FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

FACE PERCEPTION DEFICITS IN DEVELOPMENTAL PROSOPAGNOSIA

BY

ELLA MACASKILL

A thesis

submitted to the Victoria University of Wellington

in fulfilment of the requirements for the degree of

Master of Science in Psychology

specialising in Cognitive and Behavioural Neuroscience

Victoria University of Wellington

2021

Abstract

Face recognition is a fundamental cognitive function that is essential for social interaction – yet not everyone has it. Developmental prosopagnosia is a lifelong condition in which people have severe difficulty recognising faces but have normal intellect and no brain damage. Despite much research, the component processes of face recognition that are impaired in developmental prosopagnosia are not well understood. Two core processes are face perception, being the formation of visual representations of a currently seen face, and face memory, being the storage, maintenance, and retrieval of those representations. Most studies of developmental prosopagnosia focus on face memory deficits, but a few recent studies indicate that face perception deficits might also be important. Characterising face perception in developmental prosopagnosia is crucial for a better understanding of the condition. In this thesis, I addressed this issue in a large-scale experiment with 108 developmental prosopagnosics and 136 matched controls. I assessed face perception abilities with multiple measures and ran a broad range of analyses to establish the severity, scope, and nature of face perception deficits in developmental prosopagnosia. Three major results stand out. First, face perception deficits in developmental prosopagnosia were severe, and could be comparable in size to face memory deficits. Second, the face perception deficits were widespread, affecting the whole sample rather than a subset of individuals. Third, the deficits were mainly driven by impairments to mechanisms specialised for processing upright faces. Further analyses revealed several other features of the deficits, including the use of atypical and inconsistent strategies for perceiving faces, difficulties matching the same face across different pictures, equivalent impact of lighting and viewpoint variations in face images, and atypical perceptual and non-perceptual components of test performance. Overall, my thesis shows that face perception deficits are more central to developmental prosopagnosia than previously thought and motivates further research on the issue.

Acknowledgements

I am so thankful to all the wonderful people who have supported me along the way in this journey. It takes a village to raise a child, and it seems to take the same to raise a thesis!

First, to my supervisor Tirta. You are not only a brilliant mentor, but an all-round brilliant person. Thank you so much for your encouragement, generosity, and wisdom. I feel I have grown tremendously with your guidance and have learnt a great amount that will benefit my future. I am honoured to have been your student. Thank you.

My endless thanks also go to my family and friends. Mum, Dad, and Callum, thank you for your encouragement and for always having my back. To my parents especially, I owe everything I have achieved to your love and support. To my grandparents, thank you for your kindness, for reminding me what really matters when the going gets tough, and for your financial support. Conor, you have been such a source of happiness throughout this postgraduate whirlwind. I am so lucky; thank you for every moment. To my friends: what would I do without you! Special thanks to Anna, Caitlin, Nikki, Zoë, Andre, Ellie, and Liv (and the entity that is the ‘social office’). Thanks also to my friends in the Early Start Denver Model team; you always brighten my day.

Finally, I’d like to acknowledge the Marsden fund awarded to Tirta by the Royal Society of New Zealand, the associated Marsden Master’s scholarship given to me, and the Virginia Rose Burns award I received from Victoria University of Wellington. Your financial support has let me turn my research goals into reality.

Table of Contents

Abstract.....	ii
Acknowledgements	iii
Table of Contents	iv
List of Figures.....	vi
List of Tables	viii
Chapter 1: Introduction	1
1.1 Developmental prosopagnosia	2
1.2 Face perception versus face memory	5
1.3 Face perception studies of developmental prosopagnosia	11
1.4 Current study	15
Chapter 2: Method.....	17
2.1 Participants.....	17
2.2 Materials.....	18
2.3 Design.....	26
2.4 Procedure.....	28
Chapter 3: Results.....	29
3.1 Analytic approach and data pre-processing.....	29
3.2 The severity of face perception deficits in developmental prosopagnosia.....	31
3.3 The scope of face perception deficits in developmental prosopagnosia	35
3.3.1 Whole-sample analysis of scope	35
3.3.2 Single-case analysis of scope	38
3.3.3 Face perception deficits and developmental prosopagnosia diagnosis	40
3.4 The nature of face perception deficits in developmental prosopagnosia	43
3.4.1 Upright-sensitivity of face perception deficits	43
3.4.2 Associations and dissociations between face perception tests	47

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

3.4.3 Within-identity perception versus between-identity perception.....	49
3.4.4 Face perception deficits across image changes by lighting and viewpoint.....	52
3.4.5 Components of perception in the face perception tests	54
Chapter 4: Discussion	58
4.1 Empirical implications	58
4.2 Theoretical implications	63
4.3 Methodological implications.....	68
4.4 Conclusion.....	72
References	73
Appendix A: Literature review of face perception studies in developmental prosopagnosia	106
A.1 Literature review	106
A.2 Literature review references	111
Appendix B: Supplementary analyses	113
B.1 Group age and sex comparisons	113
B.2 Development of the BFRT and GFMT trial time limits	113
B.3 Response time analyses with median response time	114
B.4 Hot deck data imputation.....	116
B.5 Analyses without response time covariate.....	117
B.6 Multivariate single-case analysis of scope output	118
B.7 Hierarchical regression with face perception tests as separate predictors	120
B.8 Response time inversion effects	122
B.9 Potential floor effects in the inverted conditions	123
B.10 Signal detection analysis of the GFMT	125
B.11 Developmental prosopagnosic regression-based inversion effects compared to control variation.....	126
B.12 Predicting developmental prosopagnosia severity using face perception	127

List of Figures

Figure 1. Bruce & Young's (1986) Model of Face Recognition

Figure 2. Example Match, Lighting-change, and Viewpoint-change Trials from the BFRT

Figure 3. Example Within-identity and Between-identity Trials from the GFMT

Figure 4. An Example Trial from the CFPT

Figure 5. Example Trials from the CFMT-Aus

Figure 6. Example Inverted Face Trials for the Face Perception Tests

Figure 7. Developmental Prosopagnosic and Control Accuracy on the Upright Face Perception and Memory Tests

Figure 8. Perception Test Shift Functions and Decile Differences for Developmental Prosopagnosics and Controls

Figure 9. Developmental Prosopagnosic Z-scores for the Face Perception Tests

Figure 10. Developmental Prosopagnosic and Control Inversion Effects on the Face Perception Tests

Figure 11. Correlations Between Upright Face Perception Tests for Developmental Prosopagnosics and Controls

Figure 12. Developmental Prosopagnosic and Control Accuracy on the GFMT Within-identity and Between-identity Trials

Figure 13. Developmental Prosopagnosic and Control Accuracy on the BFRT Lighting and Viewpoint Trials

Figure 14. Developmental Prosopagnosic and Control Drift Rates (v), Thresholds (a), and Non-decision Times (t_0) on the GFMT

Figure 15. Comparing Developmental Prosopagnosic Performance to the Variation Predicted by Performance in Controls

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

Figure 16. Developmental Prosopagnosic Z-scores on their Diagnostic CFMT Assessment Compared to the CFMT-Aus and Perception Tests.

Figure B1. Prosopagnosic and Control Response Time Inversion Effects on the Face Perception Tests

Figure B2. Comparing Developmental Prosopagnosic Performance to the Variation Predicted by Performance in Controls using Regression-based Inversion Effects

List of Tables

Table 1. *Developmental Prosopagnosic and Control Mean Response Times on the Face Perception and Memory Tests*

Table 2. *Developmental Prosopagnosic and Control Accuracy on the Face Perception and Memory Tests*

Table 3. *Group Classification in the Face Perception Only Model*

Table 4. *Group Classification in the Face Perception and Face Memory Model*

Table 5. *Pearson's r and p -values for Upright Test Correlations for Developmental Prosopagnosics and Controls*

Table A1. *Summary of Prior Studies Using the BFRT, GFMT, and/or CFPT in Developmental Prosopagnosia (Sorted by Scope of Deficits)*

Table B1. *Developmental Prosopagnosic Multivariate Single-case Comparisons on All Face Perception Tests*

Table B2. *Group Classification in the BFRT, GFMT, and CFPT Model*

Table B3. *Group Classification in the BFRT, GFMT, CFPT, and CFMT-Aus Model*

Table B4. *Developmental Prosopagnosics' Inversion Effect Sizes for the Original, Easiest 50% of Trials, and Top 50% of Subjects Data*

Chapter 1: Introduction

“I can see there are eyes and a mouth and ears. ... But it’s very difficult for my brain to hold them all together”. (Evans, 2019, para. 22)

Within a day, we might see hundreds of faces and recognise those we know with relative ease. For people with developmental prosopagnosia, however, recognising faces is a challenge. Developmental prosopagnosia is a neurodevelopmental disorder defined by severe and lifelong difficulties in face recognition in the absence of broader visual problems, intellectual deficits, and brain trauma (McConachie, 1976). The causes of developmental prosopagnosia remain elusive, but it likely results from failures to develop the necessary cognitive and neural processes for face recognition (Susilo & Duchaine, 2013).

Developmental prosopagnosia has been traditionally conceptualised as a disorder of face recognition. Yet face recognition is not a single, unitary process. Leading models of face recognition describe face recognition as the outcome of multiple component processes (Bruce & Young, 1986), implying that developmental prosopagnosia may result from impairments to one or more of these processes. One core process is *face memory*, which involves linking a perceived face to a stored representation of that face (Bruce & Young, 1986). Much research has focused on investigating memory impairments in developmental prosopagnosia, and many studies have shown that developmental prosopagnosics perform poorly on memory tests such as recognising famous faces or remembering newly learned faces (for example, Bate et al., 2014; Duchaine, 2011; Duchaine & Nakayama, 2005; Duchaine & Nakayama, 2006a; Duchaine & Nakayama, 2006b; Lee et al., 2010; McKone et al., 2011; Nunn et al., 2001; Susilo & Duchaine, 2013). Indeed, diagnostic procedures for developmental prosopagnosia rely heavily on tests with significant memory demand (Bate & Tree, 2017; Susilo & Duchaine, 2013). Another core process is *face perception*. Face perception involves encoding the structural elements of a currently seen face to form a good visual representation

of that face (Bruce & Young, 1986). Good face perception is required for good face memory, but surprisingly little is known about face perception in developmental prosopagnosia. A few recent studies suggest that face perception deficits may be quite common in developmental prosopagnosia (Biotti et al., 2019; Mishra et al., 2020; Djouab et al., 2020), but the severity, scope, and nature of such deficits remain unclear. A more comprehensive investigation of face perception deficits is critical for a better understanding developmental prosopagnosia and to advance future research.

The overarching goal of my thesis is to determine the severity, scope, and nature of face perception deficits in developmental prosopagnosia. To do so I designed a study with three key features. One is the use of an unprecedentedly large developmental prosopagnosic sample ($n = 108$), which provides a greater statistical power to detect subtle deficits and reveal potential variations within the sample. I also used a study design that overcame major methodological limitations of prior studies, including circular analysis, limited task variation, and lack of response time measures. Finally, I used a broad range of analytic and modelling approaches to characterise various aspects of the face perception deficits. Overall, the results of my thesis will address major questions about face perception deficits in developmental prosopagnosia. They will also inform theories and debates about developmental prosopagnosia, yield insights for future work addressing the front-end of face processing in developmental prosopagnosia, and guide the development of rehabilitation strategies.

1.1 Developmental prosopagnosia

Faces offer rich information about identity, gender, age, and emotional state. As Oliver Sacks (2010, para. 1) wrote: *“It is with our faces that we face the world”*. Now, imagine facing the world without being able to recognise faces. This is what people with developmental prosopagnosia experience daily. As described by one of our past participants, developmental prosopagnosia is akin to *“a lifetime of wading through a sea of strangers”* (C.

E., personal communication, February 17, 2017). Developmental prosopagnosia may impact up to one in 50 adults (Bowles et al., 2009; Kennerknecht et al., 2006). The severity of developmental prosopagnosia varies, with some extreme cases describing experiences such as “*I nearly took the wrong child home from school once*” (E. R., personal communication, August 9, 2018), and “*I always joke that if my wife ever wanted to leave me all she would need is a new hair[style]*” (S. E., personal communication, July 4, 2020). Developmental prosopagnosics often describe using compensatory strategies to recognise their friends and family, such as relying on hairstyle, clothes, voice, and gait. Even with these strategies, however, difficulty recognising faces in developmental prosopagnosia can lead to chronic anxiety, reduced self-esteem, and lower self-confidence (Yardley et al., 2008). As a result, developmental prosopagnosics may avoid social situations and receive fewer job opportunities (Yardley et al., 2008). As C. E. (personal communication, February 17, 2017) described, “*It’s incredibly isolating to see every person as a stranger ... I’ve always thought I was just terribly stupid or socially incompetent*”.

Developmental prosopagnosia research has made good progress since the condition was first described by McConachie in 1976. For example, developmental prosopagnosia is selective to faces, as many developmental prosopagnosics perform normally or are only slightly impaired at recognising non-face objects (Behrmann & Avidan, 2005; Duchaine & Nakayama, 2005; Gray et al., 2019; Nunn et al., 2001, though see Gerlach et al., 2016). Some developmental prosopagnosics, however, also have difficulty recognising bodies (Biotti et al., 2017a; Righart & de Gelder, 2007), biological motion (Lange et al., 2009), and spatial layout (Bate et al., 2019a; Corrow et al., 2016). Developmental prosopagnosia can run in families, suggesting a genetic component (Duchaine et al., 2007a; Kennerknecht et al., 2006; Lee et al., 2010; Schmalzl et al., 2008a, but see Bate et al., 2009; Duchaine et al., 2008). Neural studies have linked developmental prosopagnosia with atypical event-related potential (ERP)

markers of face recognition (e.g., N170, Towler et al., 2017) and atypical fMRI responses in face-selective regions (Jiahui et al., 2018; Avidan & Behrman, 2009; Furl et al., 2011). Cognitive studies have shown that holistic face processing – the processing of faces as a whole entity rather than separate features – is likely impaired in developmental prosopagnosia (Avidan et al., 2011; Behrmann et al., 2005; Duchaine et al., 2007b; De Gelder & Rouw, 2000; Le Grand et al., 2006; Schmalzl et al., 2008a). Finally, developmental prosopagnosia may involve difficulties processing facial information beyond identity, including expression (Biotti & Cook, 2016; Schmalzl et al., 2008a), gender (Duchaine et al., 2006; but see Chatterjee & Nakayama, 2012), attractiveness (Le Grand et al., 2006), and trustworthiness (Todorov & Duchaine, 2008).

While we know quite a lot about what kinds of stimuli are subject to processing difficulties in developmental prosopagnosia, we know substantially less about the specific processes of face recognition that are impaired. Face recognition requires multiple component processes, with two core components being face perception (the ability to create a visual representation of the currently seen face) and face memory (the ability to match that visual representation to stored knowledge of that face; Bruce & Young, 1984). So far, developmental prosopagnosia research tends to focus on face memory, perhaps because developmental prosopagnosia has been traditionally studied using measures with clear memory demands. Consistently, the terms face memory and face recognition were used synonymously in early research (e.g., Bentin et al., 1999; De Haan & Campbell, 1991; Jones & Tranel, 2001).

Developmental prosopagnosics are well known to perform poorly on a broad range of face memory tests, such as memory for famous faces, newly learnt faces, and even faces seen only seconds prior to testing (for a summary, see: Duchaine, 2011; Susilo & Duchaine, 2013). Such findings are usually taken as evidence of impaired face memory, but there is a key

caveat. Failures to match a visual face representation to stored knowledge of that face may result from aberrant matching or bad storage, but it may also result from a poor visual representation to begin with. This means poor performance on face memory tests could be partially – or even entirely – explained by deficits in face perception. Little is known about face perception deficits, leaving much uncertainty about what specific processes are indeed impaired in developmental prosopagnosia. Better research with cleaner measures of face perception and face memory are needed to clarify the core deficit profile of developmental prosopagnosia. To achieve this, a clear conceptual and empirical distinction between face perception and face memory must be established.

1.2 Face perception versus face memory

The distinction between face perception and memory in modern prosopagnosia research was largely inspired by classic work on visual agnosia, namely the inability to visually recognise objects (Adler, 1944; Albert et al., 1975; Benson & Greenberg, 1969; Humphreys & Riddoch, 1987; Landis et al., 1986; Levine & Calvanio, 1989; Lissauer, 1890; Milner et al., 1991; Moscovitch et al., 1997; Ratcliff & Newcombe, 1982; Rubens & Benson, 1971). A landmark contribution was Lissauer's (1890) distinction between *apperceptive* and *associative* visual agnosia. People with apperceptive agnosia typically struggle to match shapes and objects, and might produce scribbles when attempting to copy simple drawings (Adler, 1944; Benson & Greenberg, 1969; Gelb & Goldstein, 1938). Such deficits have been taken as evidence of impairments with object perception, namely the ability to form a visual representation of the object (Lissauer, 1890). In contrast, people with associative agnosia can match and copy drawings with success, but struggle to recognise and link knowledge to objects – even for their own copied drawing of an object (Rubens & Benson, 1971). This form of deficit is considered evidence for problems with object memory, namely the ability to assign meaning to the visual representation. This distinction gives rise to the key insight that

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

object recognition is not a unitary process. Both object perception and object memory contribute to object recognition, and these component processes are separable.

The distinction between perception and memory is not only central to understanding object recognition, but also the more specific case of face recognition. This distinction is captured by the leading model of face recognition developed by Bruce and Young (1986, Figure 1). In this model, the recognition of a face follows a series of processing stages. Upon seeing a face, the first stage involves encoding elements of face structure. Invariant facial information relevant to the face's identity is encoded, such as the shape, size, and positioning of individual features. Bruce and Young (1986) coined this first stage as forming the 'structural code' of a face, or a visual representation of the face. The next stage involves matching the structural code to the stored facial memories – called 'face recognition units'. A match will occur if the face is familiar. For unfamiliar faces, the face recognition units are gradually formed and strengthened over repeated exposures to the face. The process of matching the visual representation to the stored face recognition units is the first stage of face memory. From this point, the stored recognition units can activate the associated 'person identity node' for the face. This allows episodic and semantic information about the person to be retrieved, like their name and voice.

Figure 1

Bruce and Young's (1986) Model of Face Recognition

The distinction between face perception and face memory is supported by abundant empirical evidence from a wide range of sources, including neuropsychological studies, behavioural studies, neural studies, and developmental studies with children. Below, I selectively review some of the strongest and most relevant evidence from these sources.

In terms of neuropsychological studies, face perception and face memory have been dissociated in several conditions including acquired prosopagnosia, prosopamnesia, and autism. People with acquired prosopagnosia have difficulty recognising faces following brain damage (Ellis & Florence, 1990). Similar to visual agnosia, acquired prosopagnosia can be apperceptive or associative (Davies-Thompson et al., 2014; De Renzi et al., 1991; Sergent & Signoret, 1992). Apperceptive acquired prosopagnosia is associated with poor performance

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

on face perception tests, such as matching unfamiliar faces (Barton, 2008; De Renzi, 1986; Riddoch et al., 2008; Rossion et al., 2003), as well as on face memory tests, such as identifying famous faces. Associative prosopagnosia, on the other hand, is associated with normal performance on face perception tests and difficulty with face memory tests (Barton, 2008; De Renzi et al., 1991; De Renzi & di Pellegrino, 1998; Sergent & Signoret, 1992). Apperceptive prosopagnosia is typically caused by occipital temporal lesions, while associative prosopagnosia is generally driven by anterior temporal lesions (Barton, 2008; Davies-Thompson et al., 2014; Kanwisher & Barton, 2011). A similar dissociation has been reported in prosopamnesia, a rare acquired disorder in which perception and memory for faces known pre-morbidly are preserved, but post-morbid learning of new faces is impaired (Tippett et al., 2000; Williams et al., 2007). Prosopamnesia results in impaired memory for newly learned faces, but the perception of those faces can be intact (Tippett et al., 2000). This behavioural profile is possible given the cortical structures involved in forming new memories of faces are distinct to those for retaining memories of already learnt faces (Squire, 1992; Hodges, 1995). Similarly, normal face perception with impaired face memory has been reported in autism (Weigelt et al., 2012; Weigelt et al., 2013). Participants with autism perform typically when matching faces that are shown together, but struggle when the faces are shown one after another or when they must recognise familiar faces (Weigelt et al., 2012). Even sequentially presenting a target face and the test faces with a zero second delay (Wilson et al., 2010a; Wilson et al., 2010b), or brief 250ms mask (Wallace et al., 2008), impairs performance of participants with autism compared to controls.

Face perception and face memory have also been dissociated in behavioural studies with typical participants. Hacker and Biederman (2019) had participants view a target face and then identify which of two test faces was the target. Demand on face perception was isolated by varying the similarity between the test faces, and demand on face memory was

isolated by varying the length of delay between the target and test faces. No interaction between delay length and face similarity was found across participants. In other words, participants with above-average face perception did not necessarily have above-average face memory. In a series of individual differences studies, Wilhelm and colleagues (2010; Hildebrandt et al., 2010; Hildebrandt et al., 2011; Kaltwasser et al., 2014) modelled variations in face recognition skills by administering a battery of tests to over 300 participants (see Herzmann et al., 2008). The best model fit was achieved when face perception and face memory were treated as separate components (Hildebrandt et al., 2010; Wilhelm et al., 2010). Moreover, age-related reductions in face recognition began earlier for face memory (40+ years-old) than for face perception (60+ years-old; Hildebrandt et al., 2010; Hildebrandt et al., 2011).

Turning to neural research, fMRI studies have found that three core ‘face-selective’ areas – namely the fusiform face area (FFA), occipital face area (OFA), and the superior temporal sulcus (STS) – show no difference in activation for unfamiliar versus familiar faces (di Oleggio Castello et al., 2017; Gobbini & Haxby, 2007; Haxby et al., 2000; Haxby & Gobbini, 2011, but see Davies-Thompson et al., 2009 and Davies-Thompson et al., 2013). Unfamiliar faces have no stored memory representations, meaning only perceptual processing occurs (Bruce & Young, 1986; Gobbini & Haxby, 2007). Familiar faces, on the other hand, activate both perception and memory processes (Bruce & Young, 1986; Gobbini & Haxby, 2007). The lack of activation difference for unfamiliar and familiar faces in the FFA, OFA, and STS has thus led researchers to conclude that these areas process the visual appearance of faces. Unlike the three core face-selective areas, an extended system of other frontal, parietal, and subcortical regions all show enhanced activation for familiar faces compared to unfamiliar faces. These regions are therefore thought to be critical for the retrieval of person knowledge (di Oleggio Castello et al., 2017; Gobbini & Haxby, 2007; Haxby et al., 2000;

Haxby & Gobbini, 2011). The interaction between the FFA, OFA, STS, and extended face memory network is core to understanding face recognition overall (Natu & O'Toole, 2011).

Distinct neural markers for face perception and face memory have also been identified in ERP studies using electroencephalogram (EEG). The N170 is an ERP component that peaks at approximately 170ms after stimulus onset and is larger for faces than non-face objects (Bentin et al., 1996; Eimer, 2000a; Eimer, 2000b; Eimer, 2000c; Rossion & Jacques, 2012). The N170 magnitude is similar for unfamiliar and familiar faces (Eimer, 2000a; Bentin & Deouell, 2000), suggesting that it is a marker of face perception (Eimer 2000a; Eimer, 2000b; Gosling & Eimer, 2011; Rossion & Jacques, 2012; Sagiv & Bentin, 2001). In contrast, the N250 is a later ERP component that is potentiated by familiar faces but not unfamiliar faces (Gosling & Eimer, 2011). The N250 also eventually becomes sensitive to repetitions of an unfamiliar face over time, as a memory representation for that face gradually builds. The N250 is maximally responsive to faces with long-term associative memory, such as pictures of one's own face (Herzmann et al., 2004; Tanaka et al., 2006). These findings suggest that the N250 is a marker of face memory (Gosling & Eimer, 2011).

Another neural method that has been used to dissociate face perception and face memory is transcranial magnetic stimulation (TMS). TMS to the right OFA (rOFA) disrupts performance on sequential matching tasks with faces but not objects (Bona et al., 2018; Pitcher et al., 2009; Pitcher et al., 2012). Importantly, the disruption only happens when TMS is delivered early during stimulus-viewing as opposed to later when the stimulus is no longer shown (Holiday et al., 2015; Pitcher et al., 2007). This finding suggests that TMS to the rOFA selectively disrupts the early processing of perceptual face information rather than the retention of that information in memory (Holiday et al., 2015), consistent with the notion that face perception and face memory are separable processes.

Finally, studies with children show that face perception and face memory follow different developmental trajectories. One study compared children aged 5-10 years and adults on face and object perception and memory tasks and found that face perception matured earlier than face memory (Weigelt et al., 2014). Face perception also matured at the same rate as object perception, whereas face memory took longer to mature compared to object memory. A study of children and adults with developmental prosopagnosia reported that while all children had impaired face perception and face memory, the adults showed a dissociation with over half only having impaired face memory (Dalrymple et al., 2014).

Together, the broad range of results reviewed above demonstrate that face perception is dissociable from face memory, and that the distinction is fundamental to understanding face recognition and its disorder in developmental prosopagnosia. Despite this, developmental prosopagnosia remains defined simply as disordered face *recognition*. Little work has addressed the subcomponents of face recognition in developmental prosopagnosia – particularly for face perception (Bate & Tree, 2017).

1.3 Face perception studies of developmental prosopagnosia

Most studies of developmental prosopagnosia tend to focus on face memory, with occasional inclusion of a face perception test as part of their diagnostic procedure (see Table A1 in Appendix A for a summary of the 68 reviewed studies). Only ten studies have directly focused on understanding face perception (summaries available in Table A1 in Appendix A). The results are mixed. In terms of the scope or prevalence of face perception deficits, estimates vary from a minority of the sample (Duchaine & Nakayama, 2004; Mishra et al., 2020; Stumps et al., 2020; Ulrich et al., 2017), to around half (Bate et al., 2019b; Dalrymple et al., 2014; White et al., 2017) to most of the sample (Biotti et al., 2019; Djouab et al., 2020; Huis in't Veld et al., 2012). In terms of severity, most studies conclude that face perception deficits are smaller in magnitude compared to face memory deficits (Stumps et al., 2020;

Ulrich et al., 2017) – but as I argue below, this conclusion is methodologically flawed. In terms of the nature of deficits, two studies have investigated orientation sensitivity, motivated by the classic finding that people with typical recognition are much better at perceiving upright faces compared to upside down faces (i.e., inverted; Yin, 1969). Both studies found that developmental prosopagnosics had larger deficits with upright faces than with inverted faces (Biotti et al., 2019; Huis in't Veld, 2012). A sizable literature has also investigated the role of holistic processing, showing that sensitivity to the configuration of features within a face is reduced in developmental prosopagnosia (Avidan et al., 2011; DeGutis et al., 2014; Palermo et al., 2011; Towler et al., 2018). Most of these studies, however, conflate their perceptual manipulation of face configuration with test designs that require face memory.

