

# Eye movements and retinotopic tuning in developmental prosopagnosia

**Matthew F. Peterson**

Department of Brain and Cognitive Sciences, Massachusetts  
Institute of Technology, Cambridge, MA, USA



**Ian Zaun**

Department of Brain and Cognitive Sciences, Massachusetts  
Institute of Technology, Cambridge, MA, USA



**Harris Hoke**

Center for Brain Science, Harvard University,  
Cambridge, MA, USA



**Guo Jiahui**

Department of Psychological and Brain Sciences,  
Dartmouth College, Hanover, NH, USA



**Brad Duchaine**

Department of Psychological and Brain Sciences,  
Dartmouth College, Hanover, NH, USA



**Nancy Kanwisher**

Department of Brain and Cognitive Sciences, Massachusetts  
Institute of Technology, Cambridge, MA, USA



Despite extensive investigation, the causes and nature of developmental prosopagnosia (DP)—a severe face identification impairment in the absence of acquired brain injury—remain poorly understood. Drawing on previous work showing that individuals identified as being neurotypical (NT) show robust individual differences in where they fixate on faces, and recognize faces best when the faces are presented at this location, we defined and tested four novel hypotheses for how atypical face-looking behavior and/or retinotopic face encoding could impair face recognition in DP: (a) fixating regions of poor information, (b) inconsistent saccadic targeting, (c) weak retinotopic tuning, and (d) fixating locations not matched to the individual's own face tuning. We found no support for the first three hypotheses, with NTs and DPs consistently fixating similar locations and showing similar retinotopic tuning of their face perception performance. However, in testing the fourth hypothesis, we found preliminary evidence for two distinct phenotypes of DP: (a) Subjects characterized by impaired face memory, typical face perception, and a preference to look high on the face, and (b) Subjects characterized by profound impairments to both face memory and perception and a preference to look very low on the face. Further, while all NTs and upper-looking DPs performed best when faces were presented near their preferred fixation location, this was not true for lower-looking DPs. These results suggest that

face recognition deficits in a substantial proportion of people with DP may arise not from aberrant face gaze or compromised retinotopic tuning, but from the suboptimal matching of gaze to tuning.

## Introduction

In developmental prosopagnosia (DP), individuals with no known history of brain injury exhibit striking deficits in face recognition in the absence of early visual deficits or cognitive impairment. This condition affects around 2% of the adult population (Kennerknecht et al., 2006; Kennerknecht, Ho, & Wong, 2008), with often significant consequences for everyday life (Dalrymple, Fletcher, et al., 2014; Yardley, McDermott, Pisarski, Duchaine, & Nakayama, 2008). Yet despite extensive investigation (Geskin & Behrmann, 2018; Susilo & Duchaine, 2013), the precise processing deficits underlying DP remain unknown. Here, we test four hypotheses that explain deficits in face recognition as the result of atypicalities in (a) the way faces are fixated, (b) retinotopic tuning of the face representation system, and/or (c) the relationship between the two.

Our hypotheses are based on recent work that has demonstrated a link between face recognition perfor-

Citation: 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, 1–19, <https://doi.org/10.1167/19.9.7>.

<https://doi.org/10.1167/19.9.7>

Received October 29, 2018; published August 19, 2019

ISSN 1534-7362 Copyright 2019 The Authors



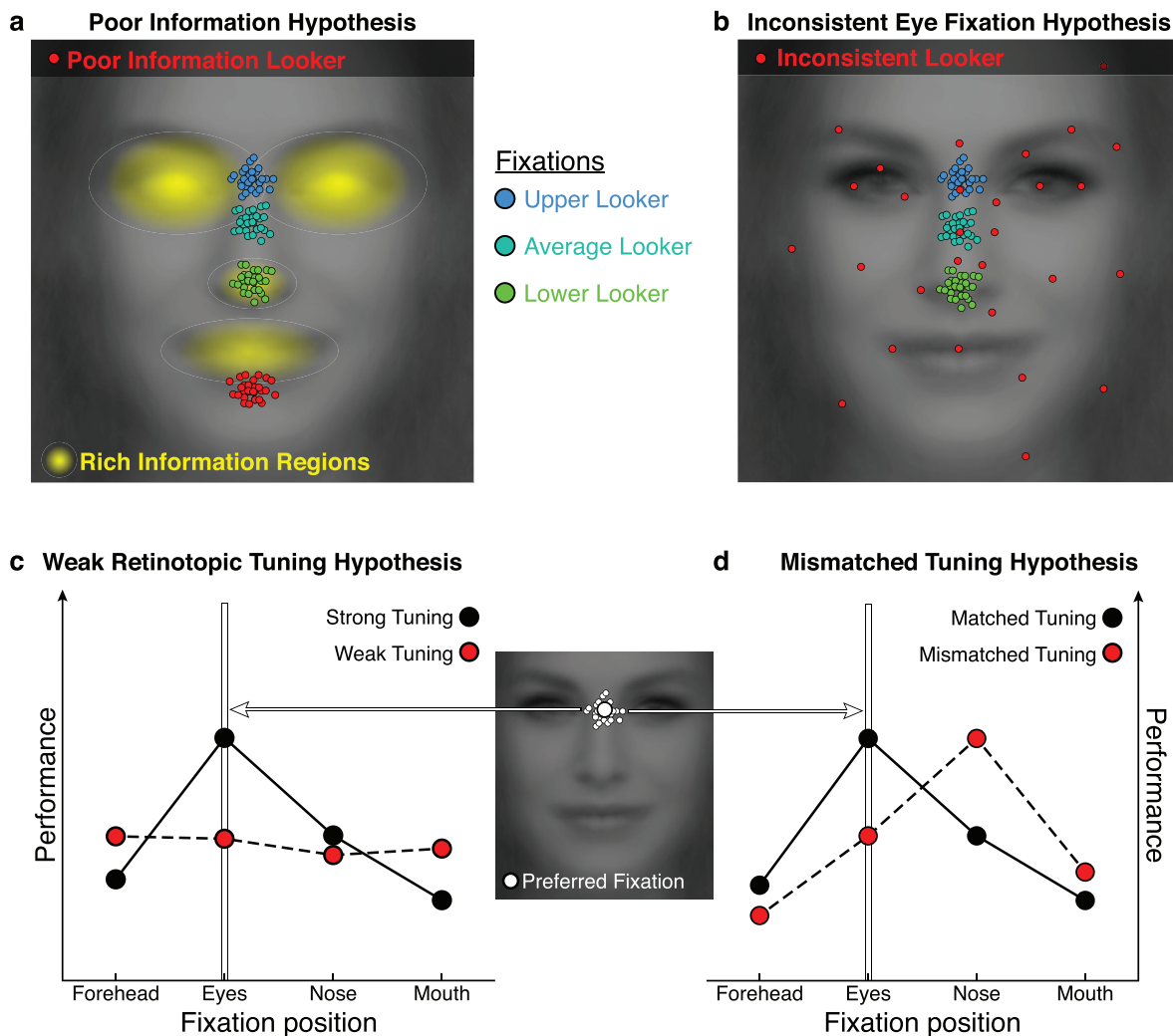


Figure 1. Hypothesized mechanisms for impaired face recognition in DP. Individuals with DP may (a) fixate locations where high quality information cannot be obtained as readily, (b) fail to fixate a consistent position on the face, (c) fail to show strong tuning to a particular retinotopic position, or (d) consistently fixate away from a strongly tuned location. Vertical white bar in (c) and (d) indicate an example subject's mean preferred fixation location.

mance and face looking behavior in the general population. Specifically, neurotypical (NT) subjects vary reliably from each other in where they look on the face, with each individual fixating their own personal preferred location with extraordinary stability and precision (Mehouadar, Arizpe, Baker, & Yovel, 2014; Peterson & Eckstein, 2013; Peterson, Lin, Zaun, & Kanwisher, 2016). Some make initial fixations either toward the tip of the nose (“lower lookers”) or between the eyes (“upper lookers”), with most looking at locations in between. This preferred fixation location is stable within an individual over years (Mehouadar et al., 2014; Peterson & Eckstein, 2013) and across face recognition tasks, including identification, sex classification, and expression categorization. Importantly, people perform much more accurately at face recognition when faces are presented at their own preferred face fixation position, showing retinotopic tuning of the

face system that is aligned with the individual's preferred fixation location (Or, Peterson, & Eckstein, 2015; Peterson & Eckstein, 2012, 2013). The systematic relationship between an individual's preferred fixation location and the fixation position where they recognize faces best suggests that joint spatial tuning of eye movement planning and face encoding plays a critical role in face recognition. Here, we test four hypotheses for how this system may malfunction in individuals with DP.

The most obvious possibility is that individuals with DP might fixate regions of the face that are not rich in discriminative information (the Poor Information Hypothesis; Figure 1a). Most NT individuals fixate somewhere between the eyes and nose tip, with computational modeling showing this to be an optimal strategy given the spatial distribution of information across the face and the reduction in processing power

from the fovea to the periphery (Or et al., 2015; Peterson & Eckstein, 2012; Tsank & Eckstein, 2017). Fixating outside this area would cause the most information-rich regions of the face to fall into the visual periphery where resolution is low. Importantly, the Poor Information Hypothesis concerns how choice of fixation modulates the *amount* of information available for cortical processing. This is distinct from hypothesized inefficiencies in how well cortex is able to *use* the information it receives, such as reports of impaired cortical processing of the eye region in acquired prosopagnosia (Caldara et al., 2005; Fiset et al., 2017), DP (Tardif et al., 2019), and low face recognition ability NTs (Royer et al., 2018). Thus, this hypothesis predicts that DP individuals will look outside the region between the eyes and the tip of the nose.

A second possibility is that individuals with DP do not fixate a single location with the same precision as NTs (Inconsistent Eye Fixation Hypothesis; Figure 1b). The narrow spatial tuning of recognition ability around an individual's optimal location means that fixating even a small distance away can result in substantial performance deficits. This is presumably why NTs saccade to their preferred fixation location with extraordinary precision when they look at a face (Kowler & Blaser, 1995; Peterson & Eckstein, 2013; Peterson et al., 2016). The key prediction of the Inconsistent Eye Fixation Hypothesis is that the variance of saccade landing points across face presentations will be larger for individuals with DP than NTs.

The third hypothesis is based on the fact that NTs show strong retinotopic tuning of their face system: Face recognition accuracy is strongly dependent on where exactly a face is fixated. Presumably, narrow spatial tuning reflects an encoding strategy where the visual system forms powerful representations over a narrow range of retinotopic positions through a preferential allocation of resources, which comes at the expense of weaker representations when faces appear at other positions. According to the Weak Retinotopic Tuning Hypothesis, this tuning is disrupted in individuals with DP. Weakened tuning, where resources are distributed across a large space of retinotopic image positions, would be expected to reduce the maximum encoding capacity. This hypothesis predicts that the dependence of face recognition performance on face fixation location will be weaker in individuals with DP (Figure 1c).

