Kent Academic Repository

Full text document (pdf)

Citation for published version

Ulrich, Philip I. N. (2017) Behavioural, Electrophysiological and Neurostimulatory Investigations into Developmental Prosopagnosia. Doctor of Philosophy (PhD) thesis, University of Kent,.

DOI

Link to record in KAR

http://kar.kent.ac.uk/61343/

Document Version

UNSPECIFIED

Behavioural, Electrophysiological and Neurostimulatory Investigations into Developmental Prosopagnosia

A thesis submitted for the Degree of Ph.D. in the faculty of Social Sciences at the University of Kent

Philip I. N. Ulrich

School of Psychology

University of Kent

April 2017

ABSTRACT

Developmental prosopagnosia (DP) is the difficulty or inability to recognise a face and may affect up to 2.9 percent of the population. There is controversy over whether these impairments are perceptual or memorial in nature, and uncertainty about their stability over time and how to remediate symptoms. In the first stage, a battery of ten tests was assembled to assess a wide range of face recognition skills in DP (n = 11) and compared to a control group (Chapter Two). The majority of DPs showed no signs of impaired face perception but profound face memory deficits. To seek electrophysiological corroboration of these impairments, the DPs (n = 8) were given three behavioural tasks known to elicit specific event related potentials (Chapter Three), assessing face perception (N170), face familiarity (N250r) and semantic access (N400). During the experiment, caloric vestibular stimulation (CVS) was also administered to see if it could reduce symptoms. The tasks revealed intact face perception and impaired accuracy in both memory based tasks, corroborated by an atypical N400. Subtle effects of CVS were observed in all measures of the face familiarity task but not at a level that was clinically relevant. To establish, for the first time, whether the impairments in DP are consistent over time, the effects in Chapter Three were replicated (n = 7)(Chapter Four). A similar pattern emerged and test-retest correlations showed high reliability overtime in the familiarity task but not the semantic access task. This implies that reliable 'diagnosis' of developmental prosopagnosia should be based on judgements of face familiarity and not associated with semantic activity. The beneficial effects of CVS were again present in the N250r behavioural measures and were limited to familiar faces only. This implies that CVS is optimising memory recall for face representations. The source of impairments was consistently shown to be memorial in nature and future studies may wish to

explore further divisions of memory in DP such as whether impairments are associated with encoding or recall. The thesis also demonstrates the potential for CVS as both a therapeutic tool and cognitive enhancer, and justify more robust trials investigation.

ACKNOWLEDGEMENTS

DISSEMINATION

Publications

Chapter Two has been published.

Ulrich, P. I. N., Wilkinson, D. T., Ferguson, H. J., Smith, L. J., Bindemann, M., Johnston, R.
A. & Schmalzl, L. (2016). Perceptual and Memorial Contributions to Developmental
Prosopagnosia. The Quarterly Journal of Experimental Psychology, 1-18.
doi:10.1080/17470218.2016.1177101

Conference Presentations

Ulrich, P. I. N. (2016). 2015 BPS Cognitive Psychology Section Conference Review.
Cognitive Psychology Bulletin, 1. Retrieved from
http://www.bps.org.uk/publications/member-network-publications/member-
publications/cognitive-psychology-bulletin

Ulrich, P. I. N. (2016). Advice for an Academic Study Visit. PsyPAG Quarterly, 98, 39-42.
Retrieved from http://www.psypag.co.uk/wp-content/uploads/2013/06/PsyPag-98.pdf

Ulrich P., Wilkinson, D. T., & Johnston, R. A. (2015). Improving semantic recall from
familiar faces with vestibular stimulation. Paper presented at the BPS Cognitive
Psychology Section Conference, Kent, UK.

Ulrich P. I. N., Johnston, R. A., & Wilkinson, D. T. (2015). Caloric Vestibular Stimulation

    Modulates High Level Face Processing. Perception, 44(S1), 281.

    doi:10.1177/0301006615598674

Ulrich P., Johnston, R. A., & Wilkinson, D. T. (2015). Caloric Vestibular Stimulation

    Modulates High Level Face Processing. Paper presented at the European Conference

    on Visual Perception, Liverpool, UK. Retrieved from

    http://ecvp.org/2015/Booklet.pdf

Ulrich P., Wilkinson, D. T., & Johnston, R. A. (2015). Superior face-recognition from

    vestibular stimulation. Poster session presented at the Experimental Psychology

    Society Meeting, Leeds, UK. Retrieved from

    http://www.eps.ac.uk/images/epsfiles/2015/leeds_programme_apr15.pdf

Ulrich, P., Wilkinson, D. T., Ferguson, H. J., Bindemann, M., & Johnston, R. A. (2014).

    Heterogeneity in Developmental Prosopagnosia. Paper presented at the 31st BPS

    Cognitive Psychology Section Conference, Nottingham, UK. Retrieved from

    http://www.ntu.ac.uk/webevents/soc/document_uploads/163523.pdf

Ulrich, P., Wilkinson, D. T., Ferguson, H. J., Bindemann, M., & Johnston, R. A. (2014).

