceptual subgroups (e.g., $N = 64$ DPs, Liu et al., 2021), though eye tracking in a sample of 22 DPs by Peterson et al. (2019) has suggested that DPs with upper versus lower preferential fixation locations have markedly better and worse face perception performance, respectively (Bobak et al., 2017 also found that lower face fixation is associated with more severe DP). It would be important to assess a larger sample of DPs to see if this represents a qualitative difference. A final limitation is that our perceptual battery did not allow us to separate match versus non-match trial performance (except for the SDFMT). A recent study of healthy controls by Bobak et al. (2023) administered a battery of face match versus non-match tasks and found subgroups who were either unbiased, prone to say match, or prone to say non-match. It would be interesting to examine if DPs cluster into similar 'bias' subgroups.

In sum, using a large group of 109 DPs and performing several cluster analyses, we found that face perception heterogeneity in DPs is primarily characterized by a unidimensional continuum rather than discrete subgroups with differential holistic versus feature processing. We also found that the majority of our DPs had mild face perception deficits (41.3%), while 24.8% had major face perception impairments, and 33.9% were perceptually unimpaired. This reinforces that face perception deficits are an integral aspect of DP and are highly heterogeneous. It also highlights the importance of future work to go beyond face perception mechanisms to fully understand developmental prosopagnosia.

Open practices

The study in this article has earned Open Data badge for transparent practices. The data and materials used in this study are available at: https://osf.io/d5kbq/.

CRediT authorship contribution statement

Joseph DeGutis: Conceptualization, Funding acquisition, Investigation, Project administration, Supervision, Writing – original draft, Writing – review & editing. **Leah Kirsch:** Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Travis C. Evans:** Formal analysis, Methodology, Supervision, Visualization, Writing – review & editing. **Regan Fry:** Data curation, Investigation, Project administration, Writing – review & editing. **Daniel J. Lee:** Formal analysis, Methodology, Supervision, Writing – review & editing. **Maruti Mishra:** Conceptualization, Writing – review & editing. **Alison Campbell:** Supervision, Writing – review & editing.

Acknowledgments

We want to thank the developmental prosopagnosics and control participants for completing our challenging battery of tasks. This study was funded by R01 from the National Eye Institute grant #R01EY032510 awarded to JD.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2024.03.011.

REFERENCES

Abu-Jamous, B., Fa, R., & Nandi, A. K. (2015). *Integrative cluster analysis in bioinformatics*. https://doi.org/10.1002/9781118906545

Abudarham, N., Shkiller, L., & Yovel, G. (2019). Critical features for face recognition. *Cognition, 182*, 73–83.

Almeida, J. A. S., Barbosa, L. M. S., Pais, A. A. C. C., & Formosinho, S. J. (2007). Improving hierarchical cluster analysis: A new method with outlier detection and automatic clustering. *Chemometrics and Intelligent Laboratory Systems, 87*(2), 208–217.

Avidan, G., Tanzer, M., & Behrmann, M. (2011). Impaired holistic processing in congenital prosopagnosia. *Neuropsychologia, 49*(9), 2541–2552.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*, 5–17.

Barton, J. J., Cherkasova, M. V., Press, D. Z., Intriligator, J. M., & O'Connor, M. (2003). Developmental prosopagnosia: A study of three patients. *Brain and Cognition, 51*(1), 12–30.

Barton, J. J., & Corrow, S. L. (2016). The problem of being bad at faces. *Neuropsychologia, 89*, 119–124.

Bate, S., Bennetts, R. J., Tree, J. J., Adams, A., & Murray, E. (2019). The domain-specificity of face matching impairments in 40 cases of developmental prosopagnosia. *Cognition, 192*, Article 104031.

Behrmann, M., Avidan, G., Marotta, J. J., & Kimchi, R. (2005). Detailed exploration of face-related processing in congenital prosopagnosia: 1. Behavioral findings. *Journal of Cognitive Neuroscience, 17*(7), 1130–1149.

Bennetts, R. J., Gregory, N. J., Tree, J., Luft, C. D. B., Banissy, M. J., Murray, E., Penton, T., & Bate, S. (2022). Face specific inversion effects provide evidence for two subtypes of developmental prosopagnosia. *Neuropsychologia, 174*, Article 108332.

Berger, A., Fry, R., Bobak, A. K., Juliano, A., & DeGutis, J. (2022). Distinct abilities associated with matching same identity faces versus discriminating different faces: Evidence from individual differences in prosopagnosics and controls. *Quarterly Journal of Experimental Psychology, 75*(12), 2256–2271.