As discussed above, accurate face recognition in NT subjects hinges critically on the alignment of retinotopic tuning with face fixation behavior: Subjects perform best at face recognition when they fixate faces in their habitual preferred location. This raises the final possibility, which is that individuals with DP may not show a similar alignment. On this Mismatched Tuning

Hypothesis, individuals with DP consistently fixate typical face locations and show typical retinotopic tuning strength of face recognition performance, but their preferred fixation location is not aligned with the retinotopic tuning of their face system. This hypothesis predicts that where a DP individual chooses to look will not predict where they perform best (Figure 1d).

To test these hypotheses, we measured eye movement behavior and face recognition performance as a function of fixation location in 22 DP and 30 NT control subjects. To characterize individuals' preferred fixation behavior, we measured the landing point of the first saccade onto a peripherally presented stimulus (initial fixation) for each of three recognition tasks: celebrity identification, expression recognition, and car recognition (Figure 2a). Recognition performance was also measured for these three tasks and on the Cambridge Face Memory Test (CFMT; Duchaine & Nakayama, 2006). To quantify how face perception ability changes across retinotopic positions, we measured perceptual performance on a sequential same/different face matching task while subjects were required to fixate at four different locations (forehead, eyes, nose, and mouth; Figure 2b). We found no support for the first three hypotheses, with subjects in both groups fixating comparable locations with similar consistency and showing strong retinotopic tuning. Instead, the results were most consistent with the Mismatched Tuning Hypothesis for a subgroup of individuals with DP who look low on the face.

## Methods

### Participants

Twenty-two participants with DP (mean age = 36.3,  $N_{female} = 16$ ) were recruited from our database of people who have reported face recognition deficits at [www.faceblind.org](http://www.faceblind.org). Participants were tested with three tests of face identity memory remotely: the CFMT (Duchaine & Nakayama, 2006), an old-new discrimination test (Duchaine & Nakayama, 2005), and a famous face test (Duchaine & Nakayama, 2005). Participants who scored two or more standard deviations below the mean on at least two tests were asked to visit the lab at the Massachusetts Institute of Technology (MIT) for the studies reported here.

Thirty NT control participants were recruited using flyers and departmental subject lists. Controls were selected to closely match the age and sex distributions of the DP group (mean age = 34.7,  $N_{female} = 22$ ).

The study was approved by the MIT Committee on the Use of Humans as Experimental Subjects, and all participants provided written informed consent. All

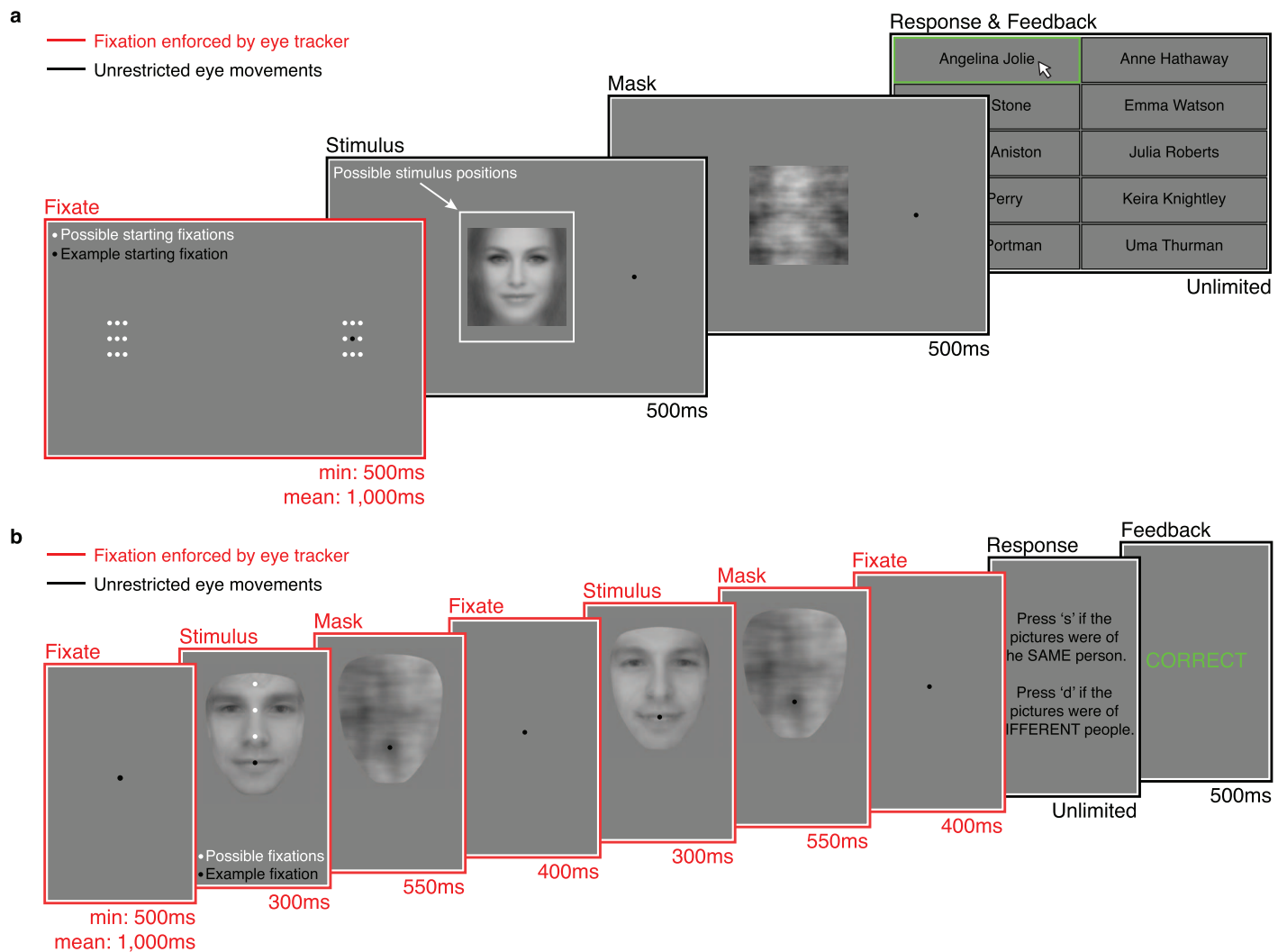


Figure 2. Experimental procedure. (a) Preferred initial fixations were measured for face identification (shown), EXP, and CAR using the same procedure. The initial fixation is defined as the landing point of the subject's saccade from a fixation dot (black dot: example location; white dots: 17 other possible locations; fixation on dot enforced with an eye tracker until stimulus onset) onto a peripheral stimulus randomly located within the central region of the display (white box). (b) Retinotopic tuning of perceptual encoding of faces was assessed by measuring performance on a same/different face discrimination task at four different retinotopic positions. Subjects maintained fixation on either the mouth (black dot), nose, eyes, or forehead (white dots) of two rapidly presented faces (fixation on dot enforced with an eye tracker) and determined whether they saw two visually distinct images of the same person or images of two different people (50% probability for each condition). Red borders indicate when subjects were required to maintain fixation on the fixation dot (enforced by an eye tracker), while black borders indicate when subjects could move their eyes freely. Face images are proxy composites (average across all stimuli) for the actual stimuli used in the study.

subjects reported normal or corrected-to-normal vision and received \$75 for participation (\$20/hour for 3 hr and 45 min).

## Eye tracking

The right eye of each participant was tracked using an SR Research EyeLink 1000 Desktop Mount sampling at 1000 Hz (SR Research Ltd., Ottawa, Ontario, Canada). A 9-point calibration and validation

were run at the beginning of the session with a mean error of no more than  $0.5^\circ$ . Saccades were classified as events where eye velocity was greater than  $22^\circ/\text{s}$  and eye acceleration exceeded  $4000^\circ/\text{s}^2$ .

## Display

All stimuli were presented on a 24-in. CRT monitor with a resolution of  $1920 \times 1200$  pixels and refresh rate of 60 Hz. Subjects sat 46 cm from the monitor, with



each pixel subtending  $0.033^\circ$ . Stimuli were presented on a mid-level gray background (RGB = [128 128 128]).

## Overview of experimental tasks

Participants in this study performed five distinct tasks: (a) CFMT (Duchaine & Nakayama, 2006), (b) celebrity identification (CELEB; Peterson et al., 2016), (c) car recognition (CAR), (d) expression recognition (EXP; Peterson & Eckstein, 2012), and (e) same/different face discrimination with unfamiliar identities (S/D). Due to its length, the S/D task was broken up into four equally-sized sections (S/D<sub>1</sub>, S/D<sub>2</sub>, S/D<sub>3</sub>, and S/D<sub>4</sub>). All NT control participants completed the tasks in the same order: CFMT → S/D<sub>1</sub> → CELEB → S/D<sub>2</sub> → CAR → S/D<sub>3</sub> → EXP → S/D<sub>4</sub>, while participants with DP—having completed the CFMT before the lab session—ran in the same order excluding the CFMT (starting with S/D<sub>1</sub>). All participants saw each image in the same order within each task. Eye movements were recorded for all tasks except the CFMT.

## Cambridge Face Memory Test

Subjects completed the CFMT with the standard protocol as described previously (Duchaine & Nakayama, 2006).

## Methods (Free eye movement tasks: CELEB, EXP, CAR)

### Overview

The free eye movement tasks were designed to measure where subjects chose to initially fixate when recognizing faces (identity or expression) or examples from a non-face object category (cars). Given that face identification is often completed within the first or second on-face fixation, we adopted a brief-presentation task structure as we have used in previous studies (Or et al., 2015; Peterson & Eckstein, 2012, 2013, 2014; Peterson et al., 2016; Tsank & Eckstein, 2017). Critically, we used an eye tracker to ensure that foveal processing of the face image was only possible after the subject made a saccade from a prestimulus peripheral fixation dot (located  $15^\circ$  off the face on average) onto the face (initial fixation). We forced fixation to begin off the face by stopping the trial if the eye tracker registered a blink or if the subject's gaze moved farther than  $1^\circ$  from the center of the peripheral fixation dot at any time before the stimulus appeared. Subjects were

then informed that their fixation was invalid and asked to try again.

## Stimuli

CELEB stimuli were 100 frontal-view images, with 10 distinct images for each of 10 well-known female Caucasian celebrities acquired using Google image search (Supplementary Figure S1a). EXP stimuli were 98 frontal-view images, with 14 distinct images for each of seven standard expressions (afraid, angry, disgusted, happy, neutral, sad, surprised) selected from a stimulus set used in previous studies (Peterson & Eckstein, 2012; Supplementary Figure S1b). CAR stimuli were 100 side-view images of cars, with 10 distinct images for each of 10 popular car models acquired using Google image search (Supplementary Figure S1c).