    Perceptual and memorial impairment in developmental prosopagnosia. In S. Bate

    (Chair), Recent advances in developmental prosopagnosia research. Symposium

    conducted at the Experimental Psychology Society Meeting, Kent, UK. Retrieved

    from http://www.eps.ac.uk/images/epsfiles/kent%20programme%20apr%20-

    %202014%20-final.pdf

Ulrich, P., Wilkinson, D. T., Ferguson, H. J., Bindemann, M., & Johnston, R. A. (2014). Vestibular Stimulation Therapy for Prosopagnosia. Poster presented at the Experimental Psychology Society Meeting, London, UK. Retrieved from http://www.eps.ac.uk/images/Jan%20-%202014%20-%20FINAL.pdf

Ulrich, P., Wilkinson, D. T., Ferguson, H. J., Bindemann, M., & Johnston, R. A. (2013). Prosopagnosia: Exploration of behaviour, EEG and CVS. Paper presented at the South East England Faces Conference, Kent, UK.

TABLE OF CONTENTS

CHAPTER II

Introduction II: Perceptual and Memorial Contributions to Developmental

CHAPTER V

General Discussion

LIST OF FIGURES

LIST OF TABLES

Table 4.2 154

CHAPTER I

Introduction I

The human face is of significant social and evolutionary importance. From the moment we are born, we are hard wired to attend to faces (Goren, Sarty, & Wu, 1975; Johnson, Dziurawiec, Ellis, & Morton, 1991; Valenza, Simion, Macchi, Cassia, & Umiltà, 1996) and as an adult, our visual attention is automatically drawn to face stimuli within our environment (Drouler & Adil, 2015; Ro, Russell, & Lavie, 2001; Theeuwes & Van der Stigchel, 2006; Vuilleumier, 2002). With minimal demand, important and complex social cues from the face such as threat and negative valence (Mogg & Bradley, 2000; Öhman, 1999) are processed. This has led some to propose that perceiving and responding to faces has adaptive value honed over years of evolution.

There are many reasons why the face may have become such an important stimulus. A static human face signals group affiliation such as gender and race, and individuating properties including age and attractiveness, and, should the face be familiar, semantic associations such as name and personality. Movement of the face is able to provide a continuous stream of additional information such as eye gaze and expression that can portray emotional state. These social cues, as well as perceiving lips and tongue movement, play an important role in communication, such that conversation proceeds differently in their absence, as illustrated during telephone use (Rutter, 1989).

The consequences of not being able to recognise faces, a condition known as prosopagnosia, can be severe. Social interactions and daily functioning are greatly affected and social situations can cause embarrassment for some. This can lead to an avoidance of social situations, which in turn can lead to career damage, feeling withdrawn, lonely or

depressed, and in rare cases to a diagnosis of social anxiety disorder. No satisfactory estimate of the prevalence of prosopagnosia following brain injury is available, but estimates of developmental prosopagnosia suggest between 1.9% and 2.9% (Bowles et al., 2009; Kennerknecht et al.,2006; Kennerknecht, Ho, & Wong, 2008) of the adult population are affected. In today's UK population alone, and including only the developmental form of prosopagnosia, the mid-point estimate would predict over 1.5 million people experience significant difficulties in face recognition.

The present thesis has three goals. Firstly, it will seek to understand the psychological and biological source of face recognition deficits in developmental prosopagnosia. Perceptual impairments, in particular configural processing, are commonly cited as being the source, but existing evidence is inconsistent, based on incomplete or weak test batteries, or derived from case- or very small group-studies. Secondly, the consistency of the face recognition deficit will be assessed. No literature exists that has specifically sought to test whether the impairments measured in participants with prosopagnosia fluctuate over time. This thesis will explore whether conclusions drawn from single session tests are reliable. Lastly, the thesis will aim to gather preliminary evidence for a novel intervention known as CVS. Current therapies are primarily based on laborious behavioural training, often with limited or no success. The possibility that vestibular stimulation may ameliorate face recognition will be examined. Together, my thesis will offer a novel examination of the impairments in developmental prosopagnosia at multiple time points, and will investigate a pioneering form of therapy.

1. Faces as special stimuli

Faces very quickly provide large quantities of information about an individual, including identity, gender, race, age, desirability, mood, well-being, threat, and gaze direction. Within one quarter of a second after seeing a face, we are able to determine whether that person is familiar or a stranger (Paller, Gonsalves, Grabowecky, Bozic, & Yamada, 2000). Within three fifths of a second, we are accessing semantic and episodic information about the person (Gosling & Eimer, 2011) such as name, occupation, nationality, when you last spoke to them, and how they are related to you. No other single stimuli is able to provide this information, thereby making face stimuli not only important, but also special. The first suggestion that faces are processed in a different way to any other object was based on the finding that inverted faces were disproportionately harder to remember than other inverted objects (Yin, 1969). This is understood to be because the face is viewed as a whole Gestalt (Moscovitch, Winocur, & Behrmann, 1997), whereas objects are first broken into their constituent parts (Palmer, 1977). Further evidence for this theory is from the composite effect, whereby one half of a face is harder to recognise when it is aligned (compared to misaligned) with a different face-half (Young, Hellawell, & Hay, 1987). This effect occurs because the two halves are processed as a whole chimeric face, inhibiting attention to only the target face half. The evidence that faces are processed holistically is what differentiates these stimuli from any other.