Biotti, F., Gray, K. L., & Cook, R. (2019). Is developmental prosopagnosia best characterised as an apperceptive or mnemonic condition? *Neuropsychologia, 124*, 285–298.

Biotti, F., Wu, E., Yang, H., Jiahui, G., Duchaine, B., & Cook, R. (2017). Normal composite face effects in developmental prosopagnosia. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 95*, 63–76.

Bobak, A. K., Jones, A. L., Hilker, Z., Mestry, N., Bate, S., & Hancock, P. J. (2023). Data-driven studies in face identity processing rely on the quality of the tests and data sets. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 166*, 348–364.

Bobak, A. K., Parris, B. A., Gregory, N. J., Bennetts, R. J., & Bate, S. (2017). Eye-movement strategies in developmental prosopagnosia and "super" face recognition. *Quarterly Journal of Experimental Psychology, 70*(2), 201–217.

Burns, E. J., Gaunt, E., Kidane, B., Hunter, L., & Pulford, J. (2022). A new approach to diagnosing and researching developmental prosopagnosia: Excluded cases are impaired too. *Behavior Research Methods*, 1–24.

Chapman, A., Bell, L., Duchaine, B., & Susilo, T. (2018). Varieties of holistic processing deficits in developmental prosopagnosia. *Journal of Vision, 18*(10), 917–917.

Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). NbClust: an R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software, 61*, 1–36.

Corrow, S. L., Davies-Thompson, J., Fletcher, K., Hills, C., Corrow, J. C., & Barton, J. J. (2019). Training face perception in developmental prosopagnosia through perceptual learning. *Neuropsychologia, 134*, Article 107196.

Dalmaijer, E. S., Nord, C. L., & Astle, D. E. (2022). Statistical power for cluster analysis. *BMC Bioinformatics, 23*(1), 1–28.

Dalrymple, K. A., Garrido, L., & Duchaine, B. (2014). Dissociation between face perception and face memory in adults, but not children, with developmental prosopagnosia. *Developmental Cognitive Neuroscience, 10*, 10–20.

DeGutis, J., Bahierathan, K., Barahona, K., Lee, E., Evans, T. C., Shin, H. M., Mishra, M., Likitlersuang, J., & Wilmer, J. B. (2023). What is the prevalence of developmental prosopagnosia? An empirical assessment of different diagnostic cutoffs. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 161*, 51–64.

DeGutis, J., Chatterjee, G., Mercado, R. J., & Nakayama, K. (2012). Face gender recognition in developmental prosopagnosia: Evidence for holistic processing and use of configural information. *Visual Cognition, 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. *Cognitive Neuropsychology, 29*(5–6), 419–446.

DeGutis, J., Cohan, S., & Nakayama, K. (2014). Holistic face training enhances face processing in developmental prosopagnosia. *Brain: a Journal of Neurology, 137*(6), 1781–1798.

DeGutis, J., Li, X., Yosef, B., & Mishra, M. V. (2022). Not so fast! Response times in the computerized Benton Face Recognition Test may not reflect face recognition ability. *Cognitive Neuropsychology, 39*(3–4), 155–169.

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*(1), 87–100.

Duchaine, B., Germine, L., & Nakayama, K. (2007). Family resemblance: Ten family members with prosopagnosia and within-class object agnosia. *Cognitive Neuropsychology, 24*(4), 419–430.

Duchaine, B. C., & Nakayama, K. (2006). Developmental prosopagnosia: A window to content-specific face processing. *Current Opinion in Neurobiology, 16*(2), 166–173.

Esterman, M., Tamber-Rosenau, B. J., Chiu, Y. C., & Yantis, S. (2010). Avoiding non-independence in fMRI data analysis: Leave one subject out. *NeuroImage, 50*(2), 572–576.

Ferreira, L., & Hitchcock, D. B. (2009). A comparison of hierarchical methods for clustering functional data. *Communications in Statistics-Simulation and Computation, 38*(9), 1925–1949.

Fisher, K., Towler, J., & Eimer, M. (2016). Reduced sensitivity to contrast signals from the eye region in developmental prosopagnosia. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 81*, 64–78.

Fitousi, D. (2023). A signal detection–based confidence–similarity model of face matching. *Psychological Review, 131*(3), 625–663.

Formann, A. K. (1984). *Die Latent-Class-Analyse: Einführung in die Theorie und Anwendung* [*Latent class analysis: Introduction to theory and application*]. Weinheim: Beltz.

