Recharacterising Face Recognition Deficits in Developmental Prosopagnosia

Ebony Murray

A thesis submitted in partial fulfilment of the requirements of Bournemouth University for the degree of Doctor of Philosophy

Submitted October 2020

Abstract

Developmental prosopagnosia (DP) is a neurodevelopmental condition characterised by a relatively selective deficit in face recognition. Current approaches to diagnose the condition vary within the field, making comparisons across studies difficult. Furthermore, theoretical inferences rely on accurate identification of people with DP and, combined with the fact that some individuals also report moderate-to-severe psychosocial consequences of the condition, the methods used to diagnose DP have been under increasing scrutiny. Thus, the present thesis aimed to address contemporary issues relating to the diagnosis of DP. Specifically, the thesis begins by investigating the utility of self-report, and presents a list of empirically-driven hallmark symptoms of the condition which can aid the identification of DP in children and adults, both in oneself and in others. The thesis then explores objective measures of face processing; Chapter Four presents, for the first known time, the test-retest reliability for the Cambridge Face Memory Test (CFMT) – the leading test of unfamiliar face recognition. Interestingly, this value falls just short of accepted psychometric protocols, and recommendations for using two separate tests of the same subprocess of face processing are suggested. With this in mind, Chapter Five supported the diagnostic utility of repeat assessment for face perception. We present a new test of face matching and, in combination with the original Benton Facial Recognition Test, advocate the use of identifying face perception deficits in DP by administering two separate tests. Finally, we come away from behavioural measures to consider more nuanced means of dissociating DP from other developmental conditions that are commonly associated with face recognition difficulties. For instance, many people with autistic spectrum conditions (ASCs) struggle with face recognition, and many DPs avoid social situations – an outward symptom that may be confused with ASCs. Given face recognition difficulties likely have different underpinnings in the two conditions, Chapter Six employs eye-tracking methods to differentially diagnose face recognition impairments in DP versus ASCs. Whilst sample size is low due to the interruption of data collection by COVID-19, we find that those with ASC look more towards non-social stimuli than DP participants, particularly when the scene is more social and interactive. Those with DP look more towards extrafacial information, such as hair and bodies, than ASC participants. The thesis concludes with a summary of recommendations for the future diagnosis of DP.

Thesis Structure

The present thesis confirms to an 'integrated thesis' format in which chapters (chapters 2 - 5) consists of articles which have been written in a style that is appropriate or publication in peer reviewed journals. Chapter 6 is not currently prepared as a manuscript. The initial and final chapters present and Introduction and Discussion of the field of research undertaken. The articles included in this thesis are at various stages of the publication and review process, and the status of each paper is summarised below. The main text in each chapter is presented as exact replications of the submitted manuscript and inevitably, there is some repetition as a consequence.

Status of articles from this thesis

Chapter 2 has been published as: Murray, E., Hills, P., Bennetts, R., & Bate, S. (2018). Identifying Hallmark Symptoms of Developmental Prosopagnosia for Non-Experts. *Scientific Reports, 8*(1). https://doi.org/10.1038/s41598-018-20089-7

Chapter 3 has been published as: Murray, E., & Bate, S. (2019). Self-ratings of face recognition ability are influenced by gender but not prosopagnosia severity. *Psychological Assessment, 31*(6), 828-832. https://doi.org/10.1037/pas0000707

Chapter 4 has been published as: Murray, E., & Bate, S. (2020). Diagnosing developmental prosopagnosia: repeat assessment using the Cambridge Face Memory Test. *Royal Society Open Science, 7*(9). https://doi.org/10.1098/rsos.200884

Chapter 5 has been submitted for publication as: Murray, E., Bennetts, R.J., Tree, J., & Bate, S. (2020). An Update of the Benton Facial Recognition Test. Manuscript submitted for publication: *Behavior Research Methods*.

Contents

List of Tables

List of Figures

Acknowledgments

Without a shadow of a doubt, I begin with my first and biggest “Thank You!” to Prof Sarah Bate. My supervisor since my second year of my BSc, I simply couldn’t have done this without you. Thank you for your patience, guidance, knowledge, and support over the last eight or nine years. It has never been unappreciated, and it has been an absolute privilege to work with you and learn from you. I also thank Dr Rachel Bennetts who has only ever been supportive of me and my studies, and an unwitting victim of my silly questions! To all of my collaborators and co-authors, for every single piece of advice and help along the way, it hasn’t ever been taken for granted. Thank you Peter, Nicola, Rachel, Jeremy and Steph. To the BU PhD office family, thank you for always helping me, encouraging me, and just generally putting up with me!

To Ryan, thank you for everything. The past few years have been a whirlwind of studying, new jobs, house moves, and loving and remembering our beautiful Goose. It’s not been the easiest of times, and I’m thankful for everything you’ve helped me achieve and get through. Big, big thank yous also to all of my friends. To the Hreod girls, Uni ‘gals’ and dancing girls, all of whom I just can’t seem to shake off after all these years, thank you for being you. To name a few, a special big thank you through my university years must go to Trice, Emma and Meghan. For the holidays and trips away, conversation and advice, unnecessary amounts of wine, and everything else that a girl needs in a friendship, I thank you all.

My family. Your everlasting support has never gone unknown or unappreciated. I am so thankful for you all. To Mum and Dad, whom I love and appreciate silly amounts, thank you for everything. I couldn’t be writing this without you. To my brothers, grandparents, and Chris, thank you for all you’ve done over the years. Your advice, quick wit and silliness has been so very welcomed and valued. And, finally, with a heavy heart, to my wonderful stepmum, Tracy. It breaks my heart you’re not here to see this and for me to thank you in person for all you did for me. You lit up every room, and we miss you every day.

Author’s Declaration

I hereby declare that the work presented in this thesis has not been, and will not be, submitted in whole or in part to another University for the award of any other degree.

Chapter One: Introduction

Faces provide information about one's gender, age, ethnicity, emotional state, and perhaps most importantly, they identify the owner; the ability to recognise an individual just by looking at their face is crucial for human social interaction. Although it was once thought that everybody is an expert at recognising faces (e.g. Carey, 1992), it is now widely acknowledged that face recognition abilities vary within the typical population (e.g. Bate et al., 2010) with some individuals recognising faces at a level that is well above average (Russell et al., 2009) and others performing very poorly. These latter individuals may have a visuo-cognitive condition known as "prosopagnosia" – a disorder characterised by a relatively selective impairment in face recognition.

Traditionally, prosopagnosia was believed to only be acquired in nature, following neurological injury or illness, typically affecting the occipitotemporal regions (e.g. Barton, 2008). However, the turn of the century saw an increase in reports of developmental prosopagnosia (DP), an apparently parallel form of the condition which occurs without any neurological damage, socio-emotional dysfunction, or lower-level visual deficits (McConachie, 1976). Currently, DP is estimated to affect approximately 2 – 2.5% of the adult population (Bowles et al., 2009) and 1.2 – 4% of those in middle childhood (Bennetts et al., 2016) (although note that, by definition, the lower end of a normal distribution would encompass 2.5% of the population and therefore, prevalence rates reflecting this may be a statistical artefact: Barton & Carrow, 2016), and recent investigations report that DP often runs in families, suggesting the condition has a genetic component (Duchaine et al., 2007; Johnen et al., 2014; Schmalzl et al., 2008), a finding that accords with the broader view that face recognition ability is a heritable trait (Shakeshaft & Plomin, 2015; Wilmer et al., 2010; Zhu et al., 2010).

Initially, however, it was believed that DP was an extremely rare disorder with very few cases being reported. Yet, with rapid increases in the reports of DP, individuals with the condition have been used to make theoretical inferences about the development and functioning of the cognitive and neural architecture of the typical and impaired face recognition system (e.g. Duchaine & Nakayama, 2005; Dalrymple et al., 2017; Daini et al., 2014; Towler et al., 2012). Given that these theoretical inferences rely on accurate identification of people with DP, and that some individuals also report moderate-to-severe psychosocial consequences of the condition (Adams et al., 2019; Dalrymple et al., 2014; Murray et al., 2018; Yardley et al., 2008)

the methods used to diagnose DP have been under increasing scrutiny. Indeed, whilst identifying cases of acquired prosopagnosia is often straightforward as the individual is aware they have lost a skill that they once had, and there is often neurological evidence to support a diagnosis, identifying and diagnosing DP is not so simple. The condition is not listed as a psychiatric disorder in the DSM-V (American Psychological Association, 2013) and no formal diagnostic criteria currently exist.

However, current approaches to diagnosis reflect and extend on the protocols used to examine face processing deficits in the acquired prosopagnosia literature, where tests of familiar/famous face recognition, unfamiliar face recognition, and face perception tests were administered. Using a combination of tasks which assess each of these three abilities, dominant recommendations suggest that DP is diagnosed when scores fall into the impaired range (typically at least two SDs from the control mean; although see DeGutis et al., 2012, 2014; Palermo et al., 2017; White et al., 2016 for using 1.7 SDs from the mean) on any two of these three tasks (Bate & Tree, 2017; Dalrymple & Palermo, 2015). This recommendation for repeated testing (i.e. not relying on a single score on a single test) reduces "the chance that it happened by chance" (Young et al., 1993, p. 945) and is vital for minimising false alarms.

Critically though, there is some contention about the precise tasks that should be used to assess each of these sub-processes. Most researchers administer different combinations of self-report, a famous faces test, an unfamiliar face memory task and a measure of unfamiliar face perception. The leading tasks used to assess these subprocesses are discussed in detail below. In addition to inconsistencies in the specific tests used to diagnose DP across the field, the approach to diagnosis can also be contentious. Inviting individuals to the lab and administering full, thorough screening sessions can be costly and time-consuming. With much larger numbers of people self-referring to research teams in the belief that they are living with DP, there is a need to develop rapid, time- and cost- effective protocols. Similarly, these protocols are needed to meet the trend for online testing; since increased accessibility of the internet, online testing has increased in popularity. However, web-based research has also been found to have larger dropout rates than lab-based studies (Birnbaum, 2004). Those taking part in online research are free of any social pressure from the researcher or potential embarrassment or confrontation in explaining that they no longer with to take part; the participant simply clicks a button to quit the test early, or completes some subsets of the experiment but not all. This is highly important for identifying cases of DP because all test results need to be considered. Moreover, in lab-based research, the fact that one person may participate on multiple occasions

has rarely been considered a problem. However, online testing increases the possibility of multiple submissions from a single person. Combined with the fact that leading tests of face recognition are readily available online, scores on tests may also not be accurate (for example, due to practice effects).

In the following sections, the leading tasks of face processing which are utilised in the diagnosis of DP are discussed. Critical limitations of the existing tasks and approaches to diagnosis are highlighted. The thesis then presents five chapters, each of which tackle a different challenge in the current protocols for DP diagnosis. The overall thesis is then discussed, and suggestions for future research are offered.

Familiar face recognition

Assessments of familiar face memory are an obvious choice of test due to their ability to reflect the hallmark symptom of DP: a deficit in recognising familiar faces. Whilst assessing familiar face recognition abilities in children often adopts a test created using personally familiar faces (e.g. Bennetts et al., 2018; Brunsdon et al., 2006; Wilson et al., 2010), these tests are seldom used in adults; it is challenging and time-consuming to acquire a large number of images of personally familiar faces, and is difficult to control for the level of familiarity for each face (e.g., whether a target is someone that is seen every day, compared to someone seen only once a month). Moreover, it is difficult to obtain normative data for personally familiar face tests, and thus, identifying an atypical score is challenging.