The counter argument to the holistic processing argument is that faces are only processed differently because we develop expertise due to high exposure, and that the inversion effects can also be seen for other stimuli given enough expertise (Diamond & Carey, 1986; Gauthier & Tarr, 1997). However, other studies have been unable to replicate expertise-like effects in experts of other stimuli (e.g. Robbins & McKone, 2007).

Furthermore, there is a possibility that expertise uses a separate mechanism to that of faces but one which functions in a similar way.

Evidence from functional imaging studies have identified two key cortical areas that are more responsive to faces than other objects; the fusiform face area (FFA; Kanwisher, McDermott, & Chun, 1997; McCarthy, Puce, Gore, & Allison, 1997) on the ventral surface of the temporal lobe, and the occipital face area (OFA; Gauthier, Tarr, Moylan, Skudlarski, Gore, & Anderson, 2000; Puce, Allison, Asgari, Gore, & McCarthy, 1996) on the lateral surface of the occipital lobe (see Figure 1.1). However, there is other evidence that suggests these areas are also active with expertise with non-face stimuli (Gauthier, Skudlarski, Gore, & Anderson, 2000). Arguments and counter arguments for the theory that faces are treated as special by the brain continue and a detailed understanding of face recognition remains remarkably elusive, despite decades of research.
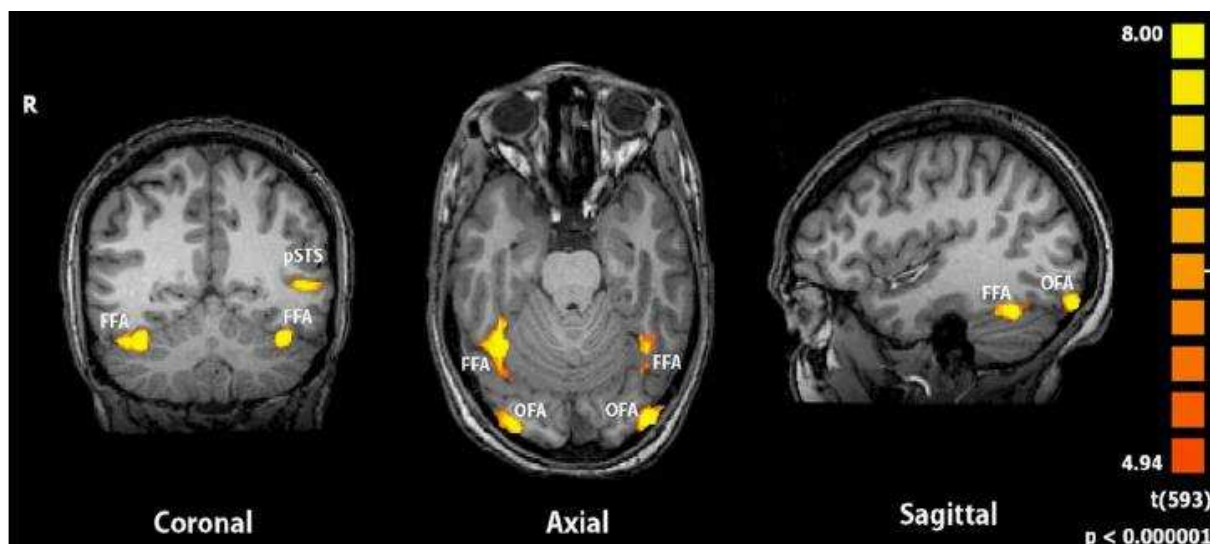
Figure 1.1: Locations of the fusiform face area (FFA), occipital face area (OFA), and posterior superior temporal sulcus (pSTS) face responsive regions (image reproduced from Lai, Pancaroglu, Oruc, Barton, & Davies-Thompson, 2014).

What is clear, however, is that face recognition can be disproportionately impaired, regardless of the underlying neural mechanism. This impairment is called prosopagnosia and is the subject of a rapidly growing research field, and this thesis.

2. Prosopagnosia

Prosopagnosia (a Greek compound word from prosopon and agnosia, meaning face and non-knowledge respectively) was first documented by Wigan (1844) but was first named by Bodamer (1947; see Ellis & Florence, 1990 for a condensed translation) after investigating two patients with impaired face recognition following severe brain injuries. Prosopagnosia is a specific type of visual agnosia, also known more colloquially as face-blindness, and is a clinical condition in which patients have a particular difficulty or inability to recognise familiar people from face information alone, but know that a face is a face. People with prosopagnosia may fail to recognise even the closest friends and relatives, and some even fail to recognise their own image in a mirror. Often patients compensate by using other identity cues such as hair, clothing, voice and gait, and are greatly assisted by the context in which the person is met.

Academic investigation of prosopagnosia not only develops understanding of the condition but also provides valuable insight into the processes involved in unimpaired face recognition systems, thereby informing and shaping functional models of face recognition (e.g. Bruce & Young, 1986; Gobbini & Haxby, 2007; Haxby, Hoffman, & Gibbini, 2000). Investigations can also direct the development of possible remedies and training programmes designed to compensate for the impairment (see Bate & Bennetts, 2014 for a detailed review of prosopagnosia rehabilitation). What has failed to gain much attention, however, is the inconsistent and often contradictory findings from research into prosopagnosia. For example,

while some claim holistic processing is impaired in prosopagnosia (e.g. Avidan, Tanzer, & Behrmann, 2011), others claim it can be intact (e.g. Susilo et al., 2010). One reason for this is that not all cases of prosopagnosia are the same.