Fry, R., Li, X., Evans, T. C., Esterman, M., Tanaka, J., & DeGutis, J. (2023). Investigating the influence of autism spectrum traits on face processing mechanisms in developmental prosopagnosia. *Journal of Autism and Developmental Disorders, 53*(12), 4787–4808.

Fry, R., Tanaka, J. W., Cohan, S., Wilmer, J. B., Germine, L. T., & DeGutis, J. (2023). Effects of age on face perception: Reduced eye region discrimination ability but intact holistic processing. *Psychology and Aging, 38*(6), 548–561.

Gerlach, C., Kühn, C. D., Mathiassen, A. B., Kristensen, C. L., & Starrfelt, R. (2023). The face inversion effect or the face upright effect? *Cognition, 232*, Article 105335.

Gerlach, C., & Mogensen, E. (2023). The face inversion effect does not provide a pure measure of holistic face processing. *Behavior Research Methods*, 1–12.

Gross, R., Matthews, I., Cohn, J., Kanade, T., & Baker, S. (2010). Multi-pie. *Image and Vision Computing, 28*(5), 807–813.

Haslam, N., McGrath, M. J., Viechtbauer, W., & Kuppens, P. (2020). Dimensions over categories: A meta-analysis of taxometric research. *Psychological Medicine, 50*(9), 1418–1432.

Holdnack, J. A., Tulsky, D. S., Brooks, B. L., Slotkin, J., Gershon, R., Heinemann, A. W., & Iverson, G. L. (2017). Interpreting patterns of low scores on the NIH toolbox cognition battery. *Archives of Clinical Neuropsychology, 32*(5), 574–584.

Kirsch, L., Fry, R., Verfaellie, M., Anderson, N., & DeGutis, J. (2023). Can face recognition/recollection in developmental prosopagnosia really be improved? Evidence from a repetition-lag training paradigm. *Journal of Vision, 23*(9), 5825–5825.

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.

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 and Cognition, 61*(2), 139–158.

Liu, X., Li, X., Song, Y., & Liu, J. (2021). Separate and shared neural basis of face memory and face perception in developmental prosopagnosia. *Frontiers in Behavioral Neuroscience, 15*, Article 668174.

Lowes, J., Hancock, P. J., & Bobak, A. K. (2024). A new way of classifying developmental prosopagnosia: Balanced Integration Score. *Cortex, 172*, 159–184.

Malcolm, G. L., Leung, C., & Barton, J. J. (2004). Regional variation in the inversion effect for faces: Differential effects for feature shape, feature configuration, and external contour. *Perception, 33*(10), 1221–1231.

Malgady, R. G., & Colon-Malgady, G. (1991). Comparing the reliability of difference scores and residuals in analysis of covariance. *Educational and Psychological Measurement, 51*(4), 803–807.

Marcus, D. K., & Barry, T. D. (2011). Does attention-deficit/hyperactivity disorder have a dimensional latent structure? A taxometric analysis. *Journal of Abnormal Psychology, 120*(2), 427.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276–282.