In light of these difficulties, personally familiar faces are often dismissed and replaced with a famous faces test. Images of famous people are plentiful on the internet and easy to access, meaning that a large stimuli set can be acquired quickly. These stimuli can then be used to create a single famous faces test, suitable for an entire sample of participants and not just a single person. However, famous faces tests do not overcome the issue of degree of familiarity, stemming from the amount of contact that person has had with each famous person, in addition to substantial age and nationality effects. Furthermore, most individuals with DP can identify at least some highly familiar faces (e.g. Bate et al., 2019). Finally, as online testing increases in popularity, famous face tests are more difficult to administer and score virtually, particularly on a largescale basis.

Self Report

Most researchers agree that anecdotal evidence of everyday difficulties with face recognition is required to support a DP diagnosis (Corrow et al., 2016; Grüter et al., 2009). Existing reports have mostly relied upon the individual to describe incidents of everyday failures of face recognition (e.g. Bennetts et al., 2015; Corrow et al., 2016), although some researchers have developed more formal questionnaires or structured interviews for this purpose (e.g. Shah et al., 2015; Kennerknecht et al., 2008; although note, that the items which make up these questionnaires and interviews are not empirically-driven, but rather, devised by the researchers).

However, a more fundamental identification of the everyday behavioural traits that are associated with DP is an important issue, irrespective of any debates surrounding the formal diagnosis of the condition. Indeed, the objective diagnosis of DP typically requires an individual to initially self-refer (or be referred by a guardian or significant other) to a laboratory for screening, yet this process requires that person to have some awareness of their face recognition difficulties relative to unaffected others. Although some self-report tools have been shown to successfully identify at least some candidates for DP (Gray et al., 2017), many anecdotal reports suggest a lack of awareness into the severity of one's face recognition difficulties (Fine, 2012; Sutton, 2016). Further, much work examining the typical population indicates that people have limited insight into their own abilities (e.g. Zell & Krizan, 2014) and specifically into their face recognition skills, reporting only weak-to-moderate correlations between subjective ratings and scores on objective tests (e.g. Bindemann et al., 2014; Laguesse et al., 2013).

Unlike those with acquired prosopagnosia, those with DP do not experience an abrupt loss of their face recognition skills and therefore, do not have a point of comparison: many individuals tested in our laboratory did not become aware of their difficulties until mid or even late adulthood (see also Fine, 2012; Sutton, 2016). It is also unclear whether participant gender influences the self-report of prosopagnosia symptoms, and there is some precedent to suggest this may be the case. In typically developing samples, males tend to rate themselves more favourably than females on traits such as intelligence and attractiveness (e.g. Cooper et al., 2018; Ehrlinger & Dunning, 2003; Sim et al., 2015). Additionally, females with various developmental conditions appear to rate their symptoms as more severe than males (e.g. autism spectrum conditions (ASCs): Lai et al., 2011; Moseley et al., 2018; attention deficit

hyperactivity disorder (ADHD): Vildalen et al., 2016). Taken together, it may be plausible to hypothesise that females with prosopagnosia will self-report more severe face recognition difficulties than their male counterparts.

In sum, whilst it is agreed that individuals who are screened must self-report difficulties with face recognition in their daily lives, it is vital that researchers continue to administer objective tests of face processing skills to overcome difficulties with self-awareness, and to overcome potential issues with self-report that are not yet known (e.g. gender effects).

Unfamiliar face memory

Unfamiliar face recognition tests are a standardised way to measure the core symptom of DP: a face *recognition* impairment. Overcoming the shortfalls of familiar face recognition tasks, unfamiliar face tests offer a common starting point for all participants (i.e. all faces are unfamiliar at the start of the test). Moreover, the level of exposure to each face is tightly controlled, allowing more accurate norming data to be collected across different samples from the typical population.

Traditionally, the Warrington Memory for Faces test (RMF: Warrington, 1984) was used to assess unfamiliar face memory and identify cases of DP (e.g. Bate et al., 2008; Bentin et al., 1999; Duchaine, 2000; Kracke, 1994; McConachie, 1976). The RMF is comprised of 50 black and white images of unfamiliar males. Each image is presented to the participant one at a time for three seconds. Following this initial presentation, participants are shown two images at the same time: an identical photograph to one they saw earlier, and a distractor photograph. Participants are asked to decide which of the two faces the familiar one is. The RMF is easy to administer, and the test places few demands on other cognitive functions such as attention, organisation and motor functioning.

Nonetheless, concerns about the lack of information about the internal-consistency or test-retest reliability of the RMF have been raised (Kapur, 1987) and the normative data for this test have also been questioned (O'Bryant et al., 2003). Issues have also been raised specifically regarding DP diagnoses; the RMF stimuli include photographs which include the models' hair, clothing, and variations of head postures and body positions, all of which vary greatly between the images. Thus, it is debatable as to whether participants are making

decisions based on the model's face, or are relying on extrafacial cues to aid recognition (for a discussion, see Duchaine & Weidenfeld, 2003).

In light of these concerns, Duchaine and Nakayama (2006) developed the Cambridge Face Memory Test (CFMT), a test of unfamiliar, short-term face memory, now considered to be the gold standard within face recognition literature. The CFMT uses images of male faces, aged in their 20s and early 30s. Each individual is displayed in the same range of poses and lighting conditions, with neutral expressions. All faces are cropped so that no hair is visible and facial blemishes are removed. Images are greyscale. The general objective of the CFMT is to introduce six unfamiliar faces to the participant and then test their recognition of those faces. It contains three test stages which increase in difficulty as the test progresses: (a) Learn stage: Participants view a target face from three viewpoints for 3-seconds per image. They are then presented with triads of faces, where one is an identical image of the target face, and two are distractors. Participants choose which of the three faces is the target. This is repeated for six faces, resulting in a maximum score of 18. (b) Test stage: Thirty triads of faces are presented, where one face is a novel image of a target identity intermixed with two distractors. Target faces are displayed under previously unseen viewpoint or lighting conditions. (c) Noise stage: Twenty-four new triads are displayed with added visual noise. Again, each trial contains any one of the targets and two distractors. The entire test is scored out of 72, and chance is a score of 24. There is no time limit for participants to elicit responses.

The CFMT is believed to overcome the drawbacks associated with the RMF. Stimuli are cropped and limited strictly to facial information (i.e. no clothing, no hair). The task is also more akin to everyday face recognition, in that the test gives the participant the chance to gradually gain knowledge of a person's face from a variety of angles; participants view each of the six target faces a total of 17 times during the test. Repetition of the distractor faces is also a benefit as the task involves discrimination of multiple familiar faces. This, once again, is a better simulation of real-world events.

Moreover, the test has high validity and high reliability (Wang et al., 2015), and whilst performance on the CFMT has been found to decline significantly after the age of 50 years, formulae for age adjusted z-scores have been provided (Bowles et al., 2009). Theoretically, there is good evidence that the CFMT is a valid measure of face memory, rather than generalised memory or memory for images. For instance, the test only has a modest correlation with other aspects of visual memory (e.g. abstract art: Wilmer et al., 2010; cars: Dennett et al.,

2012). Considering these strengths of the CFMT, it is unsurprising that it now considered the 'gold-standard' and is nearly always used to identify cases of DP (e.g. Bate et al., 2009; Burns et al., 2017; Bylemans et al., 2020; Carrow et al. 2016; Marsh et al., 2019; Nemeth et al., 2014; Palermo et al., 2011).

However, some things remain unknown about the CFMT. For instance, some individuals self-report difficulties with face recognition in their daily lives and yet score typically on the CFMT (e.g. Bate et al., 2019). Additionally, the test is readily available online and, combined with the fact that individuals likely self-refer themselves to multiple research teams and may complete the CFMT multiple times, little is known about potential practice effects. Importantly, to date, there has been no examination of the test-retest reliability of the CFMT. Aside from practice, individuals' performance on the CFMT can vary daily, and inconsistent performance has been reported in the face recognition literature for typical participants and those with proficient face-processing skills (Bindemann et al., 2012; Bate et al., 2018; Bate, Frowd et al., 2019). Aside from the influence of chance (McKone et al., 2011; Young et al., 1993) these inconsistencies could result from intrinsic factors (e.g. mood, fatigue or distraction) or the psychometric properties of the task in hand (Young & Noyes, 2019). Moreover, a recent report has suggested that the CFMT can be shortened, and maintain its diagnostic utility (Corrow et al., 2018), offering a more time-efficient but equally effective way of identifying DP. Thus, despite its popularity amongst DP researchers, further work is still required to gain a full understanding of the task's psychometric properties and diagnostic utility.

Face perception

While prosopagnosia is primarily conceived as a deficit in facial identity recognition, existing reports have identified differences in prosopagnosics' abilities to *perceive* facial identity - a stage which occurs earlier in the face processing system (e.g. Bruce & Young, 1989, Haxby et al., 2000). In the acquired prosopagnosia literature, some individuals have been found to have intact face perception skills (e.g. Barton et al., 2002; Benton, 1980) and in others, this ability is impaired (e.g. Barton et al., 2002; Gainotti et al., 2003; Young & De Haan, 1988). Such reports motivated the claim that there may be subtypes of prosopagnosia, often referred to as "apperceptive" and "associative" variants (De Renzi et al., 1991). In apperceptive prosopagnosia, the deficit is an inability to form an accurate representation of the face from the

available visual data available (i.e. a perception deficit). In contrast, in associative prosopagnosia, the individual is able to perceive the face. More precisely, the deficit lies in the process of matching this facial information to facial memory (i.e. a recognition deficit).

Similarly, some individuals with DP appear to have intact face perception skills (Bowles et al., 2009; Dalrymple et al., 2014; McKone et al., 2011; Ulrich et al., 2017) whereas others show face perception deficits (Avidan et al., 2011; Bate et al., 2019; Biotti et al., 2019; Duchaine et al., 2007; Righart & de Gelder, 2007; White et al., 2017). Consequently, current diagnostic protocols ensure that both face perception *and* face recognition skills are assessed, and permitting further investigation into these subtypes of prosopagnosia. However, as discussed above, the administration of different tests to assess these processes makes comparisons across studies difficult and, consequently, reaching a conclusion is also difficult.

Traditionally, researchers used the Benton Facial Recognition Test (BFRT: Benton & Van Allen, 1968; see Benton et al. (1983) for the formal reference of the test), a face matching task that is traditionally administered face-to-face using hard copy materials. Participants are simultaneously presented with a target face above an array of six test faces. In the first six trials, one face in the array matches the identity of the target face, and in the final 16 trials, three faces in the array match the identity of the target. All images are grayscale and display the overall shape of the face, but are cropped below the chin and beyond the hairline.

The task was originally developed for the assessment of acquired prosopagnosia (Barton, 2008; Bate & Bennetts, 2015; Van Belle et al., 2011), but has since been widely used to assess face perception skills in a number of neurological, clinical and psychiatric conditions (Annaz et al., 2009; Rabin et al., 2005; Sachse et al., 2014), including the identification of DP (e.g. Bentin et al., 1999; de Gelder & Stekelenburg, 2005; Duchaine, 2000; Kracke, 1994; McConachie, 1976).

