Search for Face Identity or Expression: Set Size Effects in Developmental Prosopagnosia

Sara Djouab[1,2*], Andrea Albonico[1*], Shanna C. Yeung[1], Manuela Malaspina[1], Anna Mogard[3], Rachel Wahlberg[3], Sherryse L. Corrow[1,3], and Jason J. S. Barton[1]

Abstract

■ The set size effect during visual search indexes the effects of processing load and thus the efficiency of perceptual mechanisms. Our goal was to investigate whether individuals with developmental prosopagnosia show increased set size effects when searching faces for face identity and how this compares to search for face expression. We tested 29 healthy individuals and 13 individuals with developmental prosopagnosia. Participants were shown sets of three to seven faces to judge whether the identities or expressions of the faces were the same across all stimuli or if one differed. The set size effect was the slope of the linear regression between the number of faces in the array and the response time. Accuracy was similar in both controls and prosopagnosic participants. Developmental prosopagnosic participants displayed increased set size effects in face identity search but not in expression search. Single-participant analyses reveal that 11 developmental prosopagnosic participants showed a putative classical dissociation, with impairments in identity but not expression search. Signal detection theory analysis showed that identity set size effects were highly reliable in discriminating prosopagnosic participants from controls. Finally, the set size ratios of same to different trials were consistent with the predictions of self-terminated serial search models for control participants and prosopagnosic participants engaged in expression search but deviated from those predictions for identity search by the prosopagnosic cohort. We conclude that the face set size effect reveals a highly prevalent and selective perceptual inefficiency for processing face identity in developmental prosopagnosia. ■

INTRODUCTION

Prosopagnosia refers to an impairment in the ability to recognize the identity of faces (Corrow, Dalrymple, & Barton, 2016). Prosopagnosic individuals lack familiarity for faces they have seen and have trouble identifying the person to whom the face belongs. It can be acquired from damage to the occipitotemporal and/or anterior temporal cortex (Davies-Thompson, Pancaroglu, & Barton, 2014), which contain components of a face processing network demonstrable with functional imaging in healthy individuals (Haxby, Hoffman, & Gobbini, 2000). A developmental form of prosopagnosia also exists, which by definition lacks any gross structural damage to the brain.

The diagnosis of prosopagnosia rests on the demonstration of impaired long-term familiarity with faces, as with various tests of famous face recognition, or short-term familiarity for recently viewed faces, as with the Cambridge Face Memory Test (Duchaine & Nakayama, 2006) and the Warrington Recognition Memory Test (Warrington, 1984). These can be supplemented by questionnaires about subjective experience in daily life (Shah, Sowden, Gaule, Catmur, & Bird, 2015) and exclusion of confounding conditions such as autism and general memory or perceptual impairment (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001; Warrington & James, 1991).

The reliability of diagnostic instruments is particularly critical for developmental prosopagnosia, which as yet lacks any objective imaging, genetic, or pathological biomarkers. Behavioral instruments, both questionnaires and even the so-called objective tests, are potentially vulnerable to subject factors such as introspection, internal criterion setting, and even malingering. For this reason, it is desirable to develop complementary methods to enhance confidence in the diagnosis.

One possibility is to consider measures that reflect perceptual efficiency rather than accuracy. Although there are both apperceptive and amnestic variants of acquired prosopagnosia (Damasio, Tranel, & Damasio, 1990; Davies-Thompson et al., 2014) and there may be similar heterogeneity in developmental prosopagnosia (Ulrich et al., 2017; Dalrymple, Garrido, & Duchaine, 2014), it has been argued that some perceptual dysfunction may be pervasive in the latter condition (Biotti, Gray, & Cook, 2019). If so, perceptual mechanisms could be characterized by reduced efficiency for faces in developmental prosopagnosia.

We recently reported on the visual search for complex objects, namely, faces, visual words, and cars (Hemström,

[1]University of British Columbia, [2]University of Auvergne, Clermont-Ferrand, France, [3]Bethel University, Minneapolis, MN
*S. D. and A. A. are joint first authors.

https://doi.org/10.1162/jocn_a_01519

Albonico, Djouab, & Barton, 2019). An important outcome variable in that study was the set size effect, the increase in search time required for every additional stimulus in the array, which indexes the effect of processing load. We showed that, with inverted faces, which healthy participants process inefficiently (Albonico, Furubacke, Barton, & Oruc, 2018), set size effects are significantly prolonged. In the current study, our primary goal was to evaluate whether the impairments with face processing in developmental prosopagnosia would be similarly reflected in exaggerated effects of processing load, resulting in an increased set size effect when participants are searching for face identity.

If there are face processing inefficiencies in developmental prosopagnosia, one important question is how specific this effect is for the identity of faces. Faces inform not only about the identity of the person but also their emotional state, direction of gaze, age, gender, and attractiveness, among others. Indeed, one of the key features of a prominent cognitive model of face processing is a divergence between identity and expression processing (Bruce & Young, 1986). Deficits in identity and expression processing can be dissociated with acquired brain lesions (Fox, Hanif, Iaria, Duchaine, & Barton, 2011). Whether this is also true about developmental prosopagnosia is not certain (Biotti & Cook, 2016; Humphreys, Avidan, & Behrmann, 2007), and most prior reports can be challenged on the basis that they did not evaluate identity and expression processing in the same manner. Our second goal was to determine, using the same methods and with tests comparable in difficulty, whether participants with developmental prosopagnosia showed a dissociation between identity and expression processing or if, on the other hand, identity and expression search performance were correlated.

Finally, one intriguing result in our prior study related to how participants searched for faces, as revealed in a comparison of *same* versus *different* trials. The search for face identity with upright faces was consistent with the predictions of a self-terminating serial search process, with a set size effect on *same* trials that was double that on *different* trials (Wolfe, 1998). However, search with inverted faces violated those predictions, with a set size ratio of *same* over *different* trials that was nearer a value of one, not two. Inversion effects are often considered an index of expert face processing, whose orientation dependency reflects a strong bias of natural experience toward upright faces (Farah, Tanaka, & Drain, 1995; Valentine, 1988). Some studies show reduced face inversion effects in acquired and developmental prosopagnosia (Klargaard, Starrfelt, & Gerlach, 2018), in keeping with a loss of this orientation-dependent proficiency for faces. Our last question was whether the search for face identity by participants with developmental prosopagnosia would show a deviation in set size ratios of *same* over *different* trials, much like that shown by healthy participants searching inverted faces. Such a result would suggest a common behavioral marker for inefficient search, whether the inefficiency stems from subject factors (developmental prosopagnosia) or stimulus properties (inverted faces).

METHODS

Participants

Thirteen individuals with developmental prosopagnosia (six men, age range = 33–67 years, mean = 48.5 years, *SD* = 11.5 years) participated. Some of them have participated in prior studies (Corrow, Albonico, & Barton, 2018; Moroz et al., 2016; Rubino, Corrow, Corrow, Duchaine, & Barton, 2016; Liu, Corrow, Pancaroglu, Duchaine, & Barton, 2015), whereas the remaining were recruited through the Department of Psychology at Bethel University, St. Paul, MN. Diagnostic criteria (Barton & Corrow, 2016) were self-reported life-long difficulty in face recognition, which, in most, was supported by the 20-item prosopagnosia index (Shah et al., 2015) and confirmation of impaired face recognition on at least two objective tests. The latter included a score at least 2 *SD*s below the previously reported control mean on the Cambridge Face Memory Test (Duchaine & Nakayama, 2006) as well as impairment on at least one of three other tests of face memory with published normative data, which were either a test of famous face identification (Duchaine, Germine, & Nakayama, 2007), an old/new test of familiarity for recently viewed faces (Duchaine, Nieminen-von Wendt, New, & Kulomäki, 2003), or a discordance between preserved word memory and impaired face memory on the Warrington Recognition Memory Test (Warrington, 1984). One participant (DP306) had extensive exposure to the Cambridge Face Memory Test, and so their diagnosis was based instead on impairment on both the old/new test and famous face identification.

Participants tested in Vancouver had best corrected visual acuity of <20/60, normal visual fields, normal general memory, attention, and perceptual abilities as determined by a neuropsychological battery (see Corrow et al., 2019; Rubino et al., 2016) that included the Wechsler Memory Scale-III (Wechsler, 1997), a visual search test (Spinnler & Tognoni, 1987), the Stars Cancellation Test (Wilson, Cockburn, & Halligan, 1987), the Visual Object and Space Perception Battery (Warrington & James, 1991), the Hooper Visual Organization Test (Hooper, 1983), and the Benton Judgment of Line Orientation. Participants tested in Minnesota had best corrected visual acuity of <20/60, no history of visual disorder, and no self-reported general memory impairment. To exclude autism spectrum disorders, all participants also scored less than 32 on the autism spectrum quotient (Baron-Cohen et al., 2001).

Twenty-nine healthy participants formed the control group, with a similar gender composition and range of ages (11 men, age range = 21–68 years, mean = 41.6 years, *SD* = 15.6 years). None had a history of neurological problems or psychiatric illness. All were right-handed and had normal or corrected-to-normal vision acuity. Before the experiment, control participants were also administered the Cambridge Face Memory Test (Duchaine & Nakayama, 2006) in the laboratory.

The protocol was reviewed and approved by the institutional review boards of the University of British Columbia, Vancouver Coastal Health, and Bethel University, and all participants gave written informed consent in accordance with the principles of the Declaration of Helsinki. Control participants were reimbursed $10 per hour for their participation.

Stimuli

We created two sets of stimuli, one for the identity search test and one for the expression search test, with each set consisting of 84 images. To reduce the influence of low-level image properties, we varied images not only in the primary dimension relevant to the task but also in a second dimension irrelevant to the task (Figure 1). Thus, the identity search test required participants to discriminate facial identity, but images slightly varied in facial expression. For the expression search test, participants reported on facial expression, but images slightly varied in identity. Our goal was to create search tests with targets that were relatively easy to discriminate for both control and prosopagnosic participants to allow us to evaluate differences in temporal measures of processing rather than accuracy.