Taken together, the studies above have begun to characterise face perception deficits in developmental prosopagnosia. Yet the overall conclusions that can be drawn are limited, because all the above studies suffer from two or more major methodological factors. These include (1) circular analysis of face memory deficits, (2) underpowered sample sizes, (3) limited variation of face perception tests, and (4) lack of response time analyses. Below I describe each limitation and how it undermines prior studies.

Circular analysis of face memory deficits. A standard diagnostic criterion for developmental prosopagnosia is a score two standard deviations or more below the control mean on several face memory tests, the most common one being the Cambridge Face Memory Test (CFMT; Duchaine & Nakayama, 2006b). Of the ten face perception studies in developmental prosopagnosia, all but two reused this diagnostic score as an experimental measure of face memory deficits and compared it to face perception measures that were obtained in the main experiment or study (see Table A1 in Appendix A1). This comparison is invalid because the size of the face memory deficit is guaranteed to be large due to the predefined diagnostic cut-off, while the size of the face perception deficit can vary. Such

comparisons are therefore biased towards finding smaller face perception deficits compared to face memory deficits. This problem is known as “circular analysis” – the reusing of data for group selection to make inferences about an effect size of interest (Kriegeskorte, et al., 2009). This limitation effectively undermines all previous studies where CFMT scores from diagnosis were reused to index face memory deficits. Strikingly, no studies to date have compared face perception and face memory using standardised measures that are both obtained as part of the main investigation.

Underpowered sample size. Underpowered samples lack the statistical power for detecting subtle deficits, producing precise effect size estimates, and identifying individual variations among the sample (Button et al., 2013; Colquhoun, 2014; Poldrack et al., 2017; Vadillo et al., 2016). They also prevent estimation of how entire distributions might differ and inflate the chance of false positives (Button et al., 2013; Rousselet et al., 2017). Most of the prior studies investigating face perception had a sample size of around 11 developmental prosopagnosics (mode = 11, median = 11, mean = 26; see Table A1 in Appendix A). Such sample sizes are likely underpowered – not only due to measurement noise, but also because developmental prosopagnosia is a heterogeneous condition with different individuals showing different profiles of deficits (Bate et al., 2019a; Duchaine et al., 2006; Lange et al., 2009; Righart & de Gelder, 2007; Schmalzl et al., 2008). Much larger samples are needed to estimate the severity and scope of face perception deficits more precisely, and to map out potential individual and subgroup variations within the sample.

Limited variation of face perception tests. Seven of the ten past studies only used one measure of face perception. This is an issue because face perception can be tested in many ways, and different methods of testing may capture different aspects of face perception. For example, some perception tests require fine-grained discrimination between similar-looking faces, such as the Cambridge Face Perception Test (Duchaine et al, 2007). Some involve

matching faces across natural changes in lighting and viewpoint, such as the Benton Face Recognition Test (Benton et al., 1983). Others require differentiating between two pictures of the same person and two pictures of different people, like the Glasgow Face Matching Test (Burton et al., 2010). These perception tests also differ in their task design and parameters (such as the face stimuli, response format, instructions, duration, and cognitive demand). To draw a more general conclusion about face perception deficits in developmental prosopagnosia, multiple measures of face perception must be used with the same sample of participants.

Lack of response times. Developmental prosopagnosics may improve their test performance by putting in more effort or using atypical compensatory strategies, both of which can result in longer response times. This is particularly relevant for tests with unlimited response windows, which applies to all prior studies directly testing face perception in developmental prosopagnosia. Not knowing whether developmental prosopagnosics improved their performance by using more effort or atypical strategies is a widespread issue for the field, as seven of the prior studies do not report response time analyses. Most studies may have therefore underestimated the size of face perception deficits in developmental prosopagnosia, and missed the opportunity to detect the use of atypical strategies. In addition, trial lengths are rarely limited in the perception tests. Capping the available response window on trials would reduce the opportunity for developmental prosopagnosics to rely on time-consuming compensatory strategies to complete the test.

Together, these methodological issues undermine current conclusions about face perception deficits in developmental prosopagnosia. Of the studies focusing on face perception in developmental prosopagnosia, six suffered from at least two limitations and the remaining four suffered from all the limitations. Further, all the 68 reviewed studies that included a face perception test in their diagnostic battery involved two of the above

limitations, and 84% of these suffered from three or more limitations (see Appendix A.1).

Studies with more rigorous methodology are required to produce robust findings and a better understanding of face perception deficits in developmental prosopagnosia. The current study aims to do exactly that.

1.4 Current study

The overall goal of my thesis is to characterise face perception deficits in developmental prosopagnosia. The thesis has three specific aims. The first aim is to determine the severity of face perception deficits in developmental prosopagnosia. I will answer this by quantifying the magnitude of developmental prosopagnosics' face perception deficits as a group. The second aim is to characterise the scope of the face perception deficits. I will achieve this by estimating how widespread the face perception deficits are at the whole-sample and single-case levels, and by testing how much face perception deficits predict developmental prosopagnosia diagnosis. The third aim is to determine the nature of face perception deficits in developmental prosopagnosia. I will do this by exploring various features of the perception deficits and investigating the underlying mechanisms that might be driving the deficits. One important facet of this aim includes repeating the face perception tests with inverted faces to examine the role of orientation-sensitive mechanisms in the face perception deficits.

To successfully achieve these three specific aims, I designed a study that overcomes the four major methodological limitations of past research. First, I avoid circular analysis by retesting face memory using the Australian Cambridge Face Memory Test (CFMT-Aus; McKone et al., 2011). The CFMT-Aus format is identical to the CFMT except it uses different individual faces, making it an ideal test for remeasuring face memory under all the same testing formats. Matched formats means that any variations in performance on the first and second face memory tests cannot be attributed to extraneous test differences. Second, I

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

enhance statistical power by collecting data from a developmental prosopagnosic sample ten-times greater ($n = 108$) than the mode of previous research ($n = 11$). This means I will be able to precisely estimate the severity and scope of the perception deficits in developmental prosopagnosia and detect any potential subgroupings. Third, I assess face perception with three standard, widely-used face perception tests: the Benton Face Recognition Test (BFRT; Benton et al., 1983), the Glasgow Face Matching Test (GFMT; Burton et al., 2010), and the Cambridge Face Perception Test (CFPT; Duchaine et al, 2007). These tests all measure face perception, but they present different-looking stimuli, use different testing formats, place different cognitive demands, and were designed with different theoretical considerations. Using these tests simultaneously with the same sample will allow me to draw more robust and general inferences about face perception deficits in developmental prosopagnosia. Finally, I limit trial time for all tests and report response time analyses throughout to account for atypical strategy use in developmental prosopagnosia.

Chapter 2: Method

2.1. Participants

Prosopagnosic participants. Participants with developmental prosopagnosia were recruited through the Prosopagnosia Research Centre (www.faceblind.org). All developmental prosopagnosics completed a standard developmental prosopagnosia diagnostic battery. The diagnostic battery included a self-report measure of prosopagnosia traits (the Prosopagnosia Index 20; PI20; Shah et al., 2015a), two face recognition tests (the Cambridge Face Memory Test (CFMT; Duchaine & Nakayama, 2006b) and a famous face recognition test (Duchaine & Nakayama, 2005), and a test of low/mid-level visual functions (the Leuven Perceptual Organisation Screening Test; L-POST; Torfs, et al., 2014). To qualify as developmental prosopagnosics, participants had to report significant lifelong difficulty recognising faces, score within the clinical range on the PI-20, and perform at least two standard deviations below mean scores from neurotypical control samples on both the CFMT and the famous faces test. Participants were excluded if they reported a history of brain damage or failed three or more subtests on the L-POST, which is suggestive of broader problems in visual processing beyond face recognition.

Altogether, 108 developmental prosopagnosics (77 women, 27 men, 4 other) participated in the experiment. The mean age was 39.01 years old ($SD = 9.29$ years, range = 21 – 55 years). Most were from the United States of America ($n = 66$), and the remaining were from the United Kingdom ($n = 15$), Canada, ($n = 10$), Australia ($n = 7$), Germany ($n = 6$), France, ($n = 2$), and South America ($n = 2$).

Control participants. Control participants were sourced from the online participant pool Testable Minds (www.minds.testable.org). Testable Minds is an online platform on which participants can complete behavioural studies posted by academic researchers for a small reward. Members of the public can create a Minds account and participate in studies of

their choosing. The Testable Minds participant pool undergoes an advanced verification process to guarantee all participants are human (i.e. not bots), unique (i.e. the same person cannot have multiple accounts), and reliable (i.e. provide high quality data; Rezlescu et al., 2020). Testable Minds allows researchers to rapidly collect large datasets from participants with varied demographics.

Overall, 136 age- and sex-matched participants (97 women, 39 men) were included in the control sample (group age and sex comparisons are available in Appendix B.1). I recruited more controls than developmental prosopagnosics to mitigate the effect of the control group possibly including a few undiagnosed developmental prosopagnosics. Controls were not excluded based on their face recognition skill because I aimed to compare face perception and face memory when both scores could vary in both groups. In addition, I did not screen for developmental prosopagnosia using the CFMT in controls because the study duration was already long (50-60 minutes). The control mean age was 38.57 years ($SD = 7.07$ years, range = 18 to 51 years). Most controls were from the United Kingdom ($n = 92$). The remaining controls were from the United States of America ($n = 36$), Canada ($n = 5$), Australia ($n = 2$) and the Philippines ($n = 1$).

2.2. Materials

The test battery for this study consisted of three face perception tests and one face memory test. The face perception tests were the Benton Face Recognition Test (BFRT; Benton et al., 1983), the Glasgow Face Matching Test (GFMT; Burton et al., 2010), and the Cambridge Face Perception Test (CFPT; Duchaine et al., 2007a). All three face perception tests remove face memory involvement by keeping all the stimuli on screen while participants make responses. The face memory test was the Australian Cambridge Face Recognition Test (CFMT-Aus; McKone et al., 2011). The CFMT-Aus measures face memory

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

by testing recognition for unfamiliar faces that participants learn at the beginning of the test. I describe each test in turn below.

Benton Face Recognition Test (BFRT; Benton et al., 1983). The BFRT measures face perception by requiring participants to select one or more test faces that match a target identity shown above them. The BFRT therefore captures the ability to perceive multiple instances of the same identity. The BFRT is a standardised, widely used paper-based test that was originally developed for clinical evaluations of various neuropsychological conditions (e.g., Benton, 1980; Benton et al., 1994). More recently, the BFRT has been used in developmental prosopagnosia research (e.g., Huis in't Veld, 2012; Mishra et al., 2020; for a comprehensive list see Table A1 in Appendix A). Several normative data for the BFRT have been published (Albonico et al., 2017; Benton et al., 1994; Rossion & Michel, 2018; Wang et al., 2020), including for the computer-based version (Rossion & Michel, 2018). In this study I used the computerised version of the test (Rossion & Michel, 2018), which has been validated in a large sample (300+ participants) and produced near identical scores to the original test.

Figure 2

Example Match, Lighting-change, and Viewpoint-change Trials from the BFRT

Note. Part A shows a match trial, part B shows a viewpoint change trial, and part C shows a lighting change trial.

As shown in Figure 2, each trial displays one target face above six test faces. Participants identify the matching test face/s by clicking the image. Participants have 12 seconds to finish each trial. If the 12 second limit is reached, the trial immediately ends and is counted as an incorrect response. I decided to use this time limit based on extensive piloting that showed controls completed 95% of upright trials within the time limit (described in Appendix B.2).

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

The BFRT has three phases. In the first phase (six trials; half the trials are male faces and half female faces) only one of the six test faces matches the target. The image used for the target face is identical to the test face image. In the second phase (eight trials; half male and half female faces) each trial contains three test faces that match the target identity. The target face is shown in a frontal view while the six test faces vary in their viewpoint. Finally, the third phase (eight trials; half male and half female faces) follows the same structure as the second phase except the test faces vary in lighting condition rather than viewpoint. A total score out of 54 is recorded for each participant. Chance performance is 34.6% accuracy.

The Glasgow Face Matching Test (GFMT; Burton et al., 2010). The GFMT is a face perception test that requires participants to decide whether two images show the same person or two different people. The GFMT has been standardised and normed using a sample of over 300 participants (Burton et al. 2010). A key feature of GFMT is that the two images were taken with different cameras at different times. By taking stimulus photos with different cameras, the GFMT introduces important image variabilities that aim to approximate how faces are seen in real life (Burton et al., 2010). The GFMT is an increasingly popular choice for assessing face perception, having been applied in research with typical recognisers (e.g., McCaffery et al., 2018), developmental prosopagnosia (e.g., White et al., 2017), so-called ‘super-recognisers’ (i.e. people with superior face recognition skills, usually identified by scoring 93% or more on the CFMT; e.g., Bobak et al., 2016; Robertson et al., 2016), and trained passport officers (e.g., White et al., 2014).

Each trial of the GFMT presents two faces side by side (shown in Figure 3). Participants respond whether the images show the same person or two different people by pressing ‘s’ or ‘d’ on the keyboard respectively. Trials showing the same person are referred to as within-identity trials, and trials showing two different people are called between-identity trials. Based on piloting described in Appendix B.2, trial length was limited to five seconds.

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

Pilot participants completed 97% of upright trials within the time limit. Trials in which the 5 second limit was reached immediately ended and were counted as incorrect. Altogether the GFMT includes 40 test trials; 20 within-identity and 20 between-identity trials.

Figure 3

Example Within-identity and Between-identity Trials from the GFMT

Note. Part A shows a within-identity trial (same person), and part B shows a between-identity trial (different people).

Aside from the image variation introduced by using different cameras, the GFMT images all have similar lighting and viewpoints across the faces. The between-identity trials were created by pairing images of similar faces together such that discrimination was still challenging (see Figure 3). Total accuracy is calculated for each participant and chance performance is 50%.

Cambridge Face Perception Test (CFPT; Duchaine et al., 2007a). The CFPT is a face perception test that measures participants' ability to perceive facial similarity between identities. The CFPT is one of the most widely used tests of face recognition in

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

developmental prosopagnosia, having been used in over 50 studies to date (see Table A1 in Appendix A). The CFPT has been standardised, with published norms for various countries and age brackets (Bowles et al., 2009; Duchaine et al., 2007a).

In the CFPT, participants rank line-ups of six test images based on their similarity to a target face. The test images are morphs that contain a unique identity mixed with a varying level of the target identity. The greater the level of the target identity morphed into the test face, the greater the similarity to the target face. Altogether, the six test images in each trial contain 88%, 76%, 64%, 52%, 40%, and 28% of the target identity. Test image starting position is in a fixed, random order for each trial. The target faces are presented with a $\frac{3}{4}$ profile view while the test images are presented with frontal views.

On each trial, one target face is shown above six test faces (Figure 4). Participants click and drag the test images to rank their similarity, with the most similar face being placed to the far left and the least similar face to the far right. Participants have 40 seconds per trial to sort the test faces, though can proceed to the next trial before 40 seconds has elapsed by pressing a button. If the 40 second limit is reached, participants can no longer re-order the faces and must click the *“I’m done”* button. Participants are encouraged to sort faces into the perfect order rather than until it appears *“good enough”*. There is one practice trial followed by eight test trials.

Figure 4

An Example Trial from the CFPT

Note. Test faces are shown in their original order at the beginning of the trial.

The CFPT is scored by calculating how far each test face deviates from its correct position in the sorted line-up. For example, if a test face is positioned two spaces away from its correct location, then this counts as two errors. An overall error score is computed for each participant, with chance performance being 93.3 errors. Accuracy is then calculated by subtracting the participant's overall error score from the maximum number of errors possible, and dividing by the maximum number of errors. Chance performance is 35.6% correct.

Cambridge Face Memory Test – Australian (CFMT-Aus; McKone et al., 2011). The CFMT-Aus measures face memory by having participants learn six unfamiliar target identities that they must later recognise. The CFMT-Aus has three main stages: learning, test with no-noise, and test with noise. An example of each of these stages is shown in Figure 5. In the learning phase participants are taught the six target identities. For each identity, three images are shown sequentially in a leftwards, direct, and rightwards viewpoint respectively for three seconds per image (Figure 5 part A). After an identity is shown, three test trials then follow in which the target face is presented in a line-up along with two distractor faces

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

(Figure 5 part B). Participants have unlimited time to identify the target face by pressing the corresponding keyboard number (1, 2, or 3). After learning all six identities, a review slide containing all six faces is shown for 20 seconds (Figure 5 part E).

Figure 5

Example Trials from the CFMT-Aus

Note. Arrows indicate the order of the learning and testing phases in the CFMT-Aus. Part A shows each of the three views of a target identity presented in the learning phase. In part B, learning of the target identity is immediately tested with identical images. Part C shows an exemplar trial from the first testing stage (no-noise), and Part D displays a trial from the second testing stage (noise). Immediately before each testing stage a review slide is presented (part E).

In the first test stage, participants complete 30 test trials (five presentations for each of the six target faces). In each trial, three test faces are displayed and participants identify which face is one of the six learnt target faces (Figure 5 part C). The test images are all novel

(i.e. not used in the learning phase) with varied lighting and pose. After these test trials, participants are again presented with the review slide for 20 seconds and then proceed to the test phase with noise. This test phase is identical to the prior phase except there are only 24 trials (four presentations for each target face) and Gaussian noise is applied to the novel face images (Figure 5 part D). Gaussian noise obscures fine-grained image information, and is therefore expected to increase participants' dependence on their holistic view of the face rather than specific features (McKone, Martini, & Nakayama, 2001). Participant scores are calculated out of 72. Chance level performance in the CFMT-Aus is 33%.

2.3. Design

The study used a within-subjects design. Participants completed all four tests in a randomised order. Test order was randomised to minimise any possible order effects once performance was averaged across participants. Trials within each test were presented in a fixed order, according to each test's standard procedures. Doing so allows for easier comparison with results from prior studies that use the same tests.

An important element of the study's design was that face orientation was manipulated. Immediately after each face perception test, participants completed a corresponding inverted face tests for the BFRT, GFMT, and CFPT. The inverted tests were identical to the original tests except the upright face stimuli were turned upside-down (see Figure 6). Instructions were repeated and the trials remained in the same fixed order. After each inverted test, a break screen encouraged participants to take a brief (~30 second) rest before continuing to the next test.

For each perception test, I decided to run the upright version first because my primary focus is on the upright face data. Completing the upright version first prevented any potential carry-over effects from the inverted version that might influence the upright face data. Running the upright version first also allows for simpler comparisons with previous studies,

which rarely test inverted faces. A potential limitation of this design feature is that I might underestimate the size of inversion effects in my data, if participants' inverted performance somehow benefits from their familiarity with the test format and stimuli in the upright version¹.

Figure 6

Example Inverted Face Trials for the Face Perception Tests

Note. The figure shows inverted faces from the BFRT (A), GFMT (B), and CFPT (C).

An inverted version of the CFMT-Aus was not included for three reasons. First, the motivation for including the CFMT-Aus in the current study was to avoid circular analysis of the diagnostic CFMT scores, not to investigate face memory deficits per se. Only the upright CFMT-Aus was necessary for producing the unbiased estimate of face memory impairment. Second, my focus when characterising underlying mechanisms was on face perception, not face memory. Running an inverted version of the CFMT-Aus would be important if I wished to examine the underlying mechanisms of face memory, but that is beyond the scope of my study. Third, the total test time was already long without an inverted CFMT-Aus (50-60 minutes), so not including an inverted CFMT-Aus helped keep the workload manageable for participants. After finishing the CFMT-Aus, participants saw the same break screen that followed the inverted versions of the face perception tests.

¹ In general, the CFPT typically intermixes upright and inverted trials. I therefore tested the possibility that inversion sizes were reduced in my thesis by comparing the size of CFPT inversion effect in my controls with that from a normative sample ($n = 128$, Bowles et al., 2009). My effect size is slightly smaller ($M = 20.72$, $SD = 12.43$) than theirs ($M = 26.43$, $SD = 14.41$). This analysis suggests that face inversion effects could be underestimated in my data, but only to a small extent.

2.4 Procedure

All participants completed the study online via Testable (www.testable.org). Testable provides a user-friendly experiment-building interface and online data collection. I used an online approach because it allowed me to substantially increase the sample size of the study by testing developmental prosopagnosics from across the world. Online studies have become mainstream across many areas in psychology and cognitive sciences (Buhrmester et al., 2011; Goodman et al., 2012), including in visual perception and face recognition (Germine et al., 2012; Rezlescu et al., 2017; Susilo et al., 2013). Online studies can produce high quality data comparable to laboratory data (Bridges et al., 2020; Germine et al., 2012), even for sensitive experiments that are typically run under strict laboratory conditions (Crump et al., 2013; Woods et al., 2015).

Developmental prosopagnosics were invited to participate in the study via email. The email included a brief study description and a Testable link that participants clicked to begin the study. Controls participated by selecting the study from a list of available experiments on Testable Minds. All participants gave informed consent before commencing the test battery. Upon completing the experiment, participants were linked to a debrief form. In compensation for their participation, developmental prosopagnosics received an Amazon gift card (www.amazon.com) worth ten US dollars and controls were directly paid five US dollars through the Testable platform. Developmental prosopagnosics also received a one-page summary of their individual performance. Ethical approval for this study was granted by the Victoria University of Wellington's Human Ethics Committee.

Chapter 3: Results

This chapter presents the results of my study. The chapter is broadly organised into three sections that are tied to the three specific aims of the thesis.

The first section addresses the severity of face perception deficits in developmental prosopagnosia (3.2). I achieve this by examining the magnitude of developmental prosopagnosics' impairment on the face perception tests relative to the face memory test.

In the second section, I characterise the scope of face perception deficits in developmental prosopagnosia (3.3). I begin by estimating how widespread face perception deficits are at the whole sample level (3.3.1) and the single case level (3.3.2). In addition, I then determine how much face perception deficits contribute to diagnosing developmental prosopagnosia (3.3.3).

Finally, in the third section, I address five issues regarding the nature of face perception deficits in developmental prosopagnosia (3.4). First, I test whether face perception deficits are driven by impairment in mechanisms specific for upright faces (3.4.1). Second, I test whether face perception deficits are linked to reduced use of perceptual mechanisms that are engaged in typical recognisers (3.4.2). Third, I explore whether face perception deficits are related to impairment in mechanisms for within-identity perception versus between-identity perception (3.4.3). Fourth, I test whether face perception deficits are disproportionately related to lighting versus viewpoint variations in face images (3.4.4). Lastly, I investigate the extent to which face perception deficits result from perceptual processes as opposed to other non-perceptual components necessary for completing the perception tests (3.4.5).

3.1. Analytic approach and data pre-processing

Statistical analyses were conducted using Jamovi version 1.6.3 (Jamovi Project, 2017). I used R (R Core Team, 2017) to generate all the figures and run the shift function

analyses. I used *fast-dm* software (Voss et al., 2015) to execute drift diffusion modelling. Test accuracy was converted from raw scores to percentage correct for all tests. Raw mean and median response times are reported in seconds. Mean response times are presented in text and median response times are available in Appendix B.3 (all results remained the same for mean and median response time analyses unless otherwise indicated). I used Bonferroni corrected p-values for multiple comparisons and applied the Greenhouse-Geisser correction whenever the assumption of sphericity was violated.

One developmental prosopagnosic's upright GFMT data was excluded for not responding and thus timing-out on all trials. Their remaining data was retained as they responded normally throughout the rest of the test battery.

Technical errors on Testable meant 25 developmental prosopagnosics were missing data for the first six trials (match trials) of the inverted BFRT. I used two approaches to address this issue. The first approach involved discarding the six match trials from all participants when comparing upright and inverted performance on the BFRT. I use this as the main approach in this chapter because the BFRT match trials use identical images for the target and test stimuli, which means the trials can be completed by image-matching and may not involve face processing (Rossion & Michel, 2018). The second approach involved using hot deck data imputation. Hot deck imputation replaces a case's missing values with values from a randomly selected, maximally similar individual based on other available data (Joensuu & Bankhofer, 2012). For each of the 25 developmental prosopagnosics' data, the missing score was replaced with the total inverted match score from a randomly selected prosopagnosic that scored within ± 1 points on the rest of the inverted BFRT trials as the prosopagnosic with missing data. This approach is reported in Appendix B.4. Both approaches led to the same pattern of results for all relevant analyses.

3.2 The severity of face perception deficits in developmental prosopagnosia

First, I estimated the severity of face perception deficits in developmental prosopagnosia. To do so, I focussed on the magnitude of face perception deficits relative to face memory deficits. This comparison is crucial for determining how much face perception impairment contributes to developmental prosopagnosia. I achieved this by comparing face perception scores from the three tests to the CFMT-Aus scores.

Before analysing accuracy, I checked response times for potential speed-accuracy trade-offs (Fitts, 1954). This is important for several reasons. Developmental prosopagnosics could show lower accuracy because they are less engaged with the tests, which would result in faster response times. Developmental prosopagnosics could also show similar accuracy to controls because they are more engaged with the tests or use atypical strategies, which may result in slower response times. Indeed, longer response times have been taken to indicate atypical face processing in developmental prosopagnosia (Behrmann et al., 2005; Dalrymple et al., 2014b). To check for these possibilities, I ran a two (group: prosopagnosic, control) by four (test: BFRT, GFMT, CFPT, CFMT-Aus) mixed-measures analysis of variance (ANOVA) on participants' raw mean response times. Group was the between-subjects factor and test was the within-subjects factor. Descriptive statistics are displayed in Table 1. There were main effects of group ($F(1, 241) = 85.78, p < .001, \eta^2_p = .26$) and test ($F(1.05, 253.05) = 1943.10, p < .001, \eta^2_p = .89$), which were qualified by a significant interaction ($F(1.05, 253.05) = 48.22, p < .001, \eta^2_p = .17$). Post hoc analyses revealed that developmental prosopagnosics took significantly longer to respond on the CFPT compared to controls ($t(939.45) = 15.06, p < .001$). There were no significant group differences in response times for the BFRT ($t(939.45) = 3.13, p = .050$), GFMT ($t(939.45) = 0.90, p = 1.000$), and CFMT-

Aus ($t(939.45) = 1.86, p = 1.000$), although developmental prosopagnosics tended to show longer response times than controls across all tests (see Table 1)².

Table 1

Developmental Prosopagnosic and Control Mean Response Times on the Face Perception and Memory Tests

Test	BFRT		GFMT		CFPT		CFMT-Aus	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Control	6.48	1.10	2.20	0.53	25.67	9.17	2.90	0.96
Prosopagnosic	8.44	1.36	2.76	0.58	34.91	10.05	4.05	1.73

Note. Response times are presented in seconds (s).