Yet, the popularity of the BFRT has reduced in recent years, particularly for the assessment of individuals suspicious for prosopagnosia. Duchaine and Weidenfeld (2003) reported that when the inner features of the faces in the BFRT were obscured, most typical participants could still achieve a typical score using the hairline and eyebrows alone. Further, many individuals with DP can also achieve typical scores on this task (e.g. Duchaine & Nakayama, 2006), in contrast to group performance on alternative tests of face perception (Bate et al., 2019; Biotti et al., 2019).

In response to these findings, the Cambridge Face Perception Test (CFPT: Duchaine et al., 2007) was developed and has since become a widely used face perception test for the diagnosis of DP (e.g. Bennetts et al., 2015; Bylemans et al., 2020; Johnen et al., 2014; Lee et al., 2010; Marsh et al., 2019; Palermo et al., 2011). The CFPT presents participants with six morphed faces that are to be organised in order of similarity to a simultaneously presented target face. The trials are timed, and the target image is presented at a rotated angle while the test faces are presented from a frontal viewpoint. Collectively, these test characteristics aimed to prevent the application of laborious atypical compensatory strategies that could be used to achieve a typical score. However, much less is known about the reliability and other psychometric properties of this test, and there are some difficulties in its administration. For instance, the task requires proficient use of a computer mouse within a strict time period, and the instructions are complex for online administration, particularly with clinical and older participants (Bate et al., 2018; Bate, Frowd et al., 2019; Bowles et al., 2009). Others query whether morphed faces are unnaturally similar (White et al., 2017), and whether the requirement for similarity judgements initiates higher-level cognitive processes than required for the simplistic identity matching of simultaneously presented naturalistic facial images (Rossion & Michel, 2018). Arguably then, taken together, there is no reliable test of face perception available, especially one which meets the needs of online testing. Consequently, the fundamental nature of DP is still not fully understood.

Other Considerations

It is now widely accepted that not one single objective test is robust enough to identify and diagnose DP. Thus, testing protocols have reflected those used in acquired prosopagnosia, and researchers often administer a combination of self-report, a test of familiar/famous face recognition, an unfamiliar face recognition test, and a face perception test. Yet, as discussed above, there are at least some issues or unknown factors associated with all the tests that are commonly used in the field. Moreover, there is no biological indicator of DP (i.e. no gene), and there is the additional complication that these assessments only measure the outer symptom (face processing difficulties) and not the underlying issue. Accordingly, alongside these assessments, there is a need to exclude alternative explanations for face-processing deficits, such as visual dysfunction or neurological injury, and the presence of other neurodevelopmental disorders such as ASCs, in which, face recognition impairments are often

reported (Barton et al., 2004; Bradshaw et al., 2011; Chawarska & Volkmar, 2007; Neil et al., 2016). Indeed, given that DP and ASC share this critical symptom, the conditions are sometimes confused. Combined with the fact that public and professional awareness of DP remains relatively low, DP can be misinterpreted as ASC, particularly in younger age groups (e.g. Dalrymple et al., 2014).

Yet, the underpinning processes in these conditions likely differ from each other and one innovative method which can potentially differentiate DP and ASC, is the use of eye-tracking technology. Eye-tracking is widely used to examine information processing (Rayner, 2009) since it is well established that eye movement patterns provide detailed insight into on-going cognitive processing (Liversedge & Findlay, 2000) and is a direct measure of visual social attention (e.g. Frank et al., 2011). Furthermore, the development of unobtrusive eye-tracking systems that are readily available to researchers has resulted in an increase in eye-tracking methods being employed in psychological sciences over the past decade. Taken together, eye-tracking offers a way to overcome artificial accuracy scores on tests on face processing, and, importantly, is suitable for young children. As such, it offers an innovative means of distinguishing DP from ASC - something which has not yet been explored.

Support for this suggestion comes from findings that ASC participants (Chawarska et al., 2016; Falck-Ytter et al., 2015; Riby et al., 2013; Sumner et al., 2018; Wilson et al., 2010; Zantinge et al., 2017) and those with DP (Bobak et al., 2017; De Luca et al., 2019; Malaspina et al, 2017; Schwarzer, 2007) both show atypical eye gaze behaviour when viewing faces and/or scenes which include faces, in contrast to control participants. While the eye movements of ASC and DP participants have not yet been directly compared, it is likely that the face-processing impairments across these groups have different cognitive and neural underpinnings that can be indexed by manipulations in eye-tracking stimuli. For instance, ASC is characterised by deficits in social communication and interaction (DSM-V: American Psychological Association, 2013) whereas those with DP do not typically report these symptoms (other than *resulting* from their face recognition impairments), and instead their difficulties appear visuo-cognitive in nature (e.g. Towler et al., 2017). Consequently, it may be that there are some distinct differences in eye-movement patterns between these conditions when the social context of facial presentation is manipulated, offering an innovative means to potentially distinguish the two conditions. Moreover, given the importance of identifying ASC and DP accurately as early in life as possible, eye-tracking methods offer a suitable task for

younger populations; eye-tracking overcomes age-related limitations of standard face processing tasks, and offers a short and engaging means of assessment.

Thesis Aims

In light of the issues discussed above, the overall aim of this thesis is to address contemporary outstanding issues related to the diagnosis of DP. Specifically, the thesis aims to:

- (a) develop an empirically-driven list of symptoms that are representative of the disorder,
- (b) evaluate the utility of existing self-report instruments (the PI20) in discriminating disorder severity and its susceptibility to gender effects,
- (c) evaluate the outstanding psychometric issue of test-retest reliability in the CFMT,
- (d) develop and evaluate a new test of face perception that overcomes issues with existing tests and offers the opportunity for repeat-assessment
- (e) investigate the utility of eye-tracking as a tool for distinguishing DP from ASC.

As a part of this integrated-thesis, three chapters are presented as published papers. Chapter Two, titled "Identifying Hallmark Symptoms of Developmental Prosopagnosia for Non-Experts", was published in *Scientific Reports* in 2018. Chapter Three was published in *Psychological Assessment* in 2019, titled "Self-Ratings of Face Recognition Ability are Influenced by Gender but not Prosopagnosia Severity". Chapter Four, titled "Diagnosing Developmental Prosopagnosia: Repeat Assessment using the Cambridge Face Memory Test" was published in 2020 in *Royal Society Open Science*. The fifth chapter is presented as a manuscript, titled "An Update of the Benton Facial Recognition Test", and is currently under review at *Behavior Research Methods*. Finally, due to data collection being interrupted by the current COVID-19 pandemic, Chapter Six presents the available data for the project titled "Using Gaze Behaviour to Differentiate Between Developmental Prosopagnosia and Autism Spectrum Conditions". Data collection for this project will be completed when face-to-face testing can resume, and the chapter will subsequently be prepared for publication.

Introduction to Chapter Two

Identifying Hallmark Symptoms of Developmental Prosopagnosia for Non-Experts

The first study reported here is concerned with the use of self-report instruments in DP diagnosis. Most researchers agree that anecdotal evidence of everyday difficulties with face recognition is needed in order to support a DP diagnosis, and some researchers have developed formal questionnaires or structured interviews for this purpose. However, the items used in these questionnaires or interviews have mainly been drawn from informal discussions with those living with the condition, and/or anecdotal reports in the available literature. Thus, no empirically-driven symptom list exists for DP that is available to the public, researchers, and other professionals who may need to identify the condition. For instance, school teachers or educational psychologists may benefit from a symptom list to distinguish DP from ASC. In addition, to our knowledge, no study has investigated the indicators of DP that could exist from the point of view of a person without the condition and, thus, how DP could be identified by others. This is imperative to consider, given that much evidence suggests that individuals have a limited insight into their abilities.

The paper presented in Chapter Two used qualitative methods to examine *when* individuals gained insight into their condition, and to identify common symptoms or characteristics that could be useful in the detection of DP. Individuals with DP, their unaffected significant others, and parents of children with the condition completed questionnaires and follow-up semi-structured interviews about their experiences of the condition. In short, findings indicate that most individuals do not become aware of their own difficulties until adulthood. Yet, symptoms of DP may still be apparent in childhood, even if the affected individual is unaware of their difficulties at the time. Consequently, there is a possibility that DP can be spotted by others, if provided with an appropriate symptom checklist. To aid the detection of DP in both adulthood and childhood, this paper lists sixteen evidence-based hallmark symptoms of the condition.

Chapter Two:

Identifying Hallmark Symptoms of Developmental Prosopagnosia for Non-Experts

Full reference:

Murray, E., Hills, P., Bennetts, R., & Bate, S. (2018). Identifying Hallmark Symptoms of Developmental Prosopagnosia for Non-Experts. *Scientific Reports, 8*(1).
https://eprints.bournemouth.ac.uk/30271/

Introduction to Chapter Three

Self-Ratings of Face Recognition Ability are Influenced by Gender but not Prosopagnosia Severity

Chapter Three continues investigation into the self-report of DP symptoms. Most researchers agree that anecdotal evidence of everyday difficulties with face recognition is required to support a DP diagnosis. However, evidence suggests that many individuals have limited insight into their abilities and, consequently, self-reported ratings of face recognition skills mostly produce only weak-to-moderate associations with objective measures of performance (e.g. Bobak et al., 2016; Bowles et al., 2009; Palermo et al., 2017). Interestingly, the 20-item Prosopagnosia Index (PI20; Shah et al., 2015) has fared more successfully in identifying people who may have DP compared to the typical population.

However, it remains unclear whether such self-report measures can be used to index the *severity* of face recognition impairments in people with prosopagnosia. While Shah and colleagues (2015) suggested that PI20 scores in the ranges of 64-74, 75-84 and 85-100 are indicative of mild, moderate and severe DP respectively, a specific analysis considering the relationship between prosopagnosia severity and PI20 scores was not reported. Moreover, a novel question that has not yet been explored is whether participant gender influences the self-report of prosopagnosia symptoms, and there is some precedent to suggest this may be the case (for example, males rate themselves more favourably than females on traits such as intelligence and attractiveness; females rate the symptoms of their developmental condition as more severe than males).

Chapter Three addressed these issues in 47 adults with DP. Comparisons between PI20 and CFMT scores suggested that (a) self-report of prosopagnosia symptoms does not index disorder severity, and (b) gender does influence self-ratings of prosopagnosia symptoms. In light of these findings, we recommend that future research needs to probe the underpinnings of this gender effect, and, as a minimum, develop separate cut-offs (or even instruments) for use with male and female participants.

Chapter Three:

Self-Ratings of Face Recognition Ability are Influenced by Gender but not Prosopagnosia Severity

Full reference:

Murray, E., & Bate, S. (2019). Self-ratings of face recognition ability are influenced by gender but not prosopagnosia severity. *Psychological Assessment*, *31*(6), 828-832. https://eprints.bournemouth.ac.uk/32362/

Introduction to Chapter Four

Diagnosing Developmental Prosopagnosia: Repeat Assessment using the Cambridge Face Memory Test

Chapter Four moves away from issues surrounding self-report, to address outstanding issues associated with the dominant test used for DP diagnosis: the Cambridge Face Memory Test (CFMT). Current approaches to DP diagnosis typically employ a combination of self-report, a famous faces test, the CFPT and the CFMT. However, several variables likely influence performance on these tests irrespective of actual ability such as mood, fatigue or distraction. Furthermore, with an increased trend towards online testing (and hence, online screening), coupled with the fact that the CFMT is readily available online, individuals may be completing these tasks multiple times. Crucially, little is known about potential practice effects and, to date, there has been no examination of the test-retest reliability of the CFMT. Additionally, with increasing numbers of people referring themselves to research teams on suspicion of DP, there is a need to develop time-effective methods to identify the condition. This concern may potentially be addressed by existing evidence that suggests the CFMT can be shortened, while retaining its reliability for DP diagnosis.