All images were converted to grayscale and rotated to an upright orientation. For faces (CELEB, EXP), images were scaled so that the center of the eyes and center of the mouth were in the same position for all photographs ( $6.0^\circ$  apart), cropped vertically from the top of the head to the chin, and cropped horizontally to achieve a square aspect ratio, with each image subtending  $16.7^\circ$  (500 pixels) in both dimensions. For cars (CAR), images were scaled so that the roof and floor were separated by the same distance as the eyes and mouth in the face tasks ( $6.0^\circ$ ), and cropped to a common size such that every car was fully visible (vertical: 260 pixels =  $8.7^\circ$ , horizontal: 640 pixels =  $21.3^\circ$ ). The contrast energy for each image was normalized to the mean contrast energy across all images for each task separately. A different mask image for each task was created by filtering zero-mean Gaussian white noise by the average amplitude spectrum of all images used in a given task and matched to the display's mean luminance.

## Procedure

Each task began with a familiarization phase in which participants were shown one example image of each category with the category label underneath the image (e.g., the name and one image of each identity in the celebrity identification task; see Supplementary Figure S1). Subjects were instructed to familiarize themselves with these example images for as long as they liked and then pressed the spacebar to proceed to the task itself. Then in the main experiment, each trial began by displaying a  $0.05^\circ$  radius black fixation dot at one of 18 possible locations on the left or right side of the screen (see Figure 2a and Supplementary Table S1). When ready, the subject fixated the center of the black fixation dot and pressed the spacebar while maintaining

fixation on the black fixation dot during a random delay period (delay = 500 ms + a random sample from a geometric distribution with mean = 500 ms). After the delay period, one of the stimulus images (face or car) was displayed for 500 ms at one of nine equally spaced positions in the central region of the display (see Figure 2a and Supplementary Table S2), and subjects were instructed to look at this stimulus when it appeared. The stimulus was then replaced by a noise mask for 500 ms followed by a response screen showing a grid of boxes containing the names of either the 10 celebrities, seven expressions, or 10 car models. The subject had unlimited time to select their answer with a mouse click on the corresponding box. The response screen was then replaced by a feedback screen for 500 ms with the correct answer highlighted in green; if the subject was incorrect, their answer would be highlighted in red. The same procedure was repeated for all trials. The location of the fixation dot and the position of the stimulus were randomly sampled on each trial. Each of the 100 stimuli were presented once with a randomized presentation order. The fixation dot location, stimulus position, and stimulus image on each trial were the same for all subjects (see Figure 2a).

## Methods (Forced fixation task: S/D)

### Overview

We used a same/different face matching task to measure perceptual accuracy as a function of fixation location. To isolate perceptual processing, we designed a task that required generalizing across different images of the same individual that differed in lighting, gaze direction, expression, etc. in order to test higher level perceptual processing, not pixel-level matching. Importantly, subjects saw each individual in only one trial: either one image of an individual if they were part of a “different” pair trial, or two different images of an individual if they were part of a “same” pair trial. By not repeating identities across trials, the task precludes any learning of face identities over the course of the experiment, providing a relatively pure measure of face perception by restricting the need for memory to a minimal ( $< 1$  s) delay between paired face images. To measure face perception performance at different retinotopic positions, we required subjects to continuously fixate a fixation dot while two face images were presented at one of four positions (fixation along the vertical midline at either the forehead, center of eyes, nose tip, or center of mouth, equally spaced  $3^\circ$  apart; see Figure 2b and Supplementary Table S2). If the subject blinked or looked more than  $0.5^\circ$  from the fixation dot at any time from the start of the trial until

the response screen, the trial was aborted and excluded from analysis (average number of trials included per subject = 390, minimum = 353, maximum = 400).

### Stimuli

Stimuli comprised 800 frontal-view images of 600 nonfamous Caucasian people taken from various face image databases used in other studies (Ara Nefian Face Recognition Page, n.d.; Computational Vision Archive, n.d.; Psychological Image Collection at Stirling, n.d.) and Google image search (two different images per person for each of 200 same identity pairs and one image for each person for 200 different identity pairs, 300 male and 300 female; Supplementary Figure S1d). Images were preprocessed and standardized and a mask image was created using the same procedure described above for the CELEB and EXP tasks. A common cropping mask was applied to remove hair, clothing, and other external features.

### Procedure

Subjects initiated each trial by fixating a  $0.05^\circ$  radius black dot located at the center of the screen and pressing the space bar. After a random delay period (sampled from the same distribution as the free eye movement tasks), the subject was shown the first face image for 300 ms, a noise mask for 550 ms, a blank fixation screen for 400 ms, the second face image for 300 ms (at the same position as the first face image), a noise mask for 550 ms, and a blank fixation screen for 400 ms, with fixation at the fixation dot enforced throughout. A response screen prompted the subject to press the ‘s’ key if they thought the two images were of the same person or the ‘d’ key if they thought they were different people (unlimited response time). After the subject responded, a 500-ms feedback screen told the subject whether they were correct or not. The positions of the stimuli were randomized across trials. The display order was fully balanced, such that each combination of S/D, face gender, and stimulus position was shown for the same number of trials (25 trials for each of 16 combinations). All subjects saw the same presentation order (see Figure 2b).

## Methods (Analyses)

### Directionality of significance tests

Statistical tests were one-tailed when directly testing predictions from our four hypotheses or for impaired

face recognition performance (i.e., DP performance lower than NT performance) and two-tailed otherwise.

### Fixation correlations within and between tasks

We used split-half correlation analysis to measure both (a) the reliability of fixation preference within a task, and (b) the similarity in fixation preference across tasks. For each free eye movement task (CELEB, EXP, CAR) and subject, we computed the mean (preferred fixation) and standard deviation (fixation inconsistency) of the location of the initial on-stimulus fixation for the first and second half of trials separately (50 trials per split for CELEB and CAR, 49 trials per split for EXP). We then correlated subjects' preferred fixation and fixation inconsistency values from the first half of trials with the values from the second half for each combination of task.

### Matched tuning

We defined matched tuning as a negative relationship (linear regression slope) between an observer's performance at each S/D forced fixation location and the distance from each forced fixation location to the observer's preferred fixation location. For each subject and each S/D forced fixation location we computed two quantities: (a) The distance, in degrees of visual angle, from the forced fixation location to the subject's preferred fixation location measured with the celebrity identification task (distance from preferred), and (b) The subject's performance at the forced fixation location minus the subject's performance averaged across all forced fixation locations (mean-centered performance). For any given group,  $g$ , we pooled distance from preferred and mean-centered performance for all forced fixation locations and subjects in the group, giving  $N_{data,g} = N_{subjects,g} N_{locations}$  data points, where  $N_{locations}$  was 4 when all locations were included and 3 when the forehead location was excluded. We then linearly regressed mean-centered performance on distance from preferred, with the (negative) regression slope quantifying how well the group matched their preferred fixations to the tuning of their retinotopic face encoding.

Confidence intervals and statistical significance for regression slopes and differences in regression slopes between groups were calculated using bootstrapping ( $N_{bootstrap} = 1,000,000$ ). For each bootstrap sample for each group of subjects (e.g., all subjects with DP, upper looking NTs, etc.), we randomly sampled (with replacement)  $N_{data,g}$  pairs of mean-centered performance and distance from preferred from the group's pooled data. We then regressed mean-centered perfor-

mance on distance from preferred, giving  $N_{bootstrap}$  regression slopes per group. For each group's regression slope, confidence intervals were defined as the values delimiting the central 95% of all samples and  $p$  values were defined as the proportion of samples with values greater than or equal to 0. For testing whether one group's slope was significantly less than another group's, we took the difference in the groups' slopes for each sample, giving  $N_{bootstrap}$  slope differences. Confidence intervals and  $p$  values for slope differences were then defined and computed in the same manner as for the group slopes.

## Results

### Performance: Face identity perception and memory are selectively impaired in DP

As expected, DP performance, in terms of proportion correct responses ( $PC$ ), was significantly impaired on all face identification tasks (CFMT, CELEB, and best performing S/D forced fixation location; all  $ps < 0.001$ ). Performance was not significantly different between the groups for either EXP ( $p = 0.509$ ) or CAR ( $p = 0.718$ ; two-tailed two-sample  $t$  tests; Figure 3a and Supplementary Table S3). Further, DP performance, in terms of the sensitivity metric  $d'$ , was significantly lower than NT performance at each of the four S/D forced fixation locations (forehead:  $p = 0.012$ ; eyes:  $p = 0.002$ ; nose:  $p < 0.001$ ; mouth:  $p < 0.001$ ; two-tailed two-sample  $t$  tests; Figure 3b and Supplementary Table S4).

### Eye movements: The locations, consistency, and domain specificity of initial fixations on faces are typical in DP

The initial fixation was defined as the landing location of the initial saccade that moved gaze from the starting fixation position near the edge of the screen onto the stimulus near the center of the screen (Figure 2a). Preferred fixation was defined as the average initial fixation location across all trials as has been used previously (Or et al., 2015; Peterson & Eckstein, 2012, 2013; Peterson et al., 2016; Tsank & Eckstein, 2017). Relevant to the Poor Information Hypothesis, preferred fixation location for CELEB was not significantly different between the groups in either the vertical ( $p = 0.977$ ; Figure 4a, left) or horizontal dimension ( $p = 0.210$ ; Figure 4a, right). Relevant to the Inconsistent Eye Fixation Hypothesis, fixation inconsistency, defined as the standard deviation of initial fixation location across trials, was not significantly different



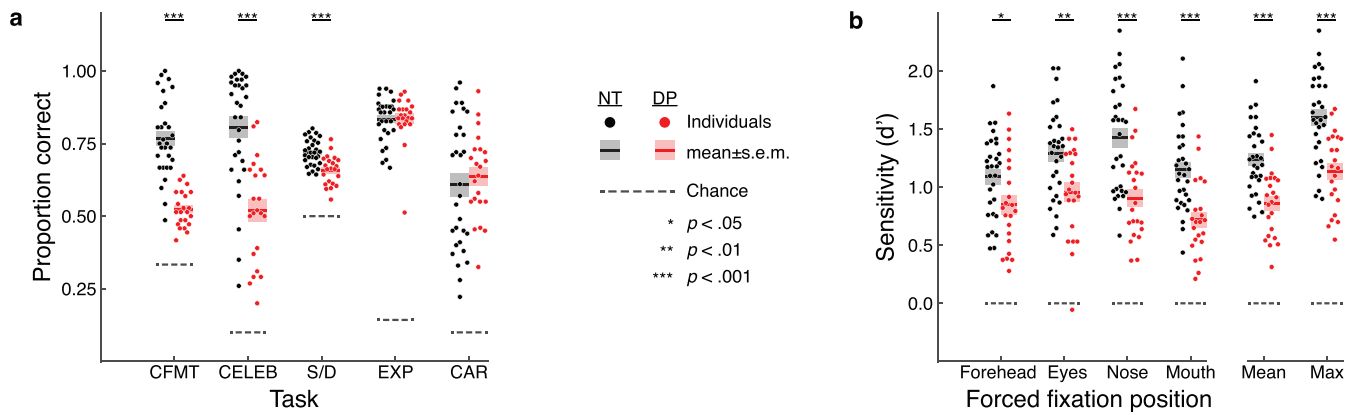


Figure 3. Recognition performance for faces and cars. (a) NT (black) and DP (red) performance for all tasks (S/D = maximum performance across the four forced fixation locations). (b) S/D performance at each forced fixation location and each subject's average and maximum performance across forced fixation locations. Dots are individual subjects, with solid lines representing the mean and shaded boxes showing the standard error of the mean across subjects. Horizontal dashed lines below each set of scores display at-chance performance.

between the groups for CELEB in either the vertical ( $p = 0.989$ ; Figure 4b, left) or horizontal ( $p = 0.838$ ; Figure 4b, right) dimension. Similarly, DPs and NTs did not differ significantly in either fixation preference or consistency for EXP. In contrast, DPs preferred to fixate significantly higher ( $p = 0.039$ ) and to the left ( $p = 0.015$ ) with greater vertical consistency ( $p = 0.027$ ) than NTs when recognizing cars (two-tailed two-sample  $t$  tests; Figure 4a and b and Supplementary Tables S5 and S6). It is not clear why NTs and DPs employed slightly different initial fixation strategies when recognizing cars, but it may reflect an adaptive strategy that attempts to compensate for the lower performance of DPs relative to controls in car recognition tasks employed in previous studies (Dalrymple, Garrido, & Duchaine, 2014; Duchaine, Germine, & Nakayama, 2007).