McKone, E., & Yovel, G. (2009). Why does picture-plane inversion sometimes dissociate perception of features and spacing in faces, and sometimes not? Toward a new theory of holistic processing. *Psychonomic Bulletin & Review, 16*(5), 778–797.
Mishra, M. V., Fry, R. M., Saad, E., Arizpe, J. M., Ohashi, Y. G. B., & DeGutis, J. M. (2021). Comparing the sensitivity of face matching assessments to detect face perception impairments. *Neuropsychologia, 163*, Article 108067.
Murphy, J., Gray, K. L., & Cook, R. (2020). Inverted faces benefit from whole-face processing. *Cognition, 194*, Article 104105.
Murray, E., Bennetts, R., Tree, J., & Bate, S. (2022). An update of the Benton facial recognition test. *Behavior Research Methods, 54*(5), 2318–2333.
Murray, E., Hills, P. J., Bennetts, R. J., & Bate, S. (2018). Identifying hallmark symptoms of developmental prosopagnosia for non-experts. *Scientific Reports, 8*(1), 1690.
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.
Pancaroglu, R., Hills, C. S., Sekunova, A., Viswanathan, J., Duchaine, B., & Barton, J. J. (2016). Seeing the eyes in acquired prosopagnosia. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 81*, 251–265.
Peterson, M. F., Zaun, I., Hoke, H., Jiahui, G., Duchaine, B., & Kanwisher, N. (2019). Eye movements and retinotopic tuning in developmental prosopagnosia. *Journal of Vision, 19*(9), 7–7.
Qiu, W., & Joe, H. (2009). clusterGeneration: random cluster generation (with specified degree of separation). *R package version, 1*(7), 75275-0122.
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*(12), 1961.
Richler, J. J., Mack, M. L., Palmeri, T. J., & Gauthier, I. (2011). Inverted faces are (eventually) processed holistically. *Vision Research, 51*(3), 333–342.
Richler, J. J., Palmeri, T. J., & Gauthier, I. (2012). Meanings, mechanisms, and measures of holistic processing. *Frontiers in Psychology, 3*, 553.
Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica, 128*(2), 274–289.
Rossion, B., & Michel, C. (2018). Normative accuracy and response time data for the computerized Benton Facial Recognition Test (BFRT-c). *Behavior Research Methods, 50*, 2442–2460.
Ruiz, S., Chen, X., Rebuschat, P., & Meurers, D. (2019). Measuring individual differences in cognitive abilities in the lab and on the web. *PLoS One, 14*(12), Article e0226217.
Sachdev, P. S., Blacker, D., Blazer, D. G., Ganguli, M., Jeste, D. V., Paulsen, J. S., & Petersen, R. C. (2014). Classifying neurocognitive disorders: The DSM-5 approach. *Nature Reviews Neurology, 10*(11), 634–642.
Shah, P., Gaule, A., Sowden, S., Bird, G., & Cook, R. (2015). The 20-item prosopagnosia index (PI20): A self-report instrument for identifying developmental prosopagnosia. *Royal Society Open Science, 2*(6), Article 140343.
Stantic, M., Brewer, R., Duchaine, B., Banissy, M. J., Bate, S., Susilo, T., Catmur, C., & Bird, G. (2022). The Oxford Face Matching Test: A non-biased test of the full range of individual differences in face perception. *Behavior Research Methods, 54*(1), 158–173.
Stumps, A., Saad, E., Rothlein, D., Verfaellie, M., & DeGutis, J. (2020). Characterizing developmental prosopagnosia beyond face perception: Impaired recollection but intact familiarity recognition. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 130*, 64–77.
Susilo, T., McKone, E., Dennett, H., Darke, H., Palermo, R., Hall, A., Pidcock, M., Dawel, A., Jeffery, L., Ellie Wilson, C., & Rhodes, G. (2010). Face recognition impairments despite normal holistic processing and face space coding: Evidence from a case of developmental prosopagnosia. *Cognitive Neuropsychology, 27*(8), 636–664.
Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition, 93*(1), B1–B9.
Tanaka, J. W., & Simonyi, D. (2016). The "parts and wholes" of face recognition: A review of the literature. *Quarterly Journal of Experimental Psychology, 69*(10), 1876–1889.
Tardif, J., Morin Duchesne, X., Cohan, S., Royer, J., Blais, C., Fiset, D., Duchaine, B., & Gosselin, F. (2019). Use of face information varies systematically from developmental prosopagnosics to super-recognizers. *Psychological Science, 30*(2), 300–308.
Torfs, K., Vancleef, K., Lafosse, C., Wagemans, J., & de-Wit, L. (2014). The Leuven Perceptual Organization Screening Test (L-POST), an online test to assess mid-level visual perception. *Behavior Research Methods, 46*, 472–487.
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. *Quarterly Journal of Experimental Psychology, 70*(2), 298–315.
Verfaillie, K., Huysegems, S., De Graef, P., & Van Belle, G. (2014). Impaired holistic and analytic face processing in congenital prosopagnosia: Evidence from the eye-contingent mask/window paradigm. *Visual Cognition, 22*(3–4), 503–521.
Wang, M. W., & Stanley, J. C. (1970). Differential weighting: A review of methods and empirical studies. *Review of Educational Research, 40*(5), 663–705.
Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58*(301), 236–244.
Yardley, L., McDermott, L., Pisarski, S., Duchaine, B., & Nakayama, K. (2008). Psychosocial consequences of developmental prosopagnosia: A problem of recognition. *Journal of Psychosomatic Research, 65*(5), 445–451.
Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*(3), 441–517.
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.
Zheng, L., Grove, R., & Eapen, V. (2019). Spectrum or subtypes? A latent profile analysis of restricted and repetitive behaviours in autism. *Research in Autism Spectrum Disorders, 57*, 46–54.