Chapter 4 therefore examined the test-retest reliability of the CFMT, and found that this fell just below psychometric standards. Single-case analyses revealed further inconsistencies in performance that were not driven by testing location (online or in-person), nor the time-lapse between attempts. In addition, later administration of an alternative version of the CFMT (the CFMT-Aus) was found to be valuable in confirming borderline cases. Finally, our data support the claim that the CFMT can be shortened; the learn and novel stages are as sensitive to identifying cases of DP as is the entire test.

Chapter Four:

Diagnosing Developmental Prosopagnosia: Repeat Assessment using the Cambridge Face Memory Test

Full reference:

Murray, E., & Bate, S. (2020). Diagnosing developmental prosopagnosia: repeat assessment using the Cambridge Face Memory Test. *Royal Society Open Science, 7*(9). https://eprints.bournemouth.ac.uk/34647/

Introduction to Chapter Five

An Update of the Benton Facial Recognition Test

Chapter Five moves on to consider issues associated with existing tests of face perception. Indeed, the many criticisms associated with these tasks were described in Chapter One, with a focus on the dominant tasks used to date: the Benton Facial Recognition Test (BFRT) and the Cambridge Face Perception Test (CFPT). Nevertheless, the advantages of the BFRT paradigm lend it to be a rapid and simple assessment of face perception - a particularly important concern in very recent times as the trend for online testing increases. However, the images in the BFRT are flawed, in that they are unusual in colour, lighting, and viewpoints (even in those images that are not manipulated) and the extrafacial information within the stimuli allow the task to be completed by means other than processing the face (Duchaine & Weidenfeld, 2003). Thus, the images in the original BFRT require updating. Moreover, Chapter Four advocates the need for multiple versions of a task to facilitate repeat-assessment; that is, face perception should be measured using two different tasks which tap the same process.

Consequently, in Chapter Five, we present a new test of face perception which adopts the procedure of the BFRT, but is created using high quality, ambient facial images to address recent theoretical progress in the face recognition literature. Norming data and reliability and validation data are presented. Investigations using a sample of participants with developmental prosopagnosia (DP) show that some of these individuals were only impaired on the BFRT-r, and some were only identified using the BFRT-c. It is important to note that for DP inclusion criteria and test performance in the present paper, we adopted more liberal cut-offs of 1.7 SDs. This was to allow for the possibility that (a) face perception may be differentially affected in different individuals and the experiment sought to capture rather than exclude this variability,

and (b) because many think that face perception is impaired rather than intact in most DPs, we were more conservative towards deeming the latter. Together, these patterns of performance highlight the importance of administering more than one task when screening for face perception deficits. Data stress the importance of monitoring completion times as well as accuracy scores on tasks of face matching. Finally, the present paper supports existing reports that online testing is an effective approach to DP testing. Upon publication, we will make the new version of the BFRT freely available to the field.

Chapter Five:

An Update of the Benton Facial Recognition Test

Full reference:

Murray, E., Bennetts, R.J., Tree, J., & Bate, S. (2020). An Update of the Benton Facial Recognition Test. Manuscript submitted for publication: *Behavior Research Methods*.
https://eprints.bournemouth.ac.uk/35553/

Introduction to Chapter Six

Using Gaze Behaviour to Differentiate Between Developmental Prosopagnosia and Autism Spectrum Condition

The final chapter moves beyond self-report and behavioural screening for developmental prosopagnosia (DP), to examine the use of eye-movement technology in differentially diagnosing face recognition deficits in different developmental conditions. Indeed, existing reports, including that of Chapter Two, suggest that DP can be misinterpreted as an autism spectrum condition (ASC), especially in younger age groups. This is because face recognition deficits manifest in both conditions, and public and professional awareness of DP remains relatively low. Yet, behavioural symptoms of both ASC and DP can be apparent in very young children, although it has not yet been explored whether the likely different underpinnings of face recognition difficulties can be detected and differentiated.

One method which has the potential to address this issue is the use of eye-tracking technology. While ASC participants and those with DP have shown atypical eye gaze behaviour when viewing faces and/or scenes with faces in contrast to control participants, the two groups have not yet been directly compared. Yet, one might predict that face recognition difficulties in ASC are more prevalent in social or interactive contexts given the common deficits in social communication and interaction, whereas those in DP are thought to be visuo-cognitive in nature, and should therefore remain stable across all social conditions.

We address this issue in Chapter Six, exploring the differences in eye gaze behaviour between those with ASC, DP, and typically developing controls. It is important to stress that data collection was interrupted by the on-going COVID-19 pandemic and, as such, the sample size presented here is small. However, we report that individuals with ASC look towards non-social

stimuli to a greater extent than DPs and controls when scenes are more social and interactive in nature. Trends in the data also suggest that those with DP look at the hair and bodies more than control and ASC participants. When examining eye gaze on the internal features of the face, DP participants dwelled on the mouth more than those with ASC and control participants. Taken together, the preliminary analyses presented here suggests that this paradigm potentially can identify group differences in eye-movements, as hypothesised, when the power increases. Consequently, when authorised to do so, data collection will resume and the present chapter will be prepared for publication.

Chapter Six:

Using Gaze Behaviour to Differentiate Between Developmental Prosopagnosia and Autism Spectrum Condition

Ebony Murray, Stephanie Chase, Nicola Gregory, Rachel Moseley and Sarah Bate

Chapter Seven: General Discussion

Present protocols for the diagnosis of developmental prosopagnosia (DP) in adulthood recommend the use of a combination of self-report, a famous faces test, a test of unfamiliar short-term memory and a face perception task (Bate & Tree, 2017; Dalrymple & Palermo, 2015). However, there are inconsistencies in the tasks which are used across research teams, and even for the tasks which most agree to use, there are some concerns. Thus, the overarching aim of the thesis was to address contemporary issues relating to the diagnosis of DP, and to develop the field's understanding of existing diagnostic protocols. More specifically, the thesis aimed to:

- (a) develop an empirically-driven list of symptoms that are representative of the disorder,
- (b) evaluate the utility of existing self-report instruments (the PI20) in discriminating disorder severity and its susceptibility to gender effects,
- (c) evaluate the outstanding psychometric issue of test-retest reliability in the CFMT,
- (d) develop and evaluate a new test of face perception that overcomes issues with existing tests and offers the opportunity for repeat-assessment
- (e) investigate the utility of eye-tracking as a tool for distinguishing DP from ASC.

First of all, given that inconsistencies in diagnosing DP is evident within the literature, it is imperative to outline how DP was classified in the present thesis, before discussing the findings in more detail. For all chapters, individuals self-referred themselves to the research team. They registered their details on our website provided anecdotal evidence of their face processing difficulties in their everyday lives. They were then contacted to either attend a screening session in the lab, or online. Typically, individuals completed the CFMT and CFPT. Those who came into the lab, or competed the tests online but were willing to complete a famous faces test over the phone, completed said test. For those who completed all three tests, they needed to show impairment on two or more of these. For those who only completed the CFMT and CFPT, they needed to show an impairment on both tests. Thus, we followed existing protocols that an individual shows atypicality on two or more of three objective measures of face processing (Bate & Tree, 2017; Dalrymple & Palermo, 2015). However, we note that many studies use impaired performance on two tests of face memory, and do not use the CFPT in the diagnosis of the condition (e.g. Biotti et al., 2017; Fry et al., 2020; Susilo & Duchaine,

2013; Ulrich et al., 2017). Presumably, this reflects the fact that difficulties with face *recognition* is the characteristic symptom of the condition.

Nevertheless, assessing both face perception and face memory acknowledges that DP can present in different ways, and the idea of there being different subtypes of DP is not new (for discussions see Bate et al., 2019; De Renzi et al., 1991; McKone et al., 2011). In addition, recent work from our lab suggests that impairments on the CFMT and CFPT, but intact famous faces test scores, is very rare (Bate et al., 2019). Indeed, most will also be impaired on a famous faces test (thus, impaired on all three tasks), or will *only* be atypical on a famous faces test. Hence, in the cases where participants only completed the CFMT and CFPT, it is probably fair to assume that they would also score within the atypical range on a famous faces test as well. Nevertheless, we point out here that during the present thesis, only a very small minority of participants were classified as DP in the absence of an atypical score on the CFMT and famous faces test. A downfall of using only the CFPT and CFMT however, is that individuals who were typical on the CFPT and atypical on the CFMT (or vice versa), may have been DP. An atypical score on a famous faces test could have confirmed their DP, and additional participants may have been available for the research. Indeed, as the popularity of online testing continues to rise, a virtual method of administering a famous faces test would be valuable. Whilst attempts have been made to do this, it is noted that the scores on these tasks may be deflated as unfamiliar famous people are not excluded from the overall scores (Aripze et al., 2019), as is typically done in the lab.

Moving on to the thesis and research presented here, an initial investigation developed a list of hallmark symptoms of DP for non-experts to use to identify the condition in themselves and in another person. Chapter Three extended this investigation of self-report tools, examining the congruency of subjective and objective face processing measures, while also exploring gender effects in self-report. The thesis then moved to objective tests of face processing. Chapter Four investigated the test-retest reliability of the leading test of unfamiliar face recognition, the Cambridge Face Memory Test (CFMT), and the importance of repeat-testing in DP screening. Chapter Five presented a new test of face perception which overcomes the limitations and builds upon the strengths of existing tasks, and investigated the diagnostic utility of repeated assessment of a single subprocess of face processing (in this case, face matching). Finally, Chapter Six presents the preliminary data from the first known study that has directly compared the eye-movements of those with autism spectrum conditions (ASCs) and those with DP. Below, I discuss each of the contemporary issues relating to the diagnosis

of DP, in light of the research presented in the current thesis. I outline what the findings have added to the current understanding of identifying the condition, and highlight area for future research.

Identifying Hallmark Symptoms of Developmental Prosopagnosia for Non-Experts

Self-report methods which aim to measure one's face processing abilities have been developed based upon anecdotal evidence and discussions in the laboratory (e.g. The PI20: Shah et al., 2015; The Cambridge Face Memory Questionnaire: Arizpe et al., 2019). While there has been some promise of subjective measures being able to identify cases of DP (e.g. Arizpe et al., 2019; Gray et al., 2017; Livingston & Shah, 2017) there is mixed evidence for the success of self-report measures elsewhere (e.g. Bate & Dudfield, 2019; Estudilo & Wong, 2021; Palermo et al., 2017). As such, there is evidence to suggest that individuals have a limited insight into their face processing skills. Chapter Two builds on this argument using qualitative data; we report that individuals with DP feel that they have a limited insight into their in(abilities) and often do not become aware of their difficulties until middle-late adulthood. Consequently, it will mean that identifying one's face recognition deficits may fall to someone else and, accordingly, a list of simple to identify symptoms of DP would be beneficial for non-experts. Thus, Chapter Two developed a list of hallmark symptoms based on empirical evidence from those living with DP, their unaffected significant others, and parents of children with DP, was developed.