Prosopagnosia can be a symptom of a brain lesion (commonly stemming from occipito-temporal damage (De Renzi, Perani, Carlesimo, Silveri, & Fazio, 1994; Gainotti & Marra, 2011) but can also arise in the seeming absence of any overt neurological abnormality or cognitive pathology. The first to identify a case of developmental prosopagnosia was McConachie (1976; retested by De Haan & Campbell, 1991) which led to the separation of prosopagnosia into two groups; face recognition impairments developed after brain tissue damage are referred to as acquired, and those apparently present from birth are developmental.

As will be described, the causes and presentations of prosopagnosia differ between the acquired and developmental types and consequently cannot be investigated as if they were the same. While both merit research, I have chosen to investigate the developmental form because there are comparably few group studies with this condition, there is a complete absence of effective treatments for it, and because of the likely availability of participants.

The term developmental is itself an umbrella term covering multiple aetiologies. Developmental prosopagnosia (DP), which in the past has also been called early onset prosopagnosia (Barton, Cherkasova, Press, Intriligator, & O'Connor, 2003), familial prosopagnosia (Barton et al., 2003) and childhood prosopagnosia (Young & Ellis, 1989) can be seen as encompassing three aetiologies (Grueter et al., 2007): Firstly, DP associated with early acquired brain injury; secondly DP associated with concurrent neurodevelopmental disorders (e.g. autism spectrum disorder; also known as symptomatic amnesia [Grüter, Grüter, & Carbon, 2008]); and thirdly, DP in the absence of brain injury and neurodevelopmental disorders, also known as congenital prosopagnosia (Ariel & Sadeh,

1996). Furthermore, the congenital type also includes a sub-group of hereditary prosopagnosia if a first-grade relative is also affected (Grueter et al., 2007). A lesser known sub-type of prosopagnosia is progressive prosopagnosia (Evans, Heggs, Antoun, & Hodges, 1995) which is the steady deterioration of face recognition skills due to posterior cortical atrophy and usually manifests in old age.

At the beginning of the 21st century, the majority of prosopagnosia research was conducted on the acquired subtype. However, researchers began to realise that the developmental form may be more common than previously suspected. Investigations into the prevalence of prosopagnosia have since revealed surprisingly high estimates. Kennerknecht and colleagues (2006) screened 689 German students, and further interviewed 56 of these. The authors diagnosed 17 of these participants as having prosopagnosia, thereby offering a prevalence estimate of 2.47%. The procedure was repeated with Chinese students and diagnosed 10 out of 533 students with prosopagnosia, a prevalence rate of 1.9%. Although these figures are widely reported in the literature, the figures are based solely on self report measures. This lack of objectivity is problematic, as highlighted by one study that explicitly tested whether self-reported face recognition ability correlates with empirically tested scores and found participants could only accurately estimate their ability to recognise celebrities (Bindemann, Attard, & Johnston, 2014; see also Bowles et al., 2009; Rotshtein, Geng, Driver, & Dolan, 2007).

Only one study has used objective tests in a population to gauge the prevalence of developmental prosopagnosia. Bowles and colleagues (2009) administered the popular and trusted Cambridge Face Memory Test (CFMT; see Figure 1.2; Duchaine & Nakayama, 2006), a task in which participants learn 6 new faces then identify them in a line-up, to 241 Australian participants and estimated a prevalence rate of at least 2.1% (5 participants impaired out of 241, but two others showed impairment in other objective tasks). Impairment

in this study was defined as scores more than two standard deviations from the mean of a control group. This would imply that the probability of being categorised as having prosopagnosia (in a two tailed test) is approximately 2.3%. Aside from the fact that the CFMT is not validated as a diagnostic test, a further problem here is that the prevalence estimate from Bowles et al. is merely an artifact of their chosen cut-off value. For example, does developmental prosopagnosia (unlike its acquired counterpart) constitute a discrete condition, or merely the lower end of a normal continuum of face recognition abilities? Is statistical significance the same as clinical significance?
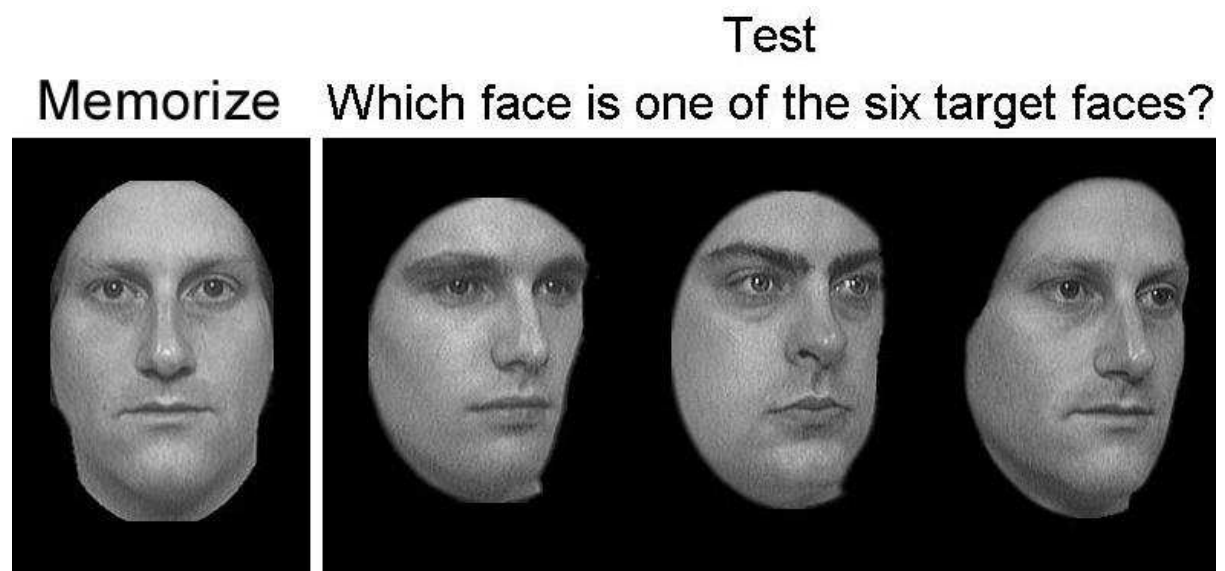
Figure 1.2: Example stimuli from the Cambridge Face Memory Test (CFMT). The target face on the left is presented independently from the test set of three faces.