Is an individual's preferred fixation when identifying faces consistent within an individual, and across face tasks, as shown previously (Peterson & Eckstein, 2012)? And does this face fixation behavior reflect a strategy specific to faces and identification, or a general strategy employed across visual categories and tasks? To assess the reliability and domain specificity of preferred fixation strategies, we correlated individuals' preferred fixations measured over the first half of trials with preferred fixations measured over the second half of trials for each pairwise combination of tasks. For both groups, first and second half preferred fixations were strongly correlated within each task along both the vertical and horizontal dimensions, indicating stable individual differences in fixation strategies within each task (Figure 4c). Across tasks, however, preferred fixations were significantly correlated between CELEB and EXP only; preferred fixations when recognizing cars were not predictive of preferred fixations when either recognizing celebrities or expressions (Figure 4c;

Supplementary Table S7). These data show similarly stable face fixation behavior in NTs and DPs, with individual differences in preferred fixation location generalizing across face tasks but not across our face and nonface (CAR) tasks in both groups.

Taken together, the results indicate that NTs and DPs prefer to fixate comparable locations on faces, and do so with similar consistency within an individual, arguing against the Poor Information and Inconsistent Eye Fixation Hypotheses.

### Retinotopic tuning of face encoding is not weaker in DP

If the encoding of stimuli by the visual system is retinotopically specific and tuned to a particular retinotopic position, then we would expect performance to depend on where a subject fixates on the stimulus and to be highest near their tuned location. To test for the presence and strength of retinotopic tuning, we used a selectivity metric,  $S$ , to measure the relative difference in sensitivity between each subject's best performing S/D forced fixation location,  $d'_{max}$ , and the average sensitivity,  $d'_{\sim max}$ , taken across each of the other  $N - 1$  forced fixation positions  $i$ ,  $d'_i$  ( $N = 4$ ):

$$S = \frac{d'_{max} - d'_{\sim max}}{d'_{max} + d'_{\sim max}} \quad (1)$$

where

$$d'_{\sim max} = \frac{1}{N-1} \sum_{i \neq max} d'_i \quad (2)$$

Retinotopic tuning strength was not significantly



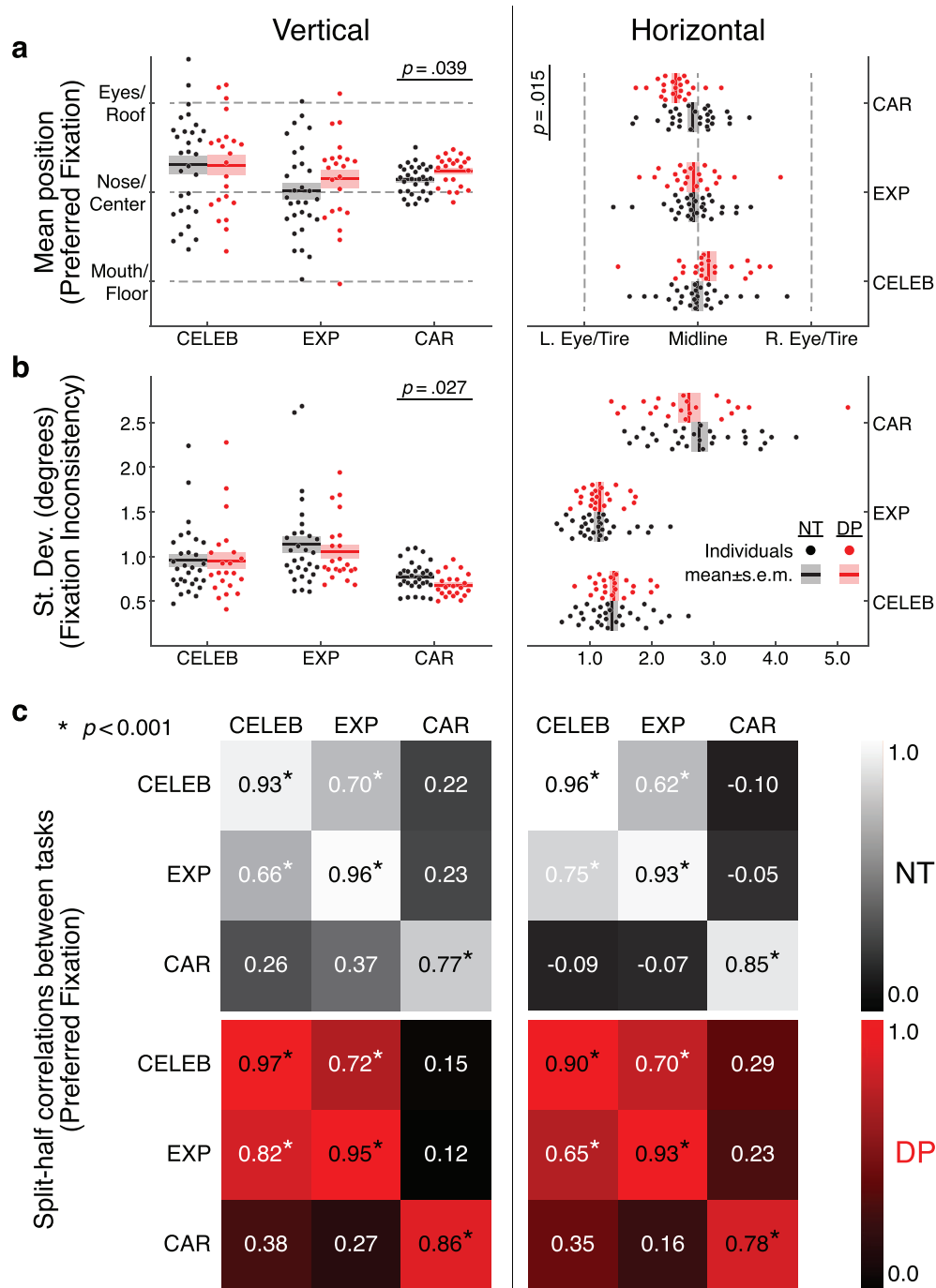


Figure 4. Preferred fixations on faces and cars. (a) NT (black) and DP (red) preferred fixations, defined as the average initial on-stimulus fixation across trials in the vertical (left) and horizontal (right) dimensions. (b) Fixation inconsistency, defined as the variance in initial fixation location across trials. (c) Correlations in subjects' preferred fixations between the first half of trials for one task and the second half of trials for a second task in the vertical (left) and horizontal (right) dimensions for both NTs (top, grayscale) and DPs (bottom, red). The diagonals show the split-half reliability for each task. For (a) and (b), dots represent individual subjects, with solid lines and shaded boxes indicating the mean and standard error of the mean across subjects.

weaker for DPs than NTs ( $p = 0.181$ ; two-tailed two-sample  $t$  test; Figure 5a; Supplementary Table S8).

In our previous work (Peterson & Eckstein, 2013), we quantified matched tuning using data from eyes and nose locations only, because we thought faces might be

processed atypically when presented at locations that are rarely or never fixated when free eye movements are allowed. Here, we included the forehead and mouth conditions because we did not know a priori whether subjects with DP would prefer to fixate outside the NT