To our knowledge, this is the first list of hallmark symptoms which have been developed based on empirical data, and are useful both for oneself and for another person, both in adulthood and childhood. However, it is not to go unappreciated that this checklist was devised by data obtained from participants who either have *confirmed* DP, are significant others of someone with confirmed DP, or are parents of DP children. It would therefore be valuable to assess how well this checklist differentiates those who have met the criteria for DP and control participants. Indeed, we have started to collect data on this matter; we sent the final list of the 18 symptoms back to those who participated in the original investigation. We asked participants to just mark whether they "agreed" or "disagreed" whether that specific symptom reflected their own experiences. Each person could return a maximum total of 18 "agrees". A total of 21 DPs, 6 SOs and 2 parents responded. We also sent this to 21 age-matched control participants to ask them to reflect upon their own face processing skills, 6 participants who

answered the questionnaire whilst thinking about their significant others' face processing abilities, and 5 parents who reflected on their child's skills. To summarise, DPs returned higher "agree" than controls (N = 310 out of a possible 378, and 68 respectively), as did significant others (N = 73 out of a possible 108 for DP significant others; N = 21 for controls) and parents (N = 30 out of a possible 36 for DP parents; N = 17 out of a possible 90 for controls). On face value, the symptoms listed in Chapter Two reflect the real-world experiences of DP.

However, we also recommend – and are working towards – future investigation that assess how well this checklist differentiates those who have met the criteria for DP and those who self-report face recognition difficulties but do not meet the current diagnostic criteria for the condition. The checklist has been translated to a questionnaire, in which participants are asked to score their (or their child's, or their significant other's) face recognition skills/experiences on a scale of 1 – 5. Individuals with confirmed DP are completing this alongside unrelated research, and individuals who refer themselves to our website are also completing it when they register. Furthermore, it is suggested that future work seeks to fully validate the checklist and determine which questions should receive a higher weighting in a self-report version of the questionnaire, or when working with children with DP. Thus, a large, new sample of DPs, significant others and parents should be sought in future investigations. It is hoped that the research which is underway will go some way to answer these questions.

Self-Ratings of Face Recognition Ability are Influenced by Gender but not Prosopagnosia Severity

As aforementioned, existing self-report methods have been developed based upon anecdotal evidence and discussions in the laboratory. To date, there is mixed evidence for the success of self-report measures. Chapter Three of the present thesis adds to this debate; we report no correlation between the CFMT and PI20 in a sample of individuals previously identified as having DP. Thus, the research strongly suggested that individuals with DP have a limited insight into their face recognition (in)abilities (a suggestion which is further supported by the qualitative data we present in Chapter Two). This may be because our DP participants were recruited using more conservative inclusion criteria than those of Livingston and Shah; whilst our DP participants scored at least two standard deviations below the mean on two or more of three tasks of face processing, the DPs in Livingston and Shah's report differed by controls – in some cases – by only one standard deviation. As a result, it is argued here that it

may be that the previously reported correlations were driven by individuals who would be better categorised as "typical perceivers". The PI20 may therefore be sufficiently sensitive to tap face recognition skills across the typical spectrum, but not to gauge severity in people with prosopagnosia. Whilst this may indeed be the case, it might also be that this reflects the fact that a sample of DPs will inevitably provide a limited range on the CFMT, given that a low CFMT score is (almost always) a prerequisite of DP classification (although see Bate et al. (2019) for cases where CFMT is intact, and CFPT and famous faces test scores are impaired).

Secondly, the paper found that males self-report their DP symptoms as less severe than female DPs. However, the direction and underpinnings of the effect reported here are unclear: while women over-reported prosopagnosia symptoms compared to men, they also showed no relationship between PI20 and CFMT scores. Data suggest that males may have more accurate insights into their face recognition skills, although it is noted that larger sample sizes are required in future work in order to further explore this effect. It should also be investigated whether gender effects are driven by certain behavioural traits (i.e. certain items in the PI20); indeed, if some items have greater sensitivity to the traits of male or female DPs, the existing instrument may require some expansion. Indeed, this is also something which is remains unanswered in relation to the checklist of symptoms that we present in Chapter Two, and future research endeavours to explore whether there are gender-differences on certain items of our own checklist as well.

Much remains unanswered or left to be debated in light of self-report and face processing skills. Importantly, no consensus exists among the research community for how self-report should be applied when identifying cases of DP. Should a DP diagnosis be given based upon exceptionally poor performance on objective measures but 'typical' scores on self-report instruments? What happens in cases where individuals return significantly low self-report scores but fall within the typical range on objective measures? The data presented in the current thesis cannot resolve this, but the checklist offers a more inclusive method of identifying face recognition difficulties, and the data in Chapter Three do go some way to add to the literature that self-report methods will not be successful, alone, to identify all cases of DP. In line with this, we suggest that self-reported everyday difficulties with face processing must be evident (and, of course, is usually the case given that these individuals self-refer themselves to the research team), but that objective face processing difficulties are needed to support and confirm a diagnosis.

Diagnosing Developmental Prosopagnosia: Repeat Assessment using the Cambridge Face Memory Test

Once individuals report that they have difficulties with face recognition in their daily lives, they are typically then invited to participate in a series of objective measures of face processing. This approach of repeated testing decreases the likelihood that a diagnosis is given (or missed) by chance. Furthermore, this approach considers the viewpoint that DP is a heterogeneous condition and may be composed of subtypes (De Renzi et al., 1991). Yet, there has been some inconsistencies in the objective measures which are used to identify cases of DP. Fortunately, the field has a "gold-standard" test of unfamiliar face recognition in the Cambridge Face Memory Test (CFMT). However, whilst much work has been done to explore the psychometric properties, validity, and reliability of the CFMT, there has been no investigation into the test-retest reliability of the task. Consequently, Chapter Four for the first known time, reported the test-retest reliability of the test. The diagnostic utility of administering a shorter version of the CFMT was explored, as was the administration of an alternative version of the CFMT (the CFMT-Aus).

Strikingly, the test-retest reliability of the CFMT was found to fall just short of accepted psychometric standards, producing a reliability co-efficient of .68. While these psychometric standards are applied to a range of psychometric tests, a co-efficient of .68 is lower than typically seen on tests of visual memory (e.g. .74 in Nakayama et al., 2014) and specifically, on alternative tests of face memory (e.g. .74 and .76 on the faces subtest of the Recognition Memory Test, in Bird et al., 2003). Single-case analyses also identified some inconsistencies in performance and interestingly, approximately 30% of participants performed better at their first attempt on the CFMT than on their second. Thus, some day-to-day inconsistencies in performance aside from practice effects are evident. These inconsistencies, and the practice effect itself, may result from individual variability in fluctuations in cognitive performance, or from intrinsic factors such as mood, fatigue, or distraction. While we did not directly assess these factors here, they should be considered in future research.

Finally, Chapter Four supports a recent proposal that the CFMT retains its sensitivity when only the learn and test stages are administered (Corrow et al., 2018). Indeed, this finding highlights that a more efficient, shorter version of the task can be utilised and that will consequently benefit the burden of repeated assessment. Taken together, as performance on tests of face processing can be affected by factors irrespective of one's face processing skills

(e.g. mood, fatigue, chance) and the test-retest reliability of the CFMT falls just short of the accepted parametric standards, the present paper argues for a new protocol for the screening of DP: a second, alternative test which assesses the same skill should be administered, preferably on a different day. To make this process more efficient, our data support the claim that the CFMT can be shortened; the learn and novel stages are as sensitive to identifying cases of DP as is the entire test. Finally, data also support that online testing is equally efficient and reliable as in-lab testing.

Alternatively, this pattern of findings may simply reflect the influence of chance and, aside from the unlikely event that a task has perfect reliability, individual scores will suffer from at least some uncertainty. Yet, this measurement error is rarely considered by researchers when identifying cases of DP. A potential solution to practice effects on the CFMT, and to consider this measurement error, is to develop different versions of the task, using the same paradigm but novel stimuli. Indeed, previous research also suggests that whilst practice effects are often large with repeated versions of the same memory test (for a discussion, see Calamia et al., 2012), alternative forms of the same test yield much smaller practice effects (e.g. Beglinger et al., 2005; Benedict & Zgaljardic, 1998). The current study investigated this possibility, via administration of the CFMT-Aus; there were inconsistencies in performance between the original CFMT and CFMT-Aus in both directions, and the paper therefore argued that repeat assessment should occur on a separate day. This would address any ambiguous cases stemming from differences in day-to-day cognitive functioning.

An Update of the Benton Facial Recognition Test

Chapter Four recommended that two separate tests of face processing which assess the same skill should be administered on different days. Chapter Five further explored this recommendation in the context of face perception. Firstly, following consideration of existing face perception tests, including the Cambridge Face Perception Test (CFPT) and the Benton Facial Recognition Test (BFRT) (although note that there are other face perception tests available to the field such as the Glasgow Face Matching Test and the Pairs Matching Test, these tasks are more widely used to explore performance of 'Super Recognisers', and reports have suggested that matching paradigms are vulnerable to responses biases; hence, these were not utilised in the present study), a new test of face perception was presented (the BFRT-r). The BFRT-r is quick to administer, with typically developing control participants taking

approximately four to six minutes to complete the task. Moreover, the task is very simple in its procedure, making it highly suitable for the growing trend for online testing. Similar to the findings reported in Chapter Four, data support online implementation: strong internal reliability and inter-item correlations derived from online samples directly support this mode of testing. Still, it is appreciated that a pivotal question remains unanswered at present, that is, whether the BFRT-r is equally vulnerable to high performance based on extrafacial information. Thus, future research is underway to explore this.

Secondly, to investigate the importance of administering more than one version of a task which assesses the same subprocess of face processing, participants with DP were asked to complete the BFRT-r *and* the BFRT-c (Rossion & Michel, 2018). Approximately half of the DP sample presented with no impairment on either version of the BFRT. While eight of the remaining 16 DPs consistently displayed deficits on both versions of the BFRT, five were only detected on the BFRT-r and three by the BFRT-c. In sum, these patterns of performance highlight the importance of administering more than one task when screening for face perception deficits, further supporting the recommendation for repeated testing in Chapter Four.

Moreover, this paper explored the role of measuring completion times for the assessment of DP. While existing tests of face perception enforce a strict time-limit to complete the test (the CFPT) or do not measure completion times at all (the original version of the BFRT), completion times are easily monitored in both the BFRT-c and BFRT-r given their electronic format. Consistent with existing work (e.g. Bukach, et al., 2006; Busigny & Rossion, 2010; Delvenne et al., 2004; Jansari et al., 2015), the finding reported here that 12 DPs were only impaired on completion time (but not accuracy) on either, or both tests, highlights the importance of assessing both measures. We acknowledge that longer completion times may reflect the use of laboured face processing strategies and methods which ultimately lead to a correct response. However, the older adult control participants took longer to complete the task than younger adults, although the same effect did not emerge for accuracy. Thus, to further the recommendations for identifying DP accurately, we strongly suggest that age-matched norms are used for identifying impaired performance on these tasks.

What this finding also brings light to is the nature of DP, that is, the argument of whether those with DP are quantitatively or qualitatively different to typically developing individuals. Indeed, research identifying individuals who are extremely good at recognising

faces – so called "super recognisers" (Russel et al., 2009) – suggests that those with DP are simply one end of the face recognition spectrum, whilst super recognisers represent the other (top) end of the spectrum. If this view is believed, the that is evidence that DP performance is quantitatively different from the majority, something which forms the basis for the quantitative data presented in Chapter Three and Chapter Four. However, these differences in response times in Chapter Five suggest that those with DP are likely qualitatively different to typical developing controls. That is, they are eliciting signs that they have different processing strategies when completing face processing tasks. This view is further supported by eye-tracking data, something which Chapter Six of the present thesis speaks to.