3. Presentation of Developmental Prosopagnosia

As well as divisions based on aetiology, prosopagnosia is also commonly divided by impairment types. In the late 19th century a distinction was proposed for two types of agnosia (Lissauer, 1890, but for an abridged translation see Shallice & Jackson, 1988). The

apperceptive stage of recognition concerns structural encoding, where the elementary shapes of a stimulus are processed and formed into a coherent percept. The later, associative stage of recognition is the use of associative or semantic links related to the stimulus. De Renzi (1986; De Renzi, Faglioni, Grossi, & Nichelli, 1991) applied this distinction to (acquired) prosopagnosia providing evidence of apperceptive prosopagnosia through impairment in face perception tasks (unknown face matching and age estimation), and of associative prosopagnosia through impairment in mnestic tasks (face familiarity and recognition) with normal perceptual performance.

Neuropsychological studies of acquired prosopagnosia have identified anatomical correlates for these functional subtypes. The apperceptive subtype is associated with damage to occipito-temporal regions, in particular the right FFA (Barton, 2008; Barton, Press, Keenan, & O'Connor, 2002; Damasio, Tranel, & Damasio, 1990). The associative subtype is associated with damage to right or bilateral anterior temporal regions (Barton, 2008; Barton, Zhao, & Keenan; Barton & Cherkasova, 2003; Evans, Heggs, Antoun, & Hodges, 1995). Typically, the severity of the impairment is related to the extent and laterality of the damage. However, the subtype that is presented is not persistently linked with the location of the lesion, with some patients showing subtle apperceptive deficits with anterior temporal lesions, and subtle associative deficits with fusiform lesions (Barton, 2008). This suggests that the two subtypes are not completely dissociable, and that the face recognition system is a network and not based on separate pathways.

The patterns of behavioural impairments are less clear in developmental prosopagnosia. While the categorisation of prosopagnosia is often based on memory based tasks such as the CFMT, most studies report impairments in both perceptual and memory tasks. Impairments in face perception are particularly varied. Most commonly, a deficit is identified in recovering configural or holistic information (Avidan et al., 2011). The

inversion effect (see Valentine, 1988), used as an indirect measure of holistic processing, is often reduced or absent in developmental prosopagnosia (Behrmann et al., 2005; Lee, Duchaine, Wilson & Nakayama, 2009; Nunn, Postma & Pearson, 2001; Schmalzl, Palermo, & Coltheart, 2008) while the composite faces test (Hole, 1994; Le Grand, Mondloch, Maurer & Brent, 2004), used as a direct measure of holistic processing, also reveals impairments in developmental prosopagnosia (Le Grand et al., 2006; Schmalzl et al., 2008; Palermo et al., 2011). These deficits may be related to the lack of global attentional bias, as demonstrated in compound letter tasks (Avidan et al., 2011; Behrmann et al., 2005; Schmalzl et al., 2008).

Tests of configural processing based on sensitivity to second order spatial relations such as spacing between features, face contours and altered features also demonstrate perceptual impairments in developmental prosopagnosia (Duchaine, Yovel, Butterworth & Nakayama, 2006; Le Grand et al., 2006; Schmalzl et al., 2008). These impairments are also present in first order spatial relations such as in detecting (Garrido, Duchaine & Nakayama, 2008) or identifying (DeHaan & Campbell, 1991; Nunn et al., 2001) Mooney faces (1957; 1960).

More general tests of face perception (i.e. which lack sensitivity to individual processes) that include only minimal memory involvement also show poor performance in prosopagnosia. These include the Cambridge Face Perception Test (Dingle, Duchaine, & Nakayama, 2005; Duchaine, Germine, Nakayama, 2007; Avidan et al., 2011; Chatterjee & Nakayama, 2012; Lee et al., 2009) and the face matching paradigm (Ariel & Sadeh, 1996; Behrmann et al., 2005; Humphreys, Avidan, & Behrmann, 2007; Lee et al., 2009; Nunn et al., 2001). Using a similar face matching paradigm, impairments are also demonstrated in changes across viewpoint (Behrmann et al., 2005; Duchaine, 2000; Duchaine et al., 2006; Lee et al., 2009; Schmalzl et al., 2008).