This analysis is consistent with the view that developmental prosopagnosics' may have improved their accuracy by taking longer to respond. I therefore used mixed-measures analyses of covariance (ANCOVAs) to analyse accuracy in subsequent sections, with response time as a covariate. Using response time as a covariate mitigates the effect of longer response times on accuracy. However, for completion and to allow for simpler comparisons with previous studies that did not analyse response time, I also ran basic ANOVAs for all ANCOVAs. All findings remain the same regardless of approach (see Appendix B.5 for ANOVA analyses).

To compare face perception deficits and face memory deficits, I ran a two (group: prosopagnosics, controls) by four (test: BFRT, GFMT, CFPT, CFMT-Aus) mixed-measures ANCOVA on accuracy (Figure 7). Test was the within-subjects factor, group was the between-subjects factor, and response time was the covariate. All descriptive statistics are available in Table 2. There were significant main effects of test ($F(3, 966) = 8.74, p < .001, \eta^2_p = .03$) and group ($F(1, 966) = 451.83, p < .001, \eta^2_p = .32$), and a significant interaction

² The same analysis with median response time found that developmental prosopagnosics took also took longer on the BFRT (details provided in Appendix B.3).

between test and group ($F(3, 966) = 3.66, p = .012, \eta^2_p = .01$). Follow-up post-hocs showed that developmental prosopagnosics performed worse than controls on the CFMT-Aus ($t(966) = 13.12, p < .001, d = 1.69$), BFRT ($t(966) = 12.51, p < .001, d = 1.62$), GFMT ($t(966) = 10.02, p < .001, d = 1.39$), and CFPT ($t(966) = 8.18, p < .001, d = 1.17$). All three effect sizes for developmental prosopagnosics' perception test impairments were large (between 1.17 – 1.62). BFRT and GFMT effect sizes (1.62 and 1.39 respectively) were similar in magnitude to the CFMT-Aus (1.69), while CFPT effect size is relatively smaller (1.17).

Figure 7

Developmental Prosopagnosic and Control Accuracy on the Upright Face Perception and Memory Tests

Note. Accuracy on the perception tests (BFRT, GFMT, and CFPT) and memory test (CFMT-Aus) is shown in percentage correct. Diamonds represent the group mean for each test and points show the individual data. Dashed lines depict chance performance.

Table 2

Developmental Prosopagnosic and Control Accuracy on the Face Perception and Memory Tests

Test	BFRT		GFMT		CFPT		CFMT-Aus	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Control	75.92	9.46	74.12	11.63	70.93	11.26	72.28	13.15
Prosopagnosic	58.20	11.94	59.79	8.78	59.47	11.43	53.63	10.72

Note. Accuracies are presented in percentage correct (%).

The significant interaction suggests that the group differences differ across tests, which is also reflected by the effect sizes. To explore this interaction, I ran three follow-up mixed-measures ANCOVAs comparing group accuracy on the CFMT-Aus against each of the three face perception tests, while adjusting for response times as a covariate. I found an interaction for the CFMT-Aus and CFPT contrast ($F(1, 483) = 5.61, p = .018, \eta^2_p = .01$) and the CFMT-Aus and GFMT contrast ($F(1, 482) = 4.05, p = .045, \eta^2_p = .01$), but not for the CFMT-Aus and BFRT contrast ($F(1, 483) = 1.34, p = .248, \eta^2_p < .01$). This finding indicates that developmental prosopagnosics had equivalently poor performance on the CFMT-Aus and BFRT, but they performed slightly better on the CFPT and GFMT. That said, the effect sizes for the developmental prosopagnosics' deficits on the CFPT and GFMT were still large at 70% and 82% the size of the CFMT-Aus deficits, suggesting severe impairments with face perception.

Altogether, developmental prosopagnosics showed large and consistent face perception deficits across all tests. The sizes of the deficits varied across tests, but they were all substantial, ranging from 70-100% of the size of face memory deficits. This result provides robust evidence that face perception deficits in developmental prosopagnosia are much more severe than previously thought.

3.3 The scope of face perception deficits in developmental prosopagnosia

Having shown that face perception deficits in developmental prosopagnosia are severe, I next used three approaches to address the second aim of my thesis, namely the scope of face perception deficits in developmental prosopagnosia.

3.3.1 Whole-sample analysis of scope

First, I examined how widespread face perception deficits are throughout the entire developmental prosopagnosic sample. Such an analysis is important because face perception deficits could either be driven by the whole developmental prosopagnosic group or instead by subgroups within the sample. To address this issue, I used Wilcox's (1995) shift function, which is a graphical and inferential tool for characterising differences between two independent distributions. Shift functions work by separating each distribution into deciles and describing how deciles in one distribution must shift to match the other distribution (Rousselet et al., 2017). As I had already found severe face perception deficits, developmental prosopagnosic deciles would generally shift toward lower scores compared to control deciles. The crucial issue is about the *pattern* of shifts across the entire sample. Widespread face perception deficits would result in a uniform, consistent shift across all developmental prosopagnosic deciles. Face perception deficits circumscribed to a subgroup of developmental prosopagnosics would result in a shift for some deciles, but not others.

Following standard practices (Rousselet et al., 2017; Wilcox & Rousselet, 2018), I applied the Harrell-Davis quantile estimator (Harrell & Davis, 1982) to calculate the deciles for each group on each face perception test. The Harrell-Davis method of estimating deciles uses the weighted average of a sample's order statistics – i.e. the sample values placed in ascending order – to calculate deciles (Harrell & Davis, 1982). The decile boundaries for each group on each test are denoted by vertical black lines in Figure 8 part A. To assess how much the developmental prosopagnosic distribution would need to shift to match the control

distribution for each test, I plotted group decile differences (developmental prosopagnosic deciles – control deciles) on the y-axis and developmental prosopagnosic deciles on the x-axis (Figure 8 part B). The more negative the y-axis value, the larger the deficits in the developmental prosopagnosic decile compared to the control decile. I also calculated the 95% confidence interval of the decile differences using a bootstrap estimation of the standard error of the deciles.

Figure 8 shows that the developmental prosopagnosic distributions are uniformly shifted toward poorer scores compared to the control distributions across all deciles and tests. None of the decile confidence intervals include zero, indicating that the decile differences are considered significant (Rousselet et al., 2017). Looking closer at the graphs, developmental prosopagnosics' performance on the BFRT and the GFMT show the largest shifts, with the majority of deciles being 15% lower than the control deciles, whereas their deciles for the CFPT were about 10% lower than control deciles. There might be a possible subgroup of extremely poor-performing developmental prosopagnosics on the BFRT, where decile differences appear larger for the lowest two deciles (about 20%) compared to the other deciles (about 15%). This pattern should be treated with caution, however, as the 95% confidence intervals for the lowest two deciles are double the size of the remaining deciles.

Figure 8

Perception Test Shift Functions and Decile Differences for Developmental Prosopagnosics and Controls

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

Note. Shift function distributions are displayed in part A for each face perception test. Developmental prosopagnosic (proso.) and control deciles are shown by the vertical black lines. The thickest black line denotes the median for each group. Points show individual performance. The decile difference values between developmental prosopagnosic and control performance for the lowest and highest performing deciles are overlaid on each graph. Finally, the grey boxes correspond to the developmental prosopagnosic deciles used for the x-axis in part B for each face perception test. All decile differences (developmental prosopagnosic decile minus the control decile) are shown in part B, where negative values indicate poorer accuracy for developmental prosopagnosics compared to controls for the given decile. The darkest circle represents the median decile, with increasing lightness representing increasingly higher/lower deciles.

Altogether, the result of the shift function analysis is clear. Across all tests, the whole developmental prosopagnosic distribution was shifted down relative to the control distribution. This finding suggests that face perception deficits in developmental prosopagnosia are not only severe, but that they are also widespread across the entire sample.

3.3.2 Single-case analysis of scope

Another way to address the scope of face perception deficits in developmental prosopagnosia is by using single-case statistical approaches that are commonly used in neuropsychology (Crawford & Howell, 1998; Crawford & Garthwaite, 2006). In contrast to the shift function approach which compares subgroups or deciles within the sample, the single case analyses isolate each developmental prosopagnosic and compare their individual performance against performance of the whole control group. In this analysis I began by converting each developmental prosopagnosic's upright accuracy on each face perception test into z-scores based on their performance relative to the control mean and standard deviation. I then counted the number of developmental prosopagnosics that met the common cut-off for impaired performance, namely a z-score of -2 or below, on one or more tests.

Figure 9*Developmental Prosopagnosic Z-scores for the Face Perception Tests*

Note. Negative z-scores indicate impaired performance. Diamonds represent the group mean for each test and points show the individual data. The dashed line depicts the z-score cut-off for clinically severe impairment.

The number of developmental prosopagnosics who were impaired on the BFRT, GFMT, and CFPT separately were 35, 16, and 19 (see Figure 9). The number of developmental prosopagnosics who were impaired on two tests was 20. Only one developmental prosopagnosic was impaired on all three tests. These numbers may seem small compared to the total sample size of 108. However, the same analysis performed with the CFMT-Aus showed that only 28 developmental prosopagnosics had impaired performance. This number can be considered a baseline, because using all 108 developmental prosopagnosics as the baseline for assessing single-case results would be committing circular analysis. The single-case approach thus suggests that the proportion of developmental

prosopagnosics with face perception deficits and those with face memory deficits are similar.

This result is consistent with the view that face perception deficits in developmental prosopagnosia are widespread.

One limitation of the above approach is that it does not consider the potential relationships between the tests, because each developmental prosopagnosic z-score for a test is compared to control performance in isolation from the other tests. To address this issue, I used a multivariate single-case comparison procedure based on Hotelling's (1931) T^2 statistic (Huizenga et al., 2007). This multivariate approach tests whether an individual performs outside the control range across multiple tests considered collectively. This is a conservative estimate of individual clinical impairment, as the individual must perform atypically across all tests to differ significantly from controls. I performed the multivariate comparison on all developmental prosopagnosics except the case with missing upright GFMT data, because the multivariate comparison required complete data from all tests. The result is presented in Table B1 in Appendix B.6. In total, 34 of the 107 prosopagnosics (32%) performed outside the range on all three tests.

Overall, both the univariate and multivariate single-case results suggest that face perception deficits were present in around just under one third of the developmental prosopagnosic sample. Considering that a similar number of individuals were impaired with face memory (CFMT-Aus), this finding shows that face perception deficits are widespread in developmental prosopagnosia, consistent with the results from the shift function approach.

3.3.3 Face perception deficits and developmental prosopagnosia diagnosis

To further establish the scope of face perception deficits, I next investigated how much face perception deficits contribute to the diagnosis of developmental prosopagnosia. This issue is important because developmental prosopagnosia is currently diagnosed using tests with heavy face memory demands. Yet if face perception deficits are highly

characteristic of developmental prosopagnosia diagnosis, this would mean our understanding of developmental prosopagnosia might need reconceptualising to include face perception deficits as a core diagnostic feature.

To address this issue, I used logistic hierarchical regression to test how well face perception scores from the BFRT, GFMT, and CFPT predicted group membership (prosopagnosic, control), and if so, whether adding face memory scores from the CFMT-Aus mattered. If face perception deficits are central to developmental prosopagnosia diagnosis, then face perception scores would be highly predictive of group membership and the addition of face memory scores would not add much predictive ability. But if face perception deficits are not central to developmental prosopagnosia diagnosis, then face perception scores would not be highly predictive of group membership.

To run the logistic hierarchical regression, I first calculated a face perception ‘composite score’ for each participant by averaging their BFRT, GFMT, and CFPT upright accuracy. I did this to simplify the interpretation of the odds ratios and group classification, as my primary focus was the predictive ability of face perception overall on diagnosis, rather than each test’s individual predictive ability³. I entered the face perception composite score as the only predictor in step one of the hierarchical regression model. The step one model was significant ($X^2(1) = 123.82, p < .001$), and the composite face perception score significantly predicted group membership ($B = 0.19, p < .001$). The odds ratio for the composite scores was 0.83 (95% CI = 0.79-0.87), meaning that with each one-point increase in the composite perception score, a participant was 17% less likely to be a developmental prosopagnosic. The model correctly classified group membership for around 80% of participants, showing high predictability (see Table 3 for the full classification table). To further test the predictive fit, I

³ For completion, I provide the regression with each perception test added to the model separately in Appendix B.7. The main finding that face perception deficits predict prosopagnosia diagnosis remained the same.

checked the area under the curve (AUC) of the model's ROC curve and the pseudo R^2 . The AUC evaluates how well the model classifies positive and negative outcomes, with positive values indicating better performance (possible range = 0.5 - 1). The AUC was good at 0.88, and the Nagelkerke pseudo R^2 was adequate at 0.53.

Table 3

Group Classification in the Face Perception Only Model

		Predicted		% correct
		Control	Prosopagnosic	
Observed	Control	113	23	83.08
	Prosopagnosic	24	84	77.78
Overall % correct				80.74

In step two of the model, I added CFMT-Aus score as another predictor. The step two model was significant ($X^2(2) = 142.44, p < .001$), and the difference between the step one and step two models was also significant ($X^2(1) = 18.60, p < .001$), meaning the step two model accounted for additional variance beyond the step one model. Both composite score ($B = 0.13, p < .001$; 95% CI = 0.83-0.92) and CFMT-Aus score ($B = 0.07, p < .001$; 95% CI = 0.90-0.96) were significant predictors. The odds ratio for composite face perception score was 0.87, meaning a participant was 13% less likely to be developmental prosopagnosic with each one-point increase in the composite score. The odds ratio for CFMT-Aus score was 0.93, meaning that with each one-point increase on CFMT-Aus score, participants were 7% less likely to be developmental prosopagnosic. However, the step two model only classified around 82% of participants correctly (Table 4), which was not much higher than the step one model. The AUC was strong at 0.90, and the Nagelkerke pseudo R^2 was adequate at 0.59.

⁴ Pseudo R^2 methods of evaluating model fit should be considered with high caution. They attempt to estimate explained variance as done in linear regression models, but are easily biased and can arrive at highly different values based on pseudo estimation method. I therefore provide two other methods of estimating model fit: classification success and AUC.

Table 4*Group Classification in the Face Perception and Face Memory Model*

		Predicted		% correct
		Control	Prosopagnosic	
Observed	Control	114	22	83.82
	Prosopagnosic	21	87	80.56
Overall % correct				82.38

The regression results show that face perception scores were highly predictive of developmental prosopagnosia diagnosis, but also that adding a face memory score mattered. However, the contribution of the face memory score was notably small. This was best illustrated by the odds ratio analysis, which showed that a one-point increase in face perception ability decreased the likelihood of being developmental prosopagnosic by double compared to a one-point increase in CFMT-Aus performance. In other words, face perception was twice as predictive of developmental prosopagnosia than face memory test performance. In addition, the CFMT-Aus does not isolate face memory, meaning the addition of the face memory score reflected the addition of both face memory and face perception. Overall, this analysis demonstrates that face perception deficits in developmental prosopagnosia are not only severe and widespread, but they are also highly predictive for diagnosing the condition.

3.4 The nature of face perception deficits in developmental prosopagnosia

The results so far have shown that face perception deficits in developmental prosopagnosia are severe, widespread, and highly predictive for diagnosis. In the following sections, I focussed on addressing a series of theoretical issues related to the third aim of my thesis, namely characterising the nature of face perception deficits.

3.4.1 Upright-sensitivity of face perception deficits

A hallmark of face perception is sensitivity to the upright orientation (DeGutis et al., 2013a; Richler & Gauthier, 2014; Tanaka & Gordon, 2011; Wang et al., 2012). This is

classically demonstrated by the face inversion effect – the much better performance with upright faces than with upside-down faces (Yin, 1969). The inversion effect is widely used as an index of typical processing of upright faces, and reduced inversion effects in developmental prosopagnosia have been reported (Behrmann et al., 2005; Biotti et al., 2019; de Gelder & Romke, 2000; Duchaine & Nakayama, 2007; Farah et al., 1995; Le Grand et al., 2006; Schmalzl et al., 2011). This finding has led to suggestions that developmental prosopagnosia is linked to impairments of mechanisms specialised for upright faces.

To examine this issue with my data, I compared the size of inversion effects in developmental prosopagnosics and controls across the three tests. I measured the inversion effect by subtracting accuracy from the inverted trials from accuracy from the upright trials. Response time analyses showed no group difference in any test (available in Appendix B.8), so I ran a two (group: prosopagnosic, control) by three (test: BFRT, GFMT, CFPT) mixed-measures ANOVA on accuracy⁵. Group was the between-subjects factor and test was the within-subjects factor. There was a main effect of group ($F(1, 241) = 141.29, p < .001, \eta^2_p = .37$), such that developmental prosopagnosics had significantly smaller inversion effect sizes ($M = 6.32\%, SD = 10.35\%$) than controls ($M = 16.83\%, SD = 11.38\%$)⁶. There was no interaction between group and test ($F(2, 482) = 0.60, p = .552, \eta^2_p < .01$). This result shows that, regardless of test, developmental prosopagnosics had reduced inversion effects compared to controls (visible in Figure 10).

⁵ Given some prosopagnosics had missing data for the inverted BFRT match trials, only the viewpoint and lighting test trials were included for the BFRT inversion effect analyses.

⁶ There was also a main effect of test ($F(2, 482) = 33.95, p < .001, \eta^2_p = .12$), but this result is not relevant to the purpose of the analysis and is not further examined.

Figure 10*Developmental Prosopagnosic and Control Inversion Effects on the Face Perception Tests*

Note. Accuracy inversion effect sizes (average upright accuracy – average inverted accuracy) are shown in percentage correct. Diamonds represent the group mean for each test and points show the individual data.

Smaller inversion effects in developmental prosopagnosia would only imply impaired upright-specific processing if they result from a larger reduction of accuracy in the upright condition than in the inverted condition. To test whether this is the case, I ran a two (group: prosopagnosic, control) by two (orientation: upright, inverted) by three (test: BFRT, GFMT, CFPT) mixed measures ANOVA on the accuracy data. Group was the between subjects factor, and orientation and test were within subjects factors. There were significant main effects for group ($F(1, 241) = 109.89, p < .001, \eta^2_p = .31$), orientation ($F(1, 482) = 691.61, p < .001, \eta^2_p = .74$), and test ($F(1.93, 475.06) = 33.52, p < .001, \eta^2_p = .12$). The three way interaction between group, orientation, and test was not significant ($F(2, 482) = 0.60, p = .551, \eta^2_p < .01$), but there was an interaction between group and orientation ($F(1, 482) =$

141.30, $p < .001$, $\eta^2_p = .37$)⁷. Post hoc analyses showed that developmental prosopagnosics had much lower upright accuracy ($M = 59.15\%$, $SD = 10.80\%$) than controls ($M = 73.65\%$, $SD = 10.99\%$; $t(355.65) = 14.71$, $p < .001$, $d = 1.33$), but only slightly lower inverted accuracy ($M = 52.83\%$, $SD = 10.99\%$) than controls ($M = 56.82\%$, $SD = 10.73\%$; $t(359.69) = 4.03$, $p < .001$, $d = 0.37$). This results confirms that the smaller inversion effects in developmental prosopagnosics are driven by a larger reduction of upright accuracy compared to inverted accuracy.

A possible explanation for the smaller reduction in inverted accuracy compared to the upright accuracy is floor effects in the inverted condition. To rule out this explanation I did two analyses. First, I ran one-sample t-tests to check whether developmental prosopagnosics' inverted accuracy on the three tests was above floor. Developmental prosopagnosics' inverted accuracy was significantly above floor for the BFRT ($t(107) = 16.76$, $p < .001$, $d = 1.61$), GFMT ($t(107) = 8.17$, $p < .001$, $d = 0.79$), and CFPT ($M = 49.96\%$, $SD = 11.69\%$; $t(107) = 13.30$, $p < .001$, $d = 1.28$). Second, I repeated the main analysis in two subsets of data where developmental prosopagnosics' inverted accuracy was clearly above floor. The first subset of data only included the easiest 50% of trials for each perception test as determined by developmental prosopagnosics' inverted performance. The second subset of data only included the best performing 50% of developmental prosopagnosics for each inverted perception test. Both subsets of data produce the same pattern of results as the original data (see Appendix B.9). These analyses together rule out floor effects as a potential explanation.

In sum, developmental prosopagnosics' smaller inversion effects suggest that their deficits are driven by impairments to mechanisms specialised for processing upright faces. This result extends prior results by providing robust evidence of reduced inversion effects

⁷ There were also significant interactions between group and test ($F(2, 482) = 15.33$, $p < .001$, $\eta^2_p = .06$) and test and orientation ($F(2, 482) = 33.96$, $p < .001$, $\eta^2_p = .12$). These were not the purpose of the analysis, however, and are not further discussed.

across multiple measures, thus clarifying an important aspect of the nature of face perception deficits in developmental prosopagnosia.

3.4.2 Associations and dissociations between face perception tests

Another way to examine the nature of the deficits is by looking at the associations and dissociations between performance on different face perception tests. While the tests have different formats, stimuli, and cognitive demands, they were all designed to consistently measure face perception, and thus they tend to show moderate to strong associations in typical participants (Mishra et al., 2020). The extent to which the tests engage common and typical mechanisms for face perception can thus be indexed by the size of correlations between the tests. If developmental prosopagnosics are using inconsistent and atypical approaches for face perception, then the correlations between the tests should be weaker in developmental prosopagnosics than in controls.

To address this issue, I ran Pearson's correlations for all possible pairs of tests for developmental prosopagnosics and controls separately, focusing on upright data (Figure 11 and Table 5). As expected, the correlations for controls were moderate to large (range 0.53 – 0.58). In contrast, the correlations for developmental prosopagnosics were small ($r = 0.29 - 0.30$). A Fisher's z transformation test confirmed that the correlations for developmental prosopagnosics were significantly smaller than those for controls across all pairs (BFRT/GFMT, $z = 2.79$, $p = .003$; BFRT/CFPT, $z = 2.26$, $p = .012$; GFMT/CFPT, $z = 2.15$, $p = .016$). Critically, the smaller correlations for developmental prosopagnosics do not appear to be driven by smaller variances in their data (see Table 2 in section 3.2).

Figure 11

Correlations Between Upright Face Perception Tests for Developmental Prosopagnosics and Controls

Note. Developmental prosopagnosics and controls are depicted using circles and crosses respectively. 95% confidence intervals for each correlation are shown by the shaded bands.

Table 5

Pearson's r and p -values for Upright Test Correlations for Developmental Prosopagnosics and Controls

Test			BFRT	GFMT	CFPT
Control	BFRT	r	-		
		p	-		
	GFMT	r	0.58	-	
		p	< .001	-	
	CFPT	r	0.54	0.53	-
		p	< .001	< .001	-
Prosopagnosic	BFRT	r	-		
		p	-		
	GFMT	r	0.29	-	
		p	.003	-	
	CFPT	r	0.30	0.30	-
		p	.002	.002	-

The moderate to large correlations between the tests in controls indicate that the tests recruit common and consistent mechanisms for face perception⁸. The smaller correlations in developmental prosopagnosics indicate the use of some inconsistent and likely aberrant mechanisms, which is expected when typical mechanisms are not working properly. In sum, this analysis suggests that face perception deficits in developmental prosopagnosia are driven by the use of atypical mechanisms for face perception.

3.4.3 Within-identity perception versus between-identity perception

Face perception tests that require participants to decide whether two face images show the same person or different people – such as the GFMT – measure two distinct aspects of face perception (Jenkins et al., 2011). Trials that show images of the same person require within-identity face perception. Participants must be able to match diagnostic information across the images and ignore differences in viewpoint, styling, lighting, and so on. In contrast, trials that show images of different people engage between-identity face perception. Participants must be able to tell faces apart to successfully distinguish between people. These different aspects of face perception have been dissociated in typical participants (Jenkins et al., 2011; Megreya & Burton, 2007; Robins et al., 2018; Ritchie & Burton, 2017). In this section I investigated whether face perception deficits in developmental prosopagnosia were disproportionate to within-identity or between-identity trials in the GFMT.

Because developmental prosopagnosics tended to take longer to respond on the GFMT than controls (see section 3.2), I ran a two (group: prosopagnosic, control) by two (trial type: within, between) mixed measures ANCOVA on the upright accuracy data, with response time as a covariate⁹. There were significant main effects of group ($F(1, 481) =$

⁸ I also checked the correlations for the inverted condition in controls. Controls had significant but small correlations for all inverted test pairs (BFRT/GFMT: $r = 0.33, p < .001$; BFRT/CFPT: $r = 0.23, p = .007$; GFMT/CFPT: $r = 0.19, p = .027$). These are all weaker than the correlations in the upright condition, suggesting that the correlations in the upright condition are capturing common mechanisms for face perception.

⁹ The ANOVA version of the analysis for GFMT same versus different trials is available in Appendix B.5

40.51, $p < .001$, $\eta^2_p = .04$) and trial type ($F(1, 481) = 22.48$, $p < .001$, $\eta^2_p = .04$), that were qualified by a significant interaction ($F(1, 481) = 39.64$, $p < .001$, $\eta^2_p = .07$). Follow-up tests revealed that developmental prosopagnosics scored poorly on within-identity trials ($M = 51.12\%$, $SD = 18.90\%$) compared to controls ($M = 75.55\%$, $SD = 16.80\%$; $t(481) = 8.86$, $p < .001$, $d = 1.22$). But on between-identity trials, developmental prosopagnosics showed comparable performance ($M = 68.46\%$, $SD = 17.00\%$) to controls ($M = 72.68\%$, $SD = 17.01\%$; $t(481) = 0.53$, $p = 1.000$, $d = 0.07$). This result suggests that face perception deficits in developmental prosopagnosia are linked to specific problems with within-identity perception (see Figure 12).

Figure 12

Developmental Prosopagnosic and Control Accuracy on the GFMT Within-identity and Between-identity Trials

Note. Accuracy is given in percentage correct. Diamonds represent the group mean and points show the individual data.

An alternative account for this finding is response bias. If developmental prosopagnosics were generally biased to respond ‘different’ rather than ‘same’, this would explain their poor performance on within-identity trials. This account does not seem likely, however, because such a bias would also result in developmental prosopagnosic accuracy for between-identity trials being close to ceiling (or at least higher than controls), which was not the case. Nevertheless, I used the signal detection analysis to address the response bias issue more formally¹⁰. Response bias or criterion is calculated by adding the inverse normal distributions of hits and false alarms and multiplying by 0.5 (Macmillan & Creelman, 1990). Hits reflect when participants respond different to between-identity trials, and false alarms reflect when participants respond different to within-identity trials. A positive criterion value indicates a “same” bias, while a negative criterion value reflects a “different” bias. An independent samples t-test confirmed that developmental prosopagnosics had significantly more negative criterions ($M = -0.26$, $SD = 0.47$) than controls ($M = 0.05$, $SD = 0.43$; $t(216.92) = 5.31$, $p < .001$, $d = 0.69$). One sample t-tests further showed that developmental prosopagnosics’ criterion significantly differed from zero ($t(106) = 5.75$, $p < .001$, $d = 0.56$), which was not the case for controls ($t(135) = 1.34$, $p = .183$, $d = 0.12$).