Using Gaze Behaviour to Differentiate Between Developmental Prosopagnosia and Autism Spectrum Condition

Finally, the thesis presents preliminary analyses of eye-tracking data, which aimed to explore eye-movement differences between those with ASC and those with DP. Existing reports, including Chapter Two of this thesis, suggest that ASC and DP can be confused, especially in childhood. It is vital to accurately differentiate these conditions so that a misdiagnosis can be avoided, and appropriate support can be provided. Chapter Six presents the first study, to our knowledge, that directly compares the eye-movements of those with DP to those with ASC with the aim of differential diagnosis.

Firstly, the data support and build upon existing reports revealing that those with ASC and DP elicit atypical eye-movements in response to scenes containing faces, as compared to control participants. Moreover, it was found that ASC participants spend more time looking at the background in scenes (i.e. non-social information) than DP participants, when the scene is more social in nature. It is vital to note that these group differences are the only ones to reach statistical significance here and this is highly likely due to the low sample size in this analysis - a consequence of interrupted data collection due to the ongoing pandemic. Yet, the data strongly suggest that those with ASC, and more importantly in light of the focus of the present thesis, DP is qualitatively different to the typically developing population. Whilst eye-tracking research in DP is scarce, these preliminary findings go some way to support the existing literature, and supporting the view that those with DP are qualitatively different to control participants; those with DP are doing something qualitatively different to typically developing individuals (and to those with ASC) when presented with faces.

Additionally, trends within the data suggest that DPs look more at extrafacial information including the hair and bodies, supporting our hypothesis. When looking at eye-movement data on the internal features of the face, both DPs and ASC participants look less at the eyes that control participants. Differences between DP and ASC participants, however, were seen in the descriptive statistics in dwell times on the mouth; those with ASC looked at the mouth similar to controls and less than those with DP. It is difficult to interpret this trend in the data at present, but it would be valuable to explore this difference further when a larger sample size has been tested. We fully acknowledge and appreciate that the present analysis requires more data to increase the power; unfortunately, the present COVID-19 pandemic interrupted data collection. Accordingly, when able to do so, data collection will continue, and the chapter will be prepared as a manuscript for publication.

Summary and Future Directions

This thesis presents a series of chapters which develop the current protocols for the diagnosis of DP, and extend existing recommendations to more reliably identify the condition. In summary, the work started by highlighting the challenges of self-report; not only do those with DP have a limited insight into their face recognition difficulties, there is also a gender difference in self-reporting behaviours which have not previously been identified. Indeed, future work is required to investigate this gender difference further and there is more to be known about the underpinnings of this effect. Based on these findings, however, it is recommended that gender-matched norms are used to identify atypical scores on self-report measures. Moreover, additional research aims to fully validate the symptom list presented in Chapter Two. Investigating gender differences in self-report using this checklist would also be invaluable; if gender differences are not apparent in the presented, empirically-driven symptom checklist, then this would raise questions about the instruments more so that gender effects (and vice versa). Similarly, additional work should explore the relationship between the checklist that we present in Chapter Two and existing self-report instruments, in order to explore the benefits of developing a symptom list based on empirical data. Taken together, more would be understood about self-report and its role in the identification of DP.

In addition, Chapters Four and Five strongly advocate the use of repeated testing in DP screening. At present, those suspicious of DP complete multiple tests, but each test typically assesses different subprocesses of face processing (that is, familiar face recognition, unfamiliar

face memory, face perception). Thus, the same process is not tested repeatedly, and this approach fails to acknowledge that scores can be achieved irrespective of actual ability; extraneous factors such as day-to-day fluctuations in cognitive abilities, intrinsic factors such as mood and fatigue, and potential issues with the psychometric properties of the task being used are overlooked. Thus, individuals can still be identified – or missed – for DP by only assessing each skill once. However, future work should further explore the diagnostic utility of the recommendation that two different tasks which assess the same skill, should be administered on separate days. Given this thesis also recommends that the efficiency of screening can be improved, repeat-testing should not be laborious for the participant. Indeed, we found that the CFMT can retain its diagnostic utility if shortened, and the BFRT-c and BFRT-r are very quick to complete. Combined with the findings that online testing is equally as effective as in-lab testing, this battery of tests would provide researchers with a more thorough examination of an individual's face recognition deficits at no additional cost (financially nor time) to the researcher.

Finally, Chapter Six presents preliminary data from the first study, to our knowledge, to directly compare DP and ASC via eye-movement analyses. While data collection will resume and more firm conclusions will be drawn from this experiment in the future, it appears that eye-tracking offers a short, engaging and low-demanding task which may successfully differentiate these conditions. Future research might also aim to explore the observed scanning differences in children with the conditions: this would exclude the possibility that the effects simply reflect coping strategies which have been developed by adulthood, and it may be that they are exacerbated in younger samples. While we await authorisation to resume this study, we hope that this work will initiate an exciting new avenue for DP diagnosis, allowing face recognition deficits to be differentiated across developmental conditions, and in increasingly younger participants. Ultimately, improved diagnostic protocols will improve the scientific study of DP, and allow those who experience the condition to access appropriate support, and perhaps intervention, in the timeliest manner.

Concluding Remarks

The present thesis has aimed to address some of the contemporary issues relating to the diagnosis of DP, and develop the field's understanding of existing diagnostic protocols. Taking all evidence together, we recommend that the empirically-driven list of hallmark symptoms

presented in Chapter Two be used to identify face recognition difficulties in oneself or in another, in adulthood or in childhood, thus overcoming the limitations that individuals lack insight into their own (in)abilities. Given that Chapter Three supports the existing literature which reports that subjective and objective evidence of face recognition impairments do not have a significant relationship, we propose that objective measures of face processing are always then administered. Whilst different research teams use different tests to identify cases of DP (for a discussion, see Corrow & Barton, 2016), we strongly advocate the use of repeat testing. Chapter Four and Five illustrate that individuals can score in the typical range on a face processing task on one day, but in the atypical range on another. Repeat testing will ensure that those who are being identified as DP are not 'false positives', but will also reduce the chance of missing 'false negatives'. Given the evidence here for practice effects and/or low test-retest reliability of face processing tasks, combined with the evidence in Chapter Five, it is proposed that different tests which tap the same subprocess of face processing are used during this repeated testing. Finally, when faced with (potentially) alternative explanations for one's face recognition difficulties, Chapter Six's preliminary data suggest that eye-tracking may offer a novel method to differentiate conditions.

References

Adams, A., Hills, P., Bennetts, R., & Bate, S. (2019). Coping strategies for developmental prosopagnosia. *Neuropsychological Rehabilitation*, 1-20. https://doi.org/10.1080/09602011.2019.1623824

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Annaz, D., Karmiloff-Smith, A., Johnson, M., & Thomas, M. (2009). A cross-syndrome study of the development of holistic face recognition in children with autism, Down syndrome, and Williams syndrome. *Journal of Experimental Child Psychology*, *102*(4), 456-486. https://doi.org/10.1016/j.jecp.2008.11.005

Arizpe, J., Saad, E., Douglas, A., Germine, L., Wilmer, J., & DeGutis, J. (2019). Self-reported face recognition is highly valid, but alone is not highly discriminative of prosopagnosia-level performance on objective assessments. *Behavior Research Methods, 51*(3), 1102-1116. https://doi.org/10.3758/s13428-018-01195-w

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

Barton, 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., & Corrow, S. (2016). The problem of being bad at faces. *Neuropsychologia*, *89*, 119-124. https://doi.org/10.1016/j.neuropsychologia.2016.06.008

Barton, J., Press, D., Keenan, J., & O'Connor, M. (2002). Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. *Neurology, 58*(1), 71-78. https://doi.org/10.1212/wnl.58.1.71

Bate, S., & Bennetts, R. (2014). The rehabilitation of face recognition impairments: a critical review and future directions. *Frontiers in Human Neuroscience, 8*. https://doi.org/10.3389/fnhum.2014.00491

Bate, S., Bennetts, R., Gregory, N., Tree, J., Murray, E., & Adams, A., Bobak, A.K., Penton, T., Yang, T., & Bannisy, M.J. (2019). *Brain Sciences*, *9*(6), 133. https://doi.org/10.3390/brainsci9060133

Bate, S., Frowd, C., Bennetts, R., Hasshim, N., Murray, E., & Bobak, A.K., Wills, H., & Richards, S. (2018). Applied screening tests for the detection of superior face

recognition. *Cognitive Research: Principles and Implications*, *3*(1). https://doi.org/10.1186/s41235-018-0116-5

Bate, S., Frowd, C., Bennetts, R., Hasshim, N., Portch, E., Murray, E., & Dudfield, G. (2019). The consistency of superior face recognition skills in police officers. *Applied Cognitive Psychology*, *33*(5), 828-842. https://doi.org/10.1002/acp.3525

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., Haslam, C., Tree, J., & Hodgson, T. (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

Bate, S., Parris, B., Haslam, C., & Kay, J. (2010). Socio-Emotional Functioning and Face Recognition Ability in the Normal Population. *Personality and Individual Differences*, 48 (2), 239-242. https://doi.org/10.1016/j.paid.2009.10.005

Bate, S., & Tree, 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

Beglinger, L., Gaydos, B., Tangphao-Daniels, O., Duff, K., Kareken, D.A., & Crawford, J., Fastenau, P.S., & Sirmers, E.R. (2005). Practice effects and the use of alternate forms in serial neuropsychological testing. *Archives of Clinical Neuropsychology*, *20*(4), 517-529. https://doi.org/10.1016/j.acn.2004.12.003

Benedict, R., & Zgaljardic, D. (1998). Practice Effects During Repeated Administrations of Memory Tests With and Without Alternate Forms. *Journal of Clinical And Experimental Neuropsychology*, *20*(3), 339-352. https://doi.org/10.1076/jcen.20.3.339.822

Bentin, S., Deouell, L., & Soroker, N. (1999). Selective visual streaming in face recognition. *Neuroreport*, *10*(4), 823-827. https://doi.org/10.1097/00001756-199903170-00029

Benton, A. (1990). Facial Recognition 1990. *Cortex*, *26*(4), 491-499. https://doi.org/10.1016/s0010-9452(13)80299-7

Benton, A. L., Sivan, A. B., Hamsher, K. D. S., Varney, N. R., & Spreen, O. (1983). Facial recognition: Stimulus and multiple choice pictures. In A. L. Benton, A. B. Sivan, K. D. S. Hamsher, N. R. Varney, & O. Spreen (Eds.), Contribution to neuropsychological assessment (pp. 30–40). Oxford University Press.