Impairments in developmental prosopagnosia are also shown in tests that require the extraction of non-featural information from a face, such as emotional expression (Ariel & Sadeh, 1996; DeHaan & Campbell, 1991; Duchaine et al., 2006; Garrido et al., 2009; Kracke, 1994; Minnebusch, Suchan, Ramon & Daum, 2007), age (Ariel & Sadeh, 1996; DeHaan & Campbell, 1991; Kracke, 1994), gender (Ariel & Sadeh, 1996, DeHaan & Campbell, 1991, Duchaine et al., 2006), and attractiveness (Duchaine et al., 2006; Le Grand et al., 2006). However, it is possible that these impairments are secondary to sub-optimal holistic or configural processing.

The varied perceptual deficits described here are in line with the proposal for an apperceptive form of developmental prosopagnosia. The difficulty or inability to recognise a famous face can be explained by difficulties in forming a percept of the face during initial encoding. The impaired face perception is assumed to negatively impact face memory, and in support of this there is no evidence of an individual with impaired face perception but intact face memory. De Renzi et al. (1991) describe a patient with higher performing associative function than perceptual function, however, actually scores in both functions were categorised as indicating impairment.

There is considerably less evidence for an associative subtype of developmental prosopagnosia; a search of the literature revealed only four studies. Two studies (Dalrymple, Garrido, & Duchaine, 2014; McKone et al., 2011) report a group of individuals with developmental prosopagnosia with no perceptual impairments, with the most recent study attempting to specifically address the proposed apperceptive-associative dissociation in developmental prosopagnosia. However, the Cambridge Face Perception Test (CFPT; Duchaine, Germine, & Nakayama, 2007), a task in which the participant arranges 6 faces in order of likeness to a target face, was the only measure of face perception. The CFPT assesses a limited number of facets of face perception, as demonstrated by its independence from scores

of other perceptual tests (e.g. Chatterjee & Nakayama, 2012) and is therefore unable to capture all perceptual deficits. Two other studies (Esins, Schultz, Stemper, Kennerknecht, & Bülthoff, 2016; Lee, Duchaine, Wilson, & Nakayama, 2009) identify four participants with prosopagnosia that show no impairment in face perception tasks. While one test battery (Esins et al., 2016) employed only four face perception tasks and may not have captured all perceptual processes, the other (Lee et al., 2009) reported only accuracy data and by neglecting to analyse reaction times, may have missed participants who were accurate but took an abnormally long time to respond. In one additional study (Palermo, Willis, Rivolta, McKone, Wilson, & Calder, 2011), eight of twelve individuals with developmental prosopagnosia showed no significant perceptual impairment, however they did still show weaker (but not statistically significant) holistic processing in comparison to a control group.

Although developmental prosopagnosia impacts explicit recognition, there is evidence that faces may still be recognised covertly. The earliest evidence of this comes from case studies of acquired prosopagnosia (Bauer, 1984; Tranel & Damasio, 1985) whereby faces of celebrities and family members could not be matched with correct names, but skin conductance responses were larger for correct face-name pairs. However, early attempts (Bentin, Deouell, & Soroker, 1999; De Haan and Campbell, 1991) were unable to find evidence for covert recognition in developmental prosopagnosia and the suggestion was made, consistent with an apperceptive deficit, that this was due to face representations having never been formed (Barton, Cherkasova, & Hefter, 2004; Barton, Cherkasova, & O'Connor, 2001), unlike in acquired prosopagnosia. However, in the last ten years there have been several case and small group studies describing individuals with covert recognition in developmental prosopagnosia (Avidan & Behrmann, 2008; Bate, Haslam, Jansari, & Hodgson, 2009; Bate, Haslam, Tree, & Hodgson, 2008; Jones & Tranel, 2010; Rivolta, Palermo, Schmalzl, & Coltheart, 2012; Rivolta, Schmalzl, Coltheart, & Palermo, 2010).

Emphasis should again be made about the heterogeneity of these findings. Measures of covert recognition can be obtained from either direct tasks such as those employing a forced choice of which face is familiar, or indirect tasks such as priming (Barton et al., 2004). Some participants with developmental prosopagnosia will show covert recognition in both tasks, whereas others will show it in one task but not the other. Some participants will not show covert recognition on any task, and there is evidence to suggest that covert recognition is positively correlated with overt recognition (Eimer, Gosling, & Duchaine, 2012; Rivolta et al., 2012), whereby those with the most profound prosopagnosia are the least likely to show covert recognition.

In summary, the above studies enable at least two key initial observations about developmental prosopagnosia. First, there appears to be a large degree of variation in impairment and for each published impairment in a facet of face perception, there are complementary papers with a different sample that show no discernible difficulty. The heterogeneity in magnitude and type of difficulty in developmental prosopagnosia is considerable (Avidan et al., 2011; De Renzi et al., 1991; Stollhoff, Jost, Elze, & Kennerknecht, 2011). Second, there is a distinct lack of strong evidence that impairments in face memory can exist with intact face perception. As I have argued however, these observations are made on the basis of studies that have employed small samples (typically less than five) and employed a narrow range of outcome measures.