Face Identity Stimuli

Frontal facial images from six male and six female entries were chosen from the Radboud Faces Database (Langner et al., 2010). For each person, we selected seven images, each with a different facial expression: neutral, happiness, sadness, fear, contempt, disgust, and anger. Using Abrosoft Fantamorph 5 (www.fantamorph.com), we created more subtle versions of the expression by blending the neutral face of a particular person with each of the emotional expressive faces of that same person (20% expression and 80% neutral mix). These six expression-morphed faces and the neutral face were the final set of images used in the test for each of the 12 identities. All images were converted to grayscale, cropped to an oval shape 195 × 225 pixels long and wide (approximately 5.3° × 6.1° of visual angle) using Adobe Photoshop CC 2014 (www.adobe.com). Moles, spots, scars, hair on the face, and other potential distractors or identity cues were removed with this software. Using MATLAB (www.mathworks.com), the luminance values of the faces were normalized so that the average luminance inside the oval shape was set to half maximum and the root-mean-squared contrast was set to 1. The 12 identities were coupled in six pairs (three female and three male pairs) that were matched in age and chosen by pilot work so that each pair was sufficiently different to give discrimination accuracy of more than 90% in healthy participants.

Face Expression Stimuli

Frontal facial images from seven male and seven female entries were chosen from the same database. These people were different from those used for the face identity stimuli. For each person, we obtained images of six different expressions: anger, sadness, disgust, happiness, contempt, and surprise. Images from one female and male person were used as base images to which images of the other people were morphed. Thus, for each expression, images of the remaining six male faces were morphed with the base male face, and images of the remaining six female faces were morphed with the base female face (mix of 80% base face and 20% of the other face). These six identity-morphed faces and the base face were the seven images used for each expression. Again, all images were converted to grayscale, cropped to an oval shape (195 × 225 pixels long and wide—approximately 5.3° × 6.1° of visual angle), and distinctive features such as spots, scars, and moles were removed. Luminance

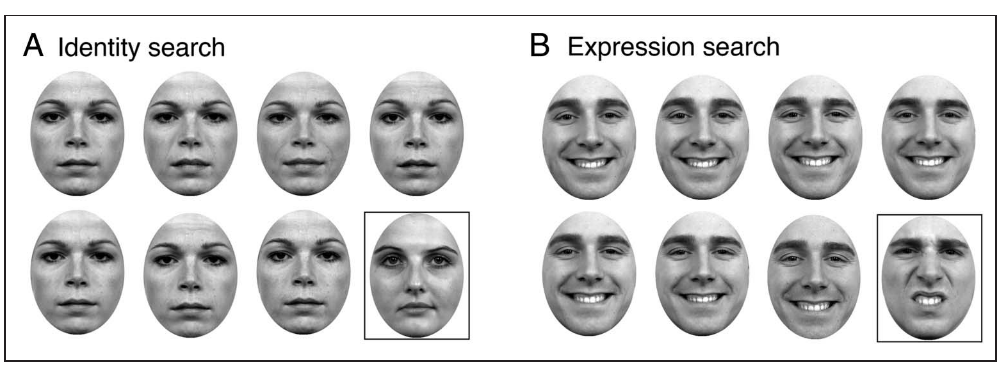

Figure 1. Sample stimuli. Example of stimuli used in the identity and expressions search tests. (A) Seven pictures depicting the same face identity with subtle variations in facial expression introduced by morphing, and in the box an image of the different person paired with this identity. (B) Seven pictures of the same facial expression with subtle variations in face identity introduced by morphing, and in the box an image of the different expression paired with this expression.

values inside the oval aperture were normalized such that average luminance was set to half maximum and root-mean-squared contrast was set to 1. Expressions were coupled in six pairs, three for each gender (angry/sad, disgust/happy, contempt/surprised), again matched by pilot work so that each pair was sufficiently different to give discrimination accuracy of more than 90% in healthy participants.

Apparatus and Procedure

The experiment was run using Experiment Builder 1.10.1630 (www.sr-research.com) and displayed on an LG monitor with a resolution of 1920 × 1080 pixels (52 × 59 cm) at a viewing distance of approximately 57 cm. The study consisted of two tests, one in which participants were required to search for facial identity and one for facial expression. Test order was counterbalanced across participants.

Participants performed a same/different task, and the procedures of each test and its trials were similar for the two tests. All stimuli were presented against a white background, and a keyboard was used to collect responses. Each trial started with a blank screen followed 250 msec later by the stimulus display. The stimulus display showed an array of faces that varied from three to seven (the set size), arranged in a cluster rather than a line (Figure 2). The array remained visible until the participant's key response. In the identity search test, the images were all from the same person on *same* trials, whereas on the *different* trials, one image was from the other member of the pair. In both *same* and *different* trials, the expression varied slightly between all images. In the expression search test, *same* trials all had the same expression whereas *different* trials had one face with the other expression of the pair, with identity varying slightly between all images in both types of trials. Participants pressed the "S" key if all the face images on the screen were of the same person in the identity search test or of the same expression in the expression search test and the "D" key if one image differed from the rest.

Each test began with a practice phase. This first showed a PowerPoint presentation that familiarized the participants with the stimuli and introduced the test. After this presentation, participants performed 12 practice trials that showed examples of at least two of each set size. During the practice but not the experimental trials, participants received auditory feedback after each response.

In each of the two tests, there were five set sizes (three, four, five, six, or seven stimuli) with 54 trials for each set size, for 270 trials total. Of these 270 trials, 90 were *same* trials and 180 were *different* trials. Participants were not informed of this proportion. The 270 trials were divided into three blocks of 90 trials to give the participants two short rest breaks. In each test, the presentation of the 84 face stimuli was balanced so that each stimulus appeared the same number of times in each test. The target on *different* trials was located as evenly as possible across all stimulus positions of the set. We recorded both accuracy and RT, defined as the time between the appearance of the stimuli and the participant's keypress.

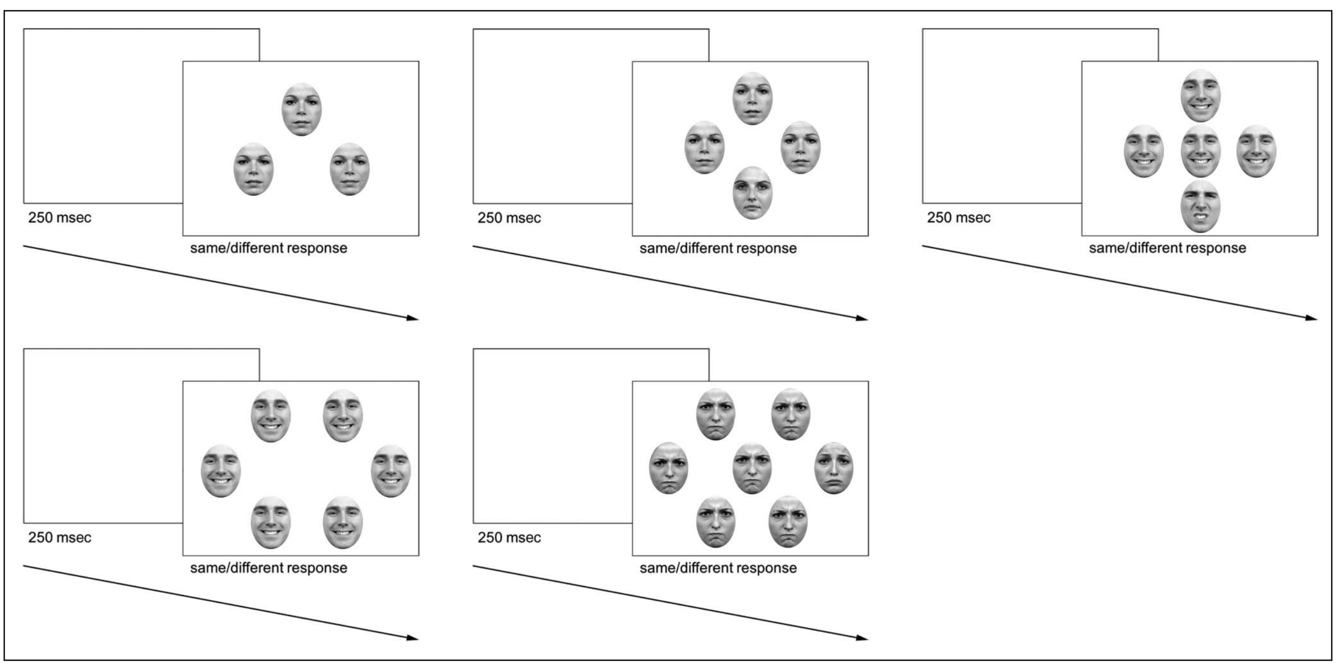

Figure 2. Sample arrays. Examples of arrays used in the identity and expression search tests. In the identity test, participants were asked to indicate whether all the faces were of the same person or one is different, whereas in the expression task, whether all the faces depicted the same expression or not. The set size of the arrays ranged from 3 to 7 stimuli. Examples of both *same* and *different* trials are shown.

Data Analysis

For each participant, we assessed performance on each test by two accuracy and two temporal variables. The accuracy variables were overall mean accuracy and d', a criterion-free index of discrimination sensitivity developed from signal detection theory (Macmillan & Creelman, 1991), including all set sizes. The temporal variables were mean RT and the set size effect, defined as the slope of the linear regression of the RT as a function of the set size. Only data from correct trials were used for these two temporal variables.

First, we conducted group-level analyses by submitting each of the four outcome variables to a mixed-design ANOVA, with Group (control vs. prosopagnosia) as between-participant factor and Test (identity vs. expression) as within-participant factor. Significant differences were explored by Bonferroni post hoc multiple comparisons (corrected p values are reported), and effect sizes were measured by computing partial eta squared.

Second, we examined the data at the level of the individual participant. We compared each prosopagnosic participant to the control group by using single-participant statistics (Crawford & Garthwaite, 2002). To address our second question, we also evaluated for dissociations between identity and expression performance with a Bayesian standardized difference test, DissocsBayes_ES.exe (Crawford, Garthwaite, & Ryan, 2011; Crawford, Garthwaite, & Porter, 2010). This single-participant approach allowed us to determine how many prosopagnosic participants were abnormal on either test and if any showed a putative classical dissociation (Gerlach, Lissau, & Hildebrandt, 2018). We then assessed the diagnostic utility of the single-participant data using signal detection theory analysis (Macmillan & Creelman, 1991).