How can developmental prosopagnosics show a bias for responding ‘different’ when their accuracy on between-identity trials is not better than controls? One possibility is that developmental prosopagnosics were biased only on within-identity trials and not between-identity trials. This selective bias might explain why developmental prosopagnosic accuracy on different-identity trials was not better than controls. If this is the case, it demonstrates that developmental prosopagnosics still have specific difficulty with within-identity trials, because they did not show bias on between-identity trials. Overall, the main result from this

¹⁰ To accompany criterion, I also calculated d' . D' is a sensitivity measure that is also derived from hits and false alarms. The d' results are available in Appendix B.10.

section remains, in that face perception deficits in developmental prosopagnosia may be selective to perceiving within-identity information rather than between-identity information. The result is preliminary, however, because the reliability of this analysis is constrained by the small number of trials in the within- and between-identity conditions ($n = 20$ in each).

3.4.4 Face perception deficits across image changes by lighting and viewpoint

Another issue concerning the nature of face perception deficits in developmental prosopagnosia is whether the deficits are associated with the complexity of changes in face images, such as those produced by lighting and viewpoint variations. Lighting variations introduce more complex changes in face shape, surface texture, and the three-dimensional structure compared to changes in face viewpoint, meaning they require more complex perceptual operations (Adini et al., 1995; Favelle et al., 2017; Favelle et al., 2011; O'Toole et al., 2006). Indeed, past studies using the BFRT showed that typical participants find lighting changes more difficult than viewpoint changes (Liu et al., 2009; Rossion & Michel et al., 2018). In this section I analysed whether face perception deficits in developmental prosopagnosia were also disproportionate for lighting change trials compared to viewpoint change trials. Such a result would suggest face perception deficits in developmental prosopagnosia arise mainly when perceptual processing operations are more complex.

According to the general pattern of developmental prosopagnosics taking longer to respond (see section 3.2), I ran a two (group: prosopagnosic, control) by two (trial type: lighting change, viewpoint change) mixed-measures ANCOVA on BFRT upright accuracy data, with group as the between-subjects factor, trial type as the within-subjects factor, and response time as the covariate¹¹. The result is shown in Figure 13. There was a main effect of group ($F(1, 483) = 112.22, p < .001, \eta^2_p = .19$), where developmental prosopagnosics ($M =$

¹¹ Match trials were not included because of the low trial number (six match trials per orientation). The ANOVA version of this analysis is available in Appendix B.5.

58.2%, $SD = 13.70\%$) scored much worse than controls ($M = 75.92\%$, $SD = 11.14\%$). There was also a main effect of trial type ($F(1, 483) = 8.24$, $p = .004$, $\eta^2_p = .02$), with higher accuracy for lighting trials ($M = 71.00\%$, $SD = 13.41\%$) than viewpoint trials ($M = 65.15\%$, $SD = 16.22\%$). The interaction between group and trial type was nearing significance ($F(1, 483) = 3.71$, $p = .055$, $\eta^2_p = .01$) in the direction that developmental prosopagnosics had more difficulty on viewpoint trials than lighting trials, but the interaction effect size is very small, suggesting that any interaction is unlikely to be meaningful.

Figure 13

Developmental Prosopagnosic and Control Accuracy on the BFRT Lighting and Viewpoint Trials

Note. Boxplots show accuracy in percentage correct. Diamonds represent the group mean for each condition. Points depict the individual data.

Overall, this analysis suggests that developmental prosopagnosics' face perception deficits are not linked to the complexity of changes in face images. However, the reliability of this result is limited and must be treated as preliminary, because there were only 16 trials in each condition. Interestingly, my control sample found viewpoint changes more difficult than lighting changes, which is the reverse of prior studies (Liu et al., 2009; Rossion & Michel, 2018).

3.4.5 Components of performance in the face perception tests

Finally, I examined the extent to which face perception deficits in developmental prosopagnosia are tied to impairments in perceptual processes themselves, rather than differences in other cognitive and motor processes demanded by the face perception tests. I addressed this issue in a modelling analysis with drift diffusion modelling (DDM, Ratcliff, 1978). DDM is a framework for analysing different components of perceptual performance (Ratcliff, 1978; Ratcliff et al., 2016; Voss & Voss, 2007; Voss et al., 2015). DDM has been widely used in many areas of visual perception involving brightness, contrast, motion, orientation, letter, objects, and faces (Palmer et al., 2005; Nosofsky et al., 2011; Nosofsky & Palmeri, 1997; Ratcliff & McKoon, 2008; Ratcliff et al., 2003; Thapar et al., 2003; Thornton & Gilden, 2007; Char et al., 2019; Meyer et al., 2019), and to compare performance in clinical groups (e.g., Powell et al., 2019; Zeguers et al., 2011).

DDM works by treating perceptual performance in binary (two-choice) tests as a noisy and continuous accumulation of evidence over time toward one of two decision boundaries (Ratcliff, 1978). The model uses accuracy and response time data to separate and model three main components of performance. The first is drift rate (v), being the efficiency and speed at which perceptual evidence is accumulated. Larger drift rates represent more efficient evidence accumulation. The second is threshold (a), which refers to the separation between the two decisional boundaries and thus the amount of evidence required before a

choice is made. The larger the separation, the more conservative the threshold. Finally, non-decisional time (t_0) indexes the average duration of all non-decisional processes, such as motor response. By modelling these parameters, I could estimate the extent to which face perception deficits in developmental prosopagnosics are driven by deficits in components relevant to perceptual processing (v and a) as opposed to non-perceptual factors (t_0).

I applied DDM analysis (*fast-dm* (v. 30.2), Voss & Voss, 2007) to the upright accuracy and response time data from the GFMT, because it was the only binary test in my study. I used the maximum-likelihood method of estimation, which calculates the sum of all logarithm density values of all trials based on the model's predicted response time distributions, as recommended for tests with low trial numbers (i.e., the GFMT had 40 trials; Voss et al., 2015). I allowed drift rate, threshold, and non-decisional time to vary based on group (prosopagnosic, control). To improve model accuracy, all other decisional parameters were set to fixed values. This is because the GFMT had too few trials to estimate any further parameters with sufficient reliability (Lerche et al., 2017). Included parameters were not split by GFMT trial type (same versus different) for the same reason.

I ran three Bonferroni corrected independent sample t-tests comparing developmental prosopagnosic and control drift rates, thresholds, and non-decisional time values (displayed in Figure 14). Developmental prosopagnosics had significantly lower drift rates ($M = 0.17$, $SD = 0.16$) than controls ($M = 0.49$, $SD = 0.28$; $t(222.47) = 11.2$, $p < .001$, $d = 1.41$), suggesting inefficient perceptual processing. Developmental prosopagnosics also had larger threshold separation ($M = 2.66$, $SD = 0.46$) than controls ($M = 2.47$, $SD = 0.52$; $t(236.23) = 3.02$, $p = .003$, $d = 0.39$), indicating that they required more information before they were willing to make a choice using the perceptual information. However, developmental prosopagnosic non-decision time ($M = 1.04$, $SD = 0.40$) was also greater than controls ($M =$

0.83, $SD = 0.37$; $t(218.82) = 4.24$, $p < .001$, $d = 0.55$), indicating a role for non-perceptual factors.

Figure 14

Developmental Prosopagnosic and Control Drift Rates (v), Thresholds (a), and Non-decision Times (t_0) on the GFMT

Note. Group comparisons for drift rates, thresholds, and non-decision times are shown in parts A, B, and C, respectively. Diamonds represent the group mean and points show the individual data.

The DDM analysis is consistent with the notion that developmental prosopagnosics' face perception deficits arise from atypical face perception processes, in this case inefficient accumulation of perceptual information and stricter criteria for making a decision about the perceptual information. However, the analysis also revealed the contribution of non-perceptual factors, being developmental prosopagnosics' longer time to execute test components unrelated to the perceptual decision, such as making a motor-response. The non-decision time finding is particularly noteworthy. It is common practice to infer that

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

developmental prosopagnosics' slower response times indicate the use of atypical strategies (Duchaine & Nakayama, 2006a). My results suggest that this inference may not always be warranted, since developmental prosopagnosics' slower response times can be explained by non-perceptual factors such as motor execution time.

Chapter 4: Discussion

“It is hard to commit a face to memory if you can’t pick out the important details”. (L.

W., personal remark after completing my study, July 18, 2020).

In my thesis, I investigated the severity, scope, and nature of face perception deficits in developmental prosopagnosia. The major results provide clear answers. First, face perception deficits in developmental prosopagnosia are severe. The magnitude of the face perception deficits ranged from 70% to 100% of the size of their face memory deficit, showing that impaired face perception in developmental prosopagnosia could be a strong contributor to their well-known difficulty with classic face recognition tests. Second, face perception deficits in developmental prosopagnosia are widespread across the whole sample, as shown by both the whole-sample and single-cases analyses, and was highly predictive of developmental prosopagnosia diagnosis. Third, face perception deficits in developmental prosopagnosia are linked to a variety of impaired mechanisms. These included disruptions to mechanisms specialised for upright faces, the use of inconsistent and atypical face perception strategies, disproportionate problems for within-identity perception than between-identity perception, similar impairments due to lighting and viewpoint changes, and abnormalities in both perceptual and non-perceptual components of test performance.

In this chapter I discuss the implications of my thesis across three broad areas. First, I discuss the empirical contributions of my results to the current literature. Second, I discuss how my results inform several theories about developmental prosopagnosia and how they help resolve key debates. Finally, I discuss the methodological implications of my results and what they mean for future research.

4.1 Empirical implications

A major contribution of my thesis is providing robust evidence that face perception deficits in developmental prosopagnosia are severe, widespread, and highly diagnostic of the

condition. Given the large size of my sample and the various methodological improvements over past studies, this result strongly suggests that face perception deficits play a much greater role in developmental prosopagnosia than currently acknowledged (e.g., Bate et al., 2019b; Bowles et al., 2009; McKone et al., 2011; Ulrich et al., 2017). My results join a line of emerging work on face perception in developmental prosopagnosia (Biotti et al., 2019; Djouab et al., 2020; White et al., 2017), that collectively suggest future research should focus more on understanding perceptual impairments in developmental prosopagnosia. My findings also suggest that future diagnosis of developmental prosopagnosia may need to include a face perception test. As face perception deficits appear core to developmental prosopagnosia, including a test designed to measure face perception might improve diagnostic sensitivity compared to only using face memory tests. Although face memory tests do inherently require face perception, if future research finds that developmental prosopagnosia is characterised by deficient face perception and *not* face memory, then using face memory tests would be contaminating diagnosis with noise.

This finding raises a fundamental question: to what extent is developmental prosopagnosia driven by face memory deficits in addition to face perception deficits? While my results suggest that face memory deficits did not contribute much to developmental prosopagnosia diagnosis, a recent study found just the opposite, with face memory being more diagnostic than face perception (Stumps et al., 2020). One possible explanation is that face memory may have been disadvantaged in my regression model because I only used one face memory test but three face perception tests, which might have improved reliability and sensitivity. More importantly, however, the face memory tests in both my study and the prior study (Stumps et al., 2020) engage not only face memory processes but also face perception processes. Future studies with measures that can isolate face memory are needed to better address how much developmental prosopagnosia is driven by face memory deficits in

addition to face perception deficits¹². One way to isolate face memory is by varying the delay between study and test sessions and subtracting performance in a long delay condition (higher memory load) from a short delay condition (lower memory load; Atkinson & Shiffrin, 1968; Peterson & Peterson, 1959). Two studies have used this approach and found no disproportionate effect of delay in developmental prosopagnosics compared to controls, suggesting a lack of face memory deficits (Biotti et al., 2019; Shah et al., 2015b). However, both studies are limited by relatively small sample sizes ($n = 16$ for both) and brief delays (1-8 seconds), and in one study the delay length had no effect in controls (Shah et al., 2015). Future studies with a much longer delay (i.e., days, McKone et al 2011) are needed to resolve this issue, especially since face recognition problems in developmental prosopagnosia may involve longer-term memory, such as when faces must be remembered hours, or even days later.

Another major result of my study is to clarify that face perception deficits in developmental prosopagnosia are much larger for upright than inverted faces, corroborating past studies (e.g., Behrmann et al., 2005; de Gelder & Romke, 2000; Duchaine & Nakayama, 2007; Farah et al., 1995; Le Grand et al., 2006; Schmalzl et al., 2011). This finding is robust across multiple tests, implying that face perception deficits in developmental prosopagnosia are rooted in impairments of mechanisms selective to upright faces. This result is consistent with studies showing aberrant holistic processing in developmental prosopagnosia (Avidan et al., 2011; DeGutis et al., 2014; Palermo et al., 2011; Towler et al., 2018, but see Biotti et al., 2017b; Susilo et al., 2010). Holistic processing is widely thought to be selective to upright faces (McKone et al 2007), and so these results dovetail with the disproportionate deficits with upright faces that I observed in my data.

¹² Some readers might suggest that I isolate face memory in my study by deducting face perception scores from face memory scores. This analysis would be problematic because the face perception and face memory tests not only differed in presence of memory demand, but also in their stimuli, number of test images, timing, instructions, and so on.

I also found that developmental prosopagnosics showed subtle deficits with inverted faces, consistent with prior work (Biotti et al., 2017b; Garrido et al., 2011; Klargaard et al., 2018; Russell et al., 2009; Ulrich et al., 2017). This result accords with the notion that inverted and upright face processing rely in part on common mechanisms. For example, TMS to the right occipital face area disrupts perception of both upright and inverted faces, but more so for upright faces (Pitcher et al., 2011). The face composite effect – a signature of holistic processing that is indexed by poorer discrimination between upper halves of faces when the lower halves are aligned versus misaligned – is strongest for upright faces, but can occur to a lesser extent with inverted faces (Susilo et al., 2013). My finding of poor inverted face perception in developmental prosopagnosia, but even poorer upright face perception, thus matches the evidence that inverted faces rely on *some* processes also used by upright faces.

Another strong indication of atypical mechanisms in developmental prosopagnosia comes from the weaker correlations between face perception tests in developmental prosopagnosics compared to controls. One possibility is that developmental prosopagnosics adopted unique feature matching or ‘guessing’ strategies to suit each perception test. This is consistent with my earlier finding that developmental prosopagnosics had reduced orientation sensitivity; alternative processing mechanisms must consequently be engaged to complete the face perception tests in developmental prosopagnosia.

My thesis also makes several preliminary contributions. One is that face perception deficits in developmental prosopagnosia may be linked more to within-identity perception (i.e., telling whether two faces show the same person) than between-identity perception (i.e., telling whether two faces show different people). Several studies have shown that the two kinds of face perception rely on partially distinct mechanisms (Jenkins et al 2011; Megreya & Burton, 2007; Robins et al., 2018; Ritchie & Burton, 2017). My finding adds to this line of

work by showing that within-identity perception can be impaired while between-identity is mostly spared. However, my analysis also revealed that response bias might have played a role (White et al., 2017). Future work could use a measure that is free of response bias, such as a sorting task where participants group a set of faces based on how many identities they think are present (Jenkins et al 2011). Another preliminary contribution is that face perception deficits in developmental prosopagnosia may not be associated with the complexity of face image changes, at least as captured by lighting and viewpoint variations. Consistent with a recent study (Mishra et al., 2020; but see Duchaine & Nakayama, 2004), this finding suggests that face perception deficits in developmental prosopagnosia may originate in early stages of face processing, before image-invariant representations are extracted, but more systematic research is needed.

A final empirical contribution of my thesis is the DDM analysis, which revealed multiple components behind face perception deficits in developmental prosopagnosia. The slower drift rates showed that the deficits are tied to poorer perceptual sensitivity, in line with the other results that indicate atypical mechanisms. The wider boundary separation shows that people with developmental prosopagnosia require more evidence to make a face perception decision. Their higher evidence requirement might be an attempt to compensate for their impairment, as developmental prosopagnosics have high self-awareness of their face recognition difficulties (Shah et al., 2015a; Palermo et al., 2017). The longer non-decision time shows that slower response times in developmental prosopagnosia can occur for reasons unrelated to face processing, such as slower motor response. This finding cautions the interpretation of slower response times as evidence of impairment. Where relevant, future work should use DDM to test whether slower response times reflect perceptual impairment as opposed to non-perceptual factors. Moreover, my study illustrates how modelling approaches can benefit developmental prosopagnosia studies and yield new insights.

4.2 Theoretical implications

My thesis has several theoretical implications. The first relates to whether Lissauer's (1890) distinction between apperceptive and associative visual agnosia applies to developmental prosopagnosia. One view says it applies, referring to the dissociation between cases who have deficits with both face perception and face memory (apperceptive) and others who have deficits only with face memory (associative; Bate et al., 2019b; Dalrymple et al., 2014b; Schmalzl et al., 2008a; Stollhoff et al., 2011). Another view says the distinction does not apply and suggests instead that developmental prosopagnosia is associative. This view refers to findings of face memory deficits in all cases but generally spared face perception in many cases (Duchaine & Nakayama, 2004; Duchaine & Nakayama, 2006b; Pertzov et al., 2020; McKone et al., 2011; Stumps et al., 2020; Ulrich et al., 2017), and considers face perception deficits in isolated cases as allied or co-morbid features that are peripheral to developmental prosopagnosia (Ulrich et al., 2017). My result of widespread face perception deficits across the whole sample of developmental prosopagnosics suggest that neither view is correct. It instead suggests a third view, namely that developmental prosopagnosia is apperceptive but to varying degrees (Biotti et al., 2019). This view dovetails with Farah's (1990) influential theory of visual agnosia. After studying many cases of visual agnosia and acquired prosopagnosia, Farah argued that all cases of visual recognition deficits always involve visual perception deficits when properly tested. This analysis led her to propose that visual perception deficits reside on a dimension rather than being a category. This view also provides a way to reconcile my group-level finding (severe and widespread deficits across the entire sample) with the single-case finding (only 32% of developmental prosopagnosics satisfying the clinical cut-off for impairment). The single-case tests have drawn a 'clinical line' down the severe end of the apperceptive spectrum, forcing a distinction between those falling below the line (and labelling them as having deficits) and those on the other side (and

labelling them as not having deficits). However, the group-level finding clearly shows that the developmental prosopagnosics on the other side of the clinical line also have face perception deficits, only to a lesser extent.

Another theoretical implication concerns the normative versus pathological views of developmental prosopagnosia (Barton & Corrow, 2016). According to the normative view, people with developmental prosopagnosia are simply individuals at the low tail-end of the normal distribution of face recognition skills, who process faces in the same way as the rest of the population, but they are doing it less efficiently. According to the pathological view, people with developmental prosopagnosia process faces differently, and so they constitute a distinct group of people whose face recognition is not just poorer, but abnormal. One way to distinguish between the two views is by testing whether a putative marker of face processing, such as the face inversion effect, is present in a weaker form in developmental prosopagnosia (normative), or whether it is abolished altogether (pathological). My results show reduced – but not abolished – inversion effects in developmental prosopagnosia, consistent with the normative view.

The two views can also be distinguished by comparing whether performance in developmental prosopagnosia matches the natural variation expected for the lower end of a control distribution (Barton & Corrow, 2016). Developmental prosopagnosic performance that matches the low-end of control distribution would suggest the use of the typical mechanisms but with reduced efficiency. This would support the normative view. But when developmental prosopagnosic performance is shifted below what is expected for natural variation in control performance, it would suggest the use of atypical processing mechanisms. This would support the pathological view, as shown by scores outside the oval in Figure 15 part A. Using the inversion effect as a marker, and then plotting its relationship with face recognition skill (indexed using CFMT-Aus scores), my data show that developmental

prosopagnosic' performance is shifted below what is predicted by natural variation (Figure 15 part B)¹³. This result supports the pathological view. Altogether, my data demonstrate that developmental prosopagnosia could be a pathological condition, as this latter approach was sufficiently sensitive to detect evidence of pathology.

Figure 15

Comparing Developmental Prosopagnosic Performance to the Variation Predicted By Performance in Controls

Note. Part A shows Barton and Corrow's (2016) hypothetical demonstration of evidence for the pathological account in developmental prosopagnosia. Developmental prosopagnosics fall below the variation expected by control performance (i.e. their scores fall below the dashed oval). Part B shows developmental prosopagnosic and control relationships between perceptual composite inversion effect sizes and face recognition skill, indexed using CFMT-Aus accuracy. As a group, developmental prosopagnosic performance is shifted down from the variation expected by control performance, consistent with the pathological view. My oval is wider because the variation in my control data is larger than in Barton and Corrow's (2016) demonstration.

¹³ I also repeated this analysis using the regression method of calculating inversion effect sizes, as this method is more robust when using inversion effect sizes in individual difference analyses such as correlations (DeGutis et al., 2013b). The pattern remained the same. More detail is available in Appendix B.11.

My results also provide evidence for several theories about face representations in developmental prosopagnosia. One theory posits that developmental prosopagnosia is caused by atypical visual-perceptual face templates (Towler et al., 2017). The theory is motivated by ERP studies showing selective impairment in early encoding of faces in developmental prosopagnosia (around 200 milliseconds after face onset; Towler et al., 2017). Unlike controls who showed different ERP responses to normal faces and faces with atypical configurations (e.g., spatially scrambling the location of the eyes, nose, and mouth), developmental prosopagnosics showed similar ERP activations to both (Itier et al, 2006; Itier & Taylor, 2002; Fisher et al., 2016a; Fisher et al., 2016b; Towler & Eimer, 2012; Towler et al., 2012; Towler et al., 2016). The template theory offers an explanation for the severe face perception deficits in my results. As the templates are said to specifically lack sensitivity to configural deviations, it also matches my finding of reduced inversion effects in developmental prosopagnosia, because inverting faces disrupts the configural order of face features (i.e., eyes below the mouth, Towler et al., 2017). The theory also suggests that more research should investigate the ‘front-end’ of face processing in developmental prosopagnosia, focusing on specific mechanisms within the first 200 milliseconds of face viewing. One candidate mechanism is atypical fixations, where developmental prosopagnosics tend to look less at the eyes and the internal facial region than controls (Bobak et al., 2017; Schwarzer et al., 2007; Wilcockson et al., 2020). Deficient face sampling would suggest that face perception deficits in developmental prosopagnosia arise from difficulty *accessing* rather than *processing* perceptual information. Another candidate is retinotopic tuning. Face recognition is optimal when the face is fixated on preferred retinotopic location, which varies across individuals (de Haas & Schwarzkopf, 2018; de Haas et al., 2016; Kaiser & Haselhuhn, 2017; Peterson et al., 2019). A recent study suggests that

people with developmental prosopagnosia have trouble matching their face fixation to their preferred retinotopic location (Peterson et al., 2019).

Another theory suggests that developmental prosopagnosia arises from disordered holistic face representations (Avidan et al., 2011; DeGutis et al., 2014; Palermo et al., 2011; Towler et al., 2018). According to this theory, faces are represented as a unified whole, without decomposition into local parts such as eyes, nose, and mouth (Maurer et al., 2002; McKone & Yovel, 2009; Tanaka & Farah, 1993). One measure of holistic face representation is the face inversion effect, with larger effects indicating more holistic representations (Piepers & Robbins, 2012). This is because inverted faces are more difficult to perceive holistically than upright faces and must instead be processed based on local parts. My finding of smaller inversion effects in developmental prosopagnosia are in line with the disordered holistic representation theory. However, it is worth noting that the inversion effect is often considered an *indirect* measure of holistic processing (Piepers & Robbins, 2012), because it does not manipulate the overall composition of the face, such as the part whole effect (poorer recognition of single features when presented in isolation versus within the whole face) and composite effect (poorer discrimination between upper face halves when bottom halves are aligned versus misaligned). To better link developmental prosopagnosia with disordered holistic representations, future research should use other holistic measures and ensure that they are run with minimal memory demands (e.g., the composite face test which is usually run with sequential presentation could be altered for simultaneous presentation).

My thesis also informs which specific perceptual process might be impaired in developmental prosopagnosia according to Bruce and Young's (1986) face recognition model. In this model, two separate forms of face perception are dissociated. View-centred processing occurs for information used to perceive facial expression and speech, while expression-independent processing occurs for information related to identity (see Figure 1 in

section 1.2). In developmental prosopagnosia, perception of facial expression is usually intact or only slightly impaired (Dobel et al., 2007; Duchaine et al., 2003; Lee et al., 2010; Nunn et al., 2001; Palermo et al. 2011, though see Biotti & Cook, 2016), while my thesis shows that perception of face identity is severely deficient. Collectively, this evidence suggests that expression-independent processing might be selectively impaired in developmental prosopagnosia, leaving view-centred processing relatively intact. Future work could test whether view-centred processing does indeed function normally by investigating speech perception in developmental prosopagnosia – an area that has not been researched to date.

My results can also inform the development of rehabilitation efforts for developmental prosopagnosia. Inspired by theories of holistic face representation, most rehabilitation efforts to date encourage participants to focus on specific face features (Brunsdon et al., 2006; Schmalzl et al., 2008b), the central eye-region (Pizzamiglio et al., 2017) or to the spatial relations between features (DeGutis et al., 2007; DeGutis et al., 2014). In contrast, a recent approach did not provide instructions about focusing on specific aspects of the face, but rather gradually increased the similarity between test faces, which was intended to increase demand on face perception (Corrow et al., 2019). This perceptual learning approach improved face perception in developmental prosopagnosia for up to three months after intervention – the most persistent to date (Corrow et al., 2019). The success of general perceptual learning accords with my finding that multiple mechanisms contribute to face perception deficits in developmental prosopagnosia. Further research using perceptual learning might thus be fruitful for advancing developmental prosopagnosia interventions, because the approach is not restricted to focusing on one specific process.

4.3 Methodological implications

My thesis is one of the first to measure both face perception and face memory as part of the main investigation of developmental prosopagnosia (Biotti et al., 2019). In contrast,

most studies reused face memory scores that were used for diagnosis (for example, Bate et al., 2019b; Dalrymple et al., 2014; Huis in't Veld et al., 2012; Mishra et al., 2020; Ulrich et al., 2017; White et al., 2017). This circular analysis often led to inferences that face perception deficits are uncommon or weak in developmental prosopagnosia. Here, I illustrate the severity of the circular analysis problem. Figure 16 compares the z-scores from the face perception tests against the z-scores from CFMT-Aus (non-circular analysis) and CFMT (circular analysis). Had I reused the CFMT scores, I would have concluded that the size of face perception deficits were substantially smaller than the size of face memory deficits (i.e., around half the size, see Figure 16). This illustration shows that future research must avoid circular analysis to make valid inferences about developmental prosopagnosia.

Figure 16

Developmental Prosopagnosic Z-scores on their Diagnostic CFMT Assessment Compared to the CFMT-Aus and Perception Tests.