4. Mechanisms of impairment: Psychological

Attempts to understand the underlying mechanism of impairment in developmental prosopagnosia often begin by fitting research findings to a relevant theoretical framework, in particular, the classic functional model for face recognition described by Bruce and Young

(1986). This model was not the first to propose a theoretical understanding (or part thereof) of the vast literature on face recognition (Bruce, 1979; Bruce & Voi, 1983; Ellis, 1981; Ellis & Young, 1983; Hay & Young, 1982; Rhodes, 1985) but has become one of the most cited. The model suggests that a face stimulus is processed through a sequential pathway. The first process is that of structural encoding, whereby abstract descriptions of features and their configuration are formed into a mental representation of the face. Successful structural encoding activates, what the authors label, a Face Recognition Unit (FRU) which represents the structural code of an individual face, essentially, a view-independent memory of the face. The higher the resemblance of the stimulus to the FRU, the more active it then later becomes. These functions then have access to so called Person Identity Nodes (PIN) which contain the semantic codes for each individual. These functions are face-independent and can be accessed through additional means such as voice recognition. Lastly, the PIN has access to name generation, a process suggested to be separate from the other semantic information of the individual. The model, in addition to face recognition, posits three other processes which can follow from the initial structural encoding stage; expression recognition, speech analysis and directed visual processing (i.e. processing of specific properties of the visual structure of the face). The centre of the model to which all processes are linked bidirectionally is a black-box type 'cognitive system'. This function contains additional semantic information such as episodic memory relating to the person, and also serves to actively direct attention to a particular face recognition function relevant to the task.

A refinement to the model proposed that it should contain four pools of units. These are FRUs (encoded structural descriptions of the face), PINs (an access point to semantic information about a known individual), a separate pool for semantic information (now separated from PINs), and a name input pool (Burton, Bruce & Johnston, 1990). Each pool can only access another via the PINs pool. The largest refinement to this model, termed an

interactive activation and competition model (IAC; McClelland & Rumelhart, 1981), is that the connections are dynamic and have changeable weights (i.e. strengths of activations). The connections between pools are excitatory, and connections within pools are inhibitory. The inhibition creates competition within pools, which then ensures only one PIN (for example) reaches threshold. The excitation means the signal from one pool activates associated networks for access to semantic information (for example). This model moves forward from a traditional 'box and arrow' design, and is able to model more complex phenomena such as associative and repetition priming.

The model was further developed by Ellis and Young (1990) by adding an affective route. They claimed the original model could only explain overt recognition, yet struggled to explain the covert recognition observed in acquired prosopagnosia. As mentioned, there is evidence that individuals with acquired prosopagnosia show a greater autonomic response to familiar faces, despite not overtly recognising them (Bauer, 1984; 1986). In a similar but opposite pattern, individuals with Capgras' syndrome (Capgras & Reboul-Lachaux, 1923, but see 1994 for a translation) successfully recognise faces but claim they depict identical doubles of the original person and show the same autonomic response to that of unknown faces (Ellis & Lewis, 2001; Ellis, Young, Quayle, & Pauw, 1997; Hirstein & Ramachandran, 1997). Based on this double dissociation a two route process to recognition was suggested, one of overt recognition and one of covert recognition.

The resulting two route IAC model is capable of explaining the mechanism behind general deficits observed in prosopagnosia. The onset of apperceptive and associative subtypes, rests on the integrity of the first process, structural encoding. If this stage is impaired, the structure of the observed face will not be formed sufficiently to move forward and activate a face recognition unit. It is also possible that in the more severe cases of developmental prosopagnosia, no face recognition units exist because no face has been

successfully encoded. Impaired structural encoding, whether this is due to anatomical damage or atypical development, is deemed to lead to apperceptive prosopagnosia (De Renzi et al., 1991). According to the refinement made by Bruce et al. (1990) that posits an additional semantic pool, associative prosopagnosia arises from an impairment at either the FRU or PIN stage because access to semantic information is prevented. However, the feeling of familiarity with the face that would be associated with FRU activation is not typical in prosopagnosia and suggests the impairment immediately follows structural encoding (Young & Burton, 1999). In the less severe cases of prosopagnosia, a measure of familiarity as well as full semantic recollection may be advantageous.

According to the model, covert recognition is achieved following successful structural encoding and FRU activation. The signal is then hypothesised to divide into two routes, one for directing affective responses to the face, and the other to the PIN. In the case of prosopagnosia with covert recognition, only the affective route is intact (Burton, Young, Bruce, Johnston, & Ellis, 1991). Others have suggested that the face recognition route is intact but fails to reach consciousness (de Haan, Bauer, & Greve, 1992), presumably a process contained within the 'cognitive system' black box.

What the model is unable to explain, however, is the more specific impairments observed in prosopagnosia. For example, apperceptive deficits can be related to anomalous feature detection, age perception, or holistic processing, but these are only parts contained within the generic black box of 'structural encoding'. Likewise, adding a second route for affective recognition is not able to delineate direct (forced choice) and indirect (priming) types of covert recognition (Barton et al., 2004). The model also lacks detail in predicting impaired memory processes, such as how sub-optimal but successful structure encoding may impact recognition processes. The most commonly accepted cognitive models are therefore only able to explain the mechanism of more general face recognition impairments, and

currently lack enough specificity to explain finer grained impairments. Current understanding of the more specific processes such as face inversion effects can be explained with theories of holistic processing, but these are yet to be integrated into one more current grand model. Accordingly, while the model can guide our investigation into the source of deficit in developmental prosopagnosia, it lacks specificity in some areas.