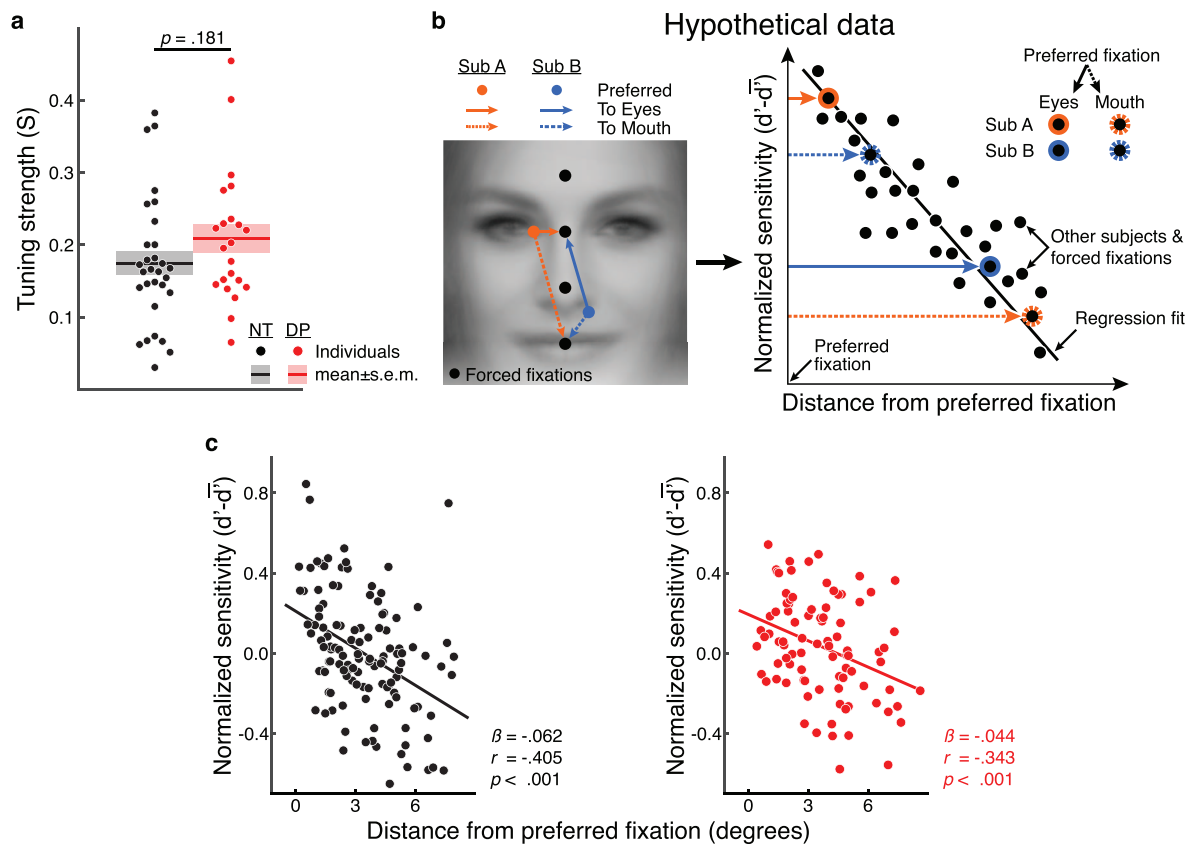


Figure 5. Retinotopic tuning strength and matching of eye movements to tuning. (a) No significant differences between DPs (red) and NTs (black) in retinotopic tuning strength, defined as subjects' performance at their best-performing forced fixation location relative to performance averaged across all other locations. Dots indicate individual subjects, with solid lines and shaded boxes indicating the mean and standard error of the mean across subjects. (b) Matched tuning predicts that subjects' performance will decrease the further they are forced to fixate from their preferred location, as described in this illustrative example. Left, Subject A (orange) prefers to fixate high on the face, and thus should perform well when forced to fixate the eyes (small distance from preferred, solid arrow) and poorly when forced to fixate the mouth (large distance from preferred, dashed arrow). Subject B (blue) prefers to fixate low on the face and should show the opposite pattern. Right, this relationship is quantified by the slope (black line) when regressing normalized performance (centered on each subject's mean performance separately) on the distance from preferred fixation for all subjects and forced fixation locations. Here, performance when forced to fixate the eyes (dots with solid borders) is high for upper-looking Subject A but poor for lower-looking Subject B. When forced to fixate the mouth (dots with broken borders), Subject B now outperforms Subject A. Black dots represent other strongly matched hypothetical subjects, while the black line represents a best linear regression fit. (c) The observed significant correlations between normalized performance and absolute distance from preferred fixation for NTs (left) and DPs (right). Larger slope magnitudes (linear regression coefficients,  $\beta$ ) indicate more strongly matched tuning. Dots are individual subjects at different forced fixation locations.

range. While more subjects in this study preferred to fixate close to the mouth than in previous reports (Mehouadar et al., 2014; Or et al., 2015; Peterson & Eckstein, 2012, 2013, 2014; Peterson et al., 2016; Tsank & Eckstein, 2017), no subjects preferred to fixate close the forehead forced fixation location (minimum distance from subjects' preferred fixations to forehead =  $2.38^\circ$ ). Thus, it could be argued that we should follow the analysis method we used previously, excluding the forehead condition. Indeed, when the forehead condition is excluded, retinotopic tuning was significantly *stronger* for DPs ( $p = 0.023$ ; two-tailed, two-sample  $t$  test; Supplementary Figure S3a; Supplementary Table

S8). Taken together, the results indicate that retinotopic tuning is at least as strong, and maybe even stronger, in DP, failing to support the Weak Retinotopic Tuning Hypothesis.

### Preferred fixations are matched to retinotopic tuning of face perception in both NTs and DPs

If matching preferred fixation to retinotopic tuning is a strategy that optimizes face recognition performance, then performance should be worse when subjects fixate away from their preferred fixation, as we have found

previously for NTs (Or et al., 2015; Peterson & Eckstein, 2012, 2013). Here, we used the S/D face discrimination task to measure performance (sensitivity) at four fixation locations by forcing subjects to fixate at either the forehead, eyes, nose tip, or mouth. Matched tuning predicts that each subject's performance should decrease the farther they were forced to fixate from their preferred fixation. We tested for matched tuning by regressing S/D performance on distance from preferred fixation; with four forced fixation locations, this resulted in 120 data points for the NT group and 88 data points for the DP group. To test for matched tuning at the group level, we controlled for individual variability in overall performance by subtracting the mean sensitivity across the four locations from the sensitivity at each forced fixation location separately for each subject, resulting in a normalized sensitivity value at each forced fixation location for each subject. We then regressed normalized sensitivity on distance from preferred across all subjects and forced fixation locations for each group (Figure 5b). Performance was significantly and negatively correlated with the distance from a forced fixation location to a subject's preferred fixation for both NTs (mean [95% confidence interval]; slope =  $\beta = -0.062$  [-0.030, -0.092],  $p < 0.001$ ; Figure 5c, left) and DPs ( $\beta = -0.044$  [-0.019, -0.068],  $p < 0.001$ ; Figure 5c, right), with the slope for DPs not significantly lower than the slope for NTs ( $p = 0.184$ ; bootstrapping, one-tailed in accordance with predictions from the Mismatched Tuning Hypothesis, see Methods; NT and DP slopes were not significantly different when the forehead condition was excluded; Supplementary Figure S3b; Supplementary Table S8). These results indicate that DPs and NTs show comparable matching of fixations to tuning, arguing against the Mismatched Tuning Hypothesis.

### Similar perceptual processing and matched tuning across different face looking behaviors in NTs

Are differences in face fixation preference associated with differences in the way in which faces are processed? Previous studies have found no differences in face recognition *memory* performance between NT subjects who either looked high (near the eyes) or low (near the nose tip) on the face (Mehoudar et al., 2014; Peterson & Eckstein, 2013). However, whether different face looking behaviors in NTs are associated with differences in the *perceptual* processing of faces is unknown. Here, we used our new perceptual face matching task (S/D) to test for differences in perceptual face processing between those who looked high versus low on the face. Using preferred fixation data

aggregated across several hundred subjects (Or et al., 2015; Peterson & Eckstein, 2012, 2013, 2014; Peterson et al., 2016) and classification criteria (Peterson & Eckstein, 2013) from previous studies, we classified subjects as either “upper lookers” (ULs, who looked higher on the face than 75% of the population;  $N = 6$ ) or “lower lookers” (LLs, who looked lower on the face than 75% of the population;  $N = 14$ ). Performance was not significantly different between UL and LL NTs for any face task, including S/D perceptual face matching (CFMT:  $p = 0.786$ ; CELEB:  $p = 0.612$ ; S/D:  $p = 0.964$ ; EXP:  $p = 0.838$ ; CAR:  $p = 0.079$ ; two-tailed, two-sample  $t$  tests; Figure 6a, left; Supplementary Table S3). Further, UL and LL NTs did not significantly differ in the strength of their retinotopic tuning ( $p = 0.501$ , Figure 6b; no significant difference with forehead excluded,  $p = 0.930$ , Supplementary Figure S3c; two-tailed, two-sample  $t$  tests; see Supplementary Table S8). Finally, both UL NTs ( $\beta = -0.097$  [-0.051, -0.151],  $p < 0.001$ ,  $N = 24$ , Figure 6b, upper left) and LL NTs ( $\beta = -0.060$  [-0.012, -0.107],  $p = 0.008$ ,  $N = 56$ , Figure 6c, lower left) matched their preferred fixations to their retinotopic tuning, with no difference in slopes between the groups ( $p = 0.775$ ; similar results with the forehead excluded, see Supplementary Figure S3d, left; Supplementary Table S8). These results are consistent with similar perceptual processing, both quantitatively and qualitatively, across the wide range of face fixation behaviors in NTs.

### Impaired perceptual processing and mismatched tuning in subjects with DP who look low, but not high, on the face

Do UL ( $N = 4$ ) and LL ( $N = 10$ ) DPs also show similar performance to each other across tasks? As expected given our diagnostic criteria for DP (see Methods), both DP groups were strongly impaired at face memory relative to the NT group (CFMT:  $p_{ULs} = 0.004$ ,  $p_{LLs} < 0.001$ ; CELEB:  $p_{ULs} < 0.001$ ,  $p_{LLs} < 0.001$ ), with UL DP performance slightly but significantly better than LL DP performance on the CFMT ( $p = 0.018$ ) but not CELEB ( $p = 0.436$ ). This face memory impairment was domain specific for both UL and LL DPs, with performance not significantly lower relative to NTs for either EXP (task-selective;  $p_{ULs} = 0.602$ ,  $p_{LLs} = 0.386$ ) or CAR (stimulus-selective;  $p_{ULs} = 0.691$ ,  $p_{LLs} = 0.654$ ). In contrast, UL DPs performed much better than LL DPs at face perception (S/D:  $p = 0.013$ ). In fact, while LL DPs were strongly impaired at face perception relative to both the entire NT group ( $p < 0.001$ ) and LL NTs specifically ( $p < 0.001$ ), UL DPs did not perform significantly lower than either the whole NT group ( $p = 0.106$ ) or UL NTs in particular ( $p = 0.385$ ; two-tailed, two sample  $t$  tests; Figure 6a, right;

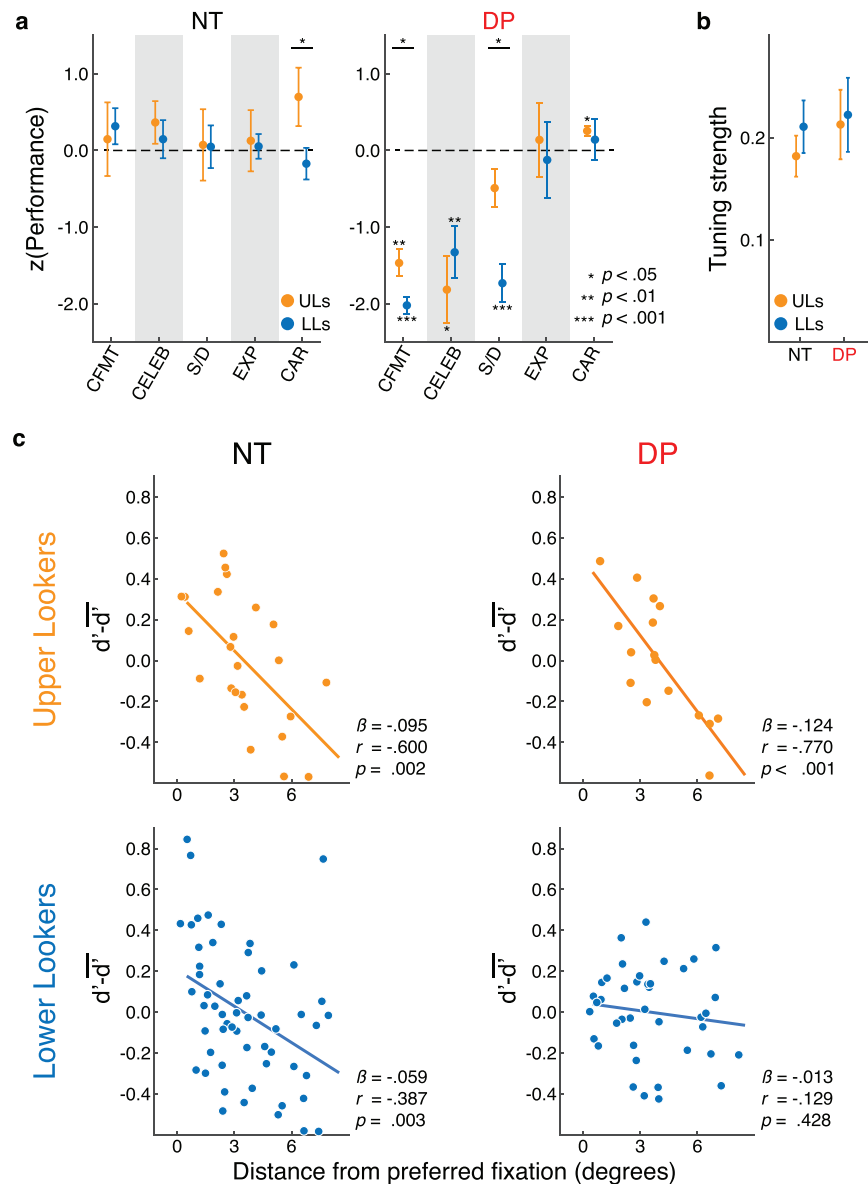


Figure 6. Evidence for distinct subgroups of DP. (a) UL (orange) and LL (blue) performance, converted to z scores relative to performance across all NTs, for each task for NTs (left) and DPs (right). Dots are means and error bars are 1 SEM across subjects. (b) Tuning strength for ULs and LLs. (c) S/D performance (centered on each subject's average performance across forced fixation locations, separately for each subject) as a function of the distance from forced fixation locations to subjects' preferred fixation location for NTs (left) and DPs (right) separated by ULs (top) and LLs (bottom). Dots are individual subjects at different forced fixation locations and lines are linear regression fits.