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

Note. To qualify as having developmental prosopagnosia, participants had to score at least 2 standard deviations below the control average on the CFMT. Had the CFMT diagnostic scores been reused to index face memory performance, this would have forced perception deficits to appear weaker and less common than memory deficits. Remeasuring face memory using the CFMT-Aus demonstrates that this is not the case.

This analysis also raised a critical issue for the diagnosis of developmental prosopagnosia. Despite the CFMT being one of the ‘gold standard’ tests for diagnosis (Bowles et al., 2009; Duchaine & Nakayama, 2006b), only 25% of the 108 developmental prosopagnosics in my sample re-met the diagnostic cut-off when retested with the CFMT-Aus. A similar finding was reported recently (Murray & Bate, 2020). While this issue can perhaps be partly accounted for by practice effects and regression to the mean (i.e., when extreme scores become less extreme upon remeasurement, see Barnett et al., 2005), it is striking that only one quarter of my sample re-met a key diagnostic criterion for developmental prosopagnosia. This finding raises serious concerns about the psychometric properties of the CFMT and other similar tests, and suggests that repeated assessments might be necessary for more robust diagnosis (Murray & Bate, 2020).

Relatedly, my results also suggest that current face memory tests might not be suitable for capturing individual differences within the developmental prosopagnosia sample (i.e., the severity of developmental prosopagnosia). While the tests successfully predicted group membership and distinguished between developmental prosopagnosics and controls, they failed to predict the severity of developmental prosopagnosia as indexed by CFMT and PI20 scores of the developmental prosopagnosics (analyses in Appendix B.12). This failure may suggest that scores on the face memory tests are only spread out enough to detect differences in group membership and not individual differences within developmental prosopagnosia. Designing new face memory tests with the capacity to distinguish individual differences within developmental prosopagnosia will allow future research to determine whether face perception deficits are associated with developmental prosopagnosia severity.

Another methodological contribution of my thesis was the large sample size, which was ten times larger than the mode sample size. The large sample allowed me to accurately estimate the severity and scope of perception deficits in developmental prosopagnosia. Historically low sample sizes have been offered as a possible explanation for the high instance of mixed results regarding cognitive performance in developmental prosopagnosia (Corrow et al., 2016; Murray et al., 2018; Susilo & Duchaine, 2013). Future research with even greater sample sizes (~300-400 prosopagnosics) could run cluster analyses to unpack whether face perception deficits in developmental prosopagnosia are best characterised as residing on a spectrum of severity or as a dichotomous present-versus-not present. This exact approach has been used to characterise the autism spectrum (e.g., Cholemkery et al., 2016; Spiker et al., 2002; Veatch et al., 2014).

A further advantage demonstrated by my thesis was including multiple face perception tests. The alternative test formats recruit different types of perceptual processing, and include other variations important to face processing, such as whether faces had external features or changed viewpoint. Despite all these differences, developmental prosopagnosics showed face perception deficits across the board. Incorporating three face perception tests has shown that face perception deficits in developmental prosopagnosia are pervasive, and cannot be discounted as only existing when certain test parameters are present.

Finally, my thesis shows that advanced analytic methods such as shift functions and DDM can generate important insights about developmental prosopagnosia. The shift functions yielded the key insight that face perception deficits in developmental prosopagnosia were present across the entire sample and were not driven by particular subgroups. The shift functions also offered further confidence that the central tendency analyses represented the prosopagnosics sample well and were not obscuring any other important information such as skewness. Similarly, the DDM analysis revealed that face perception deficits in

developmental prosopagnosia were driven by multiple components including inefficient perceptual processing, stricter decision criteria, and slower motor response. DDM can therefore be a valuable tool for further characterising face perception and other functions in developmental prosopagnosia, similar to how it has been used to better understand dyslexia (Zeguers et al., 2011) and autism (Powell et al., 2019). As my current DDM analysis is limited by the relatively low trial number of the GFMT, future studies could use the extended 168-trial version of the GFMT to compare parameters for within-identity versus between-identity trials. Doing so would yield insights about whether poor within-identity trial performance in developmental prosopagnosia is driven by response bias or impaired within-identity processing mechanisms, as diffusion models with more trials can estimate criterion (i.e., a parameter called starting point (z)).

4.4 Conclusion

The overarching goal of my thesis was to investigate face perception deficits in developmental prosopagnosia. I addressed this issue by administering a battery of standardised face perception and memory tests to a large sample of developmental prosopagnosics and controls. My main results show that face perception deficits in developmental prosopagnosia are profound in magnitude, widespread across the whole sample, and are highly predictive of developmental prosopagnosia diagnosis. Other results reveal various features of the deficits that further elucidate the nature of face perception impairments in developmental prosopagnosia. Altogether, my thesis provides strong evidence that face perception deficits play a much greater role in developmental prosopagnosia than previously assumed, and opens up new avenues for future research in the field.

References

- Adler, A. (1944). Disintegration and restoration of optic recognition in visual agnosia: Analysis of a case. *Archives of Neurology & Psychiatry*, 51(3), 243–259.
<https://doi.org/10.1001/archneurpsyc.1944.02290270032004>
- Albert, M. L., Reches, A., & Silverberg, R. (1975). Associative visual agnosia without alexia. *Neurology*, 25(4), 322–322. <https://doi.org/10.1212/WNL.25.4.322>
- 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: Official Journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*, 38.
<https://doi.org/10.1007/s10072-017-3030-6>
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 89–195). New York, NY: Academic Press.
- 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., 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>
- Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology*, 34(1), 215–220.
<https://doi.org/10.1093/ije/dyh299>

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

- Barton, J. J. S. (2008). Structure and function in acquired prosopagnosia: Lessons from a series of 10 patients with brain damage. *Journal of Neuropsychology*, 2(1), 197–225. <https://doi.org/10.1348/174866407X214172>
- Barton, J. J. S., & Corrow, S. L. (2016). The problem of being bad at faces. *Neuropsychologia*, 89, 119–124. <https://doi.org/10.1016/j.neuropsychologia.2016.06.008>
- Bate, S., Adams, A., Bennetts, R., & Line, H. (2019a). Developmental prosopagnosia with concurrent topographical difficulties: A case report and virtual reality training programme. *Neuropsychological Rehabilitation*, 29(8), 1290–1312. <https://doi.org/10.1080/09602011.2017.1409640>
- Bate, S., Bennetts, R. J., Gregory, N., Tree, J. J., Murray, E., Adams, A., Bobak, A. K., Penton, T., Yang, T., & Banissy, M. J. (2019b). Objective patterns of face recognition deficits in 165 adults with self-reported developmental prosopagnosia. *Brain Sciences*, 9(6), 133. <https://doi.org/10.3390/brainsci9060133>
- Bate, S., Cook, S. J., Duchaine, B., Tree, J. J., Burns, E. J., & Hodgson, T. L. (2014). Intranasal inhalation of oxytocin improves face processing in developmental prosopagnosia. *Cortex*, 50, 55–63. <https://doi.org/10.1016/j.cortex.2013.08.006>
- Bate, S., Haslam, C., Jansari, A., & Hodgson, T. (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., & Tree, J. J. (2017). The definition and diagnosis of developmental prosopagnosia. *Quarterly Journal of Experimental Psychology*, 70(2), 193–200. <https://doi.org/10.1080/17470218.2016.1195414>
- Behrman, M., & Avidan, G. (2005). Congenital prosopagnosia: face-blind from birth. *Trends in Cognitive Science*, 9(4), 180–187. <https://doi.org/10.1016/j.tics.2005.02.011>

- Behrmann, M., Avidan, G., Marotta, J. J., & Kimchi, R. (2005). Detailed exploration of face-related processing in congenital prosopagnosia: 1. Behavioral findings. *Journal of Cognitive Neuroscience*, 17(7), 1130–1149. <https://doi.org/10.1162/0898929054475154>
- Benson, D. F., & Greenberg, J. P. (1969). Visual form agnosia: A specific defect in visual discrimination. *Archives of Neurology*, 20(1), 82–89. <https://doi.org/10.1001/archneur.1969.00480070092010>
- Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, 8(6), 551–565. <https://doi.org/10.1162/jocn.1996.8.6.551>
- Bentin, S., & Deouell, L. Y. (2000). Structural encoding and identification in face processing: ERP evidence for separate mechanisms. *Cognitive Neuropsychology*, 17(1–3), 35–54. <https://doi.org/10.1080/026432900380472>
- Bentin, S., Deouell, L. Y., & Soroker, N. (1999). Selective visual streaming in face recognition evidence from developmental prosopagnosia. *NeuroReport*, 10(4), 823–827.
- Benton, Arthur L. (1980). The neuropsychology of facial recognition. *American Psychologist*, 35(2), 176–186. <https://doi.org/10.1037/0003-066X.35.2.176>
- Benton AL, Hamsher K, Varney NR, Spreen O. (1983) *Contributions to neuropsychological assessment*. New York, NY: Oxford University Press.
- Biotti, F., & Cook, R. (2016). Impaired perception of facial emotion in developmental prosopagnosia. *Cortex*, 81, 126–136. <https://doi.org/10.1016/j.cortex.2016.04.008>
- Biotti, F., Gray, K. L. H., & Cook, R. (2017a). Impaired body perception in developmental prosopagnosia. *Cortex*, 93, 41–49. <https://doi.org/10.1016/j.cortex.2017.05.006>
- 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>

- Biotti, F., Wu, E., Yang, H., Jiahui, G., Duchaine, B., & Cook, R. (2017b). Normal composite face effects in developmental prosopagnosia. *Cortex*, 95, 63–76.
<https://doi.org/10.1016/j.cortex.2017.07.018>
- Bobak, Anna K., Parris, B. A., Gregory, N. J., Bennetts, R. J., & Bate, S. (2017). Eye-movement strategies in developmental prosopagnosia and “super” face recognition. *Quarterly Journal of Experimental Psychology*, 70(2), 201–217.
<https://doi.org/10.1080/17470218.2016.1161059>
- Bobak, Anna Katarzyna, Dowsett, A. J., & Bate, S. (2016). Solving the border control problem: evidence of enhanced face matching in individuals with extraordinary face recognition skills. *PLOS ONE*, 11(2), e0148148.
<https://doi.org/10.1371/journal.pone.0148148>
- Bona, S., Silvanto, J., & Cattaneo, Z. (2018). TMS over right OFA affects individuation of faces but not of exemplars of objects. *Neuropsychologia*, 117, 364–370.
<https://doi.org/10.1016/j.neuropsychologia.2018.06.024>
- Bowles, D. C., McKone, E., Dawel, A., Duchaine, B., Palermo, R., Schmalzl, L., Rivolta, D., Wilson, C. E., & Yovel, G. (2009). Diagnosing prosopagnosia: Effects of ageing, sex, and participant-stimulus ethnic match on the Cambridge Face Memory Test and Cambridge Face Perception Test. *Cognitive Neuropsychology*, 26(5), 423–455.
<https://doi.org/10.1080/02643290903343149>
- Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ: Brain and Cognition*. <https://doi.org/10.7717/peerj.9414>
- Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, 77(3), 305–327. <https://doi.org/10.1111/j.2044-8295.1986.tb02199.x>

- Brunsdon, R., Coltheart, M., Nickels, L., & Joy, P. (2006). Developmental prosopagnosia: A case analysis and treatment study. *Cognitive Neuropsychology*, 23(6), 822–840.
<https://doi.org/10.1080/02643290500441841>
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6(1), 3–5. <https://doi.org/10.1177/1745691610393980>
- Burton, A. M., White, D., & McNeill, A. (2010). The Glasgow Face Matching Test. *Behavior Research Methods*, 42(1), 286–291. <https://doi.org/10.3758/BRM.42.1.286>
- Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5), 365–376.
<https://doi.org/10.1038/nrn3475>
- Chatterjee, G., & Nakayama, K. (2012). Normal facial age and gender perception in developmental prosopagnosia. *Cognitive Neuropsychology*, 29(5–6), 482–502.
<https://doi.org/10.1080/02643294.2012.756809>
- Cholemkery, H., Medda, J., Lempp, T., & Freitag, C. M. (2016). Classifying autism spectrum disorders by ADI-R: Subtypes or severity gradient? *Journal of Autism and Developmental Disorders*, 46(7), 2327–2339. <https://doi.org/10.1007/s10803-016-2760-2>
- Colquhoun, D. (2014). An investigation of the false discovery rate and the misinterpretation of p-values. *Royal Society Open Science*, 1(3), 140216.
<https://doi.org/10.1098/rsos.140216>
- Corbett, E. A., & Smith, P. L. (2020). A diffusion model analysis of target detection in near-threshold visual search. *Cognitive Psychology*, 120.
<https://doi.org/10.1016/j.cogpsych.2020.101289>

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

- Corrow, J. C., Corrow, S. L., Lee, E., Pancaroglu, R., Burles, F., Duchaine, B., Iaria, G., & Barton, J. J. S. (2016). Getting lost: Topographic skills in acquired and developmental prosopagnosia. *Cortex*, 76, 89–103. <https://doi.org/10.1016/j.cortex.2016.01.003>
- Corrow, S. L., Dalrymple, K. A., & Barton, J. J. (2016). Prosopagnosia: Current perspectives. *Eye and Brain*, 8, 165–175. <https://doi.org/10.2147/EB.S92838>
- 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. <https://doi.org/10.1016/j.neuropsychologia.2019.107196>
- Crawford, J. R., & Garthwaite, P. H. (2006). Methods of testing for a deficit in single-case studies: Evaluation of statistical power by Monte Carlo simulation. *Cognitive Neuropsychology*, 23(6), 877–904. <https://doi.org/10.1080/02643290500538372>
- Crawford, J. R., & Howell, D. C. (1998). Comparing an individual's test score against norms derived from small samples. *The Clinical Neuropsychologist*, 12(4), 482–486. <https://doi.org/10.1076/clin.12.4.482.7241>
- Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's mechanical turk as a tool for experimental behavioural research. *PLoS ONE*, 8(3). <https://doi.org/10.1371/journal.pone.0057410>
- Dahl, C. D., Logothetis, N. K., & Hoffman, K. L. (2007). Individuation and holistic processing of faces in rhesus monkeys. *Proceedings of the Royal Society B: Biological Sciences*, 274(1622), 2069–2076. <https://doi.org/10.1098/rspb.2007.0477>
- Dahl, C. D., Rasch, M. J., Tomonaga, M., & Adachi, I. (2013). The face inversion effect in non-human primates revisited – An investigation in chimpanzees (*Pan troglodytes*). *Scientific Reports*, 3(1), 2504. <https://doi.org/10.1038/srep02504>

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

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>

Davies-Thompson, J., Gouws, A., & Andrews, T. J. (2009). An image-dependent representation of familiar and unfamiliar faces in the human ventral stream.

Neuropsychologia, 47(6), 1627–1635.

<https://doi.org/10.1016/j.neuropsychologia.2009.01.017>

Davies-Thompson, J., Newling, K., & Andrews, T. J. (2013). Image-invariant response in face-selective regions do not explain the perceptual advantage for familiar face recognition. *Cerebral Cortex*, 23(2), 370–377. <https://doi.org/10.1093/cercor/bhs024>

Davies-Thompson, J., Pancaroglu, R., & Barton, J. (2014). Acquired prosopagnosia:

Structural basis and processing impairments. *Frontiers in Bioscience (Elite Edition)*, 6, 159–174.

DeGutis, J., Bentin, S., Robertson, L. C., & D’Esposito, M. (2007). Functional plasticity in ventral temporal cortex following cognitive rehabilitation of a congenital prosopagnosic.

Journal of Cognitive Neuroscience, 19(11), 1790–1802.

<https://doi.org/10.1162/jocn.2007.19.11.1790>

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

<https://doi.org/10.1093/brain/awu062>

DeGutis, J., Mercado, R. J., Wilmer, J., & Rosenblatt, A. (2013a). Individual differences in holistic processing predict the own-race advantage in recognition memory. *PLoS One*,

8(4). <https://doi.org/10.1371/journal.pone.0058253>

- DeGutis, J., Wilmer, J., Mercado, R. J., & Cohan, S. (2013b). Using regression to measure holistic face processing reveals a strong link with face recognition ability. *Cognition*, 126(1), 87–100. <https://doi.org/10.1016/j.cognition.2012.09.004>
- de Gelder, B., & Rouw, R. (2000). Configural face processes in acquired and developmental prosopagnosia: Evidence for two separate face systems? *NeuroReport*, 11(14), 3145–3150.
- De Haan, E. H. F., & Campbell, R. (1991). A fifteen year follow-up of a case of developmental prosopagnosia. *Cortex*, 27(4), 489–509. [https://doi.org/10.1016/S0010-9452\(13\)80001-9](https://doi.org/10.1016/S0010-9452(13)80001-9)
- de Haas, B., & Schwarzkopf, D. S. (2018). Feature–location effects in the Thatcher illusion. *Journal of Vision*, 18(4), 16. <https://doi.org/10.1167/18.4.16>
- 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>
- De Renzi, E. (1986). Prosopagnosia in two patients with CT scan evidence of damage confined to the right hemisphere. *Neuropsychologia*, 24(3), 385–389. [https://doi.org/10.1016/0028-3932\(86\)90023-0](https://doi.org/10.1016/0028-3932(86)90023-0)
- De Renzi, E., & di Pellegrino, G. (1998). Prosopagnosia and alexia without object agnosia. *Cortex*, 34(3), 403–415. [https://doi.org/10.1016/S0010-9452\(08\)70763-9](https://doi.org/10.1016/S0010-9452(08)70763-9)
- De Renzi, E., Faglioni, P., Grossi, D., & Nichelli, P. (1991). Apperceptive and associative forms of prosopagnosia. *Cortex*, 27, 213–221. [https://doi.org/10.1016/S0010-9452\(13\)80125-6](https://doi.org/10.1016/S0010-9452(13)80125-6)
- di Oleggio Castello, M. V., Halchenko, Y. O., Guntupalli, J. S., Gors, J. D., & Gobbini, M. I. (2017). The neural representation of personally familiar and unfamiliar faces in the

distributed system for face perception. *Scientific Reports*, 7(1), 12237.

<https://doi.org/10.1038/s41598-017-12559-1>

Djouab, S., Albonico, A., Yeung, S. C., Malaspina, M., Mogard, A., Wahlberg, R., Corrow, S. L., & Barton, J. J. S. (2020). Search for face identity or expression: Set size effects in developmental prosopagnosia. *Journal of Cognitive Neuroscience*, 32(5), 889–905.

https://doi.org/10.1162/jocn_a_01519

Dobel, C., Bölte, J., Aicher, M., & Schweinberger, S. R. (2007). Prosopagnosia without apparent cause: Overview and diagnosis of six cases. *Cortex*, 43(6), 718–733.

[https://doi.org/10.1016/S0010-9452\(08\)70501-X](https://doi.org/10.1016/S0010-9452(08)70501-X)

Duchaine, B. C. (2008). Comment on prevalence of hereditary prosopagnosia (HPA) in Hong Kong Chinese population. *American Journal of Medical Genetics Part A*, 146(22), 2860–2862. <https://doi.org/10.1002/ajmg.a.32548>

Duchaine, B. (2011). Developmental prosopagnosia: Cognitive, neural, and developmental investigations. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *Oxford handbook of face perception* (pp. 821–838). New York, NY: Oxford University Press.

<https://doi.org/10.1093/oxfordhb/9780199559053.013.0042>

Duchaine, B. C., & Nakayama, K. (2004). Developmental prosopagnosia and the Benton Facial Recognition Test. *Neurology*, 62(7), 1219–1220.

<https://doi.org/10.1212/01.WNL.0000118297.03161.B3>

Duchaine, B. C., & Nakayama, K. (2005). Dissociations of face and object recognition in developmental prosopagnosia. *Journal of Cognitive Neuroscience*, 17(12), 249–261.

<https://doi.org/10.1162/0898929053124857>

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

<https://doi.org/10.1016/j.conb.2006.03.003>

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

- Duchaine, B. C., Parker, H., & Nakayama, K. (2003). Normal recognition of emotion in a prosopagnosic. *Perception*, 32(7), 827–838. <https://doi.org/10.1068/p5067>
- Duchaine, B. C., Yovel, G., Butterworth, E. J., & Nakayama, K. (2006). Prosopagnosia as an impairment to face-specific mechanisms: Elimination of the alternative hypotheses in a developmental case. *Cognitive Neuropsychology*, 23(5), 714–747.
<https://doi.org/10.1080/02643290500441296>
- Duchaine, Brad, Germine, L., & Nakayama, K. (2007a). Family resemblance: Ten family members with prosopagnosia and within-class object agnosia. *Cognitive Neuropsychology*, 24, 419–430. <https://doi.org/10.1080/02643290701380491>
- Duchaine, Brad, & 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, Brad, & Nakayama, K. (2006b). 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>
- Duchaine, Brad, & Yovel, G. (2015). A revised neural framework for face processing. *Annual Review of Vision Science*, 1, 393–416. <https://doi.org/10.1146/annurev-vision-082114-035518>
- Duchaine, Bradley, Yovel, G., & Nakayama, K. (2007b). No global processing deficit in the Navon task in 14 developmental prosopagnosics. *Social Cognitive and Affective Neuroscience*, 2(2), 104–113. <https://doi.org/10.1093/scan/nsm003>
- Eimer, M. (2000a). Event-related brain potentials distinguish processing stages involved in face perception and recognition. *Clinical Neurophysiology*, 111(4), 694–705.
[https://doi.org/10.1016/S1388-2457\(99\)00285-0](https://doi.org/10.1016/S1388-2457(99)00285-0)

- Eimer, M. (2000b). The face-specific N170 component reflects late stages in the structural encoding of faces. *NeuroReport*, 11(10), 2319–2324.
- Eimer, M. (2000c). Effects of face inversion on the structural encoding and recognition of faces: Evidence from event-related brain potentials. *Cognitive Brain Research*, 10(1), 145–158. [https://doi.org/10.1016/S0926-6410\(00\)00038-0](https://doi.org/10.1016/S0926-6410(00)00038-0)
- Ellis, H. D., & Florence, M. (1990). Bodamer's (1947) paper on prosopagnosia. *Cognitive Neuropsychology*, 7(2), 81–105. <https://doi.org/10.1080/02643299008253437>
- Evans, M. (2019, March 6). *Face blindness: 'I can't recognise my loved-ones'*. BBC News. <https://www.bbc.com/news/uk-wales-47304678>
- Farah, M. J. (1990). *Visual Agnosia* (2nd ed.). The MIT Press.
- Favelle, S., Hill, H., & Claes, P. (2017). About face: Matching unfamiliar faces across rotations of view and lighting. *i-Perception*, 8(6). <https://doi.org/10.1177/2041669517744221>
- Fisher, K., Towler, J., & Eimer, M. (2016a). Effects of contrast inversion on face perception depend on gaze location: Evidence from the N170 component. *Cognitive Neuroscience*, 7(1–4), 128–137. <https://doi.org/10.1080/17588928.2015.1053441>
- Fisher, K., Towler, J., & Eimer, M. (2016b). Reduced sensitivity to contrast signals from the eye region in developmental prosopagnosia. *Cortex*, 81, 64–78. <https://doi.org/10.1016/j.cortex.2016.04.005>
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47(6), 381–391. <https://doi.org/10.1037/h0055392>
- Freiwald, W. A., & Tsao, D. Y. (2010). Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science*, 330(6005), 845–851. <https://doi.org/10.1126/science.1194908>

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

- Freiwald, W. A., Tsao, D. Y., & Livingstone, M. S. (2009). A face feature space in the macaque temporal lobe. *Nature Neuroscience*, 12(9), 1187–1196.
<https://doi.org/10.1038/nn.2363>
- Freiwald, W., Duchaine, B., & Yovel, G. (2016). Face processing systems: From neurons to real world social perception. *Annual Review of Neuroscience*, 39, 325–346.
<https://doi.org/10.1146/annurev-neuro-070815-013934>
- Furl, N., Garrido, L., Dolan, R. J., Driver, J., & Duchaine, B. (2011). Fusiform gyrus face selectivity relates to individual differences in facial recognition ability. *Journal of Cognitive Neuroscience*, 23(7), 1723–1740. <https://doi.org/10.1162/jocn.2010.21545>
- Garrido, L., Duchaine, B., & Nakayama, K. (2008). Face detection in normal and prosopagnosic individuals. *Journal of Neuropsychology*, 2(1), 119–140.
<https://doi.org/10.1348/174866407X246843>
- Gelb, A., & Goldstein, K. (1938). Analysis of a case of figural blindness. In W. D. Ellis (Eds.), *A source book of Gestalt psychology* (pp. 315–325). Kegan Paul, Trench, Trubner & Company. <https://doi.org/10.1037/11496-026>
- Gerlach, C., Klargaard, S. K., & Starrfelt, R. (2016). On the relation between face and object recognition in developmental prosopagnosia: No dissociation but a systematic association. *PLoS ONE*, 11(10). <https://doi.org/10.1371/journal.pone.0165561>
- Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review*, 19(5), 847–857.
<https://doi.org/10.3758/s13423-012-0296-9>
- Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia*, 45(1), 32–41.
<https://doi.org/10.1016/j.neuropsychologia.2006.04.015>

- Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioural Decision Making*, 26(3), 213–224. <https://doi.org/10.1002/bdm.1753>
- Gosling, A., & Eimer, M. (2011). An event-related brain potential study of explicit face recognition. *Neuropsychologia*, 49(9), 2736–2745.
<https://doi.org/10.1016/j.neuropsychologia.2011.05.025>
- Gothard, K. M., Erickson, C. A., & Amaral, D. G. (2004). How do rhesus monkeys (Macaca mulatta) scan faces in a visual paired comparison task? *Animal Cognition*, 7(1), 25–36.
<https://doi.org/10.1007/s10071-003-0179-6>
- Gray, K. L. H., Biotti, F., & Cook, R. (2019). Evaluating object recognition ability in developmental prosopagnosia using the Cambridge Car Memory Test. *Cognitive Neuropsychology*, 36(1-2), 89–96. <https://doi.org/10.1080/02643294.2019.1604503>
- Griffin, J. W. (2020). Quantifying the face inversion effect in nonhuman primates: A phylogenetic meta-analysis. *Animal Cognition*, 23(2), 237–249.
<https://doi.org/10.1007/s10071-019-01340-8>
- Hacker, C., & Biederman, I. (2019). The proficiency for distinguishing faces is independent of the proficiency for remembering them. *PsyArXiv*.
<https://doi.org/10.31234/osf.io/9bwct>
- Harrell, F. E., & Davis, C. E. (1982). A new distribution-free quantile estimator. *Biometrika*, 69(3), 635–640. <https://doi.org/10.1093/biomet/69.3.635>
- Harris, A. M., Duchaine, B. C., & Nakayama, K. (2005). Normal and abnormal face selectivity of the M170 response in developmental prosopagnosics. *Neuropsychologia*, 43(14), 2125–2136. <https://doi.org/10.1016/j.neuropsychologia.2005.03.017>
- Haxby, J. V., & Gobbini, M. I. (2011). Distributed neural systems for face perception. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *Oxford handbook of face*

perception (pp. 93–110). New York, NY: Oxford University Press.