Benton, A. L., & Van Allen, M. W. (1968). Impairment in facial recognition in patients with cerebral disease. *Transactions of the American Neurological Association, 93*, 38–42. https://doi.org/10.1016/s0010-9452(68)80018-8

Bennetts, R., Butcher, N., Lander, K., Udale, R., & Bate, S. (2015). Movement cues aid face recognition in developmental prosopagnosia. *Neuropsychology*, *29*(6), 855-860. https://doi.org/10.1037/neu0000187

Bennetts, R., Murray, E., Boyce, T. & Bate, S. (2018). Prevalence of face recognition deficits in middle childhood. *The Quarterly Journal of Experimental Psychology*. 70(2), 234-258. https://doi.org/10.1080/17470218.2016.1167924

Bindemann, M., Attard, J., & Johnston, R. (2014). Perceived ability and actual recognition accuracy for unfamiliar and famous faces. *Cogent Psychology, 1*(1). https://doi.org/10.1080/23311908.2014.986903

Bindemann, M., Avetisyan, M., & Rakow, T. (2012). Who can recognize unfamiliar faces? Individual differences and observer consistency in person identification. *Journal Of Experimental Psychology: Applied, 18*(3), 277-291. https://doi.org/10.1037/a0029635

Biotti, F., Gray, K., & 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. (2017). Normal Composite Face Effects in Developmental Prosopagnosia. *Cortex, 95*, 63-76. https://doi.org/10.1016/j.cortex.2017.07.018

Bird, C., Papadopoulou, K., Ricciardelli, P., Rossor, M., & Cipolotti, L. (2003). Test-retest reliability, practice effects and reliable change indices for the recognition memory test. *British Journal of Clinical Psychology*, *42*(4), 407-425. https://doi.org/10.1348/014466503322528946

Birnbaum, M. (2004). Human Research and Data Collection via the Internet. *Annual Review of Psychology*, *55*(1), 803-832. https://doi.org/10.1146/annurev.psych.55.090902.141601

Bobak, A., Pampoulov, P., & Bate, S. (2016). Detecting Superior Face Recognition Skills in a Large Sample of Young British Adults. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01378

Bobak, A., Parris, B., Gregory, N., Bennetts, R., & 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

Bowles, D., McKone, E., Dawel, A., Duchaine, B., Palermo, R., & Schmalzl, L., Rivolta, D., Wilson, 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

Bradshaw, J., Shic, F., & Chawarska, K. (2011). Brief Report: Face-Specific Recognition Deficits in Young Children with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *41*(10), 1429-1435. https://doi.org/10.1007/s10803-010-1150-4

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

Bukach, C. M., Bub, D. N., Gauthier, I., & Tarr, M. J. (2006). Perceptual expertise effects are not all or none: Spatially limited perceptual expertise for faces in a case of prosopagnosia. *Journal of Cognitive Neuroscience*, *18*, 48–63. https://doi.org/10.1162/089892906775250094

Burns, E., Bennetts, R., Bate, S., Wright, V., Weidemann, C., & Tree, J. (2017). Intact word processing in developmental prosopagnosia. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-01917-8

Busigny, T., & Rossion, B. (2010). Acquired prosopagnosia abolishes the face inversion effect. *Cortex*, 46, 965–981. https://doi.org/10.1016/j.cortex.2009.07.004

Bylemans, T., Vrancken, L., & Verfaillie, K. (2020). Developmental Prosopagnosia and Elastic Versus Static Face Recognition in an Incidental Learning Task. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.02098

Calamia, M., Markon, K., & Tranel, D. (2012). Scoring Higher the Second Time Around: Meta-Analyses of Practice Effects in Neuropsychological Assessment. *The Clinical Neuropsychologist*, *26*(4), 543-570. https://doi.org/10.1080/13854046.2012.680913

Carey, S. (1992). Becoming a face expert. *Philosophical Transactions of The Royal Society B Biological Sciences*, *335*(1273). https://doi.org/10.1098/rstb.1992.0012

Chawarska, K., Ye, S., Shic, F., & Chen, L. (2015). Multilevel Differences in Spontaneous Social Attention in Toddlers with Autism Spectrum Disorder. *Child Development*, *87*(2), 543-557. https://doi.org/10.1111/cdev.12473

Chawarska, K., & Volkmar, F. (2007). Impairments in monkey and human face recognition in 2-year-old toddlers with Autism Spectrum Disorder and Developmental Delay. *Developmental Science*, *10*(2), 266-279. https://doi.org/10.1111/j.1467-7687.2006.00543.x

Cooper, K.M., Krieg, A., & Brownell, S.E. (2018). Who Perceives They Are Smarter? Exploring the Influence of Student Characteristics on Student Academic Self-Concept in Physiology. *Advances in Physiology Education*, *42*(2), 200-208. https://doi.org/10.1152/advan.00085.2017

Corrow, S., Albonico, A., & Barton, J. (2018). Diagnosing Prosopagnosia: The Utility of Visual Noise in the Cambridge Face Recognition Test. *Perception*, *47*(3), 330-343. https://doi.org/10.1177/0301006617750045

Corrow, S.L., Dalrymple, K.A. & Barton, J.J. (2016). Prosopagnosia: current perspectives. *Eye and Brain*, 26(8), 165-175. https://doi.org/10.2147/EB.S92838

Daini, R., Comparetti, C.M. & Ricciardeli, P. (2014). Behavioral dissociation between emotional and non-emotional facial expressions in congenital prosopagnosia. *Frontiers in Human Neuroscience*, *8*, 974. https://doi.org/10.3389/fnhum.2014.00974

Dalrymple, K.A., Elison, J.T. & Duchaine, B. (2017). Face-selective -and domain-general visual processing deficits in children with developmental prosopagnosia. *Quarterly Journal of Experimental Psychology*, *70(2),* 259-275. https://doi.org/10.1080/17470218.2015.1122642

Dalrymple, K.A., Fletcher, K., Corrow, S., das Nair, R., Barton, J.J., Yonas, A., & Duchaine, B. (2014). "A Room Full of Strangers Every Day": The Psychosocial Impact of Developmental Prosopagnosia on Children and Their Families. *Journal of Psychosomatic Research*, *77*(2), 144-150. https://doi.org/10.1016/j.jpsychores.2014.06.001

Dalrymple, K., 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

de Gelder, B., & Stekelenburg, J. (2005). Naso-temporal asymmetry of the N170 for processing faces in normal viewers but not in developmental

prosopagnosia. *Neuroscience Letters, 376*(1), 40-45. https://doi.org/10.1016/j.neulet.2004.11.047

De Luca, M., Pizzamiglio, M., Di Vita, A., Palermo, L., Tanzilli, A., Dacquino, C., & Piccardi, L. (2019). First the nose, last the eyes in congenital prosopagnosia: Look like your father looks. *Neuropsychology*, *33*(6), 855-861. https://doi.org/10.1037/neu0000556

De Renzi, E., Faglioni, P., Grossi, D., & Nichelli, P. (1991). Apperceptive and Associative Forms of Prosopagnosia. *Cortex, 27*(2), 213-221. https://doi.org/10.1016/s0010-9452(13)80125-6

DeGutis, J., Chatterjee, G., Mercado, R., & Nakayama, K. (2012). 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., & Nakayama, K. (2014). Holistic face training enhances face processing in developmental prosopagnosia. *Brain*, *137*(6), 1781-1798. https://doi.org/10.1093/brain/awu062

Delvenne, J.F., Seron, X., Coyette, F., & Rossion, B. (2004). Evidence for perceptual deficits in associative visual (prosop)agnosia: A single-case study. *Neuropsychologia, 42*, 597–612. https://doi.org/10.1016/j.neuropsychologia.2003.10.008

Dennett, H., McKone, E., Tavashmi, R., Hall, A., Pidcock, M., Edwards, M., & Duchaine, B. (2011). The Cambridge Car Memory Test: A task matched in format to the Cambridge Face Memory Test, with norms, reliability, sex differences, dissociations from face memory, and expertise effects. *Behavior Research Methods, 44*(2), 587-605. https://doi.org/10.3758/s13428-011-0160-2

Duchaine, B. (2000). Developmental prosopagnosia with normal configural processing. *Neuroreport*, *11*(1), 79-83. https://doi.org/10.1097/00001756-200001170-00016

Duchaine, B., Germine, L., & Nakayama, K. (2007). Family resemblance: Ten family members with prosopagnosia and within-class object agnosia. *Cognitive Neuropsychology, 24*(4), 419-430. https://doi.org/10.1080/02643290701380491

Duchaine, D., & Nakayama, K. (2005). Dissociations of face and object recognition in developmental prosopagnosia. *Journal of Cognitive Neuroscience*, *17*(2), 249-261. https://doi.org/10.1162/0898929053124857

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, *44*(4), 576-585. https://doi.org/10.1016/j.neuropsychologia.2005.07.001

Duchaine, B., & Weidenfeld, A. (2003). An evaluation of two commonly used tests of unfamiliar face recognition. *Neuropsychologia*, *41*(6), 713-720. https://doi.org/10.1016/s0028-3932(02)00222-1

Ehrlinger, J. & Dunning, D. (2003). How Chronic Self-Views Influence (And Potentially Mislead) Estimates of Performance. *Journal of Personality & Social Psychology*, *84*(1), 5-17. https://doi.org/10.1037/0022-3514.84.1.5

Falck-Ytter, T., Thorup, E., & Bölte, S. (2014). Brief Report: Lack of Processing Bias for the Objects Other People Attend to in 3-Year-Olds with Autism. *Journal of Autism And Developmental Disorders*, *45*(6), 1897-1904. https://doi.org/10.1007/s10803-014-2278-4

Fine, D.R. (2012). A life with prosopagnosia. *Cognitive Neuropsychology, 29*(5-6). 354-359. https://doi.org/10.1080/02643294.2012.736377

Fletcher-Watson, S., Leekam, S., Benson, V., Frank, M., & Findlay, J. (2009). Eye-movements reveal attention to social information in autism spectrum disorder. *Neuropsychologia*, *47*(1), 248-257. https://doi.org/10.1016/j.neuropsychologia.2008.07.016

Frank, M., Vul, E., & Saxe, R. (2011). Measuring the Development of Social Attention Using Free-Viewing. *Infancy*, *17*(4), 355-375. https://doi.org/10.1111/j.1532-7078.2011.00086.x

Freeth, M., Chapman, P., Ropar, D., & Mitchell, P. (2010). Do Gaze Cues in Complex Scenes Capture and Direct the Attention of High Functioning Adolescents with ASD? Evidence from Eye-tracking. *Journal of Autism and Developmental Disorders*, *40*(5), 534-547. https://doi.org/10.1007/s10803-009-0893-2

Fry, R., Wilmer, J., Xie, I., Verfaeille, M., & DeGutis, J. (2020). Evidence for Normal Novel Object Recognition Abilities in Developmental Prosopagnosia. *Royal Society Open Science, 7*(9). https://doi.org/10.1098/rsos.200988

Gainotti, G., Barbier, A., & Marra, C. (2003). Slowly progressive defect in recognition of familiar people in a patient with right anterior temporal atrophy. *Brain*, *126*(4), 792-803. https://doi.org/10.1093/brain/awg092

Gray, K., Bird, G., & Cook, R. (2017). Robust associations between the 20-item prosopagnosia index and the Cambridge Face Memory Test in the general population. *Royal Society Open Science*, *4*(3), 160923. https://doi.org/10.1098/rsos.160923

Grüter, T., Grüter, M., Bell, V. & Carbon, C.C. (2009). Visual mental imagery in congenital prosopagnosia. *Neuroscience Letters*, *453*(3), 135-140. https://doi.org/10.1016/j.neulet.2009.02.021

Haxby, J., Hoffman, E., & Gobbini, M. (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