Supplementary Table S3). This disparity in face perception ability between UL and LL DPs was not associated with differences in retinotopic tuning strength ( $p = 0.881$ , Figure 6b; no significant difference with forehead location excluded,  $p = 0.845$ , Supplementary Figure S3c; two-tailed, two-sample  $t$  tests; see Supplementary Table S8).

We found that this contrast in perceptual face processing extended to the matching of fixations to retinotopic tuning, as performance declined significantly with distance from preferred fixation location for

UL DPs ( $\beta = -0.123 [-0.076, -0.169]$ ,  $p = 0.001$ ,  $N = 16$ ; Figure 6c, upper right) but not LL DPs ( $\beta = -0.013 [0.016, -0.039]$ ,  $p = 0.180$ ,  $N = 40$ ; Figure 6c, lower right). The slope for LL DPs was significantly smaller than the slope for LL NTs ( $p = 0.047$ ; one-tailed test in accordance with the predictions of the Mismatched Tuning Hypothesis), UL DPs ( $p = 0.002$ , two-tailed), and the NT group as a whole ( $p = 0.010$ , one-tailed; the same pattern was observed with the forehead condition excluded; Supplementary Figure S3d, right; Supplementary Table S8). Thus, this preliminary evidence



suggests that looking high on the face may distinguish one DP phenotype in which strong, retinotopically matched perceptual representations are generated but memory-related impairments lead to poor face recognition. Looking low on the face may mark a second phenotype where deficits arise from poor retinotopic matching at early stages of perceptual encoding. While the sample sizes generated by our classification criteria preclude strong conclusions, the stark differences between LL DPs and each other group warrant future investigation.

## Discussion

Here we measured face fixation behavior and face recognition performance as a function of fixation location in individuals with DP in order to test four hypotheses for why face recognition is impaired in DP. Face processing deficits in DP subjects could not be explained by atypical fixation strategies, as they preferred similar fixation locations to NT subjects (mean; Figure 4a and b) and showed similar consistency across trials in selecting those fixation locations (variance; Figure 4c and d). These findings fail to support the Poor Information and Inconsistent Eye Fixation Hypotheses (Figure 1a and b). The strength of retinotopic tuning of perceptual face processing in DPs was either similar or stronger than in NTs, contradicting the Weak Retinotopic Tuning Hypothesis (Figure 5a; Supplementary Figure S3a). Across all subjects, both groups matched their preferred fixations to their retinotopic tuning to a similar degree, failing to support the Mismatched Tuning Hypothesis for the DP group as a whole (Figure 5c). However, we found evidence for two distinct subgroups of DP: those who looked high on the face showed typical face perception ability and strong matching between preferred fixation locations and retinotopic tuning, whereas those who looked low on the face were impaired in both face memory and perception and showed no significant relationship between fixation preference and retinotopic tuning (Figure 6b, right). Overall, the findings argue strongly against three of our hypotheses, and support the Mismatched Tuning Hypothesis as a mechanism of impaired face recognition for almost half of the DP group. Next we discuss these conclusions in more detail.

### Poor Information Hypothesis

Face recognition performance is maximized when eye movements most closely align the variation in information density across the face with the variation

in resolution across the visual field. Computational models show that the theoretically optimal alignment occurs when fixation is directed toward a region between the eyes and the nose tip (Or et al., 2015; Peterson & Eckstein, 2012, 2013, 2014; Tsank & Eckstein, 2017). NT individuals preferentially fixate this region, and both recognition performance (de Haas & Schwarzkopf, 2018; Or et al., 2015; Peterson & Eckstein, 2012, 2013; Peterson et al., 2016) and the response of face-selective cortical regions (de Haas et al., 2016; Zerouali, Lina, & Jemel, 2013; Stacchi, Ramon, Lao, & Caldara, 2019) are maximized when fixating near the theoretical optimum location. The Poor Information Hypothesis holds that individuals with DP fixate outside the optimal region, directly impairing face recognition accuracy through a reduction in the quality of information entering cortex. Our results do not support this hypothesis: NT and DP fixations were indistinguishable from each other, with both groups preferring to fixate about 70% of the distance downward from the eyes to the nose tip on average (Figure 4a).

Although some papers have found typical fixation behavior on faces in DP (Bate, Haslam, Jansari, & Hodgson, 2009; Bate, Haslam, Tree, & Hodgson, 2008), the similarity of NT and DP gaze patterns we found appears to contradict several other previous reports of increased fixation of lower or external face regions in DP (Bobak, Parris, Gregory, Bennetts, & Bate, 2017; Pizzamiglio et al., 2017; Schwarzer et al., 2007). This discrepancy may be due to the speeded nature of the current tasks, which allowed for only one or two eye movements, whereas faces were displayed for longer periods in studies finding abnormal fixations in DP (Bobak et al., 2017; Pizzamiglio et al., 2017; Schwarzer et al., 2007). While NT face recognition performance generally asymptotes after one or two fixations (Hsiao & Cottrell, 2008; Or et al., 2015) so that they no longer need to search the face for information, DPs' impaired face processing causes them to take longer to make decisions about facial identity than NTs (Avidan, Tanzer, & Behrmann, 2011; Palermo et al., 2011), and during this additional time they may sample lower and external face regions. Thus, it is possible that with a longer stimulus presentation, we would have found that after the initial fixations, eye movements in the DPs shift to more atypical locations (e.g., very low on the face).

It is worth noting that NT individuals looked substantially lower in this study than previously reported (~70% of the distance from the eyes to the nose tip versus ~40%; Mehoudar et al., 2014; Or et al., 2015; Peterson & Eckstein, 2012, 2013, 2014; Peterson et al., 2016; Tsank & Eckstein, 2017). This discrepancy may be due to differences in the ages of the participants: Subjects in previous studies were almost

exclusively undergraduate students, while the average age of the current participants was 35. Splitting the current NT group into under- and over-25 age brackets, we found that younger subjects indeed fixated significantly higher than older subjects, and not significantly different than previous findings. In contrast, preferred fixation was not systematically related to age in subjects with DP; DPs under the age of 25 looked as low as their over-25 counterparts (see Supplementary Figure S2). Thus, lower face looking in the NT group appears to reflect a migration from a preference to look higher on the face early in life, potentially as an adaptation to hearing loss through increased reliance on visual information from the mouth for speech perception (Gurler, Doyle, Walker, Magnotti, & Beauchamp, 2015), while individuals with DP look lower from a younger age. However, even if (younger) DPs look somewhat lower on the face, the results do not strongly support the Poor Information Hypothesis. Both groups looked lower on average than their younger NT cohorts measured in previous studies, but well within the good information region between the eyes and the nose tip.

### Inconsistent Eye Fixation Hypothesis

The narrow retinotopic tuning of the face system leads to performance penalties when individuals fixate even a small distance from their optimal point. Do recognition deficits in DP arise from a failure to consistently target the same location with high-precision saccades? Our results do not support this hypothesis: The variance in initial fixation location across image presentations for the DP and NT groups was quite similar (Figure 4b). In fact, for both groups, the precision of saccades onto the face approached the precision of saccades to highly salient peripheral point targets (Kowler & Blaser, 1995).

### Weak Retinotopic Tuning Hypothesis

Excellent face recognition in NT subjects is accomplished with narrow retinotopic tuning combined with precise eye movements. Might recognition deficits in DP then arise from broadened retinotopic tuning, where a more even distribution of resources across a wider range of retinotopic positions reduces the resource allocation, and thus the encoding capacity, of face representations at the best-tuned location? Our results provide no support for this hypothesis, as the difference in performance between subjects' best-performing forced fixation location and neighboring fixation locations was at least as large for the DP group (Figure 5a; Supplementary Figure S3a).

### Mismatched Tuning Hypothesis

Do recognition deficits in DP arise from a tendency to fixate poorly tuned locations? As expected from prior findings (Or et al., 2015; Peterson & Eckstein, 2012, 2013; Tsank & Eckstein, 2017), the farther NT subjects were forced to fixate from their preferred location, the worse they performed on the S/D face discrimination task (Figure 5c, left). The same relationship between performance and distance from preferred fixation was found across the DP group as a whole (Figure 5c, right). However, our post hoc separation of the groups into categories used in a previous studies (Peterson & Eckstein, 2013; Peterson et al., 2016), ULs and LLs, provides preliminary evidence for two distinct groups of DPs. UL DPs showed strong matching of their preferred fixation to their retinotopic tuning (Figure 6b, upper right), while LL DPs did not systematically fixate their optimal tuning location (Figure 6b, lower right). Critically, this distinction was not present for NTs, with strong matching for ULs and LLs alike (Figure 6b, left). Thus, a failure to match eye movements to the tuning of retinotopic face encoding may contribute to face recognition deficits in DPs who look low on the face, whereas DPs who look high have deficits due to other factors.

In sum, our main planned analyses find evidence against our first three hypotheses. Our post hoc analyses support the Mismatched Tuning Hypothesis for a subgroup of DP. In addition to this main conclusion, the rich data set collected here provide insights into several other facets of DP, as we discuss next.