<https://doi.org/10.1093/oxfordhb/9780199559053.013.0042>

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)

Hay, D. C., & Young, A. W. (1982). The human face. In A. W. Ellis (Eds.), *Normality and pathology in cognitive functions* (pp. 173–202). London: Academic Press.

Herzmann, G., Danthiir, V., Schacht, A., Sommer, W., & Wilhelm, O. (2008). Toward a comprehensive test battery for face cognition: Assessment of the tasks. *Behaviour Research Methods*, 40, 840–857. <https://doi-org./10.3758/BRM.40.3.840>

Herzmann, G., Schweinberger, S. R., Sommer, W., & Jentsch, I. (2004). What's special about personally familiar faces? A multimodal approach. *Psychophysiology*, 41(5), 688–701. <https://doi.org/10.1111/j.1469-8986.2004.00196.x>

Hildebrandt, A., Sommer, W., Herzmann, G., & Wilhelm, O. (2010). Structural invariance and age-related performance differences in face cognition. *Psychology and Aging*, 25(4), 794–810. <https://doi.org/10.1037/a0019774>

Hildebrandt, A., Wilhelm, O., Schmiedek, F., Herzmann, G., & Sommer, W. (2011). On the specificity of face cognition compared with general cognitive functioning across adult age. *Psychology and Aging*, 26(3), 701–715. <https://doi.org/10.1037/a0023056>

Hodges, J.R. (1995). Retrograde amnesia. In A. Baddeley, B. Wilson, & F. Watts (Eds.), *Handbook of memory disorders* (pp. 81–107). Chichester, UK: Wiley

Holiday, K., Pitcher, D., & Ungerleider, L. (2015). Temporal dynamics of memory and maintenance of faces in visual cortex: An on-line TMS study. *Journal of Vision*, 15(12), 294–294. <https://doi.org/10.1167/15.12.294>

- Hotelling, H. (1931). The economics of exhaustible resources. *Journal of Political Economy*, 39(2), 137–175.
- Huis in't Veld, E., Van den Stock, J., & DeGelder, B. (2012). Configuration perception and face memory, and face context effects in developmental prosopagnosia. *Cognitive Neuropsychology*, 29(5–6), 464–481. <https://doi.org/10.1080/02643294.2012.732051>
- Huizenga, H. M., Smeding, H., Grasman, R. P. P. P., & Schmand, B. (2007). Multivariate normative comparisons. *Neuropsychologia*, 45(11), 2534–2542. <https://doi.org/10.1016/j.neuropsychologia.2007.03.011>
- Humphreys, G. W., & Riddoch, M. J. (1987). *To see but not to see: A case study of visual agnosia*. Psychology Press.
- Issa, E. B., & DiCarlo, J. J. (2012). Precedence of the eye region in neural processing of faces. *The Journal of Neuroscience*, 32(47), 16666–16682. <https://doi.org/10.1523/jneurosci.2391-12.2012>
- Itier, R. J., Latinus, M., & Taylor, M. J. (2006). Face, eye and object early processing: What is the face specificity? *NeuroImage*, 29(2), 667–676. <https://doi.org/10.1016/j.neuroimage.2005.07.041>
- Itier, R. J., & Taylor, M. J. (2002). Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: A repetition study using ERPs. *NeuroImage*, 15(2), 353–372. <https://doi.org/10.1006/nimg.2001.0982>
- Jenkins, R., White, D., Van Montfort, X., & Mike Burton, A. (2011). Variability in photos of the same face. *Cognition*, 121(3), 313–323. <https://doi.org/10.1016/j.cognition.2011.08.001>
- Jiahui, G., Yang, H., & Duchaine, B. (2018). Developmental prosopagnosics have widespread selectivity reductions across category-selective visual cortex. *PNAS*, 115(28), 6418–6427. <https://doi.org/10.1073/pnas.1802246115>

- Joenssen, D. W., & Bankhofer, U. (2012). Hot deck methods for imputing missing data. In P. Perner (Ed.), *Machine learning and data mining in pattern recognition* (pp. 63–75). Springer. https://doi.org/10.1007/978-3-642-31537-4_6
- Jones, R. D., & Tranel, D. (2001). Severe developmental prosopagnosia in a child with superior intellect. *Journal of Clinical and Experimental Neuropsychology*, 23(3), 265–273. <https://doi.org/10.1076/jcen.23.3.265.1183>
- Kaiser, D., & Haselhuhn, T. (2017). Facing a regular world: How spatial object structure shapes visual processing. *Journal of Neuroscience*, 37(8), 1965–1967. <https://doi.org/10.1523/jneurosci.3441-16.2017>
- Kaltwasser, L., Hildebrandt, A., Recio, G., Wilhelm, O., & Sommer, W. (2013). Neurocognitive mechanisms of individual differences in face cognition: A replication and extension. *Cognitive, Affective, & Behavioural Neuroscience*, 14, 861–878. <https://doi.org/10.3758/s13415-013-0234-y>
- Kanwisher, N., & Barton, J. J. S. (2011). The functional architecture of the face system: Integrating evidence from fMRI and patient studies. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *Oxford handbook of face perception* (pp. 111–129). New York, NY: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199559053.013.0042>
- 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>
- 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. <https://doi.org/10.1016/j.neuropsychologia.2018.03.026>

- Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., & Baker, C. I. (2009). Circular analysis in systems neuroscience: The dangers of double dipping. *Nature Neuroscience*, 12(5), 535–540. <https://doi.org/10.1038/nn.2303>
- Landis, T., Cummings, J. L., Benson, D. F., & Palmer, E. P. (1986). Loss of topographic familiarity: An environmental agnosia. *Archives of Neurology*, 43(2), 132–136. <https://doi.org/10.1001/archneur.1986.00520020026011>
- Lange, J., Lussanet, M. de, Kuhlmann, S., Zimmermann, A., Lappe, M., Zwitserlood, P., & Dobel, C. (2009). Impairments of biological motion perception in congenital prosopagnosia. *PLOS ONE*, 4(10). <https://doi.org/10.1371/journal.pone.0007414>
- Le Grand, R., Cooper, P. A., Mondloch, C. J., Lewis, T. L., Sagiv, N., de Gelder, B., & Maurer, D. (2006). What aspects of face processing are impaired in developmental prosopagnosia? *Brain and Cognition*, 61(2), 139–158. <https://doi.org/10.1016/j.bandc.2005.11.005>
- Lee, Y., Duchaine, B., Wilson, H. R., & Nakayama, K. (2010). Three cases of developmental prosopagnosia from one family: Detailed neuropsychological and psychophysical investigation of face processing. *Cortex*, 46(8), 949–964. <https://doi.org/10.1016/j.cortex.2009.07.012>
- Lerche, V., Voss, A., & Nagler, M. (2017). How many trials are required for parameter estimation in diffusion modeling? A comparison of different optimization criteria. *Behavior Research Methods*, 49(2), 513–537. <https://doi.org/10.3758/s13428-016-0740-2>
- Levine, D. N., & Calvanio, R. (1989). Prosopagnosia: A defect in visual configural processing. *Brain and Cognition*, 10(2), 149–170. [https://doi.org/10.1016/0278-2626\(89\)90051-1](https://doi.org/10.1016/0278-2626(89)90051-1)

- Lissauer, H. (1890). Ein Fall von Seelenblindheit nebst einem Beitrage zur Theorie derselben [A case of visual agnosia with a contribution to theory]. *Archiv für Psychiatrie und Nervenkrankheiten*, 21(2), 222–270. <https://doi.org/10.1007/BF02226765>
- Liu, C. H., Bhuiyan, M. A.-A., Ward, J., & Sui, J. (2009). Transfer between pose and illumination training in face recognition. *Journal of Experimental Psychology. Human Perception and Performance*, 35(4), 939–947. <https://doi.org/10.1037/a0013710>
- Macmillan, N. A., & Creelman, C. D. (1990). Response bias: Characteristics of detection theory, threshold theory, and “nonparametric” indexes. *Psychological Bulletin*, 107(3), 401–413. <https://doi.org/10.1037/0033-2909.107.3.401>
- Maurer, D., Grand, R. L., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, 6(6), 255–260. [https://doi.org/10.1016/s1364-6613\(02\)01903-4](https://doi.org/10.1016/s1364-6613(02)01903-4)
- McCaffery, J. M., Robertson, D. J., Young, A. W., & Burton, A. M. (2018). Individual differences in face identity processing. *Cognitive Research: Principles and Implications*, 3(1), 21. <https://doi.org/10.1186/s41235-018-0112-9>
- McConachie, H. R. (1976). Developmental prosopagnosia. A single case report. *Cortex*, 12(1), 76–82. [https://doi.org/10.1016/S0010-9452\(76\)80033-0](https://doi.org/10.1016/S0010-9452(76)80033-0)
- McKone, E., Hall, A., Pidcock, M., Palermo, R., Wilkinson, R. B., Rivolta, D., Yovel, G., Davis, J. M., & O’Connor, K. B. (2011). Face ethnicity and measurement reliability affect face recognition performance in developmental prosopagnosia: Evidence from the Cambridge Face Memory Test–Australian. *Cognitive Neuropsychology*, 28(2), 109–146. <https://doi.org/10.1080/02643294.2011.616880>
- McKone, E., & Yovel, G. (2009). Why does picture-plane inversion sometimes dissociate perception of features and spacing in faces, and sometimes not? Toward a new theory of

holistic processing. *Psychonomic Bulletin & Review*, 16(5), 778–797.

<https://doi.org/10.3758/PBR.16.5.778>

Megreya, A. M., & Burton, A. M. (2007). Hits and false positives in face matching: A familiarity-based dissociation. *Perception & Psychophysics*, 69(7), 1175–1184.

<https://doi.org/10.3758/BF03193954>

Meyer, K., Schmitz, F., Wilhelm, O., & Hildebrandt, A. (2019). Perceiving faces: Too much, too fast?—Face specificity in response caution. *Journal of Experimental Psychology: Human Perception and Performance*, 45(1), 16–38. <https://doi.org/10.1037/xhp0000582>

Meyers, E. M., Borzello, M., Freiwald, W. A., & Tsao, D. (2015). Intelligent information loss: the coding of facial identity, head pose, and non-face information in the macaque face patch system. *The Journal of Neuroscience*, 35(18), 7069–7081.

<https://doi.org/10.1523/jneurosci.3086-14.2015>

Milner, A. D., Perrett, D. I., Johnston, R. S., Benson, P. J., Jordan, T. R., Heeley, D. W., Bettucci, D., Mortara, F., Mutani, R., Terazzi, E., & Davidson, D. L. W. (1991). Perception and action in ‘visual form agnosia’. *Brain*, 114(1), 405–428.

<https://doi.org/10.1093/brain/114.1.405>

Mishra, M., Fry, R., Saad, E., Arizpe, J., Ohashi, Y.-G., & DeGutis, J. (2020). Comparing the sensitivity of face matching assessments to detect face perception deficits. *PsyArXiv*.

<https://doi.org/10.31234/osf.io/68gbm>

Moeller, S., Freiwald, W. A., & Tsao, D. Y. (2008). Patches with links: A unified system for processing faces in the macaque temporal lobe. *Science*, 320(5881), 1355–1359.

<https://doi.org/10.1126/science.1157436>

Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia

but normal face recognition. *Journal of Cognitive Neuroscience*, 9(5), 555–604.

<https://doi.org/10.1162/jocn.1997.9.5.555>

Murray, E., & Bate, S. (2020). Diagnosing developmental prosopagnosia: Repeat assessment using the Cambridge Face Memory Test. *Royal Society Open Science*, 7(9), 200884.

<https://doi.org/10.1098/rsos.200884>

Murray, E., Hills, P. J., Bennetts, R. J., & Bate, S. (2018). Identifying hallmark symptoms of developmental prosopagnosia for non-experts. *Scientific Reports*, 8(1), 1690.

<https://doi.org/10.1038/s41598-018-20089-7>

Natu, V., & O'Toole, A. (2011). The neural processing of familiar and unfamiliar faces: A review and synopsis. *British Journal of Psychology*, 102, 726–747.

<https://doi.org/10.1111/j.2044-8295.2011.02053.x>

Neiworth, J. J., Gleichman, A. J., Olinick, A. S., & Lamp, K. E. (2006). Global and local processing in adult humans (*Homo sapiens*), 5-year-old children (*Homo sapiens*), and adult cotton-top tamarins (*Saguinus oedipus*). *Journal of Comparative Psychology*,

120(4), 323–330. <https://doi.org/10.1037/0735-7036.120.4.323>

Nosofsky, R. M., Little, D. R., Donkin, C., & Fific, M. (2011). Short-term memory scanning viewed as exemplar-based categorization. *Psychological Review*, 118(2), 280–315.

<https://doi.org/10.1037/a0022494>

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, 104(2), 266–300. [https://doi.org/10.1037/0033-](https://doi.org/10.1037/0033-295X.104.2.266)

[295X.104.2.266](https://doi.org/10.1037/0033-295X.104.2.266)

Nunn, J. A., Postma, P., & Pearson, R. (2001). Developmental prosopagnosia: Should it be taken at face value? *Neurocase*, 7(1), 15–27. <https://doi.org/10.1093/neucas/7.1.15>

Palermo, R., Rossion, B., Rhodes, G., Laguesse, R., Tez, T., Hall, B., Albonico, A.,

Malaspina, M., Daini, R., Irons, J., Al-Janabi, S., Taylor, L. C., Rivolta, D., & McKone,

- E. (2017). Do people have insight into their face recognition abilities? *Quarterly Journal of Experimental Psychology*, 70(2), 218–233.
<https://doi.org/10.1080/17470218.2016.1161058>
- 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>
- Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision*, 5(5), 1–1.
<https://doi.org/10.1167/5.5.1>
- Parr, L. A., Hecht, E., Barks, S. K., Preuss, T. M., & Votaw, J. R. (2009). Face processing in the chimpanzee brain. *Current Biology*, 19(1), 50–53.
<https://doi.org/10.1016/j.cub.2008.11.048>
- Parr, L. A., & Heintz, M. (2006). The perception of unfamiliar faces and houses by chimpanzees: Influence of rotation angle. *Perception*, 35(11), 1473–1483.
<https://doi.org/10.1068/p5455>
- Parr, L. A., Heintz, M., & Akamagwuna, U. (2006). Three studies on configural face processing by chimpanzees. *Brain and Cognition*, 62(1), 30–42.
<https://doi.org/10.1016/j.bandc.2006.03.006>
- Parr, L. A., Winslow, J. T., Hopkins, W. D., & de Waal, F. B. M. (2000). Recognizing facial cues: Individual discrimination by chimpanzees (*Pan troglodytes*) and rhesus monkeys (*Macaca mulatta*). *Journal of Comparative Psychology*, 114(1), 47–60.
<https://doi.org/10.1037/0735-7036.114.1.47>
- Parr, L., & Hecht, E. (2011). Face perception in non-human primates. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *Oxford handbook of face perception* (pp.

177–194). New York, NY: Oxford University Press.

<https://doi.org/10.1093/oxfordhb/9780199559053.013.0042>

Pertsov, Y., Krill, D., Weiss, N., Lesinger, K., & Avidan, G. (2020). Rapid forgetting of faces in congenital prosopagnosia. *Cortex*, 129, 119–132.

<https://doi.org/10.1016/j.cortex.2020.04.007>

Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items.

Journal of Experimental Psychology, 58(3), 193–198. <https://doi.org/10.1037/h0049234>

Peterson, M. F., Zaun, I., Hoke, H., Jiahui, G., Duchaine, B., & Kanwisher, N. (2019). Eye movements and retinotopic tuning in developmental prosopagnosia. *Journal of Vision*, 19(9), 7–7. <https://doi.org/10.1167/19.9.7>

Piepers, D., & Robbins, R. (2012). A review and clarification of the terms “holistic,” “configural,” and “relational” in the face perception literature. *Frontiers in Psychology*, 3. <https://doi.org/10.3389/fpsyg.2012.00559>

Pitcher, D., Charles, L., Devlin, J. T., Walsh, V., & Duchaine, B. (2009). Triple dissociation of faces, bodies, and objects in extrastriate cortex. *Current Biology*, 19(4), 319–324. <https://doi.org/10.1016/j.cub.2009.01.007>

Pitcher, D., Duchaine, B., Walsh, V., Yovel, G., & Kanwisher, N. (2011). The role of lateral occipital face and object areas in the face inversion effect. *Neuropsychologia*, 49(12), 3448–3453. <https://doi.org/10.1016/j.neuropsychologia.2011.08.020>

Pitcher, D., Goldhaber, T., Duchaine, B., Walsh, V., & Kanwisher, N. (2012). Two critical and functionally distinct stages of face and body perception. *Journal of Neuroscience*, 32(45), 15877–15885. <https://doi.org/10.1523/jneurosci.2624-12.2012>

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>
- Pokorny, J. J., & de Waal, F. B. M. (2009). Face recognition in capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology*, 123(2), 151–160. <https://doi.org/10.1037/a0014073>
- Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafò, M. R., Nichols, T. E., Poline, J.-B., Vul, E., & Yarkoni, T. (2017). Scanning the horizon: Towards transparent and reproducible neuroimaging research. *Nature Reviews Neuroscience*, 18(2), 115–126. <https://doi.org/10.1038/nrn.2016.167>
- Powell, G., Jones, C., Hedge, C., Charman, T., Happe, F., Simonoff, E., & Sumner, P. (2019). Face processing in autism spectrum disorder re-evaluated through diffusion models. *Neuropsychology*, 33. <https://doi.org/10.1037/neu0000524>
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59–108. <https://doi.org/10.1037/0033-295X.85.2.59>
- Ratcliff, R., & McKoon, G. (2007). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20(4), 873–922. <https://doi.org/10.1162/neco.2008.12-06-420>
- Ratcliff, G., & Newcombe, F. (1982). Object recognition: Some deductions from the clinical evidence. In A. W. Ellis (Eds.), *Normality and pathology in cognitive functions* (pp. 147–171). London, UK: Academic Press.
- Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, 20(4), 260–281. <https://doi.org/10.1016/j.tics.2016.01.007>

- Ratcliff, R., Thapar, A., & Mckoon, G. (2003). A diffusion model analysis of the effects of aging on brightness discrimination. *Perception & Psychophysics*, 65(4), 523–535.
<https://doi.org/10.3758/BF03194580>
- Renzi, C., Schiavi, S., Carbon, C.-C., Vecchi, T., Silvanto, J., & Cattaneo, Z. (2013). Processing of featural and configural aspects of faces is lateralized in dorsolateral prefrontal cortex: A TMS study. *NeuroImage*, 74, 45–51.
<https://doi.org/10.1016/j.neuroimage.2013.02.015>
- Rezlescu, C., Danaila, I., Miron, A., & Amariei, C. (2020). Chapter 13 - More time for science: Using Testable to create and share behavioral experiments faster, recruit better participants, and engage students in hands-on research. In B. L. Parkin (Ed.), *Progress in Brain Research* (Vol. 253, pp. 243–262). Elsevier.
<https://doi.org/10.1016/bs.pbr.2020.06.005>
- Rezlescu, C., Susilo, T., Wilmer, J. B., & Caramazza, A. (2017). The inversion, part-whole, and composite effects reflect distinct perceptual mechanisms with varied relationships to face recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 43(12), 1961. <https://doi.org/10.1037/xhp0000400>
- Richler, J. J., & Gauthier, I. (2014). A meta-analysis and review of holistic face processing. *Psychological Bulletin*, 140(5), 1281–1302. <https://doi.org/10.1037/a0037004>
- Riddoch, M. J., Johnston, R. A., Bracewell, R. M., Boutsen, L., & Humphreys, G. W. (2008). Are faces special? A case of pure prosopagnosia. *Cognitive Neuropsychology*, 25(1), 3–26. <https://doi.org/10.1080/02643290801920113>
- Righart, R., & Gelder, B. de. (2007). Impaired face and body perception in developmental prosopagnosia. *Proceedings of the National Academy of Sciences*, 104(43), 17234–17238. <https://doi.org/10.1073/pnas.0707753104>

Ritchie, K. L., & Burton, A. M. (2017). Learning faces from variability. *Quarterly Journal of Experimental Psychology*, 70(5), 897–905.

<https://doi.org/10.1080/17470218.2015.1136656>

Rivolta, D., Palermo, R., Schmalzl, L., & Williams, M. A. (2012). Investigating the features of the M170 in congenital prosopagnosia. *Frontiers in Human Neuroscience*, 6.

<https://doi.org/10.3389/fnhum.2012.00045>

Robertson, D. J., Noyes, E., Dowsett, A. J., Jenkins, R., & Burton, A. M. (2016). Face recognition by metropolitan police super-recognisers. *PLOS ONE*, 11(2).

<https://doi.org/10.1371/journal.pone.0150036>

Robins, E., Susilo, T., Ritchie, K., & Devue, C. (2018). Within-person variability promotes learning of internal facial features and facilitates perceptual discrimination and memory.

Open Science Framework. <https://doi.org/10.31219/osf.io/5scnm>

Rossion, B., Caldara, R., Seghier, M., Schuller, A., Lazeyras, F., & Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain*, 126(11), 2381–2395.

<https://doi.org/10.1093/brain/awg241>

Rossion, B., & Jacques, C. (2012). The N170: Understanding the time course of face perception in the human brain. In S. J. Luck., & E. S. Kappenman (Eds.), *Oxford handbook of event-related potential components* (pp.115–141). Oxford University Press.

Rossion, B., & Michel, C. (2018). Normative accuracy and response time data for the computerized Benton Facial Recognition Test (BFRT-c). *Behavior Research Methods*,

50(6), 2442–2460. <https://doi.org/10.3758/s13428-018-1023-x>

Rossion, B., & Taubert, J. (2019). What can we learn about human individual face recognition from experimental studies in monkeys? *Vision Research*, 157, 142–158.

<https://doi.org/10.1016/j.visres.2018.03.012>

- Rousselet, G. A., Pernet, C. R., & Wilcox, R. R. (2017). Beyond differences in means: Robust graphical methods to compare two groups in neuroscience. *The European Journal of Neuroscience*, 46(2), 1738–1748. <https://doi.org/10.1111/ejn.13610>
- Rubens, A. B., & Benson, D. F. (1971). Associative visual agnosia. *Archives of Neurology*, 24(4), 305–316. <https://doi.org/10.1001/archneur.1971.00480340037003>
- Russell, R., Duchaine, B., & Nakayama, K. (2009). Super-recognizers: People with extraordinary face recognition ability. *Psychonomic Bulletin & Review*, 16(2), 252–257. <https://doi.org/10.3758/PBR.16.2.252>
- Sacks, O. (2010, August 23). *Face-Blind*. The New Yorker. <https://www.newyorker.com/magazine/2010/08/30/face-blind>
- Sagiv, N., & Bentin, S. (2001). Structural encoding of human and schematic faces: Holistic and part-based processes. *Journal of Cognitive Neuroscience*, 13(7), 937–951. <https://doi.org/10.1162/089892901753165854>
- Schmalzl, L., Palermo, R., & Coltheart, M. (2008a). Cognitive heterogeneity in genetically based prosopagnosia: A family study. *Journal of Neuropsychology*, 2(1), 99–117. <https://doi.org/10.1348/174866407X256554>
- Schmalzl, L., Palermo, R., Green, M., Brunsdon, R., & Coltheart, M. (2008b). Training of familiar face recognition and visual scan paths for faces in a child with congenital prosopagnosia. *Cognitive Neuropsychology*, 25(5), 704–729. <https://doi.org/10.1080/02643290802299350>
- 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>
- Sergent, J., & Signoret, J.-L. (1992). Varieties of functional deficits in prosopagnosia. *Cerebral Cortex*, 2(5), 375–388. <https://doi.org/10.1093/cercor/2.5.375>

- Shah, P., Gaule, A., Gaigg, S. B., Bird, G., & Cook, R. (2015b). Probing short-term face memory in developmental prosopagnosia. *Cortex*, 64, 115–122.
<https://doi.org/10.1016/j.cortex.2014.10.006>
- Shah, P., Gaule, A., Sowden, S., Bird, G., & Cook, R. (2015a). The 20-item prosopagnosia index (PI20): A self-report instrument for identifying developmental prosopagnosia. *Royal Society Open Science*, 2(6). <https://doi.org/10.1098/rsos.140343>
- Spiker, D., Lotspeich, L. J., Dimiceli, S., Myers, R. M., & Risch, N. (2002). Behavioral phenotypic variation in autism multiplex families: Evidence for a continuous severity gradient. *American Journal of Medical Genetics*, 114(2), 129–136.
<https://doi.org/10.1002/ajmg.10188>
- Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, 99(2), 195–231.
<https://doi.org/10.1037/0033-295x.99.2.195>
- Stollhoff, R., Jost, J., Elze, T., & Kennerknecht, I. (2011). Deficits in long-term recognition memory reveal dissociated subtypes in congenital prosopagnosia. *PLOS ONE*, 6(1), <https://doi.org/10.1371/journal.pone.0015702>
- Stumps, A., Saad, E., Rothlein, D., Verfaellie, M., & DeGutis, J. (2020). Characterizing developmental prosopagnosia beyond face perception: Impaired recollection but intact familiarity recognition. *PsyArXiv*. <https://doi.org/10.31234/osf.io/xt6y4>
- 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>
- Susilo, T., McKone, E., Dennett, H., Darke, H., Palermo, R., Hall, A., Pidock, M., Dawel, A., Jeffery, L., Wilson, E., & Rhodes, G. (2010). Face recognition impairments despite normal holistic processing and face space coding: Evidence from a case of

developmental prosopagnosia. *Cognitive Neuropsychology*, 27(8), 636–664.

<https://doi.org/10.1080/02643294.2011.613372>

Tanaka, J. W., Curran, T., Porterfield, A. L., & Collins, D. (2006). Activation of preexisting and acquired face representations: The N250 event-related potential as an index of face familiarity. *Journal of Cognitive Neuroscience*, 18(9), 1488–1497.

<https://doi.org/10.1162/jocn.2006.18.9.1488>

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology Section A*, 46(2), 225–245.

<https://doi.org/10.1080/14640749308401045>

Tanaka, J. W., & Gordon, I. (2011). Features, configuration, and holistic face processing. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *Oxford handbook of face perception* (pp. 177–194). New York, NY: Oxford University Press.

<https://doi.org/10.1093/oxfordhb/9780199559053.013.0042>

Thapar, A., Ratcliff, R., & McKoon, G. (2003). A diffusion model analysis of the effects of aging on letter discrimination. *Psychology and Aging*, 18(3), 415–429.