Third, we performed correlational analyses to determine whether the results for identity search were correlated with those for the expression search test.

Fourth, we compared the results of *same* (i.e., target absent) and *different* (i.e., target present) trials. For each participant, we computed the natural logarithm of the ratio of the set size effect on *target-absent* over *target-present* trials (Hemström et al., 2019). A standard serial self-terminating search model predicts a ratio of 2 when exhaustive search of all stimuli is needed to reach a correct decision on *same* trials, whereas search on *different* trials can be terminated by discovery of the target, which will occur on average about halfway through searching (Cousineau & Shiffrin, 2004; Townsend & Wenger, 2004; Wolfe, 1998). Therefore, we assessed for the two groups of participants whether this value was significantly different from ln(2).

Finally, at the suggestion of a reviewer, we analyzed the decision processes of our participants using a linear ballistic accumulator model (Donkin, Brown, & Heathcote, 2011; Donkin, Averell, Brown, & Heathcote, 2009; Brown & Heathcote, 2008) in a preliminary attempt to identify the origins of the prolonged RTs in our prosopagnosic cohort (see Appendix A).

RESULTS

Group-level Analyses

Mean accuracy was similar for controls and prosopagnosic participants in both the identity (controls 0.93, SD = 0.05; prosopagnosics 0.92, SD = 0.05) and the expression (controls 0.93, SD = 0.03; prosopagnosics 0.92, SD = 0.03) visual search tests. There was no main effect of Group, $F(1, 40) = 0.35$, $p = .560$, $\eta_p^2 = .009$, or Test, $F(1, 40) = 0.15$, $p = .699$, $\eta_p^2 = .004$, and no interaction, $F(1, 40) = 0.01$, $p = .910$, $\eta_p^2 = .000$ (Figure 3A).

The d' data were also similar between the groups for identity (controls 3.23, SD = 0.63; prosopagnosics 3.20, SD = 0.60) and expression search (controls 3.14, SD = 0.43; prosopagnosics 2.97, SD = 0.51). There was no main effect of Group, $F(1, 40) = 0.23$, $p = .635$, $\eta_p^2 = .006$, and no interaction between Group and Test, $F(1, 40) = 0.73$, $p = .397$, $\eta_p^2 = .018$ (Figure 3B). There was a trend to a main effect of Test, $F(1, 40) = 3.46$, $p = .070$, $\eta_p^2 = .080$, with the identity search test being slightly easier.

These accuracy and d' results established the achievement of two important aims of our test design: first, that searches for identity and expression were similar in difficulty, and second, that the face pairs used were sufficiently distinct to allow prosopagnosic participants to also achieve reasonable accuracy. Accuracy rates above 90% for both tests in both groups allowed similar and sufficient numbers of correct trials for our main analysis of temporal variables.

For mean RTs, there was a significant main effect of Group, $F(1, 40) = 33.75$, $p < .001$, $\eta_p^2 = .455$, and Test, $F(1, 40) = 60.96$, $p < .001$, $\eta_p^2 = .604$. There were longer RTs for prosopagnosic participants (controls 2310 msec, SD = 665 msec; prosopagnosics 4175 msec, SD = 1594 msec) and for the identity search test (identity 4089 msec, SD = 2053 msec; expression 2396 msec, SD = 652 msec). Critically, the interaction between the two main factors was significant, $F(1, 40) = 23.83$, $p < .001$, $\eta_p^2 = .373$ (Figure 3C). Prosopagnosic participants had longer RTs on the identity test than on the expression search test (identity 5550 msec, SD = 2498 msec; expression 2799 msec, SD = 690 msec; $p < .001$), whereas controls had RTs that did not significantly differ between the two conditions (identity 2627 msec, SD = 872 msec; expression 1993 msec, SD = 458 msec; $p = .062$). Prosopagnosic participants were also slower than controls on both tests (both ps < .001).

For the set size effect, the main effects of Group, $F(1, 40) = 35.24$, $p < .001$, $\eta_p^2 = .468$, and Test, $F(1, 40) = 82.92$, $p < .001$, $\eta_p^2 = .675$, were significant. As found for the mean RTs, set size effects were larger for prosopagnosic participants (controls 183 msec/face, SD = 72 msec/face;

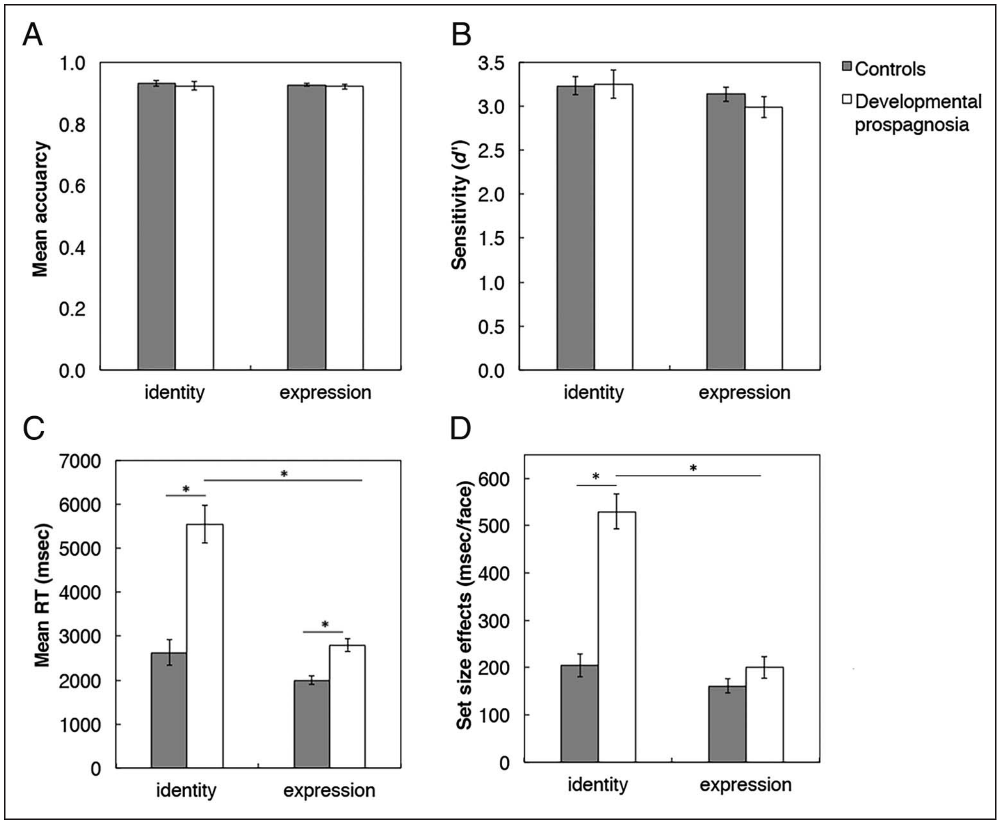

Figure 3. Group-level analyses. (A) Mean accuracy, (B) *d′*, (C) mean RT, and (D) set size effects for the two groups (control, developmental prosopagnosia) in the face identity and expression search tests. Error bars indicate ± 1 *SE*, and horizontal lines indicate significant pairwise comparisons (*$p < .05$).

prosopagnosics 365 msec/face, *SD* = 161 msec/face) and for the identity search test (identity 367 msec/face, *SD* = 22 msec/face; expression 181 msec/face, *SD* = 14 msec/face). The interaction between Group and Test was significant, $F(1, 40) = 48.67$, $p < .001$, $\eta_p^2 = .549$ (Figure 3D): Prosopagnosic participants had larger set size effects for the identity search test than for the expression search test (identity 530 msec/face, *SD* = 213 msec/face; expression 200 msec/face, *SD* = 110 msec/face; $p < .001$), whereas controls had set size effects that did not significantly differ between the two tests (identity 205 msec/face, *SD* = 75 msec/face; expression 161 msec/face, *SD* = 68 msec/face; $p = .063$). Whereas set size effects for prosopagnosic participants were larger than for controls on identity search ($p < .001$), the set size effects of the two groups did not differ on expression search ($p = .164$).

Set size effects were estimated by computing the slope of the linear regression of the RT as a function of the set size. This procedure assumes the existence of a linear relationship between set size and RTs. To confirm the goodness of fit of our set size estimation, we computed for each participant the coefficient of determination (R^2) of the linear regression between RTs and set size (individual data in Supplementary Table 1). In controls, average R^2 was .75 (*SD* = .2) in the identity and .82 (*SD* = .18) in the expression task. For prosopagnosic participants, it was .80 (*SD* = .16) in the identity and .64 (*SD* = .28) in the expression task. Slopes of the linear regression were significant in 81% of controls and 85% of prosopagnosic participants. These results suggest that set size effects are reasonably derived by linear regression.

Correlation between Expression and Identity Search

Control participants showed a modest correlation between identity and expression search tests in accuracy ($r = .374$, $p = .045$) and *d′* ($r = .373$, $p = .047$). They showed a stronger correlation in RT ($r = .620$, $p < .001$), but not in set size effect ($r = .244$, $p = .201$). In the prosopagnosic group, RTs were significantly correlated between the two tests ($r = .638$, $p = .019$), whereas set size effects showed a trend ($r = .544$, $p = .055$), with no correlation in accuracy or *d′*.

Single-participant Level Analyses

Mean accuracy was reduced for two prosopagnosic participants (DP032 and DP301) on the identity search test, but for none on the expression search test. Two prosopagnosic participants (DP035 and DP301) also showed a dissociation between the two tests; however, although DP301 was more accurate on expression search, DP035 did better on identity search. No prosopagnosic participant had a low *d′* for identity and only one (DP035) on the expression search test.