Jansari, A., Miller, S., Pearce, L., Cobb, S., Sagiv, N., Williams, A. L., Tree, J., & Hanley, J. R. (2015). The man who mistook his neuropsychologist for a popstar: When configural processing fails in acquired prosopagnosia. *Frontiers in Human Neuroscience*, *9*, 390. https://doi.org/10.3389/fnhum.2015.00390

Johnen, A., Schmukle, S., Hüttenbrink, J., Kischka, C., Kennerknecht, I., & Dobel, C. (2014). A family at risk: Congenital prosopagnosia, poor face recognition and visuoperceptual deficits within one family. *Neuropsychologia*, *58*, 52-63. https://doi.org/10.1016/j.neuropsychologia.2014.03.013

Kapur, N. (1987). Some comments on the technical acceptability of Warrington's Recognition Memory Test. *British Journal of Clinical Psychology*, *26*(2), 144-146. https://doi.org/10.1111/j.2044-8260.1987.tb00741.x

Kennerknecht, I., Ho, N.Y., & Wong, V.C. (2008). Prevalence of Hereditary Prosopagnosia (HPA) in Hong Kong Chinese Population. *American Journal of Medical Genetics Part A*, *146A*(22), 2863-2870. https://doi.org/10.1002/ajmg.a.32552

Kracke, I. (2008). Developmental prosopagnosia in Asperger syndrome: presentation and discussion of an individual case. *Developmental Medicine & Child Neurology*, *36*(10), 873-886. https://doi.org/10.1111/j.1469-8749.1994.tb11778.x

Laguesse, R., Tez, T., Hall, B., Irons, J., McKone, E., Daini, R., Albonico, A., Malaspina, M., Taylor, E., Rhodes, G., Charpentier, A., Rossion, B., & Palermo, R. (2013). Subjective self-assessment of face recognition ability is only weakly related to objective measures of face recognition performance. *Journal of Vision*, *13*(9), 979-979. https://doi.org/10.1167/13.9.979

Lai, M., Lombardo, M.V., Pasco, G., Ruigrok, A.N.V., Wheelwright, S.J., Sadek, S.A., Chakrabarti, B., MRC AIMS Consortium., & Baron-Cohen, S. (2011). A Behavioral

Comparison of Male and Female Adults with High Functioning Autism Spectrum Conditions. *PLoS ONE*, *6*(6). https://doi.org/10.1371/journal.pone.0020835

Lee, Y., Duchaine, B., Wilson, H., & 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

Liversedge, S., & Findlay, J. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences*, *4*(1), 6-14. https://doi.org/10.1016/s1364-6613(99)01418-7

Livingston, L., & Shah, P. (2017). People with and without prosopagnosia have insight into their face recognition ability. *Quarterly Journal of Experimental Psychology*, *71*(5), 1260-1262. https://doi.org/10.1080/17470218.2017.1310911

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., Biotti, F., Cook, R., & Gray, K. (2019). The discrimination of facial sex in developmental prosopagnosia. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-55569-x

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

McKone, E., Hall, A., Pidcock, M., Palermo, R., Wilkinson, R., Rivolta, D., Yovel, G., David, 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

Moseley, R.L., Hitchiner, R., & Kirkby, J.A. (2018). Self-Reported Sex Differences in High-Functioning Adults with Autism: A Meta-Analysis. *Molecular Autism*, *9*(1). https://doi.org/10.1186/s13229-018-0216-6

Murray, E., Hills, P.J., Bennetts, R.J., & Bate, S. (2018). Identifying Hallmark Symptoms of Developmental Prosopagnosia for Non-Experts. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-20089-7

Nakayama, Y., Covassin, T., Schatz, P., Nogle, S., & Kovan, J. (2014). Examination of the Test-Retest Reliability of a Computerized Neurocognitive Test Battery. *The American*

Journal of Sports Medicine, *42*(8), 2000-2005. https://doi.org/10.1177/0363546514535901

Neil, L., Cappagli, G., Karaminis, T., Jenkins, R., & Pellicano, E. (2016). Recognizing the same face in different contexts: Testing within-person face recognition in typical development and in autism. *Journal of Experimental Child Psychology*, *143*, 139-153. https://doi.org/10.1016/j.jecp.2015.09.029

Németh, K., Zimmer, M., Schweinberger, S., Vakli, P., & Kovács, G. (2014). The Background of Reduced Face Specificity of N170 in Congenital Prosopagnosia. *Plos ONE*, *9*(7), e101393. https://doi.org/10.1371/journal.pone.0101393

Neumann, D., Spezio, M., Piven, J., & Adolphs, R. (2006). Looking you in the mouth: abnormal gaze in autism resulting from impaired top-down modulation of visual attention. *Social Cognitive and Affective Neuroscience*, *1*(3), 194-202. https://doi.org/10.1093/scan/nsl030

O'Bryant, S., Hilsabeck, R., McCaffrey, R., & Gouvier, W. (2003). The Recognition Memory Test Examination of ethnic differences and norm validity. *Archives of Clinical Neuropsychology*, *18*(2), 135-143. https://doi.org/10.1093/arclin/18.2.135

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

Rabin, L., Barr, W., & Burton, L. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA Division 40 members. *Archives of Clinical Neuropsychology*, *20*(1), 33-65. https://doi.org/10.1016/j.acn.2004.02.005

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*(8), 1457-1506. https://doi.org/10.1080/17470210902816461

Riby, D., Hancock, P., Jones, N., & Hanley, M. (2013). Spontaneous and cued gaze-following in autism and Williams syndrome. *Journal of Neurodevelopmental Disorders*, *5*(1), 13. https://doi.org/10.1186/1866-1955-5-13

Righart, R., & de Gelder, B. (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

Russell, R., Duchaine, B., & Nakayama, K. (2009). Super-recognizers: People with extraordinary face recognition ability. People with extraordinary face recognition ability. *Psychonomic Bulletin & Review, 16*(2), 252-257. https://doi.org/10.3758/PBR.16.2.252

Rutherford, M., & Towns, A. (2008). Scan Path Differences and Similarities During Emotion Perception in those With and Without Autism Spectrum Disorders. *Journal Of Autism And Developmental Disorders*, *38*(7), 1371-1381. https://doi.org/10.1007/s10803-007-0525-7

Sachse, M., Schlitt, S., Hainz, D., Ciaramidaro, A., Walter, H., & Poustka, F., Bolte, S., & Freitag, C.M. (2014). Facial emotion recognition in paranoid schizophrenia and autism spectrum disorder. *Schizophrenia Research*, *159*(2-3), 509-514. https://doi.org/10.1016/j.schres.2014.08.030

Schmalzl, L., Palermo, R., & Coltheart, M. (2008). Cognitive heterogeneity in genetically based prosopagnosia: A family study. *Journal of Neuropsychology, 2*(1), 99-117. https://doi.org/10.1348/174866407x256554

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

Shah, P., Sowden, S., Gaule, A., Catmur, C., & Bird, G. (2015). The 20 item prosopagnosia index (PI20): relationship with the Glasgow face-matching test. *Royal Society Open Science*, *2*(11), 150305. https://doi.org/10.1098/rsos.150305

Shakeshaft, N., & Plomin, R. (2015). Genetic specificity of face recognition. *Proceedings of The National Academy of Sciences*, *112*(41), 12887-12892. https://doi.org/10.1073/pnas.1421881112

Sim, S.Y, Saperia, J., Brown, J.A., & Bernieri, F.J. (2015). Judging Attractiveness: Biases Due to Raters' Own Attractiveness and Intelligence. *Cogent Psychology, 2*(1). https://doi.org/10.1080/23311908.2014.996316

Sumner, E., Leonard, H., & Hill, E. (2018). Comparing Attention to Socially-Relevant Stimuli in Autism Spectrum Disorder and Developmental Coordination Disorder. *Journal of Abnormal Child Psychology*, *46*(8), 1717-1729. https://doi.org/10.1007/s10802-017-0393-3

Susilo, T., & Duchaine, B. (2013). Advances in Developmental Prosopagnosia. *Current Opinion in Neurobiology, 23*(3), 423-429. https://doi.org/10.1016/j.conb.2012.12.011

Sutton, A. Face blindness (prosopagnosia) and me. (2008). http://www.faceblind.me/andy/faceblindandme.html

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., 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

Trevisan, D., Roberts, N., Lin, C., & Birmingham, E. (2017). How do adults and teens with self-declared Autism Spectrum Disorder experience eye contact? A qualitative analysis of first-hand accounts. *PLOS ONE*, *12*(11), e0188446. https://doi.org/10.1371/journal.pone.0188446

Ulrich, P., 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

Van Belle, G., De Graef, P., Verfaillie, K., Busigny, T., & Rossion, B. (2010). Whole not hole: Expert face recognition requires holistic perception. *Neuropsychologia*, *48*(9), 2620-2629. https://doi.org/10.1016/j.neuropsychologia.2010.04.034

Vildalen, V.U., Brevik, E.J., Haavik, J., & Lundervold, A.J. (2016). Females with ADHD Report More Severe Symptoms Than Males on the Adult ADHD Self-Report Scale. *Journal of Attention Disorders*. https://doi.org/10.1177/1087054716659362

Warrington, E. K. (1984).Recognition Memory Test: Manual. NFER-Nelson.

White, D., Rivolta, D.A., Burton, 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

Wilmer, J.B., Germine, L., Chabris, C.F., Chatterjee, G., Williams, M., Loken, E., Nakayama, K., & Duchaine, B. (2010). Human face recognition ability is specific and highly heritable. *Proceedings of The National Academy of Sciences*, *107*(11), 5238-5241. https://doi.org/10.1073/pnas.0913053107

Wilmer, J., Germine, L., & Nakayama, K. (2014). Face recognition: a model specific ability. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00769

Wilson, C., 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. E., Palermo, R., Schmalzl, L., & Brock, J. (2010). Specificity of impaired facial identity recognition in children with suspected developmental prosopagnosia. *Cognitive Neuropsychology, 27*(1), 30–45. https://doi.org/10.1080/02643294.2010.490207

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

Young, A., & De Haan, E. (1988). Boundaries of covert recognition in prosopagnosia. *Cognitive Neuropsychology, 5*(3), 317-336. https://doi.org/10.1080/02643298808252938

Young, A. W., Newcombe, F., de Haan, E. H., Small, M., & Hay, D. C. (1993). Face perception after brain injury. Selective impairments affecting identity and expression. *Brain, 116*, 941–959. https://doi.org/10.1093/brain/116.4.941

Young, A., & Noyes, E. (2019). We need to talk about super-recognizers Invited commentary on: Ramon, M., Bobak, A. K., & White, D. Super-recognizers: From the lab to the world and back again. *British Journal of Psychology*, *110*(3), 492-494. https://doi.org/10.1111/bjop.12395

Zantinge, G., van Rijn, S., Stockmann, L., & Swaab, H. (2017). Physiological Arousal and Emotion Regulation Strategies in Young Children with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *47*(9), 2648-2657. https://doi.org/10.1007/s10803-017-3181-6

Zell, E., & Krizan, Z. (2014). Do People Have Insight into Their Abilities? A Metasynthesis. *Perspectives on Psychological Science*, *9*(2), 111-125. https://doi.org/10.1177/1745691613518075

Zhu, Q., Song, Y., Hu, S., Li, X., Tian, M., Zhen, Z., Dong, Q., Kanwisher, N., & Liu, J. (2010). Heritability of the Specific Cognitive Ability of Face Perception. *Current Biology*, *20*(2), 137-142. https://doi.org/10.1016/j.cub.2009.11.067