### Implications of UL and LL DP subgroups

Beyond differences in preferred fixation behavior and matched tuning, our preliminary post hoc analyses further suggest that the two distinct DP subgroups also differ in the stages of visual processing where face recognition impairments may originate. ULs with DP showed impaired face memory but normal face perception, while LLs with DP showed impairments to both face memory and perception (see Results and Figure 6). Although the existence of these two subgroups will need to be further tested in the future, we consider here how this distinction may be related to the neural basis of face processing in NTs and DPs (Haxby et al., 2001; Haxby, Hoffman, & Gobbini, 2000; Kanwisher, 1998; Kanwisher, McDermott, & Chun, 1997), and how it might help explain the inconsistent reports of functional and structural atypicalities in the DP imaging literature. Specifically, we might expect posterior face-selective regions impli-

cated in initial structural encoding (e.g., OFA, FFA) to respond typically to faces in UL DPs, as has been found in several studies (Avidan & Behrmann, 2009; Avidan, Hasson, Malach, & Behrmann, 2005; Hasson, Avidan, Deouell, Bentin, & Malach, 2003). Face memory impairments in UL DPs might result from compromised response properties of higher order downstream regions implicated in face memory (e.g., anterior temporal lobe, medial temporal lobe; Damasio, Tranel, & Damasio, 1990; Haxby et al., 2000; Quiroga, Reddy, Kreiman, Koch, & Fried, 2005), or from disruptions to their afferent connections from lower order regions (Avidan & Behrmann, 2009; Kravitz, Saleem, Baker, Ungerleider, & Mishkin, 2013). The perceptual deficits for DPs, on the other hand, may predict disruptions to early stages of face processing, consistent with studies finding atypical structural (Garrido et al., 2009; Gomez et al., 2015; Song et al., 2015) and response (Bentin, DeGutis, D’Esposito, & Robertson, 2007; Furl, Garrido, Dolan, Driver, & Duchaine, 2010; Jiahui, Yang, & Duchaine, 2018; Lohse et al., 2016; Thomas et al., 2009; Towler, Gosling, Duchaine, & Eimer, 2012; Towler, Parketny, & Eimer, 2016) properties within and between posterior face selective regions. Moreover, the possibly narrowed spatial tuning of DP face perception is consistent with a report of smaller receptive fields in face selective regions for individuals with DP (Witthoft et al., 2016), and emphasizes the need for careful measurement and control of fixation in neuroimaging studies.

### Gaze specificity in NTs and DP

Are eye movements on faces in NTs directed by a general-purpose eye movement system or by a subsystem specialized for faces, and might a lack of specialization contribute to recognition impairments in DP? Extensive evidence suggests that face recognition in NTs engages distinct cognitive mechanisms from recognition of other stimulus classes (Haxby et al., 2001; Haxby et al., 2000; Kanwisher, 1998; Kanwisher et al., 1997; Tsao, Moeller, & Freiwald, 2008). The existence of face-specific mechanisms in the occipital and temporal lobe raises the question of whether eye movements to faces are controlled by face-specific processes. Previous studies that have reported stable individual differences in preferred face fixations (Mehouard et al., 2014; Or et al., 2015; Peterson & Eckstein, 2013; Peterson et al., 2016) lacked nonface control conditions, and thus were unable to determine whether individuals’ distinct fixation preferences on faces reflect a property of either a face-specific eye movement process or of the eye movement system in general. Here, we measured preferred fixations for two face tasks (CELEB, EXP) and one nonface task (CAR)

within individual subjects using a common experimental protocol. In both the NT and DP groups, individuals’ preferred fixations for face identification were strongly predictive of their preferred fixations when recognizing expressions but not cars (Figure 4c). The conservation of individuals’ distinct gaze behavior across face tasks, but not across stimulus categories, is consistent with an eye movement subsystem specialized for the distinct and stereotyped visual structures of faces. Further, the finding of a similar degree of face-selectivity of fixation behavior in both NTs and DPs does not support the hypothesis that reliance on nonspecific eye movement mechanisms contributes to recognition deficits in DP.

### Encoding and tuning for perception and memory

Because our S/D task was designed to isolate perceptual processing by minimizing memory requirements, we can use it to answer two open questions. In this task, participants saw each unfamiliar identity just once, eliminating any possible effects of memory across trials that have confounded some previous tests of face perception and minimizing memory requirements to the brief (950 ms) interstimulus interval. In contrast, the CFMT task entails face memory (learning new faces during the course of the experiment) and the CELEB task tests participants’ memory for famous faces that they bring to the lab from real-world experience. A comparison of performance across these three tasks enables us to ask: (a) Is matched tuning in NTs a perceptual effect or a memory effect? and (b) To what extent do impairments in DP result from deficits in face perception, memory, or their interaction?

The first question concerns our prior finding that NT participants perform best at face recognition when they fixate at their preferred location (Or et al., 2015; Peterson & Eckstein, 2013). In these studies, participants studied a set of faces while freely fixating and attempted to recognize the faces in a subsequent forced fixation task. We interpreted the superior performance at the preferred fixation location to result from superior face representation at that location, but this finding could reflect a benefit of matching fixation location at test to fixation location at study since it is likely that participants fixated those faces at study using their preferred fixation location. However, because each identity was seen on only one trial while fixation was controlled in our S/D task, the results in NTs show that perceptual encoding contributes to the retinotopic specificity of face recognition.

With regard to the second question, while LL DPs were impaired at both face memory and face perception, UL DPs were impaired at face memory



only. These results support the hypothesis that face recognition deficits in DP can be associated with a selective disruption to early perceptual encoding stages (e.g., for LLs), and/or to a later stage where effectively encoded perceptual information is not properly encoded, retrieved, or compared for memory processes (e.g., ULs). The finding of typical perceptual encoding in only a small subgroup of our DP subjects is consistent with a recent paper by Biotti, Gray, and Cook (2019) arguing that impaired perceptual encoding is a pervasive feature of DP. They found that DP performance on a delayed match-to-sample task was impaired to the same degree for both short (1 s) and long (6 s) retention intervals, suggesting a typical ability to retain information in short-term face memory, at least across brief spans of time. Matching performance was also highly correlated with performance on the Cambridge Face Perception Test (Duchaine, Germine, & Nakayama, 2007), suggesting impairments to face memory are often inherited from impaired perceptual encoding. The authors conclude that it remains possible that some DPs have typical face perception and impaired face memory; our findings suggest that this form of DP, if it occurs, may be associated with a preference to look high on the face. While the dissociation between memory and perception suggested by the UL versus LL DP distinction is compelling, it is also preliminary, and emphasizes the need for care when designing paradigms that isolate perception or memory.

## Conclusion

We found no differences between NTs and subjects with DP in either the location or consistency of preferred face fixation behavior or in the strength of their retinotopic tuning, results that are inconsistent with the Poor Information, Inconsistent Eye Fixation, and Weak Retinotopic Tuning Hypotheses. In post hoc analyses, we found that deficits in DPs who looked high on the face were isolated to face memory and that these subjects matched their fixations to the tuning of their retinotopic perceptual encoding to the same degree as NTs. In contrast, DPs who looked low on the face had profound impairments to both face memory and face perception and they did not match their fixations to their retinotopic tuning, supporting the Mismatched Tuning Hypothesis as a mechanism of impaired face recognition in this subgroup of DP.

*Keywords:* developmental prosopagnosia, face recognition, face perception, eye tracking, eye movements, retinotopic encoding

## Acknowledgments

This work was supported by the Center for Brains, Minds and Machines (CBMM), funded by National Science Foundation Science and Technology Center award Computing and Communication Foundations – 1231216. The authors have no financial or proprietary interests.

Commercial relationships: none.

Corresponding author: Matthew F. Peterson.

Email: mfpeters@mit.edu.

Address: Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA.

## References

- Ara Nefian face recognition page. (n.d.). Retrieved from [http://www.anefian.com/research/face\\_reco.htm](http://www.anefian.com/research/face_reco.htm)
- Avidan, G., & Behrmann, M. (2009). Functional MRI reveals compromised neural integrity of the face processing network in congenital prosopagnosia. *Current Biology*, *19*(13), 1146–1150, <https://doi.org/10.1016/j.cub.2009.04.060>.
- Avidan, G., Hasson, U., Malach, R., & Behrmann, M. (2005). Detailed exploration of face-related processing in congenital prosopagnosia: II. Functional neuroimaging findings. *Journal of Cognitive Neuroscience*, *17*(7), 1150–1167, <https://doi.org/10.1162/0898929054475145>.
- Avidan, G., Tanzer, M., & Behrmann, M. (2011). Impaired holistic processing in congenital prosopagnosia. *Neuropsychologia*, *49*(9), 2541–2552, <https://doi.org/10.1016/j.neuropsychologia.2011.05.002>.
- Bate, S., Haslam, C., Jansari, A., & Hodgson, T. L. (2009). Covert face recognition relies on affective valence in congenital prosopagnosia. *Cognitive Neuropsychology*, *26*(4), 391–411, <https://doi.org/10.1080/02643290903175004>.
- Bate, S., Haslam, C., Tree, J. J., & Hodgson, T. L. (2008). Evidence of an eye movement-based memory effect in congenital prosopagnosia. *Cortex*, *44*(7), 806–819, <https://doi.org/10.1016/j.cortex.2007.02.004>.
- Bentin, S., DeGutis, J. M., D’Esposito, M., & Robertson, L. C. (2007). Too many trees to see the forest: Performance, event-related potential, and functional magnetic resonance imaging manifestations of integrative congenital prosopagnosia.