In contrast, 9 of 13 prosopagnosic participants had abnormally long RTs on the identity search test, 7 of whom showed a dissociation from their better expression score

Table 1. Single Participant Comparisons for the Four Outcome Variables in the Identity and Expression Tests

	Identity			*Expression*			*Identity–Expression*			
	Raw Score	*z Score*	*p (One-Tailed)*	*Raw Score*	*z Score*	*p (One-Tailed)*	*z Difference*	*p (One-Tailed)*	*Z-DCC*	*95% CI Z-DCC*
Accuracy										
	Controls Average 0.93, SD = 0.05			*Controls Average 0.93, SD = 0.03*						
DP003	0.92	−0.172	.434	0.95	0.623	.272	−0.795	.242	−0.712	−1.133, −0.305
DP024	0.98	1.019	.162	0.96	1.052	.155	−0.033	.488	−0.030	−0.513, 0.453
DP032	**0.83**	**−2.012**	**.029**	0.89	−1.176	.129	−0.836	.237	−0.748	−1.404, −0.127
DP033	0.96	0.641	.267	0.89	−1.006	.166	1.647	.076	1.475	0.959, 2.025
DP035	0.96	0.641	.267	0.88	−1.605	.063	**2.246**	**.027**	**2.010**	**1.383**, **2.695**
DP039	0.91	−0.333	.373	0.89	−1.262	.112	0.929	.208	0.831	0.355, 1.333
DP044	0.96	0.641	.267	0.91	−0.577	.287	1.218	.143	1.091	0.640, 1.564
DP301	**0.82**	**−2.175**	**.021**	0.94	0.454	.329	**−2.629**	**.013**	**−2.353**	**−3.151**, **−1.629**
DP302	0.95	0.316	.379	0.96	0.883	.196	−0.567	.309	−0.507	−0.942, −0.086
DP303	0.95	0.316	.379	0.94	0.367	.360	−0.051	.482	−0.046	−0.425, 0.332
DP304	0.89	−0.766	.229	0.93	0.025	.490	−0.791	.243	−0.708	−1.139, −0.293
DP305	0.95	0.425	.340	0.93	0.194	.425	0.231	.419	0.206	−0.172, 0.587
DP306	0.91	−0.495	.315	0.91	−0.404	.347	−0.091	.468	−0.081	−0.471, 0.307
Sensitivity (d′)										
	Controls Average 3.23, SD = 0.63			*Controls Average 3.14, SD = 0.43*						
DP003	2.92	−0.497	.314	3.37	0.531	.303	−1.028	.183	−0.918	−1.362, −0.492
DP024	3.81	0.918	.187	3.60	1.064	.152	−0.146	.449	−0.131	−0.607, 0.342
DP032	2.29	−1.507	.075	2.56	−1.337	.100	−0.170	.442	−0.152	−0.723, 0.413
DP033	3.38	0.236	.409	2.56	−1.334	.100	1.570	.087	1.401	0.878, 1.963
DP035	3.49	0.410	.345	**2.32**	**−1.893**	**.037**	**2.303**	**.025**	**2.056**	**1.400**, **2.774**
DP039	2.81	−0.672	.257	2.47	−1.548	.070	0.876	.224	0.782	0.259, 1.335
DP044	3.51	0.446	.332	2.72	−0.969	.174	1.415	.109	1.263	0.780, 1.775
DP301	3.28	0.067	.474	3.29	0.343	.369	−0.276	.404	−0.247	−0.623, 0.125

Table 1. *(continued)*

Sensitivity (d′)										
	Controls Average 3.23, SD = 0.63			*Controls Average 3.14, SD = 0.43*						
DP302	3.58	0.551	.296	3.72	1.341	.099	−0.790	.246	−0.706	−1.217, −0.217
DP303	3.49	0.412	.344	3.24	0.246	.405	0.166	.442	0.148	−0.230, 0.528
DP304	2.84	−0.626	.272	3.12	−0.043	.483	−0.583	.304	−0.520	−0.927, −0.125
DP305	3.73	0.795	.220	3.07	−0.144	.444	0.939	.205	0.839	0.411, 1.283
DP306	3.10	−0.208	.420	2.81	−0.758	.231	0.550	.314	0.492	0.084, 0.910
RT										
	Controls Average 2627, SD = 872			*Controls Average 1993, SD = 458*						
DP003	**5145**	**2.888**	**.004**	**3188**	**2.609**	**.008**	0.279	.387	0.319	−0.658, 1.312
DP024	3445	0.938	.182	2043	0.109	.458	0.829	.176	0.950	0.490, 1.436
DP032	**6554**	**4.503**	**.000**	2500	1.107	.143	**3.396**	**.001**	**3.894**	**2.550**, **5.394**
DP033	**5758**	**3.591**	**.001**	2394	0.876	.198	**2.715**	**.003**	**3.113**	**2.018**, **4.329**
DP035	**11750**	**10.462**	**.000**	**3478**	**3.242**	**.002**	**7.220**	**.000**	**8.278**	**5.303**, **2.15E+09**
DP039	3354	0.834	.210	**3198**	**2.631**	**.008**	**−1.797**	**.029**	**−2.061**	**−2.964**, **−1.234**
DP044	**4459**	**2.101**	**.024**	2258	0.579	.287	**1.522**	**.050**	**1.746**	**1.044**, **2.509**
DP301	**4599**	**2.261**	**.017**	2277	0.62	.274	**1.641**	**.039**	**1.882**	**1.141**, **2.691**
DP302	**6154**	**4.045**	**.000**	**3952**	**4.277**	**.000**	−0.232	.414	−0.267	−1.700, 1.158
DP303	3741	1.278	.110	2380	0.845	.207	0.433	.315	0.496	−0.019, 1.027
DP304	**5826**	**3.669**	**.001**	2671	1.48	.078	**2.189**	**.014**	**2.509**	**1.426**, **3.696**
DP305	**8858**	**7.146**	**.000**	**4011**	**4.406**	**.000**	**2.740**	**.015**	**3.141**	**1.101**, **5.332**
DP306	2514	−0.130	.450	2041	0.105	.459	−0.235	.395	−0.269	−0.641, 0.099

Set size effects										
	Controls Average 205, SD = 75			*Controls Average 161, SD = 68*						
DP003	**778**	**7.640**	**.000**	262	1.485	.078	**6.155**	**.000**	**5.007**	**3.238**, **2.15E+09**
DP024	330	1.667	.056	147	−0.206	.421	1.873	.071	1.523	0.966, 2.125
DP032	**562**	**4.760**	**.000**	45	−1.706	.052	**6.466**	**.000**	**5.260**	**3.854**, **2.15E+09**
DP033	**525**	**4.267**	**.000**	236	1.103	.144	**3.164**	**.011**	**2.574**	**1.546**, **3.730**
DP035	**901**	**9.280**	**.000**	**300**	**2.044**	**.027**	**7.236**	**.000**	**5.886**	**3.763**, **2.15E+09**
DP039	**366**	**2.147**	**.022**	85	−1.118	.141	**3.265**	**.007**	**2.655**	**1.894**, **3.491**
DP044	**452**	**3.293**	**.002**	216	0.809	.217	**2.484**	**.031**	**2.021**	**1.193**, **2.944**
DP301	**392**	**2.493**	**.010**	174	0.191	.426	**2.302**	**.038**	**1.873**	**1.184**, **2.632**
DP302	**798**	**7.907**	**.000**	**478**	**4.662**	**.000**	**3.245**	**.030**	**2.640**	**0.742**, **4.713**
DP303	**489**	**3.787**	**.000**	204	0.632	.270	**3.155**	**.010**	**2.566**	**1.628**, **3.621**
DP304	**348**	**1.907**	**.036**	121	−0.588	.284	**2.495**	**.027**	**2.030**	**1.387**, **2.729**
DP305	**738**	**7.107**	**.000**	132	−0.426	.339	**7.533**	**.000**	**6.128**	**4.343**, **2.15E+09**
DP306	212	0.093	.464	205	0.647	.265	−0.554	.328	−0.450	−0.852, −0.058

Bold indicates significant result. Z-DCC = the effect size for the score difference between the two tasks.

(Table 1). In the reverse direction, five had long RTs on the expression search test, one of whom showed a dissociation. Set size effects were more consistent: 11 of the 13 prosopagnosic participants had abnormally large set size effects on identity search, all of whom showed a dissociation from better expression set size effects, which were increased in only two (Figure 4).

Signal Detection Theory Analysis

From the single-participant data, we assessed the diagnostic utility of the set size effect for identity search and the difference in set size effects for identity and expression search (Figure 5). In particular, we used individual set size effect data from controls and prosopagnosic participants to compute the receiver operating characteristic (ROC) curve and the area under the curve (AUC). The ROC curve illustrates the diagnostic ability of the set size effects in discriminating between controls and prosopagnosic participants by plotting true positive rate (sensitivity) against false positive rate (1 − specificity). It is then possible to estimate sensitivity and specificity for any chosen cutoff value. The AUC, instead, represent the probability of correctly classifying a participant as control