<https://doi.org/10.1037/0882-7974.18.3.415>

Thornton, T. L., & Gilden, D. L. (2007). Parallel and serial processes in visual search.

Psychological Review, 114(1), 71–103. <https://doi.org/10.1037/0033-295X.114.1.71>

Tippett, L. J., Miller, L. A., & Farah, M. J. (2000). Prosopamnesia: A selective impairment in face learning. *Cognitive Neuropsychology*, 17(1–3), 241–255.

<https://doi.org/10.1080/026432900380599>

Todorov, A., & Duchaine, B. (2008). Reading trustworthiness in faces without recognizing faces. *Cognitive Neuropsychology*, 25(3), 395–410.

<https://doi.org/10.1080/02643290802044996>

- Torfs, K., Vancleef, K., Lafosse, C., Wagemans, J., & de-Wit, L. (2014). The Leuven Perceptual Organization Screening Test (L-POST), an online test to assess mid-level visual perception. *Behavior Research Methods*, 46(2), 472–487.
<https://doi.org/10.3758/s13428-013-0382-6>
- Towler, J., & Eimer, M. (2012). Electrophysiological studies of face processing in developmental prosopagnosia: Neuropsychological and neurodevelopmental perspectives. *Cognitive Neuropsychology*, 29(5–6), 503–529.
<https://doi.org/10.1080/02643294.2012.716757>
- Towler, J., Fisher, K., & Eimer, M. (2017). The cognitive and neural basis of developmental prosopagnosia. *Quarterly Journal of Experimental Psychology*, 70(2), 316–344.
<https://doi.org/10.1080/17470218.2016.1165263>
- Towler, J., Fisher, K., & Eimer, M. (2018). Holistic face perception is impaired in developmental prosopagnosia. *Cortex*, 108, 112–126.
<https://doi.org/10.1016/j.cortex.2018.07.019>
- 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>
- Ulrich, P. I. N., Wilkinson, D., Ferguson, H., Smith, L., Bindemann, M., Johnston, R., & Schmalzl, L. (2017). Perceptual and memorial contributions to developmental prosopagnosia. *Quarterly Journal of Experimental Psychology*, 70(2), 298–315.
<https://doi.org/10.1080/17470218.2016.1177101>
- Vadillo, M. A., Konstantinidis, E., & Shanks, D. R. (2016). Underpowered samples, false negatives, and unconscious learning. *Psychonomic Bulletin & Review*, 23(1), 87–102.
<https://doi.org/10.3758/s13423-015-0892-6>

- Veatch, O. J., Veenstra-VanderWeele, J., Potter, M., Pericak-Vance, M. A., & Haines, J. L. (2014). Genetically meaningful phenotypic subgroups in autism spectrum disorders. *Genes, Brain and Behavior*, 13(3), 276–285. <https://doi.org/10.1111/gbb.12117>
- Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory and Cognition*, 32, 1206–1220. <https://doi.org/10.3758/BF03196893>
- Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods*, 39(4), 767–775. <https://doi.org/10.3758/BF03192967>
- Voss, A., Voss, J., & Lerche, V. (2015). Assessing cognitive processes with diffusion model analyses: A tutorial based on fast-dm-30. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00336>
- Wallace, S., Coleman, M., & Bailey, A. (2008). Face and object processing in autism spectrum disorders. *Autism Research*, 1(1), 43–51. <https://doi.org/10.1002/aur.7>
- Wang, L. A. L., Herrington, J. D., Tunç, B., & Schultz, R. T. (2020). Bayesian regression-based developmental norms for the Benton Facial Recognition Test in males and females. *Behavior Research Methods*, 52(4), 1516–1527. <https://doi.org/10.3758/s13428-019-01331-0>
- Wang, R., Li, J., Fang, H., Tian, M., & Liu, J. (2012). Individual differences in holistic processing predict face recognition ability. *Psychological Science*, 23(2), 169–177. <https://doi.org/10.1177/0956797611420575>
- Weigelt, S., Koldewyn, K., Dilks, D. D., Balas, B., McKone, E., & Kanwisher, N. (2014). Domain-specific development of face memory but not face perception. *Developmental Science*, 17(1), 47–58. <https://doi.org/10.1111/desc.12089>

- Weigelt, S., Koldewyn, K., & Kanwisher, N. (2012). Face identity recognition in autism spectrum disorders: A review of behavioral studies. *Neuroscience & Biobehavioral Reviews*, 36(3), 1060–1084. <https://doi.org/10.1016/j.neubiorev.2011.12.008>
- Weigelt, S., Koldewyn, K., & Kanwisher, N. (2013). Face recognition deficits in autism spectrum disorders are both domain specific and process specific. *PLOS ONE*, 8(9). <https://doi.org/10.1371/journal.pone.0074541>
- White, D., Kemp, R. I., Jenkins, R., Matheson, M., & Burton, A. M. (2014). Passport officers' errors in face matching. *PLOS ONE*, 9(8). <https://doi.org/10.1371/journal.pone.0103510>
- White, D., Rivolta, D., Burton, A. M., Al-Janabi, S., & Palermo, R. (2017). Face Matching Impairment in Developmental Prosopagnosia. *Quarterly Journal of Experimental Psychology*, 70(2), 287–297. <https://doi.org/10.1080/17470218.2016.1173076>
- Wilcockson, T. D. W., Burns, E. J., Xia, B., Tree, J., & Crawford, T. J. (2020). Atypically heterogeneous vertical first fixations to faces in a case series of people with developmental prosopagnosia. *Visual Cognition*, 28(4), 311–323. <https://doi.org/10.1080/13506285.2020.1797968>
- Wilcox, R. R. (1995). Comparing two independent groups via multiple quantiles. *Journal of the Royal Statistical Society*, 44, 91–99.
- Wilcox, R. R., & Rousselet, G. A. (2018). A guide to robust statistical methods in neuroscience. *Current Protocols in Neuroscience*, 82(1), 8.42.1–8.42.30
- Wilhelm, O., Herzmann, G., Kunina, O., Danthiir, V., Schacht, A., & Sommer, W. (2010). Individual differences in perceiving and recognizing faces – One element of social cognition. *Journal of Personality and Social Psychology*, 99(3), 530–548. <https://doi.org/10.1037/a0019972>

- Williams, M. A., Berberovic, N., & Mattingley, J. B. (2007). Abnormal fMRI adaptation to unfamiliar faces in a case of developmental prosopamnesia. *Current Biology*, 17(14), 1259–1264. <https://doi.org/10.1016/j.cub.2007.06.042>
- Wilson, C. E., Brock, J., & Palermo, R. (2010). Attention to social stimuli and facial identity recognition skills in autism spectrum disorder. *Journal of Intellectual Disability Research*, 54(12), 1104–1115. <https://doi.org/10.1111/j.1365-2788.2010.01340.x>
- Wilson, C. Ellie, Freeman, P., Brock, J., Burton, A. M., & Palermo, R. (2010). Facial identity recognition in the broader autism phenotype. *PLOS ONE*, 5(9). <https://doi.org/10.1371/journal.pone.0012876>
- Wilson, D. A., & Tomonaga, M. (2018). Visual discrimination of primate species based on faces in chimpanzees. *Primates*, 59(3), 243–251. <https://doi.org/10.1007/s10329-018-0649-8>
- Woods, A. T., Velasco, C., Levitan, C. A., Wan, X., & Spence, C. (2015). Conducting perception research over the internet: a tutorial review. *PeerJ*, 3. <https://doi.org/10.7717/peerj.1058>
- 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>
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, 81(1), 141–145. <https://doi.org/10.1037/h0027474>
- Yovel, G., & Freiwald, W. A. (2013). Face recognition systems in monkey and human: Are they the same thing? *F1000prime Reports*, 5, 10. <https://doi.org/10.12703/P5-10>
- Zeguers, M., Snellings, P., Tijms, J., Weeda, W., Tamboer, P., Bexkens, A., & Huizenga, H. m. (2011). Specifying theories of developmental dyslexia: A diffusion model analysis of

word recognition. *Developmental Science*, 14, 1340–1354.

<https://doi.org/10.1111/j.1467-7687.2011.01091.x>

Appendix A: Literature review of face perception studies in developmental prosopagnosia

A.1 Literature review

I conducted a selective literature review on face perception studies in developmental prosopagnosia (Table A1). The purpose of the review was to motivate the use of BFRT, GFMT, and CFPT in my thesis, so I only included articles using these tests. I excluded articles if: (1) data from any of the same developmental prosopagnosics had been reported in an earlier paper, (2) only average performance was provided *without* any statistical significance testing, (3) the article was a thesis or conference abstract, or (4) it was not written in English. Overall, 68 articles were included. References for articles not cited in the main text are available in section A.2.

The results from the studies are mixed. At the single-case level, deficits in a minority of the sample were reported in 26 studies, in about half the sample in 27 studies, and in the majority of the sample in two studies. At the group level, a complete lack of deficits were reported in two studies, and widespread deficits across the whole sample were reported in 11 studies. The mode sample size was 11 developmental prosopagnosics, the median was 11, and the mean was 14.62. Only five studies used multiple tests, and only four studies analysed response times. Of the nine studies comparing face perception deficits relative to face memory deficits, only two retested face memory using independent tests, and both used in-house tests rather than standardised tests. Overall, my review exemplified the critical need for more studies with much larger sample size, independent testing of face memory with standardised measures, the use of multiple face perception tests, and response time analyses.

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

Table A1

Summary of Prior Studies Using the BFRT, GFMT, and/or CFPT in Developmental Prosopagnosia (Sorted by Scope of Deficits)

Authors	Year	Tests	Scope of deficits	Prosopagnosic sample size	Reaction time data?	Independent face memory test?
Bate et al.	2008	BFRT	All (1/1) ^c	1	No	Yes
Biotti et al.	2019	CFPT	All (significant difference)	72	No	
de Gelder & Rouw	2000	BFRT	All (1/1) ^c	1	No	
Duchaine et al.	2009	CFPT	All (significant difference)	7	No	
Duchaine et al.	2007	CFPT	All (significant difference)	14	No	
Duchaine et al.	2010	CFPT	All (2/2) ^b	2	No	No
Huis in't Veld et al.	2012	BFRT	All (significant difference)	10	No	
Logan et al.	2016	GFMT	All (1/1)	1	No	
Russell et al.	2012	CFPT	All (significant difference)	10	No	
Shah, et al.	2015b	CFPT	All (significant difference)	16	No	
Susilo et al.	2010	CFPT	All (1/1)	1	No	No
Biotti et al.	2016	CFPT	Majority (16/17)	17	No	
Chatterjee & Nakayama	2012	CFPT	Majority (13/18)	18	No	
White et al	2017	GFMT; CFPT	Mixed (GFMT: 0/6; CFPT: 2/6)	6	GFMT: yes; CFPT: no	
Bate et al.	2019b	CFPT	Mixed (32/63) ^b	63	No	
Biotti, et al.	2017	CFPT	Mixed (13/20)	13	No	No
Bobak et al.	2017	CFPT	Mixed (5/10)	10	No	
Bowles et al.	2009	CFPT	Mixed (3/5) ^a	5	No	
Burns et al.	2014	CFPT	Mixed (5/8)	8	No	
Cenac et al.	2019	CFPT	Mixed (8/15)	15	No	

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

Corrow et al.	2019	CFPT	Mixed (5/10)	10	No	
Daini et al.	2014	BFRT	Mixed (4/6)	6	No	
Dalrymple et al.	2014	CFPT (or DFPT)	Mixed (children: 8/8; adults: 5/16) ^b	8	Adults: yes; children: no	No
DeGutis et al.	2012a	CFPT	Mixed (24/38)	38	No	
DeGutis, et al.	2012b	CFPT	Mixed (6/10)	10	No	
DeGutis et al.	2014	CFPT	Mixed (14/24)	24	No	
Garrido et al.	2011	CFPT	Mixed (6/14) ^b	14	No	
Gerlach et al	2016	CFPT	Mixed (6/10)	10	No	
Gerlach et al.	2019	CFPT	Mixed (8/15) ^b	15	No	
Golan et al	2013	CFPT	Mixed (7/12)	12	No	
Jackson et al.	2017	CFPT	Mixed (4/10)	10	No	
Liu et al.	2015	CFPT	Mixed (5/12)	13	No	
Marsh et al.	2019	CFPT	Mixed (9/17)	17	No	
Murray et al.	2018	CFPT	Mixed (34/50)	50	No	
Nunn et al.	2001	BFRT	Mixed (0/1) ^c	1	Yes	
Rubino et al.	2016	CFPT	Mixed (4/10)	10	No	
Shah et al.	2015a	CFPT	Mixed (9/18)	18	No	
Starrfelt et al.	2016	CFPT	Mixed (6/10)	10	No	
Tanzer et al.	2016	CFPT	Mixed (6/12)	12	No	
Yamanashi Leib et al.	2012	BFRT & CFPT	Mixed (BFRT: 2/4; CFPT: 4/4) ^b	4	No	
Albonico et al.	2017	BFRT	Minority (4/32)	32	No	
Avidan et al.	2011	CFPT	Minority (3/11)	14	No	
Awasthi et al	2012	CFPT	Minority (3/7)	7	No	
Bate et al.	2019a	CFPT	Minority (14/40)	40	No	
Biotti, et al.	2017	CFPT	Minority (Exp1: 5/16; Exp 2: 14/24)	16; 24	No	
Burns et al.	2017	CFPT	Minority (2/10)	10	No	
Cattaneo et al.	2016	BFRT	Minority (8/23)	23	No	
Duchaine & Nakayama	2004	BFRT	Minority (1/11)	11	No	No

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

Duchaine & Nakayama	2006	BFRT	Minority (1/6) ^a	6	No	
Eimer et al.	2012	CFPT	Minority (2/12)	12	No	
Fisher et al.	2016	CFPT	Minority (4/11)	11	No	
Fisher et al.	2017	CFPT	Minority (3/12)	12	No	
Klargaard, et al.	2018	CFPT	Minority (3/16)	16	No	
Lee et al.	2010	CFPT	Minority (1/3)	3	No	
Malaspina et al.	2016	BFRT	Minority (1/10)	10	No	
Malaspina et al.	2017	BFRT	Minority (3/12)	12	No	
Mishra et al.	2020	BFRT & CFPT	Minority (BFRT: 6/30; CFPT: 7/30)	30	Yes	No
Palermo et al.	2011	CFPT & GFMT	Minority (CFPT: 3/12; GFMT: 0/12)	12	No	
Palermo et al.	2011	CFPT	Minority (4/14)	14	No	
Parke et al.	2015	CFPT	Minority (2/10)	10	No	
Rivolta et al.	2012	CFPT	Minority (2/11)	11	No	
Rivolta et al.	2014	CFPT	Minority (2/7)	7	No	
Robson et al.	2018	CFPT	Minority (3/11)	11	No	
Stumps et al.	2020	BFRT & CFPT	Minority (composite: 3/30 impaired)	30	No	Yes
Tsantani et al.	2020	CFPT	Minority (8/22)	22	No	
Ulrich et al.	2017	CFPT	Minority (1/11)	11	No	No
McKone et al.	2011	CFPT	None (0/6)	6	No	
Tree & Wilkie	2010	BFRT	None (0/4)	4	No	

Note. Bolded studies aimed to investigate face perception deficits in developmental prosopagnosia. Non-bolded studies ran the tests as part of diagnostic testing for developmental prosopagnosics. The bracketed values in the scope of deficits column show the number of impaired individuals out of the total sample. For studies with group analyses (ANOVAs or t-tests) that did not run single-case statistics, I noted whether the group difference was significant ($p < .05$) or not. The remaining studies used z-scores of two standard deviations below control mean as the cut-off for impairment unless otherwise marked. Studies marked with an **a** used z-scores between 1-1.9 standard deviations

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

below the control mean. Studies marked with a **b** used Crawford's t-tests (Crawford & Howell, 1998). Studies marked with a **c** used raw scores below the typical test range as impairment cut-offs.

A.2 Literature review references

- Awasthi, B., Friedman, J., & Williams, M. A. (2012). Reach trajectories reveal delayed processing of low spatial frequency faces in developmental prosopagnosia. *Cognitive Neuroscience*, 3(2), 120–130. <https://doi.org/10.1080/17588928.2012.673482>
- 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>
- 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. <https://doi.org/10.1016/j.neuropsychologia.2017.06.020>
- Burns, E. J., Tree, J. J., & Weidemann, C. T. (2014). Recognition memory in developmental prosopagnosia: Electrophysiological evidence for abnormal routes to face recognition. *Frontiers in Human Neuroscience*, 8. <https://doi.org/10.3389/fnhum.2014.00622>
- Cattaneo, Z., Daini, R., Malaspina, M., Manai, F., Lillo, M., Fermi, V., Schiavi, S., Suchan, B., & Comincini, S. (2016). Congenital prosopagnosia is associated with a genetic variation in the oxytocin receptor (OXTR) gene: An exploratory study. *Neuroscience*, 339, 162–173. <https://doi.org/10.1016/j.neuroscience.2016.09.040>
- Cenac, Z., Biotti, F., Gray, K. L. H., & Cook, R. (2019). Does developmental prosopagnosia impair identification of other-ethnicity faces? *Cortex*, 119, 12–19. <https://doi.org/10.1016/j.cortex.2019.04.007>
- Daini, R., Comparetti, C. M., & Ricciardelli, P. (2014). Behavioral dissociation between emotional and non-emotional facial expressions in congenital prosopagnosia. *Frontiers in Human Neuroscience*, 8. <https://doi.org/10.3389/fnhum.2014.00974>
- DeGutis, J., Chatterjee, G., Mercado, R. J., & Nakayama, K. (2012b). Face gender recognition in developmental prosopagnosia: Evidence for holistic processing and use of configural information. *Visual Cognition*, 20(10), 1242–1253. <https://doi.org/10.1080/13506285.2012.744788>
- DeGutis, J., Cohan, S., Mercado, R. J., Wilmer, J., & Nakayama, K. (2012a). Holistic processing of the mouth but not the eyes in developmental prosopagnosia. *Cognitive Neuropsychology*, 29(5–6), 419–446. <https://doi.org/10.1080/02643294.2012.754745>
- Duchaine, B., Jenkins, R., Germine, L., & Calder, A. J. (2009). Normal gaze discrimination and adaptation in seven prosopagnosics. *Neuropsychologia*, 47(10), 2029–2036. <https://doi.org/10.1016/j.neuropsychologia.2009.03.011>
- Duchaine, B., Murray, H., Turner, M., White, S., & Garrido, L. (2009). Normal social cognition in developmental prosopagnosia. *Cognitive Neuropsychology*, 26(7), 620–634. <https://doi.org/10.1080/02643291003616145>
- Eimer, M., Gosling, A., & Duchaine, B. (2012). Electrophysiological markers of covert face recognition in developmental prosopagnosia. *Brain*, 135(2), 542–554. <https://doi.org/10.1093/brain/awr347>
- Fisher, K., Towler, J., & Eimer, M. (2017). Face identity matching is selectively impaired in developmental prosopagnosia. *Cortex*, 89, 11–27. <https://doi.org/10.1016/j.cortex.2017.01.003>
- Gerlach, C., Klargaard, S. K., Alnæs, D., Kolskår, K. K., Karstoft, J., Westlye, L. T., & Starrfelt, R. (2019). Left hemisphere abnormalities in developmental prosopagnosia when looking at faces but not words. *Brain Communications*, 1. <https://doi.org/10.1093/braincomms/fcz034>
- Golan, T., Bentin, S., DeGutis, J. M., Robertson, L. C., & Harel, A. (2014). Association and dissociation between detection and discrimination of objects of expertise: Evidence from visual search. *Attention, Perception & Psychophysics*, 76(2), 391–406. <https://doi.org/10.3758/s13414-013-0562-6>
- 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. <https://doi.org/10.1016/j.neuropsychologia.2017.09.003>
- Leib, A. Y., Puri, A. M., Fischer, J., Bentin, S., Whitney, D., & Robertson, L. (2012). Crowd perception in prosopagnosia. *Neuropsychologia*, 50(7), 1698–1707. <https://doi.org/10.1016/j.neuropsychologia.2012.03.026>

- Liu, R. R., Corrow, S. L., Pancaroglu, R., Duchaine, B., & Barton, J. J. S. (2015). The processing of voice identity in developmental prosopagnosia. *Cortex*, 71, 390–397. <https://doi.org/10.1016/j.cortex.2015.07.030>
- Logan, A. J., Wilkinson, F., Wilson, H. R., Gordon, G. E., & Loffler, G. (2016). The Caledonian face test: A new test of face discrimination. *Vision Research*, 119, 29–41. <https://doi.org/10.1016/j.visres.2015.11.003>
- Malaspina, M., Albonico, A., & Daini, R. (2016). Right perceptual bias and self-face recognition in individuals with congenital prosopagnosia. *Laterality*, 21(2), 118–142. <https://doi.org/10.1080/1357650X.2015.1084312>
- 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(5), 546–563. <https://doi.org/10.1037/neu0000362>
- Marsh, J. E., Biotti, F., Cook, R., & Gray, K. L. H. (2019). The discrimination of facial sex in developmental prosopagnosia. *Scientific Reports*, 9(1), 19079. <https://doi.org/10.1038/s41598-019-55569-x>
- Palermo, R., Rivolta, D., Wilson, C. E., & Jeffery, L. (2011). Adaptive face space coding in congenital prosopagnosia: Typical figural aftereffects but abnormal identity aftereffects. *Neuropsychologia*, 49(14), 3801–3812. <https://doi.org/10.1016/j.neuropsychologia.2011.09.039>
- Parketny, J., Towler, J., & Eimer, M. (2015). The activation of visual face memory and explicit face recognition are delayed in developmental prosopagnosia. *Neuropsychologia*, 75, 538–547. <https://doi.org/10.1016/j.neuropsychologia.2015.07.009>
- Rivolta, D., Lawson, R. P., & Palermo, R. (2017). More than just a problem with faces: Altered body perception in a group of congenital prosopagnosics. *Quarterly Journal of Experimental Psychology*, 70(2), 276–286. <https://doi.org/10.1080/17470218.2016.1174277>
- Rivolta, D., Palermo, R., Schmalzl, L., & Coltheart, M. (2012). Covert face recognition in congenital prosopagnosia: A group study. *Cortex*, 48(3), 344–352. <https://doi.org/10.1016/j.cortex.2011.01.005>
- Rivolta, D., Woolgar, A., Palermo, R., Butko, M., Schmalzl, L., & Williams, M. A. (2014). Multi-voxel pattern analysis (MVPA) reveals abnormal fMRI activity in both the “core” and “extended” face network in congenital prosopagnosia. *Frontiers in Human Neuroscience*, 8. <https://doi.org/10.3389/fnhum.2014.00925>
- Robson, M. K., Palermo, R., Jeffery, L., & Neumann, M. F. (2018). Ensemble coding of face identity is present but weaker in congenital prosopagnosia. *Neuropsychologia*, 111, 377–386. <https://doi.org/10.1016/j.neuropsychologia.2018.02.019>
- Rubino, C., Corrow, S. L., Corrow, J. C., Duchaine, B., & Barton, J. J. S. (2016). Word and text processing in developmental prosopagnosia. *Cognitive Neuropsychology*, 33(5–6), 315–328. <https://doi.org/10.1080/02643294.2016.1204281>
- Russell, R., Chatterjee, G., & Nakayama, K. (2012). Developmental prosopagnosia and super-recognition: No special role for surface reflectance processing. *Neuropsychologia*, 50(2), 334–340. <https://doi.org/10.1016/j.neuropsychologia.2011.12.004>
- Starrfelt, R., Klargaard, S. K., Petersen, A., & Gerlach, C. (2018). Reading in developmental prosopagnosia: Evidence for a dissociation between word and face recognition. *Neuropsychology*, 32(2), 138–147. <https://doi.org/10.1037/neu0000428>
- Tanzer, M., Weinbach, N., Mardo, E., Henik, A., & Avidan, G. (2016). Phasic alertness enhances processing of face and non-face stimuli in congenital prosopagnosia. *Neuropsychologia*, 89, 299–308. <https://doi.org/10.1016/j.neuropsychologia.2016.06.032>
- Tree, J. J., & Wilkie, J. (2010). Face and object imagery in congenital prosopagnosia: A case series. *Cortex*, 46(9), 1189–1198. <https://doi.org/10.1016/j.cortex.2010.03.005>
- Tsantani, M., Gray, K. L. H., & Cook, R. (2020). Holistic processing of facial identity in developmental prosopagnosia. *Cortex*, 130, 318–326. <https://doi.org/10.1016/j.cortex.2020.06.003>

Appendix B: Supplementary analyses

B.1 Group age and sex comparisons

Developmental prosopagnosic and control gender distributions ($\chi^2(2) = 5.34, p = .069$) and ages ($t(195.31) = 0.40, p = .687$) did not significantly differ, showing that the control group was age- and sex-matched to the developmental prosopagnosic group.

B.2 Development of the BFRT and GFMT trial time limits

BFRT and GFMT trials are traditionally administered with unlimited time. This may allow developmental prosopagnosics to use time-consuming and atypical strategies to improve performance on the tests. To mitigate this issue, I ran multiple pilot studies with control participants to determine an appropriate time limit for both tests, namely a time limit that would allow control participants to (1) produce performance similar to performance with unlimited time and (2) generate substantial face inversion effect.

I started by running the upright and inverted versions to 17 participants from Amazon Mechanical Turk (<https://www.mturk.com/>) with unlimited time. Upright accuracy was good ($M = 83.02\%$, $SD = 8.62\%$) and the median time was 12.55 seconds per trial. There was a substantial inversion effect (inverted $M = 66.82\%$, $SD = 8.09\%$), and the median time for inverted trials was 9.93 seconds. I then tried 8 seconds as the limit for the next pilot with 28 new participants. This limit reduced both upright ($M = 74.14\%$, $SD = 10.84\%$) and inverted ($M = 58.47\%$, $SD = 5.20\%$) accuracy substantially, and led to participants timing-out on around 3 trials in the upright condition (median = 2.5, $M = 3.46$, $SD = 3.00$). I therefore increased the limit to 12 seconds with a third pilot sample ($n = 29$). With this limit, upright ($M = 78.51\%$, $SD = 8.04\%$) and inverted accuracy ($M = 59.85\%$, $SD = 7.66\%$) became more similar to unlimited performance. Response time was 6.70 seconds ($SD = 1.3s$) on upright trials and 6.3 seconds ($SD = 1.8s$) on inverted trials. For upright trials, participants only

timed-out on a median of 1 trial and a mean of 1.75 trials ($SD = 2.06$). These results suggest that 12 seconds was an appropriate time limit.