- Journal of Cognitive Neuroscience*, 19(1), 132–146, <https://doi.org/10.1162/jocn.2007.19.1.132>.
- Biotti, F., Gray, K. L. H., & Cook, R. (2019). Is developmental prosopagnosia best characterised as an apperceptive or mnemonic condition? *Neuropsychologia*, 124, 285–298, <https://doi.org/10.1016/j.neuropsychologia.2018.11.014>.
- 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. *The Quarterly Journal of Experimental Psychology*, 70(2), 201–217, <https://doi.org/10.1080/17470218.2016.1161059>.
- Caldara, R., Schyns, P., Mayer, E., Smith, M. L., Gosselin, F., & Rossion, B. (2005). Does prosopagnosia take the eyes out of face representations? Evidence for a defect in representing diagnostic facial information following brain damage. *Journal of Cognitive Neuroscience*, 17(10), 1652–1666, <https://doi.org/10.1162/089892905774597254>.
- Computational Vision: Archive*. (n.d.). Retrieved from <http://www.vision.caltech.edu/html-files/archive.html>
- Dalrymple, K. A., Fletcher, K., Corrow, S., das Nair, R., Barton, J. J. S., Yonas, A., & Duchaine, B. (2014). “A room full of strangers every day”: The psychosocial impact of developmental prosopagnosia on children and their families. *Journal of Psychosomatic Research*, 77(2), 144–150, <https://doi.org/10.1016/j.jpsychores.2014.06.001>.
- 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, <https://doi.org/10.1016/j.dcn.2014.07.003>.
- Damasio, A. R., Tranel, D., & Damasio, H. (1990). Face agnosia and the neural substrates of memory. *Annual Review of Neuroscience*, 13(1), 89–109, <https://doi.org/10.1146/annurev.ne.13.030190.000513>.
- de Haas, B., & Schwarzkopf, D. S. (2018). Feature–location effects in the Thatcher illusion. *Journal of Vision*, 18(4):16, 1–12, <https://doi.org/10.1167/18.4.16>. [PubMed] [Article]
- de Haas, B., Schwarzkopf, D. S., Alvarez, I., Lawson, R. P., Henriksson, L., Kriegeskorte, N., & Rees, G. (2016). Perception and processing of faces in the human brain is tuned to typical feature locations. *Journal of Neuroscience*, 36(36), 9289–9302, <https://doi.org/10.1523/JNEUROSCI.4131-14.2016>.
- 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, <https://doi.org/10.1080/02643290701380491>.
- Duchaine, B., & Nakayama, K. (2005). Dissociations of face and object recognition in developmental prosopagnosia. *Journal of Cognitive Neuroscience*, 17(2), 249–261, <https://doi.org/10.1162/0898929053124857>.
- Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, 44(4), 576–585, <https://doi.org/10.1016/j.neuropsychologia.2005.07.001>.
- Fiset, D., Blais, C., Royer, J., Richoz, A.-R., Dugas, G., & Caldara, R. (2017). Mapping the impairment in decoding static facial expressions of emotion in prosopagnosia. *Social Cognitive and Affective Neuroscience*, 12(8), 1334–1341, <https://doi.org/10.1093/scan/nsx068>.
- Furl, N., Garrido, L., Dolan, R. J., Driver, J., & Duchaine, B. (2010). Fusiform gyrus face selectivity relates to individual differences in facial recognition ability. *Journal of Cognitive Neuroscience*, 23(7), 1723–1740, <https://doi.org/10.1162/jocn.2010.21545>.
- Garrido, L., Furl, N., Draganski, B., Weiskopf, N., Stevens, J., Tan, G. C.-Y., . . . Duchaine, B. (2009). Voxel-based morphometry reveals reduced grey matter volume in the temporal cortex of developmental prosopagnosics. *Brain*, 132(12), 3443–3455, <https://doi.org/10.1093/brain/awp271>.
- Geskin, J., & Behrmann, M. (2018). Congenital prosopagnosia without object agnosia? A literature review. *Cognitive Neuropsychology*, 35(1–2), 4–54, <https://doi.org/10.1080/02643294.2017.1392295>.
- 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, <https://doi.org/10.1016/j.neuron.2014.12.027>.
- Gurler, D., Doyle, N., Walker, E., Magnotti, J., & Beauchamp, M. (2015). A link between individual differences in multisensory speech perception and eye movements. *Attention, Perception, & Psychophysics*, 77(4), 1333–1341, <https://doi.org/10.3758/s13414-014-0821-1>.
- Hasson, U., Avidan, G., Deouell, L. Y., Bentin, S., & Malach, R. (2003). Face-selective activation in a congenital prosopagnosic subject. *Journal of Cog-*

- nitive Neuroscience*, 15(3), 419–431, <https://doi.org/10.1162/089892903321593135>.
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001, September 28). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293(5539), 2425–2430, <https://doi.org/10.1126/science.1063736>.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, 4(6), 223–233, [https://doi.org/10.1016/S1364-6613\(00\)01482-0](https://doi.org/10.1016/S1364-6613(00)01482-0).
- Hsiao, J. H., & Cottrell, G. (2008). Two fixations suffice in face recognition. *Psychological Science*, 19(10), 998–1006, <https://doi.org/10.1111/j.1467-9280.2008.02191.x>.
- Jiahui, G., Yang, H., & Duchaine, B. (2018). Developmental prosopagnosics have widespread selectivity reductions across category-selective visual cortex. *Proceedings of the National Academy of Sciences*, 115(28), E6418–E6427, <https://doi.org/10.1073/pnas.1802246115>.
- Kanwisher, N. (1998). The modular structure of human visual recognition: Evidence from functional imaging. In M. Sabourin, M. Robert, & F. Craik (Eds.), *Advances in psychological science, Volume 2: Biological and cognitive aspects* (pp. 199–213). Hove, UK: Psychology Press.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, 17(11), 4302–4311.
- Kennerknecht, I., Grueter, T., Welling, B., Wentzek, S., Horst, J., Edwards, S., & Grueter, M. (2006). First report of prevalence of non-syndromic hereditary prosopagnosia (HPA). *American Journal of Medical Genetics Part A*, 140A(15), 1617–1622, <https://doi.org/10.1002/ajmg.a.31343>.
- Kennerknecht, I., Ho, N. Y., & Wong, V. C. N. (2008). Prevalence of hereditary prosopagnosia (HPA) in Hong Kong Chinese population. *American Journal of Medical Genetics Part A*, 146A(22), 2863–2870, <https://doi.org/10.1002/ajmg.a.32552>.
- Kowler, E., & Blaser, E. (1995). The accuracy and precision of saccades to small and large targets. *Vision Research*, 35(12), 1741–1754, [https://doi.org/10.1016/0042-6989\(94\)00255-K](https://doi.org/10.1016/0042-6989(94)00255-K).
- Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2013). The ventral visual pathway: An expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences*, 17(1), 26–49, <https://doi.org/10.1016/j.tics.2012.10.011>.
- Lohse, M., Garrido, L., Driver, J., Dolan, R. J., Duchaine, B., & Furl, N. (2016). Effective connectivity from early visual cortex to posterior occipitotemporal face areas supports face selectivity and predicts developmental prosopagnosia. *Journal of Neuroscience*, 36(13), 3821–3828, <https://doi.org/10.1523/JNEUROSCI.3621-15.2016>.
- Mehouadar, E., Arizpe, J., Baker, C. I., & Yovel, G. (2014). Faces in the eye of the beholder: Unique and stable eye scanning patterns of individual observers. *Journal of Vision*, 14(7):6, 1–11, <https://doi.org/10.1167/14.7.6>. [PubMed] [Article]
- Or, C. C.-F., Peterson, M. F., & Eckstein, M. P. (2015). Initial eye movements during face identification are optimal and similar across cultures. *Journal of Vision*, 15(13):12, 1–25, <https://doi.org/10.1167/15.13.12>. [PubMed] [Article]
- 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, <https://doi.org/10.1016/j.neuropsychologia.2011.02.021>.
- Peterson, M. F., & Eckstein, M. P. (2012). Looking just below the eyes is optimal across face recognition tasks. *Proceedings of the National Academy of Sciences*, 109(48), E3314–E3323, <https://doi.org/10.1073/pnas.1214269109>.
- Peterson, M. F., & Eckstein, M. P. (2013). Individual differences in eye movements during face identification reflect observer-specific optimal points of fixation. *Psychological Science*, 24(7), 1216–1225, <https://doi.org/10.1177/0956797612471684>.
- Peterson, M. F., & Eckstein, M. P. (2014). Learning optimal eye movements to unusual faces. *Vision Research*, 99, 57–68, <https://doi.org/10.1016/j.visres.2013.11.005>.
- Peterson, M. F., Lin, J., Zaun, I., & Kanwisher, N. (2016). Individual differences in face-looking behavior generalize from the lab to the world. *Journal of Vision*, 16(7):12, 1–18, <https://doi.org/10.1167/16.7.12>. [PubMed] [Article]
- Pizzamiglio, M. R., Luca, M. D., Vita, A. D., Palermo, L., Tanzilli, A., Dacquino, C., & Piccardi, L. (2017). Congenital prosopagnosia in a child: Neuropsychological assessment, eye movement recordings and training. *Neuropsychological Rehabilitation*, 27(3), 369–408, <https://doi.org/10.1080/09602011.2015.1084335>.
- Psychological Image Collection at Stirling. (n.d.). Retrieved from <http://pics.psych.stir.ac.uk/>

- Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005, June 23). Invariant visual representation by single neurons in the human brain. *Nature*, *435*(7045), 1102–1107, <https://doi.org/10.1038/nature03687>.
- Royer, J., Blais, C., Charbonneau, I., Déry, K., Tardif, J., Duchaine, B., . . . Fiset, D. (2018). Greater reliance on the eye region predicts better face recognition ability. *Cognition*, *181*, 12–20, <https://doi.org/10.1016/j.cognition.2018.08.004>.
- Schwarzer, G., Huber, S., Grüter, M., Grüter, T., Groß, C., Hipfel, M., & Kennerknecht, I. (2007). Gaze behaviour in hereditary prosopagnosia. *Psychological Research*, *71*(5), 583–590, <https://doi.org/10.1007/s00426-006-0068-0>.
- 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, <https://doi.org/10.1016/j.neuropsychologia.2015.10.010>.
- Stacchi, L., Ramon, M., Lao, J., & Caldara, R. (2019). Neural representations of faces are tuned to eye movements. *Journal of Neuroscience*, *39*(21), 4113–4123.
- Susilo, T., & Duchaine, B. (2013). Advances in developmental prosopagnosia research. *Current Opinion in Neurobiology*, *23*(3), 423–429, <https://doi.org/10.1016/j.conb.2012.12.011>.
- Tardif, J., Morin Duchesne, X., Cohan, S., Royer, J., Blais, C., Fiset, D., . . . Gosselin, F. (2019). Use of face information varies systematically from developmental prosopagnosics to super-recognizers. *Psychological Science*, *30*(2), 300–308, <https://doi.org/10.1177/0956797618811338>.
- Thomas, C., Avidan, G., Humphreys, K., Jung, K., Gao, F., & Behrmann, M. (2009). Reduced structural connectivity in ventral visual cortex in congenital prosopagnosia. *Nature Neuroscience*, *12*(1), 29–31, <https://doi.org/10.1038/nn.2224>.
- Towler, J., Gosling, A., Duchaine, B., & Eimer, M. (2012). The face-sensitive N170 component in developmental prosopagnosia. *Neuropsychologia*, *50*(14), 3588–3599, <https://doi.org/10.1016/j.neuropsychologia.2012.10.017>.
- Towler, J., Parketny, J., & Eimer, M. (2016). Perceptual face processing in developmental prosopagnosia is not sensitive to the canonical location of face parts. *Cortex*, *74*, 53–66, <https://doi.org/10.1016/j.cortex.2015.10.018>.
- Tsank, Y., & Eckstein, M. P. (2017). Domain specificity of oculomotor learning after changes in sensory processing. *Journal of Neuroscience*, *1208–1217*, <https://doi.org/10.1523/JNEUROSCI.1208-17.2017>.
- Tsao, D. Y., Moeller, S., & Freiwald, W. A. (2008). Comparing face patch systems in macaques and humans. *Proceedings of the National Academy of Sciences*, *105*(49), 19514–19519, <https://doi.org/10.1073/pnas.0809662105>.
- Witthoft, N., Poltoratski, S., Nguyen, M., Golarai, G., Liberman, A., LaRocque, K. F., . . . Grill-Spector, K. (2016). Reduced spatial integration in the ventral visual cortex underlies face recognition deficits in developmental prosopagnosia. *BioRxiv*, 051102, <https://doi.org/10.1101/051102>.
- 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, <https://doi.org/10.1016/j.jpsychores.2008.03.013>.
- Zerouali, Y., Lina, J.-M., & Jemel, B. (2013). Optimal eye-gaze fixation position for face-related neural responses. *PLoS One*, *8*(6), e60128, <https://doi.org/10.1371/journal.pone.0060128>.