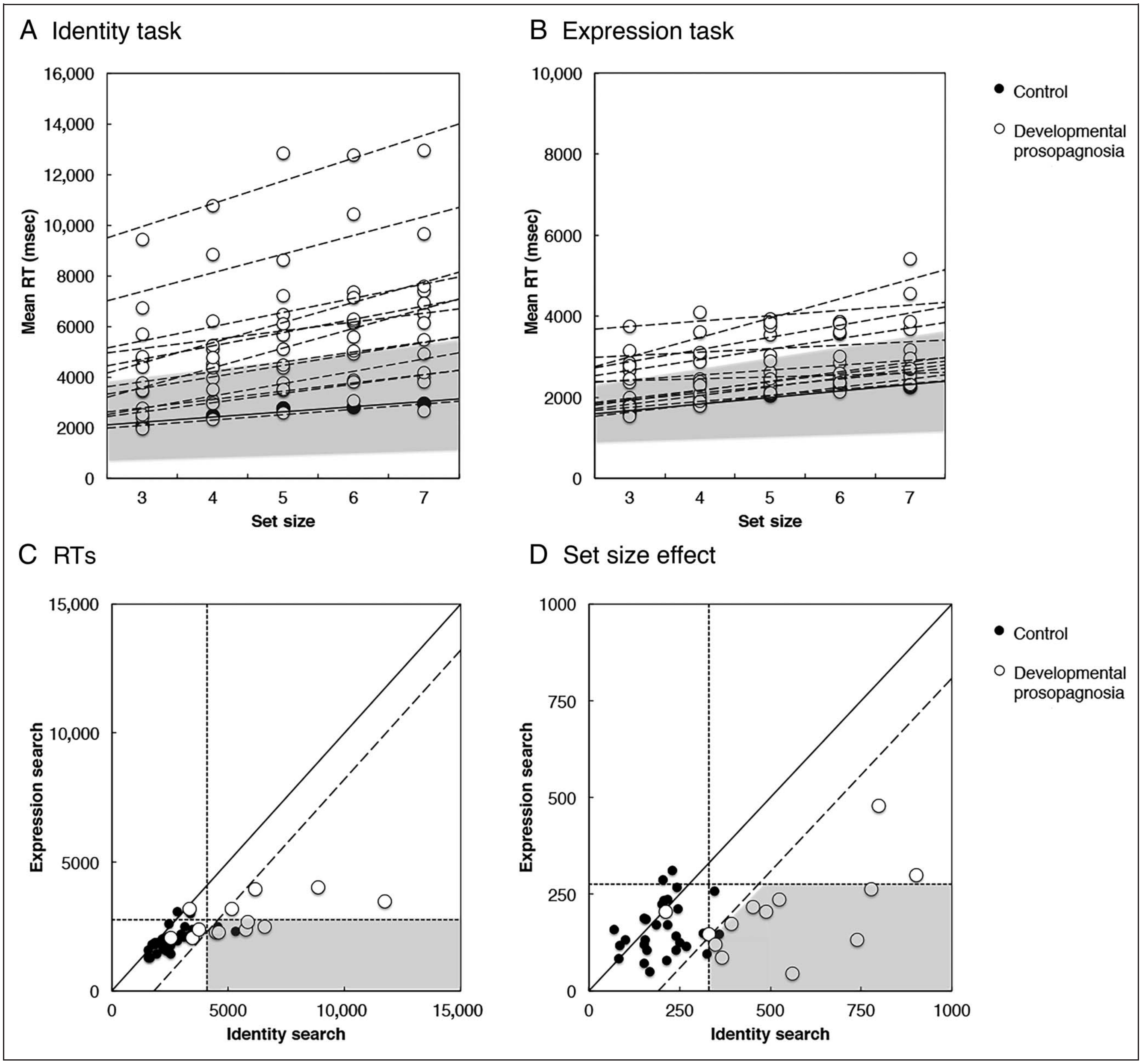

Figure 4. Single-participant data for mean RT and set size effects. Mean RT as a function of the set size for the (A) identity and (B) expression task; linear regressions are fitted on the control average (black line) and individually for the prosopagnosic participants (dashed lines). The gray zone indicates the 95% prediction limit for the controls' linear regression. (C) RTs and (D) set size effects for the identity search tests are plotted against that of the expression search. In C and D, vertical and horizontal dotted lines show 95% prediction limits for the individual test, whereas the dashed oblique line indicates the 95% prediction limit for the difference between identity and expression search scores; participants whose scores fall in the gray zones have a putative classical dissociation between face identity and expression search scores.

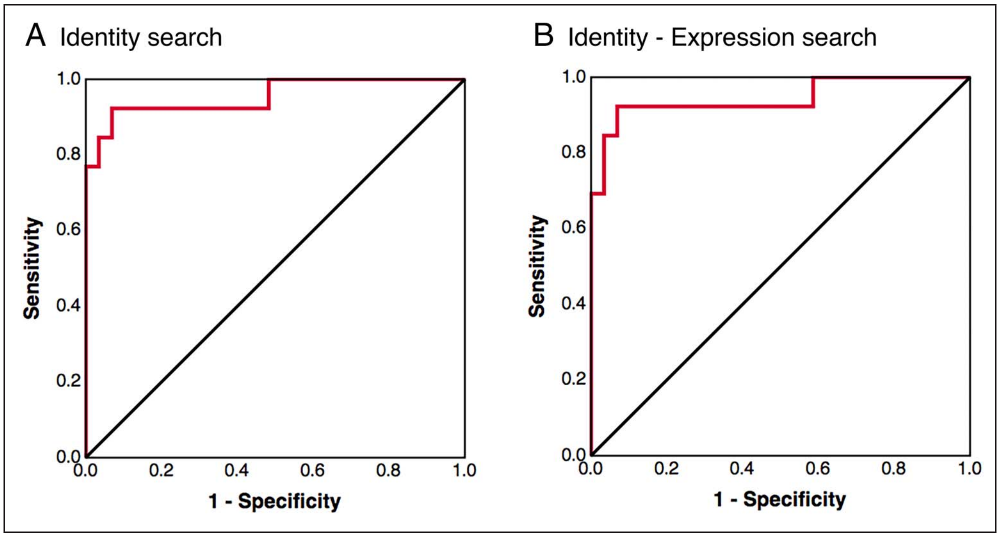

Figure 5. Signal detection analyses of set size effects. (A) ROC curves for the identity search and (B) the difference between identity and expression search.

or prosopagnosic according to their set size effect: An AUC value of .5 is representative of a random classifier, whereas a value of 1 represents a perfect model whose predictions are 100% correct.

For identity search, the AUC was .96, with a lower 95% confidence limit of .88. For the difference between identity and expression search, the AUC was .94, with a lower 95% confidence limit of .86. Thus, in our small sample, the ability of either measure to discriminate developmental prosopagnosia from healthy controls was very good.

Set Size Effects on *Target-absent* (*Same*) Versus *Target-present* (*Different*) Trials

Healthy controls showed a *target-absent*/*target-present* ratio of 1.79 for identity search and 1.87 for expression search, neither of which differ from the predicted ratio of 2 (Table 2). Although prosopagnosic participants had a similar ratio of 2.15 for expression search, they showed an abnormally low ratio of 0.78 for identity search. Thus, for identity search, prosopagnosic participants take just as much or even slightly more time to reach a decision on target present trials as they do on target absent trials.

However, single-participant data showed that only two prosopagnosic participants had abnormal *target-absent*/*target-present* ratios for the identity search (Figure 6A). Signal detection theory showed that the AUC was .74, with a lower 95% confidence limit of .58 (Figure 6B). Thus the diagnostic utility of the target-absent/target-present ratio is only modest.

Table 2. Results for "Same" and "Different" Trials

	Face Identity		*Face Expression*	
	Mean	*SD*	*Mean*	*SD*
Controls				
Different trials	162	86	125	75
Same trials	289	168	234	101
Ratio (same/different)	1.79		1.87	
p(1)	.190		.437	
Developmental Prosopagnosia				
Different trials	571	319	146	127
Same trials	447	259	314	199
Ratio (same/different)	0.78		2.15	
p(1)	.0004		.584	

p(1) = probability that ln(ratio) differs from ln(2) according to one-sample *z* test.

DISCUSSION

On tests constructed to give good and similar accuracy rates for visual search for face identity and expression in both groups, we found significant differences in the temporal outcome measures of visual search. Even though the analysis of mean RT and set size effects produced similar results, of the two temporal variables, set size effects may more accurately reflect the perceptual processing load specific to faces. Indeed, mean RT inevitably includes factors such as general processing speed and motor latency that are not specific to face processing. We found that set size effects were increased in developmental prosopagnosia, but only in search for face identity, not in search for face expression. At an individual level, this was reflected in a dissociation between impaired identity search and better expression search in 11 of 13 prosopagnosic participants. Finally, a more fine-grained analysis of search on *same* versus *different* trials showed that expression and identity searches by controls were consistent with the predictions of self-terminated serial search when the target was present. This was also

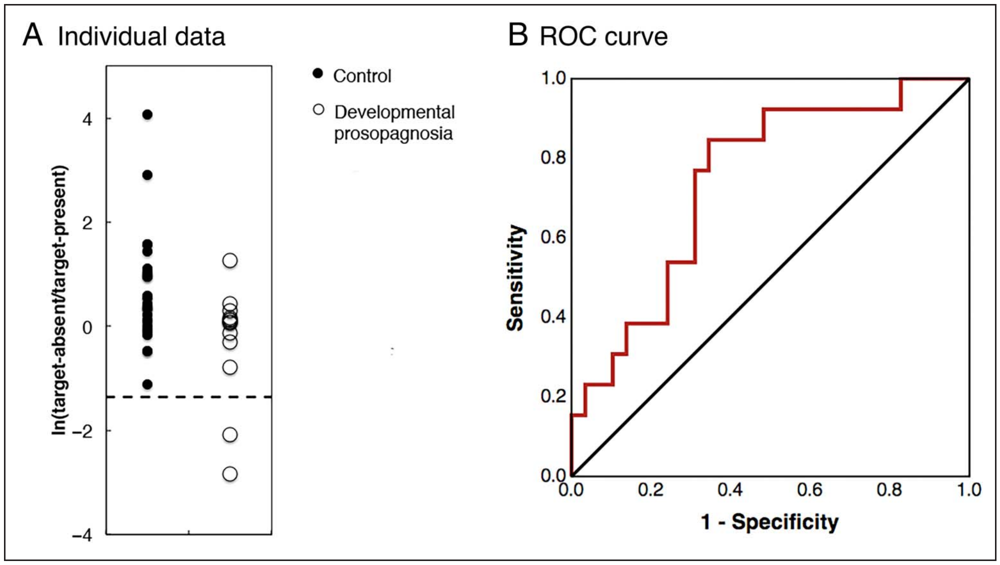

Figure 6. Target-absent/target-present ratios for set size effects on identity search. (A) Individual data, with dashed line indicating lower 95% prediction limit. (B) ROC curve.

true for expression search by prosopagnosic participants, but not for their identity search.

Our tests of visual search emphasized perceptual discrimination rather than memory for faces. However, the hallmark of prosopagnosia is impaired face familiarity, and face perception need not be impaired to make the diagnosis. Indeed, in acquired prosopagnosia, there are both apperceptive and amnestic variants (Davies-Thompson et al., 2014; Damasio et al., 1990). Whether a similar heterogeneity exists in developmental prosopagnosia is debated. The Cambridge Face Perception Test is often used to measure face discrimination in these cohorts, though its reliability is not as good as that of the Cambridge Face Memory Test (Bowles et al., 2009). Nevertheless, this has shown perceptual impairments in half of 16 adults with developmental prosopagnosia (Dalrymple et al., 2014). Another study of 11 participants found that six performed normally on all face perception tests, although only one performed poorly on the Cambridge Face Perception Test (Ulrich et al., 2017).