With the GFMT, I started with 31 participants who did the upright and inverted versions of the test with unlimited time. Upright accuracy was 76.88% ($SD = 18\%$) and inverted accuracy was 62.58% ($SD = 22.04\%$). Mean response time was 2.74 seconds ($SD = 1.51s$) for upright trials and 2.56 seconds ($SD = 1.51s$) for inverted trials. I then tested a 5 second limit with 28 new participants. Performance was similar to that with unlimited time for both upright ($M = 77.13\%$, $SD = 9.45\%$) and inverted ($M = 60.09\%$, $SD = 5.26\%$) trials. Mean response time for upright trials was 2.07 seconds ($SD = 0.85s$) and for inverted trials it was 1.89 seconds ($SD = 0.71s$). For upright trials, participants timed-out on a median of 1 trial and a mean of 0.96 trials ($SD = 1.16s$). These results showed the 5 second time-limit was appropriate.

B.3 Response time analyses with median response time

Medians are less susceptible to outliers than means. Despite all face perception tests having limited trial durations, which reduces the impact of outlier response times, I reanalysed all response time data using medians in addition to the mean analysis that I reported in the main text.

I ran a two (group: prosopagnosic, control) by four (test: BFRT, GFMT, CFPT, CFMT-Aus) ANOVA on median response times to test for speed-accuracy trade-offs. The main effects of group ($F(1, 241) = 95.44$, $p < .001$, $\eta^2_p = .28$) and test ($F(1.07, 258.39) = 2133.30$, $p < .001$, $\eta^2_p = .90$) were qualified by a significant interaction ($F(1.07, 258.39) = 51.87$, $p < .001$, $\eta^2_p = .18$). Post-hoc analyses revealed that developmental prosopagnosics took longer to respond on both the BFRT ($M = 8.86s$; $SD = 1.94s$) and CFPT ($M = 34.4s$; $SD = 8.52s$) compared to controls on the BFRT ($M = 6.40s$; $SD = 1.35s$; $t(935.02) = 4.15s$, $p = .001$) and CFPT ($M = 25.30s$; $SD = 9.46s$; $t(935.02) = 52.15$, $p < .001$). There was no

difference between developmental prosopagnosics ($M = 2.62\text{s}$; $SD = 0.65\text{s}$) and controls on the GFMT ($M = 2.01\text{s}$; $SD = 0.55\text{s}$; $t(935.02) = 1.04$, $p = 1.000$), and for developmental prosopagnosics ($M = 3.46\text{s}$; $SD = 1.41\text{s}$) and controls on the CFMT-Aus ($M = 2.57\text{s}$; $SD = 0.82\text{s}$; $t(935.02) = 1.52$, $p = 1.000$). The general pattern of developmental prosopagnosics taking longer to respond was therefore replicated using median response time, so I used median response time as a covariate in the following accuracy analyses.

I ran a two (group: prosopagnosic, control) by four (test: BFRT, GFMT, CFPT, CFMT-Aus) ANCOVA on accuracy with median response time as a covariate. All descriptive statistics are available in Table 2 in section 3.2. There were significant main effects of test ($F(3, 966) = 8.97$, $p < .001$, $\eta^2_p = .03$) and group ($F(1, 966) = 450.88$, $p < .001$, $\eta^2_p = .32$), and a significant interaction between test and group ($F(3, 966) = 3.56$, $p = .014$, $\eta^2_p = .01$). Follow-up post-hocs showed that developmental prosopagnosics performed poorly compared to controls on the CFMT-Aus ($t(966) = 13.10$, $p < .001$, $d = 1.69$), BFRT ($t(966) = 12.56$, $p < .001$, $d = 1.63$), GFMT ($t(966) = 10.03$, $p < .001$, $d = 1.30$), and CFPT ($t(966) = 8.24$, $p < .001$, $d = 1.18$). To explore this interaction, I ran three two (group: prosopagnosic, control) by two (test: CFMT-Aus, one of the three perception tests) ANCOVAs comparing accuracy on the CFMT-Aus and each of the three face perception tests with median response times as a covariate. I found an interaction for the CFMT-Aus/CFPT contrast ($F(1, 483) = 4.46$, $p = .035$, $\eta^2_p < .01$) and the CFMT-Aus/GFMT contrast ($F(1, 482) = 4.42$, $p = .036$, $\eta^2_p < .01$), but not for the CFMT-Aus/BFRT contrast ($F(1, 483) = 3.55$, $p = .060$, $\eta^2_p < .01$). These results mean that developmental prosopagnosics had equivalently poor performance across the CFMT-Aus and BFRT, but they performed slightly better on the CFPT and GFMT. Altogether, median response times produce the same pattern as the original analyses.

B.4 Hot deck data imputation

I used hot deck data imputation to address the issue of missing upright match trials ($n = 6$) in the BFRT for 25 developmental prosopagnosics. Hot deck data imputation allowed me to rerun all analyses involving the inverted BFRT with the match trials included.

Size of inversion effects. I ran a two (group: prosopagnosics, control) by three (test: BFRT, GFMT, CFPT) mixed-measures ANOVA for accuracy inversion effect sizes. There was a significant main effect of group ($F(1, 241) = 142.94, p < .001, \eta^2_p = .37$), such that prosopagnosics had significantly smaller inversion effects ($M = 6.80\%, SD = 10.32\%$) than controls ($M = 17.22\%, SD = 11.08\%$). There was also a significant main effect of test ($F(2, 482) = 38.30, p < .001, \eta^2_p = .14$), such that inversion effects were smaller on the GFMT ($M = 8.40\%, SD = 11.62\%$) compared to the BFRT ($M = 13.68\%, SD = 9.74\%$; $t(482) = 6.20, p < .001$) and CFPT ($M = 15.76\%, SD = 13.02\%$; $t(482) = 8.45, p < .001$). This time, inversion effect size did not differ between the BFRT and CFPT ($t(482) = 2.25, p = .074$). Finally, the interaction between group and test was not significant ($F(2, 482) = 0.82, p = .440, \eta^2_p < .01$). Overall, hot deck data imputation led to the same pattern of prosopagnosics having smaller inversion effects than controls, consistent with the original analysis.

Performance on upright and inverted trials. I ran a two (group: prosopagnosics) by two (orientation: upright, inverted) by three (test: BFRT, GFMT, CFPT) mixed ANOVA on participants' accuracy data to see if differences in upright or inverted trials – or both – were driving prosopagnosics' inversion effects. I obtained similar results as the original analysis. The interaction between group, orientation, and test was not significant ($F(1.94, 467.30) = 0.82, p = .436, \eta^2_p < .01$), but there was a significant interaction between group and orientation ($F(1, 467.30) = 142.49, p < .001, \eta^2_p = .37$). Developmental prosopagnosics had much poorer performance than controls on upright trials ($M_{\text{prosopagnosic}} = 60.22\%, SD_{\text{prosopagnosic}} = 11.66\%$; $M_{\text{control}} = 74.36\%, SD_{\text{control}} = 10.99\%$; $t(353.38) = 14.37, p < .001$) but only slightly poorer

performance on inverted trials ($M_{\text{prosopagnosic}} = 53.42\%$, $SD_{\text{prosopagnosic}} = 10.69\%$; $M_{\text{control}} = 57.13\%$, $SD_{\text{control}} = 10.73\%$; $t(353.38) = 3.74$, $p < .001$).

B.5 Analyses without response time covariate

I ran parallel traditional ANOVA analyses for all ANCOVA analyses that used response time as a covariate. Some readers might appreciate ANOVA analyses for easier comparison of my results with prior literature, as using response time as a covariate is not common in past research.

First, I ran a two (group: prosopagnosic, control) by four (test: BFRT, GFMT, CFPT, CFMT-Aus) ANOVA on accuracy. All descriptive statistics are available in Table 2 in section 3.2. The outcome was the same as the original ANCOVA. There were significant main effects of test ($F(3, 726) = 27.56$, $p < .001$, $\eta^2_p = .10$) and group ($F(1, 242) = 196.58$, $p < .001$, $\eta^2_p = .45$), and a significant interaction between test and group ($F(3, 726) = 7.98$, $p < .001$, $\eta^2_p = .03$). Follow-up post-hocs showed that developmental prosopagnosics performed worse than controls on the CFMT-Aus ($t(612.29) = 12.93$, $p < .001$, $d = 1.55$), BFRT ($t(612.29) = 11.52$, $p < .001$, $d = 1.64$), GFMT ($t(612.29) = 10.32$, $p < .001$, $d = 1.39$), and CFPT ($t(612.29) = 7.94$, $p < .001$, $d = 1.01$). To explore this interaction, I ran three two (group: prosopagnosic, control) by two (test: CFMT-Aus, one of the three perception tests) ANOVAs comparing accuracy on the CFMT-Aus and each of the three face perception tests. I found an interaction for the CFMT-Aus/CFPT contrast ($F(1, 242) = 23.49$, $p = .035$, $\eta^2_p = .09$) and the CFMT-Aus/GFMT contrast ($F(1, 242) = 5.25$, $p = .023$, $\eta^2_p = .02$), but not for the CFMT-Aus/BFRT contrast ($F(1, 242) = 1.67$, $p = .198$, $\eta^2_p < .01$). These results mean that developmental prosopagnosics had equivalently poor performance across the CFMT-Aus and BFRT, but they performed slightly better on the CFPT and GFMT.

I then investigated the effect of within- versus between-identity trials on the GFMT. I ran a two (group: prosopagnosic, control) by two (trial type: within, between) mixed

measures ANOVA on the GFMT upright accuracy data. All the descriptive statistics are available in section 3.4.3. There were significant main effects of group ($F(1, 241) = 112.19, p < .001, \eta^2_p = .32$) and trial type ($F(1, 241) = 16.27, p < .001, \eta^2_p = .06$), that were qualified by a significant interaction ($F(1, 241) = 31.72, p < .001, \eta^2_p = .12$). Follow-up tests revealed that developmental prosopagnosics scored poorly on same-identity trials compared to controls ($t(448.15) = 10.87, p < .001, d = 1.37$). On different-identity trials, developmental prosopagnosics again showed similar performance to controls ($t(448.15) = 1.88, p = .364, d = 0.25$). Consistent with the original analyses, this result suggests that face perception deficits in developmental prosopagnosia may be selective to within-identity trials.

Lastly, I analysed whether face perception deficits in developmental prosopagnosia were evident in both viewpoint changes and lighting changes on the BFRT. I ran a two (group: prosopagnosic, control) by two (trial type: lighting change, viewpoint change) mixed-measures ANOVA on BFRT accuracy. Descriptive statistics are available in section 3.4.4. There was a main effect of group ($F(1, 242) = 167.46, p < .001, \eta^2_p = .41$), where developmental prosopagnosics scored worse than controls. There was also a main effect of trial type ($F(1, 242) = 70.50, p < .001, \eta^2_p = .23$), with higher accuracy for lighting trials than viewpoint trials. The interaction between group and trial type was again not significant ($F(1, 242) = 3.59, p = .059, \eta^2_p = .01$). Altogether, the lighting versus viewpoint results remained the same as the original analysis.

B.6 Multivariate single-case analysis of scope output

To determine the proportion of developmental prosopagnosics with impaired face perception across all three face perception tests, I ran a multivariate single-case comparison for each developmental prosopagnosic against control performance. Table B1 shows the mean difference between control performance and each developmental prosopagnosic, and

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

the corresponding significance value. In total, 34 developmental prosopagnosics had face perception deficits across all perception tests.

Table B1

Developmental Prosopagnosic Multivariate Single-case Comparisons on All Face Perception Tests

Participant	1	2	3	4	5	6	7
Mean difference	-6.27	-9.82	-5.98	-5.84	-8.61	-5.86	-7.23
P-value	< .001	< .001	< .001	< .001	< .001	< .001	< .001
	8	9	10	11	12	13	14
	-8.12	-6.71	-4.53	-5.03	-7.39	-8.60	-6.63
	< .001	< .001	< .001	.001	.001	.001	.001
	15	16	17	18	19	20	21
	-9.84	-5.38	-6.26	-5.47	-8.24	-7.46	-0.87
	.002	.002	.003	.004	.005	.006	.006
	22	23	24	25	26	27	28
	-7.34	-7.41	-7.73	-5.61	-6.85	0.01	-4.51
	.006	.009	.014	.016	.016	.019	.023
	29	30	31	32	33	34	35
	-3.64	-5.74	-6.32	-3.74	-6.65	-5.97	-6.53
	.024	.028	.034	.038	.045	.045	.05
	36	37	38	39	40	41	42
	-3.67	-4.97	-5.62	-3.68	-3.65	-5.18	-4.18
	.055	.059	.062	.064	.067	.078	.101
	43	44	45	46	47	48	49
	-4.94	-6.33	-0.76	-4.73	-5.50	-5.59	-4.49
	.102	.102	.118	.129	.136	.138	.159
	50	51	52	53	54	55	56
	-1.90	-2.48	-4.45	-5.18	-3.74	-3.74	-4.20
	.16	.17	.194	.198	.2	.2	.212
	57	58	59	60	61	62	63
	-5.22	-3.75	-1.50	-3.94	-3.40	-5.10	-4.08
	.217	.219	.228	.233	.247	.247	.255
	64	65	66	67	68	69	70
	-4.94	-3.46	-1.60	-4.28	-3.07	-1.45	-4.36
	.263	.272	.274	.275	.287	.289	.289
	71	72	73	74	75	76	77
	-4.89	-4.15	-2.75	-3.44	-4.42	-3.09	-2.64

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

.294	.334	.341	.345	.353	.356	.382
78	79	80	81	82	83	84
-3.38	-2.85	-3.56	-1.30	-2.93	-3.27	-3.23
.397	.416	.417	.426	.435	.447	.482
85	86	87	88	89	90	91
0.99	-0.22	-3.88	-1.78	-3.66	0.46	-1.21
.491	.505	.508	.521	.528	.61	.619
92	93	94	95	96	97	98
-2.74	-0.68	-0.68	-2.26	-2.96	-2.63	-1.79
.62	.663	.663	.675	.678	.688	.688
99	100	101	102	103	104	105
-2.50	-2.96	-1.79	-2.16	-0.32	-2.53	-1.03
.694	.712	.741	.76	.775	.803	.84
106	107					
-1.56	-0.30					
.872	.944					

Note. Participant numbers in bold index developmental prosopagnosics with significant individual impairment on all three face perception tests compared to control performance. Developmental prosopagnosic mean differences ranged from -11.79 to 0.99, where negative values denote poorer scores relative to controls. Developmental prosopagnosics are ordered by multivariate test-statistic significance.

B.7 Hierarchical regression with face perception tests as separate predictors

To determine whether any of the face perception tests has a larger contribution toward predicting developmental prosopagnosia diagnosis, I entered accuracy on the BFRT, GFMT, and CFPT as separate predictors in step one of a hierarchical regression model. The step one model was significant ($X^2(3) = 138.39, p < .001$). The BFRT ($B = 0.14, p < .001$) and GFMT ($B = 0.06, p = .004$) significantly predicted group membership, but the CFPT did not ($B < 0.01, p = .717$). The BFRT and GFMT odds ratios were 0.86 (95% CI = 0.82-0.91) and 0.94 (95% CI = 0.91-0.98) respectively. The model classified group membership for 80% of participants, had a good AUC at 0.90, and an adequate R^2 at 0.58. Classification by group is provided in Table B2. Altogether, face perception remained highly predictive of

developmental prosopagnosia diagnosis, but this was driven by BFRT and GFMT scores as the CFPT failed to predict group membership.

Table B2

Group Classification in the BFRT, GFMT, and CFPT Model

		Predicted		% correct
		Control	Prosopagnosic	
Observed	Control	114	22	83.82
	Prosopagnosic	26	81	75.70
Overall % correct				80.25

Adding CFMT-Aus scores into step two of the model again returned a significant result ($X^2(4) = 156.94, p < .001$), which also accounted for additional variance compared to step one ($X^2(1) = 18.55, p < .001$). BFRT ($B = 0.12, p < .001$), GFMT ($B = 0.5, p = .032$), and CFMT-Aus ($B = 0.08, p < .001$) scores all predicted group membership, while CFPT scores did not ($B = 0.02, p = .379$; 95% CI = 0.98-1.06). The BFRT, GFMT, and CFMT-Aus odds ratios were 0.89 (95% CI = 0.82-0.91), 0.96 (95% CI = 0.91-0.98), and 0.93 (95% CI = 0.89-0.96) respectively. In other words, the likelihood a participant was developmental prosopagnosic reduced by 11% (BFRT), 4% (GFMT), and 7% (CFMT-Aus) with a one-point increase on the respective test. Overall, 81.89% of participants were correctly classified (see Table B3). The AUC was strong at 0.92, and the Nagelkerke pseudo R^2 was adequate at 0.64. Consistent with the original results, the gain from adding CFMT-Aus scores was minimal. Face perception as indexed by BFRT scores had nearly double the predictive power over CFMT-Aus scores, while prediction using GFMT was similar to CFMT-Aus.

Table B3*Group Classification in the BFRT, GFMT, CFPT, and CFMT-Aus Model*

		Predicted		% correct
		Control	Prosopagnosic	
Observed	Control	115	21	84.56
	Prosopagnosic	23	84	78.51
Overall % correct				81.89

B.8 Response time inversion effects

To determine if mean response time inversion effects differed across groups, I ran a two (group: prosopagnosic, control) by three (test: BFRT, GFMT, CFPT) mixed-measures ANOVA. Response time inversion effects were calculated by subtracting average inverted response time from average upright response time for each participant. The outcome is shown in Figure B1. There was a significant main effect of test ($F(1.04, 249.99) = 35.55, p < .001, \eta^2_p = .13$), such that response time inversion effects were larger on the CFPT ($M = 3.33s, SD = 8.10s$) compared to the BFRT ($M = 0.43s, SD = 1.33s; t(482) = 7.00, p < .001$) and GFMT ($M = 0.17s, SD = 0.45s; t(482) = 7.57, p < .001$), but did not differ between the BFRT or GFMT ($t(482) = 0.58, p = 1.000$). Most importantly, the main effect of group ($F(1, 241) = 0.92, p = .339, \eta^2_p < .01$) and interaction between the two factors ($F(1.04, 249.99) = 0.41, p = .532, \eta^2_p < .01$) were not significant. Because response time inversion effect sizes did not differ between groups across the tasks, I did not use response time inversion effects as a covariate when analysing group accuracy inversion effects.

Figure B1

Developmental Prosopagnosic and Control Response Time Inversion Effects on the Face Perception Tests

Note. Response time inversion effects (upright average response time – inverted average response time) are shown in seconds. Diamonds represent the group mean for each test and points show the individual data.

B.9 Potential floor effects in the inverted conditions

Floor effects could explain developmental prosopagnosics' smaller inversion effect sizes. To further test the impact of potential floor effects on the inversion effect results, I reran the accuracy inversion effect analyses on two subsets of data that were chosen to lift developmental prosopagnosics' inverted accuracy to be as far away from floor as possible.

Table B4

Developmental Prosopagnosics' Inversion Effect Sizes for the Original, Easiest 50% of Trials, and Top 50% of Subjects Data

			Original	Easiest 50% trials	Top 50% prosopagnosics
BFRT	Upright	<i>M</i>	58.20	69.25	62.08
		<i>SD</i>	11.94	14.12	9.99
	Inverted	<i>M</i>	51.06	57.87	55.32
		<i>SD</i>	10.20	11.61	8.51
GFMT	Upright	<i>M</i>	59.79	62.69	63.10
		<i>SD</i>	8.78	12.92	8.70
	Inverted	<i>M</i>	57.48	66.85	61.76
		<i>SD</i>	9.51	11.85	9.29
CFPT	Upright	<i>M</i>	59.47	62.86	64.35
		<i>SD</i>	11.43	14.30	9.28
	Inverted	<i>M</i>	49.96	55.92	57.79
		<i>SD</i>	11.69	13.11	8.00

Note. Inversion effects are shown in percentages (upright accuracy – inverted accuracy).

The first subset of data contains the easiest 50% of inverted trials for each face perception test as determined by developmental prosopagnosics' inverted accuracy. Doing so lifted developmental prosopagnosic inverted scores away from floor (see Table B4 for comparison means and standard deviations). Using only these trials, I repeated the two (group: developmental prosopagnosic, control) by three (test: BFRT, GFMT, CFPT) ANOVA on accuracy inversion effects. The same result emerged. There was a significant main effect of group, such that prosopagnosics had significantly smaller inversion effects ($M = 4.72\%$, $SD = 15.44\%$) than controls ($M = 15.14\%$, $SD = 15.45\%$), $F(1, 242) = 78.47$, $p < .001$, $\eta^2_p = .24$. There was also a main effect of test ($F(2, 484) = 70.49$, $p < .001$, $\eta^2_p = .23$), with larger inversion effects for BFRT ($M = 16.87\%$, $SD = 13.57\%$) compared to the GFMT ($M = 2.68\%$, $SD = 15.96\%$; $t(484) = 11.66$, $p < .001$) and the CFPT ($M = 12.03\%$, $SD = 15.94\%$; $t(484) = 3.90$, $p < .001$), and greater on the CFPT compared to the GFMT ($t(484) = 7.76$, $p < .001$). There was no significant interaction between group and test for inversion effects ($F(2, 484) = 0.91$, $p = .403$, $\eta^2_p < .01$).

The second subset of data contains the 50% top performing developmental prosopagnosics as determined by their average inverted accuracy across all face perception tests. Table B4 displays the raised means and standard deviations. I reran the two (group: developmental prosopagnosic, control) by three (test: BFRT, GFMT, CFPT) ANOVA on this subset of data, and again I obtained results similar to the original analysis. There was a significant main effect of group, such that prosopagnosics had much smaller inversion effects ($M = 4.88\%$, $SD = 9.84\%$) than controls ($M = 16.83\%$, $SD = 11.38\%$), $F(1, 188) = 122.42$, $p < .001$, $\eta^2_p = .39$). There was also a main effect of test ($F(2, 376) = 19.36$, $p < .001$, $\eta^2_p = .09$). Accuracy inversion effects were larger on the BFRT ($M = 14.61\%$, $SD = 9.84\%$) compared to the GFMT ($M = 9.84\%$, $SD = 11.79\%$; $t(376) = 4.62$, $p < .001$). They were also larger on the CFPT ($M = 16.70\%$, $SD = 13.41\%$) than the GFMT ($t(376) = 5.92$, $p < .001$). Inversion effect sizes did not differ across the BFRT and CFPT ($t(376) = 1.31$, $p = .193$). There was no interaction between group and test ($F(2, 376) = 1.17$, $p = .311$, $\eta^2_p < .01$).

Altogether, both subsets of data produced the same general finding, namely that developmental prosopagnosics have smaller inversion effect sizes across all perception tests, and that this result cannot be explained by potential floor effects.

B.10 Signal detection analysis of the GFMT

I calculated d' to further analyse the GFMT data. Derived from signal detection theory, d' is a measure of perceptual sensitivity that takes response bias into account. To calculate d' , I first calculated participants' hit and false alarm rates. The hit and false alarm rates were converted to z-scores, and then each participant's false alarm z-score was subtracted from their hit z-score to give their d' . I ran a two (group: prosopagnosic, control) by two (orientation: upright, inverted) ANOVA on d' data. There was a significant main effect of group ($F(1, 241) = 64.56$, $p < .001$, $\eta^2_p = 0.21$) and orientation ($F(1, 241) = 114.06$, $p < .001$, $\eta^2_p = 0.32$), that were qualified by a significant interaction ($F(1, 241) = 60.52$, $p <$

.001, $\eta^2_p = 0.20$). Developmental prosopagnosics had much poorer upright sensitivity ($M = 0.59$, $SD = 0.52$) than controls ($M = 1.50$, $SD = 0.81$; $t(398.88) = 10.91$, $p < .001$), but developmental prosopagnosics also had slightly poorer inverted sensitivity ($M = 0.46$, $SD = 0.59$) than controls ($M = 0.70$, $SD = 0.58$; $t(398.88) = 2.80$, $p = .005$). This suggests that developmental prosopagnosics' poor perceptual sensitivity on the GFMT is disproportionately related to upright-specific processes.

B.11 Developmental prosopagnosic regression-based inversion effects compared to control variation

To distinguish between the normative and pathological accounts of developmental prosopagnosia, I compared group relationships between average inversion effects (calculating using the standard subtraction method) and face recognition skill in section 4.2.

Developmental prosopagnosic performance was shifted below what would be expected by variation within the low-end of control performance, thus supporting the pathological view.

In this analysis, inversion effects were calculated using the common subtraction method (upright accuracy – inverted accuracy). The subtraction method, however, is limited when investigating individual differences such as in correlations (DeGutis et al., 2013b). This is due to the inversion effect estimate being confounded with the variation within the inverted control condition. One solution is to calculate inversion effects by regressing the inverted scores from the upright scores and taking the residuals. Doing so removes all variance explained by inverted performance and leaves a pure index of upright-specific performance (DeGutis et al., 2013b). I therefore repeated the comparison using the regression method for calculating inversion effect sizes, provided in Figure B2. Consistent with the original result, developmental prosopagnosic performance was shifted away from what would be expected within control variance, thus supporting the pathological view.

Figure B2

Comparing Developmental Prosopagnosic Performance to the Variation Predicted By Performance in Controls using Regression-based Inversion Effects

B.12 Predicting developmental prosopagnosia severity using face perception

I investigated whether face perception scores predicted developmental prosopagnosia severity as indexed by original CFMT scores and PI20 scores. First, I ran a multiple linear regression with the face perception tests as predictors of original CFMT accuracy within developmental prosopagnosia. The model was not significant ($F(3, 104) = 1.45, p = .234$) and only explained 4% of variance in developmental prosopagnosic CFMT scores ($R^2 = 0.04$). None of the perception tests were significant predictors of developmental prosopagnosics' CFMT scores (BFRT: $B = 0.33, p = .625$; GFMT: $B = 0.75, p = .306$; CFPT: $B = 0.07, p = .315$). I then repeated the same analysis, except this time with the face perception tests as predictors of PI20 scores within developmental prosopagnosia. Again, the model was not

FACE PERCEPTION IN DEVELOPMENTAL PROSOPAGNOSIA

significant ($F(3, 104) = 0.82, p = .484$) and only explained 2% of variance in developmental prosopagnosics' PI20 scores ($R^2 = 0.02$). None of the face perception tests predicted PI20 scores in developmental prosopagnosics (BFRT: $B = 0.03, p = .296$; GFMT: $B = 0.08, p = .499$; CFPT: $B = 0.09, p = .188$). These results imply that face perception deficit severity does not track developmental prosopagnosia severity as indexed by either CFMT or PI20 scores. Face perception deficits did, however, predict group membership, as described in section 3.3.3. Collectively, these results suggest that the face recognition tests have good discriminability at the middle of the spectrum of face recognition skill (i.e., discriminating developmental prosopagnosics from controls), but might have poor discriminability at the lower end of the spectrum (i.e., discriminating mild versus severe developmental prosopagnosia). In other words, the CFMT and PI20 might be unable to sufficiently spread-out developmental prosopagnosics' scores to track their deficit severity.