At a group level, one study of working memory for faces in 10 participants looked at the effect of the number of faces being memorized on the accuracy of recall 1 sec later. Accuracy was reduced in developmental prosopagnosia, but the effect of number of faces was similar to controls (Jackson, Counter, & Tree, 2017). Although they inferred that problems with perceptual encoding could not fully account for the working memory deficit, others argued that the results actually support an encoding rather than a working memory problem (Biotti et al., 2019). This latter study reported a group-wide decrease in performance on the Cambridge Face Perception Test across a cohort of 72 participants and concluded that perceptual deficits played a significant role in developmental prosopagnosia.

Our results for visual search would be consistent with this conclusion. Temporal measures such as set size effects may be particularly useful in demonstrating perceptual deficits in this population. Although the accuracy of developmental prosopagnosic participants on face discrimination tests like the Benton Facial Recognition Test (Benton, 1983) can be equivalent to that of control participants (Albonico, Malaspina, & Daini, 2017; Duchaine & Nakayama, 2006; Duchaine & Nakayama, 2004), impairments can be shown when RTs are considered (Rossion & Michel, 2018) or time constraints added (Özbek & Bindemann, 2011). Although the number of participants is small, our preliminary signal detection theory analysis suggests that the diagnostic accuracy of the set size effect may be quite high.

Our findings also point to a dissociation between identity and expression perception in developmental prosopagnosia, with the latter generally spared. Our comparison of identity and expression processing has several advantages. Because performance on the identity search test is not part of the diagnostic criteria for our prosopagnosic group, it avoids the problem of "double dipping" (Kriegeskorte, Simmons, Bellgowan, & Baker, 2009). Second, the methods for assessing identity and expression processing are similar, and the results for accuracy and d' indicate that the two tests are comparable in difficulty. Third, the single-participant analysis showed that most of our participants met criteria for a putative classical dissociation (Gerlach et al., 2018; Crawford, Garthwaite, & Gray, 2003; Shallice, 1988), with a difference between identity and expression performance exceeding that seen in controls. These three points reduce the likelihood that the finding of a dissociation is due to a methodological artifact.

There is a report of a double dissociation between expression and identity processing after acquired cerebral lesions (Fox et al., 2011). This study linked expression deficits to damage to the STS and identity deficits to lesions of the fusiform gyrus, consistent with predictions from functional imaging of the face processing network in healthy participants (Fox, Moon, Iaria, & Barton, 2009; Haxby et al., 2000). For developmental prosopagnosia, formal establishment of an identity/expression dissociation has not been made, though there are some claims of spared processing of expression perception. This includes two case reports with extensive testing of expression processing (Duchaine, Parker, & Nakayama, 2003; Jones & Tranel, 2001) and a small series of three participants that reported normal expression discrimination, even when the stimuli were made more ambiguous by morphing (Humphreys et al., 2007). A larger study found normal identification of more obvious expressions in 11 participants (Ulrich et al., 2017). However, subtler expression-related deficits have occasionally been reported. In one study of 17 participants, only those with poorer performance on the Cambridge Face Perception Test had some difficulty, but even these had normal discrimination thresholds for most expressions, with the exception of differentiating fear from surprise as well as some problem judging emotions from the eyes alone (Biotti & Cook, 2016). Another study of 10 participants reported anomalies in aftereffects for happiness (Burns, Martin, Chan, & Xu, 2017). Taken together, our results and this body of work suggest that the processing of face expression is much less affected than the processing of identity in developmental prosopagnosia.

On the other hand, we also found some correlations between identity and expression search results. The implications of these correlations are debatable. Correlations in accuracy and d' between identity and expression search in control participants could reflect some shared processing, but shared processes could be either general perceptual abilities or face-specific mechanisms. These correlations were weak and need to be treated with caution. The lack of correlation of these variables in prosopagnosic participants would be expected if their perceptual deficit is specific for identity, as indicated by the analysis for dissociations, but the small number of participants limits the strength of the inference. Small samples are associated with increased variability and less reliable correlation coefficients (Schönbrodt & Perugini, 2013). We found larger correlations in both groups for mean RTs, but as stated above, this temporal variable could be dominated by general factors like processing speed and motor RTs, which could explain such correlations. Set size effect is more likely to isolate a temporal effect more specific to face processing, and for this variable, there is no correlation between identity and expression search for the control group and only a weak trend for the prosopagnosic group. However, this interesting pattern of results needs replication in larger sample.

Our last question was whether prosopagnosic participants showed any difference in their performance between *target-absent (same)* and *target-present (different)* trials. The ratio of set size effects between the two is informative of the termination rules adopted. As stated in the methods, a standard serial self-terminating search model predicts a ratio of 2 (Cousineau & Shiffrin, 2004; Townsend & Wenger, 2004; Wolfe, 1998). In our previous report (Hemström et al., 2019), controls showed a ratio of 1.95 for face identity search, which was not significantly different from the prediction of the serial self-terminating search model. The data of the control participants in the current study replicated that finding and also showed a similar result for expression search. Prosopagnosic participants demonstrated a similar behavior for expression search but deviated significantly in identity search, where the ratio was near a value of 1. This resembles the result we found for control participants performing identity search with inverted faces (Hemström et al., 2019).

A set size ratio of 1 violates the prediction for self-terminating serial search, but why this should be is not clear. One suggested possibility is more efficient search on *same* trials, perhaps through premature termination of search when search is enabled by parallel preattentive mechanisms (Cousineau & Shiffrin, 2004). However, more efficient search on *same* trials for face identity seems implausible in participants who have difficulty in processing face identity, and indeed, the set size effect for prosopagnosic participants on those trials is increased rather than decreased compared with either control participants or prosopagnosic expression search (Table 2). Rather, it is more likely that their search on *different* trials for face identity is less efficient. If search is serial—and we note that there are debates about whether set size effects are best explained by parallel or serial models of search (Moran, Zehetleitner, Liesefeld, Müller, & Usher, 2016; Townsend & Wenger, 2004; Wolfe, 1998)—a set size ratio of 1 suggests that prosopagnosic search for facial identity may not be terminated by discovery of the target. In the face of perceptual uncertainty, participants may either continue to complete the same exhaustive search they perform on *same* trials or, having found the target, continue to use time checking and validating their decision until the difference between *different* and *same* trials is erased.

Finally, it is worth mentioning that set size effects in visual search have a long history of being used with much simpler stimuli to study visual attention (Treisman & Gelade, 1980; Wolfe, 1998, 2003, 2019). However, although it is perhaps stating the obvious, visual perception has as much of a role as attention in determining performance on visual search. For example, search is much less efficient in patients with hemianopia (Hardiess, Papageorgiou, Schiefer, & Mallot, 2010; Zihl, 1995), and studies with virtual hemianopia confirm that this is attributable to the field defect (Simpson, Abegg, & Barton, 2011).

Although face perception deficits are frequently found in developmental prosopagnosia (Susilo & Duchaine, 2013), one can still ask whether problems with attention play a role, particularly in our study of visual search for faces. All the Vancouver prosopagnosic participants had been given a neuropsychology battery that had assessments of attention (see Methods), including a visual search test, and did well. Hence, our results cannot be attributed to a general attentional failure. This does not exclude the possibility of face-specific attentional deficits in developmental prosopagnosia, though. For example, recent evidence in network analyses of acquired prosopagnosia have shown a link to both a network involving the right fusiform face area as well as one involving left frontal regions (Cohen et al., 2019). This led the authors to speculate that prosopagnosia may be characterized by both impaired face perception from damage to the right fusiform face network, as well as inability to attend to facial regions useful for identification, such as the eyes, which may be mediated by the left frontal network. Indeed, other studies of acquired prosopagnosia have found a failure to process eye information, particularly under attentionally demanding conditions (Pancaroglu et al., 2016; Caldara et al., 2005), whereas some but not all studies of developmental prosopagnosia have shown reduced scanning of the eyes (Lee, Corrow, Pancaroglu, & Barton, 2019; Bobak, Parris, Gregory, Bennetts, & Bate, 2017).

In conclusion, our results indicate that set size effects in visual search for face identity are prolonged in participants with developmental prosopagnosia. As an index of the effects of the processing load being demanded of face perception, it demonstrates inefficiencies in their perception of face identity, which is not true of their perception of face expression. A larger study would help clarify the discriminative power of this measure and whether it proves to be a useful diagnostic complement to measures of accuracy on tests of face familiarity.

APPENDIX A

As an additional analysis, we investigated the origins of the slower responses by prosopagnosic participants with the linear ballistic accumulator (LBA) model (Donkin et al., 2009; 2011; Brown & Heathcote, 2008). LBA is a cognitive model of decision processes that, like other similar models (e.g., Van Zandt, Colonius, & Proctor, 2000; Ratcliff & Rouder, 1998; Ratcliff, 1978), takes into account the interaction between speed and accuracy in the decision being made (Donkin et al., 2009). In particular, these models assume that when making a decision, an observer has to sample evidence from the environment and that, as soon as the evidence reaches a threshold, the response is made. These models have been used to investigate the mechanisms underlying simple decision-making (e.g., Carpenter, 2004; Reddi, 2001; Hanes & Carpenter, 1999), recognition memory (Ratcliff, 1978), and visual discrimination (Smith & Ratcliff, 2009; Ratcliff & Tuerlinckx, 2002).

LBA provides estimates of five parameters: (i) the drift rate, v, at which evidence accumulates (with the underlying assumption that this is linear); (ii) the between-trial variability in drift rate (s); (iii) how much evidence is required to lead to a response (response threshold, b); (iv) the maximum value of the starting evidence for each drift rate (A); and (v) the amount of time taken for nondecisional aspects (nondecision time, t_0). Drift rates are assumed to indicate the quality of the stimulus and to be influenced by the task demand (Donkin et al., 2009). However, sensory, memory, and attentional processes can modulate drift rates and thresholds (e.g., Prinzmetal, Whiteford, Austerweil, & Landau, 2015; Schmiedek, Oberauer, Wilhelm, Süß, & Wittmann, 2007).

Methods

Data from each participant were fitted using the LBA (using R, www.R-project.org/). We estimated drift rates (v), drift rates' variability (s), response thresholds (b), and starting evidence point (A) for the two tasks, whereas nondecision time (t_0) was kept fixed. Mixed ANOVAs with Task (identity vs. expression) as within-participant factor and Group (controls vs. developmental prosopagnosia) as between-participant factor were conducted on the four free parameters.

Results

No significant effect was found in drift rate (v) or its between-trial variability (s). The threshold (b) data showed a main effect of Group, $F(1, 40) = 26.0$, $p < .001$, $\eta_p^2 = .393$, and Task, $F(1, 40) = 45.7$, $p < .001$, $\eta_p^2 = .533$.

Prosopagnosic participants had higher thresholds (b) than controls (controls 2674, $SD = 1621$; prosopagnosics 5426, $SD = 1619$), and thresholds were higher for the identity than the expression task (identity 5251, $SD = 2735$; expression 2848, $SD = 1141$ msec). Of greater interest, there was an interaction between Group and Task, $F(1, 39) = 16.5$, $p < .001$, $\eta_p^2 = .292$, due to the difference between the two groups being larger for the identity than the expression task.

For starting point evidence (A), both the main effect of Group, $F(1, 40) = 24.7$, $p < .001$, $\eta_p^2 = .382$, and Task, $F(1, 40) = 47.2$, $p < .001$, $\eta_p^2 = .541$, were significant. Starting point evidence was higher for prosopagnosic participants than controls (controls 2211, $SD = 1562$; prosopagnosics 4800, $SD = 1561$), and for the identity than the expression task (identity 4726, $SD = 2664$; expression 2284, $SD = 1115$ msec). The interaction between Task and Group was significant, $F(1, 40) = 18.6$, $p < .001$, $\eta_p^2 = .317$, again due to the difference between the two groups being larger for the identity than the expression task (Figure A1).

Conclusion

Drift rates (v) are thought to reflect the quality of the stimulus and to be influenced by the task demand (Donkin et al., 2009). For instance, frequently used words are associated with higher drift rates compared with low-frequency

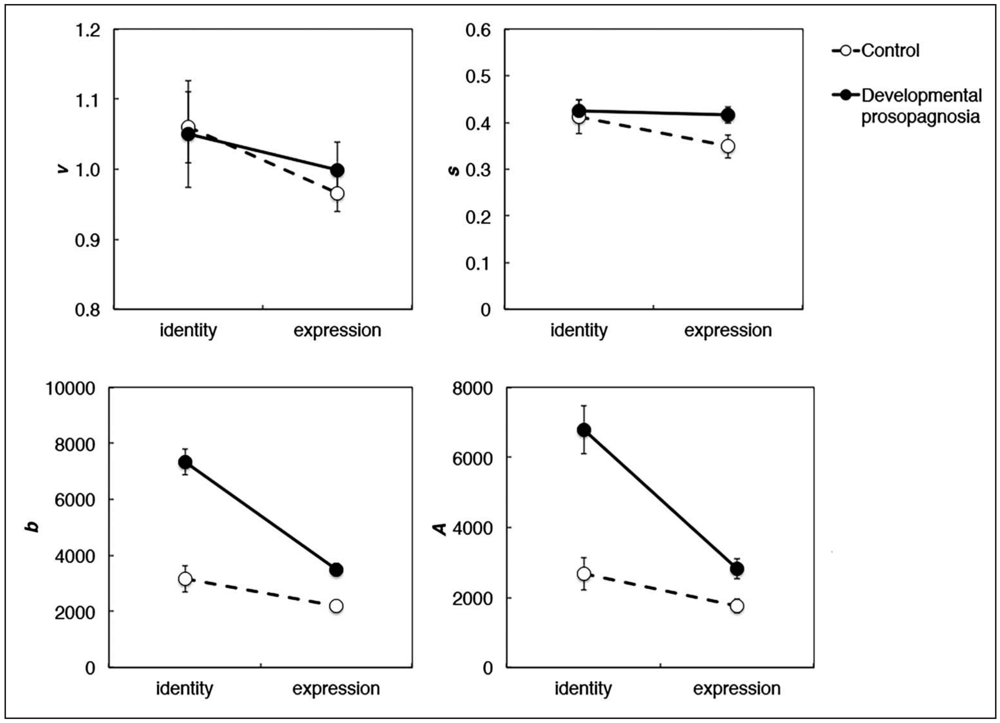

Figure A1. Average of individual parameter estimates for the LBA model of decision-making. Each graph shows the estimates for the identity and the expression search tasks separately for both control and prosopagnosic participants. Top left shows mean drift rate (*v*), top right shows the between-trial variability (*s*), bottom left shows the response threshold (*b*), bottom right shows the starting point evidence (*A*). For both *b* and *A*, prosopagnosics differ from control participants, more on the identity than on the expression task. Error bars indicate 1 *SE*.

words. The lack of differences between the identity and expression tasks in the *v* and *s* data suggests that the identity and expression tasks were of similar difficulty and had similar task demands. Prosopagnosic participants had increased evidence thresholds (*b*) compared with controls, particularly for the identity task. The relative values of *b* can be considered an index of response bias: setting a smaller value of *b* might reveal a preference for a particular response (Donkin et al., 2009, 2011). However, higher thresholds in prosopagnosic participants could also reflect their need to acquire more information to make facial judgments than controls. This could account for the fact that prosopagnosic participants spend more time and make more fixations when judging the identity of a face (Malaspina, Albonico, Lao, Caldara, & Daini, 2018; Malaspina, Albonico, Toneatto, & Daini, 2017; Schmalzl, Palermo, Green, Brunsdon, & Coltheart, 2008; Schwarzer et al., 2007). Their greater starting point evidence (*A*) may support this second possibility. A higher starting point evidence could suggest that the amount of evidence needed to recognize a face is by default higher for prosopagnosic participants compared with controls, and it is independent from any possible response bias.

Although the LBA can model choice RT data to parse perceptual and decisional components, we stress that our study was not designed for this purpose, and that the ability to model our data may have limitations. For instance, diffusion models may be more appropriate for paradigms requiring simple rapid decisions with average RTs around 1 sec (Donkin et al., 2011), whereas our tasks with their complex stimuli generated much longer RTs. Hence, the present exercise should be interpreted with caution.

Acknowledgments

J. B. was supported by Canada Research Chair 950-228984 and the Marianne Koerner Chair in Brain Diseases. S. C. was supported by the National Eye Institute under award number F32 EY023479-02. We thank B. Duchaine for assistance in participant recruitment.

Reprint requests should be sent to Jason Barton, Departments of Medicine (Neurology), Ophthalmology and Visual Sciences, University of British Columbia, Vancouver, Canada, or via e-mail: jasonbarton@shaw.ca.

REFERENCES

Albonico, A., Furubacke, A., Barton, J. J. S., & Oruc, I. (2018). Perceptual efficiency and the inversion effect for faces, words and houses. *Vision Research*, *153*, 91–97.

Albonico, A., Malaspina, M., & Daini, R. (2017). Italian normative data and validation of two neuropsychological tests of face recognition: Benton facial recognition test and Cambridge face memory test. *Neurological Sciences*, *38*, 1637–1643.

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

Benton, A. L. (1983). *Test of Facial Recognition, form SL: Manual*. Iowa City: University of Iowa Press.

Biotti, F., & Cook, R. (2016). Impaired perception of facial emotion in developmental prosopagnosia. *Cortex*, *81*, 126–136.

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

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*, 201–217.
Bowles, D. C., McKone, E., Dawel, A., Duchaine, B., Palermo, R., Schmalzl, L., et al. (2009). Diagnosing prosopagnosia: Effects of ageing, sex, and participant-stimulus ethnic match on the Cambridge face memory test and cambridge face perception test. *Cognitive Neuropsychology*, *26*, 423–455.
Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: linear ballistic accumulation. *Cognitive Psychology*, *57*, 153–178.
Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*, 305–327.
Burns, E. J., Martin, J., Chan, A. H. D., & Xu, H. (2017). Impaired processing of facial happiness, with or without awareness, in developmental prosopagnosia. *Neuropsychologia*, *102*, 217–228.
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*, 1652–1666.
Carpenter, R. H. (2004). Contrast, probability, and saccadic latency: Evidence for independence of detection and decision. *Current Biology*, *14*, 1576–1580.
Cohen, A. L., Soussand, L., Corrow, S. L., Martinaud, O., Barton, J. J. S., & Fox, M. D. (2019). Looking beyond the face area: Lesion network mapping of prosopagnosia. *Brain*, *142*, 3975–3990.
Corrow, S. L., Albonico, A., & Barton, J. J. (2018). Diagnosing prosopagnosia: The utility of visual noise in the Cambridge Face Recognition Test. *Perception*, *47*, 330–343.
Corrow, S. L., Dalrymple, K. A., & Barton, J. J. S. (2016). Prosopagnosia: Current perspectives. *Eye and Brain*, *8*, 165–175.
Corrow, S. L., Davies-Thompson, J., Fletcher, K., Hills, C., Corrow, J. C., & Barton, J. J. S. (2019). Training face perception in developmental prosopagnosia through perceptual learning. *Neuropsychologia*, *134*, 107196.
Cousineau, D., & Shiffrin, R. M. (2004). Termination of a visual search with large display size effects. *Spatial Vision*, *17*, 327–352.
Crawford, J. R., & Garthwaite, P. H. (2002). Investigation of the single case in neuropsychology: Confidence limits on the abnormality of test scores and test score differences. *Neuropsychologia*, *40*, 1196–1208.
Crawford, J. R., Garthwaite, P. H., & Gray, C. D. (2003). Wanted: Fully operational definitions of dissociations in single-case studies. *Cortex*, *39*, 357–370.
Crawford, J. R., Garthwaite, P. H., & Porter, S. (2010). Point and interval estimates of effect sizes for the case-controls design in neuropsychology: Rationale, methods, implementations, and proposed reporting standards. *Cognitive Neuropsychology*, *27*, 245–260.
Crawford, J. R., Garthwaite, P. H., & Ryan, K. (2011). Comparing a single case to a control sample: Testing for neuropsychological deficits and dissociations in the presence of covariates. *Cortex*, *47*, 1166–1178.
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.
Damasio, A. R., Tranel, D., & Damasio, H. (1990). Face agnosia and the neural substrates of memory. *Annual Review of Neuroscience*, *13*, 89–109.
Davies-Thompson, J., Pancaroglu, R., & Barton, J. (2014). Acquired prosopagnosia: Structural basis and processing impairments. *Frontiers in Bioscience*, *6*, 159–174.
Donkin, C., Averell, L., Brown, S., & Heathcote, A. (2009). Getting more from accuracy and response time data: Methods for fitting the linear ballistic accumulator. *Behavior Research Methods*, *41*, 1095–1110.
Donkin, C., Brown, S., & Heathcote, A. (2011). Drawing conclusions from choice response time models: A tutorial using the linear ballistic accumulator. *Journal of Mathematical Psychology*, *55*, 140–151.
Donkin, C., Brown, S., Heathcote, A., & Wagenmakers, E. J. (2003). Diffusion versus linear ballistic accumulation: Different models but the same conclusions about psychological processes? *Psychonomic Bulletin & Review*, *18*, 61–69.
Duchaine, B. C., Germine, L., & Nakayama, K. (2007). Family resemblance: Ten family members with prosopagnosia and within-class object agnosia. *Cognitive Neuropsychology*, *24*, 419–430.
Duchaine, B. C., & Nakayama, K. (2004). Developmental prosopagnosia and the Benton Facial Recognition Test. *Neurology*, *62*, 1219–1220.
Duchaine, B. C., & Nakayama, K. (2006). The Cambridge face memory test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, *44*, 576–585.
Duchaine, B. C., Nieminen-von Wendt, T., New, J., & Kulomäki, T. (2003). Dissociations of visual recognition in a developmental agnosic: Evidence for separate developmental processes. *Neurocase*, *9*, 380–389.
Duchaine, B. C., Parker, H., & Nakayama, K. (2003). Normal recognition of emotion in a prosopagnosic. *Perception*, *32*, 827–838.
Farah, M., Tanaka, J., & Drain, H. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 628–634.
Fox, C. J., Hanif, H. M., Iaria, G., Duchaine, B. C., & Barton, J. J. (2011). Perceptual and anatomic patterns of selective deficits in facial identity and expression processing. *Neuropsychologia*, *49*, 3188–3200.
Fox, C. J., Moon, S. Y., Iaria, G., & Barton, J. J. (2009). The correlates of subjective perception of identity and expression in the face network: An fMRI adaptation study. *Neuroimage*, *44*, 569–580.
Gerlach, C., Lissau, C. H., & Hildebrandt, N. K. (2018). On defining and interpreting dissociations. *Cognitive Neuropsychology*, *35*, 66–69.
Hanes, D. P., & Carpenter, R. H. (1999). Countermanding saccades in humans. *Vision Research*, *39*, 2777–2791.
Hardiess, G., Papageorgiou, E., Schiefer, U., & Mallot, H. A. (2010). Functional compensation of visual field deficits in hemianopic patients under the influence of different task demands. *Vision Research*, *50*, 1158–1172.
Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*, 223–233.
Hemström, J., Albonico, A., Djouab, S., & Barton, J. J. (2019). Visual search for complex objects: Set-size effects for faces, words and cars. *Vision Research*, *162*, 8–19.
Hooper, E. H. (1983). *Hooper Visual Organization Test (VOT)*. Los Angeles, CA: Western Psychological Services.
Humphreys, K., Avidan, G., & Behrmann, M. (2007). A detailed investigation of facial expression processing in congenital prosopagnosia as compared to acquired prosopagnosia. *Experimental Brain Research*, *176*, 356–373.
Jackson, M. C., Counter, P., & Tree, J. J. (2017). Face working memory deficits in developmental prosopagnosia: Tests of encoding limits and updating processes. *Neuropsychologia*, *106*, 60–70.

Jones, R., & Tranel, D. (2001). Severe developmental prosopagnosia in a child with superior intellect. *Journal of Clinical and Experimental Neuropsychology*, *23*, 265–273.
Klargaard, S. K., Starrfelt, R., & Gerlach, C. (2018). Inversion effects for faces and objects in developmental prosopagnosia: A case series analysis. *Neuropsychologia*, *113*, 52–60.
Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: The dangers of double dipping. *Nature Neuroscience*, *12*, 535–540.
Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & Van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, *24*, 1377–1388.
Lee, D. H., Corrow, S. L., Pancaroglu, R., & Barton, J. J. S. (2019). The scanpaths of subjects with developmental prosopagnosia during a face memory task. *Brain Sciences*, *9*, 188.
Liu, R. R., Corrow, S. L., Pancaroglu, R., Duchaine, B., & Barton, J. J. (2015). The processing of voice identity in developmental prosopagnosia. *Cortex*, *71*, 390–397.
MacMillan, N., & Creelman, C. (1991). *Detection theory: A user's guide*. Cambridge: Cambridge University Press.
Malaspina, M., Albonico, A., Lao, J., Caldara, R., & Daini, R. (2018). Mapping self-face recognition strategies in congenital prosopagnosia. *Neuropsychology*, *32*, 123–137.
Malaspina, M., Albonico, A., Toneatto, C., & Daini, R. (2017). What do eye movements tell us about the visual perception of individuals with congenital prosopagnosia? *Neuropsychology*, *31*, 546–563.
Moran, R., Zehetleitner, M., Liesefeld, H. R., Müller, H. J., & Usher, M. (2016). Serial vs. parallel models of attention in visual search: Accounting for benchmark RT-distributions. *Psychonomic Bulletin & Review*, *23*, 1300–1315.
Moroz, D., Corrow, S. L., Corrow, J. C., Barton, A. R. S., Duchaine, B., & Barton, J. J. S. (2016). Localization and patterns of cerebral dyschromatopsia: A study of subjects with prospagnosia. *Neuropsychologia*, *89*, 153–160.
Özbek, M., & Bindemann, M. (2011). Exploring the time course of face matching: Temporal constraints impair unfamiliar face identification under temporally unconstrained viewing. *Vision Research*, *51*, 2145–2155.
Pancaroglu, R., Hills, C. S., Sekunova, A., Viswanathan, J., Duchaine, B., & Barton, J. J. (2016). Seeing the eyes in acquired prosopagnosia. *Cortex*, *81*, 251–265.
Prinzmetal, W., Whiteford, K. L., Austerweil, J. L., & Landau, A. N. (2015). Spatial attention and environmental information. *Journal of Experimental Psychology: Human Perception and Performance*, *41*, 1396–1408.
Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.
Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*, 347–356.
Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*, 438–481.
Reddi, B. A. (2001). Decision making: The two stages of neuronal judgement. *Current Biology*, *11*, R603–R606.
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.
Rubino, C., Corrow, S. L., Corrow, J. C., Duchaine, B., & Barton, J. J. (2016). Word and text processing in developmental prosopagnosia. *Cognitive Neuropsychology*, *33*, 315–328.
Schmalzl, L., Palermo, R., Green, M., Brunsdon, R., & Coltheart, M. (2008). Training of familiar face recognition and visual scan paths for faces in a child with congenital prosopagnosia. *Cognitive Neuropsychology*, *25*, 704–729.
Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H. M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General*, *136*, 414–429.
Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, *47*, 609–612.
Schwarzer, G., Huber, S., Grüter, M., Grüter, T., Gross, C., Hipfel, M., et al. (2007). Gaze behaviour in hereditary prosopagnosia. *Psychological Research*, *71*, 583–590.
Shah, P., Sowden, S., Gaule, A., Catmur, C., & Bird, G. (2015). The 20 item prosopagnosia index (PI20): Relationship with the Glasgow Face-Matching Test. *Royal Society Open Science*, *2*, 150305.
Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge: Cambridge University Press.
Simpson, S. A., Abegg, M., & Barton, J. J. (2011). Rapid adaptation of visual search in simulated hemianopia. *Cerebral Cortex*, *21*, 1593–1601.
Smith, P. L., & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychological Review*, *116*, 283–317.
Spinnler, H., & Tognoni, G. (1987). Standardizzazione e taratura italiana di test neuropsicologici. *Italian Journal of Neurological Sciences*, *8(Suppl.)*, 1–120.
Susilo, T., & Duchaine, B. (2013). Advances in developmental prosopagnosia research. *Current Opinion in Neurobiology*, *23*, 423–429.
Townsend, J. T., & Wenger, M. J. (2004). The serial-parallel dilemma: A case study in a linkage of theory and method. *Psychonomic Bulletin & Review*, *11*, 391–418.
Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136.
Ulrich, P. I., Wilkinson, D. T., Ferguson, H. J., Smith, L. J., Bindemann, M., Johnston, R. A., et al. (2017). Perceptual and memorial contributions to developmental prosopagnosia. *Quarterly Journal of Experimental Psychology*, *70*, 298–315.
Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, *79*, 471–491.
Van Zandt, T., Colonius, H., & Proctor, R. W. (2000). A comparison of two response time models applied to perceptual matching. *Psychonomic Bulletin & Review*, *7*, 208–256.
Warrington, E. (1984). *Manual for the recognition memory test for words and faces*. Windsor, United Kingdom: NFER-Nelson.
Warrington, E. K., & James, M. (1991). *The visual object and space perception battery: VOSP*. London: Pearson.
Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-Third Edition (WAIS-III)*. San Antonio, TX: The Psychological Corporation.
Wilson, B., Cockburn, J., & Halligan, P. W. (1987). *Behavioural Inattention Test*. Titchfield, United Kingdom: Thames Valley Test Company.
Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, *9*, 33–39.
Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends in Cognitive Sciences*, *7*, 70–76.
Wolfe, J. M. (2019). Visual attention: The multiple ways in which history shapes selection. *Current Biology*, *29*, R155–R156.
Zihl, J. (1995). Visual scanning behavior in patients with homonymous hemianopia. *Neuropsychologia*, *33*, 287–303.