5 Mechanisms of impairment: Physiological

i. Spatial

Multiple physiological measures have been employed to understand how faces are recognised, and how failures of this system might underlie the cause of prosopagnosia. One of the most influential models of the physiology of face processing, and one largely compatible with the cognitive model described above, is the model of the distributed human neural system for face perception (Haxby, Hoffman, & Gobbini, 2000). They propose a two system model, with a core system analysing the visual configuration of a face, and an extended system which analyses the meaning of that configuration, incorporating emotion to speech perception.

The regions showing the earliest face responsive activity are the inferior occipital gyri (OFA) which the authors suggest are responsible for early visual processing of face features. This then feeds into two regions not differentiated in the cognitive model (Bruce & Young, 1986; Burton et al., 1990; Ellis & Young, 1990). Here, the lateral fusiform gyrus (i.e. the FFA) processes the invariant aspects of faces, that is, features that are generally unchanging over time. It is during this stage that individual identity is represented, thereby evaluating face familiarity but before access to semantic information. The inferior occipital gyri also feed through to the superior temporal sulcus where changeable aspects of faces are processed

such as eye and lip movements. This completes the core system, after which information is processed in the extended system. Signals relating to changeable aspects of faces can progress to other regions for processing directed attention (intraparietal sulcus), speech perception (auditory cortex), and emotion (amygdala, insula and the limbic system). Signals relating to invariant aspects of faces progress to anterior temporal regions where semantic information is accessed, including personal identity and name recollection. In a modification of the model (Gobbini & Haxby, 2007), the core system is largely unchanged but the semantic information of the extended system is separated into four nearby regions of anterior paracingulate (attitudes), posterior superior temporal sulcus (intentions), anterior temporal cortex (biographical information) and posterior cingulate (episodic memories).

Consistent with the model of Haxby et al. (2000), perceptual deficits in acquired prosopagnosia are typically associated with occipito-temporal lesions and associate deficits (with intact face perception) are linked with more anterior lesions (Barton, 2008; Barton & Cherkasova, 2003; Barton et al., 2002; Damasio et al., 1990; Davies-Thompson et al., 2014). However, there is far less evidence for anatomical differences in developmental prosopagnosia.

Investigations from participants with developmental prosopagnosia regarding abnormal activity in regions associated with the core system provide heterogeneous results, leading some to argue that there is no strong evidence that activity is reliably different from those with typical face recognition in the FFA, OFA or superior temporal sulcus (Avidan & Behrmann, 2014). While some studies show largely spared activity in core system regions (Avidan & Behrmann, 2009; Avidan, Hasson, Malach, & Behrmann, 2005), others show zero activity (Bentin, Degutis, D'Esposito, & Robertson, 2007; Hadjikhani & De Gelder, 2002), and others show more subtle reductions in activity (Dinkelacker et al., 2011; Furl, Garrido, Dolan, Driver, & Duchaine, 2011). An additional group study found abnormal activity in

only some participants and only in some conditions (while activity was enhanced in other conditions; Minnebusch, Suchan, Koster, & Daum, 2009), suggesting that one possible reasons for the mixed findings may be due to individual differences.

Evidence from participants with developmental prosopagnosia showing abnormality in the extended system, specifically the anterior temporal sulcus, is also scarce. However, some studies have found either reduced (Avidan & Behrmann, 2009; Furl et al., 2011) or absent activity in this area (Avidan et al., 2013) and reduced connectivity with the FFA (von Kriegstein, Kleinschmidt & Giraud, 2006). Heterogeneity still exists in these studies regarding the laterality of the effects.

In sum, the functional model is able to explain the deficits observed in acquired prosopagnosia but its explanatory power in developmental prosopagnosia is weak. As with psychological models, functional anatomical models of developmental prosopagnosia remain undeveloped.

ii. Temporal

Electroencephalography (EEG) provides excellent temporal sensitivity (typically rates of 500-2000Hz are reported) to the processes involved in recognising a face. More specifically, it is the investigation into event related potentials (ERPs) following face stimuli that have received the most attention in the literature. While some researchers choose an explorative approach, such as exploring changes in ERP waveforms for each 50ms window following stimulus onset, several face ERPs have become established and seem to offer the most hope in helping understand how the processes involved in developmental prosopagnosia differ from unimpaired face recognition.

The first ERP that emerges following the onset of face stimuli is the P1, a positive deflecting peak occurring 80-120 ms after stimulus onset (for an example see Figure 1.3) recorded over the occipital lobes. However, while some reports have shown the P1 to be larger for faces compared to other objects (Halgren et al., 1999; Itier & Taylor, 2004a; 2004b; 2004c; 2004d; Liu, Harris, & Kanwisher, 2002), the difference may be because of low-level visual properties. As Rossion and Caharel (2011) demonstrate, the effect is seen even when the images are phase scrambled because it is not properties from the face itself that are affecting the P1.
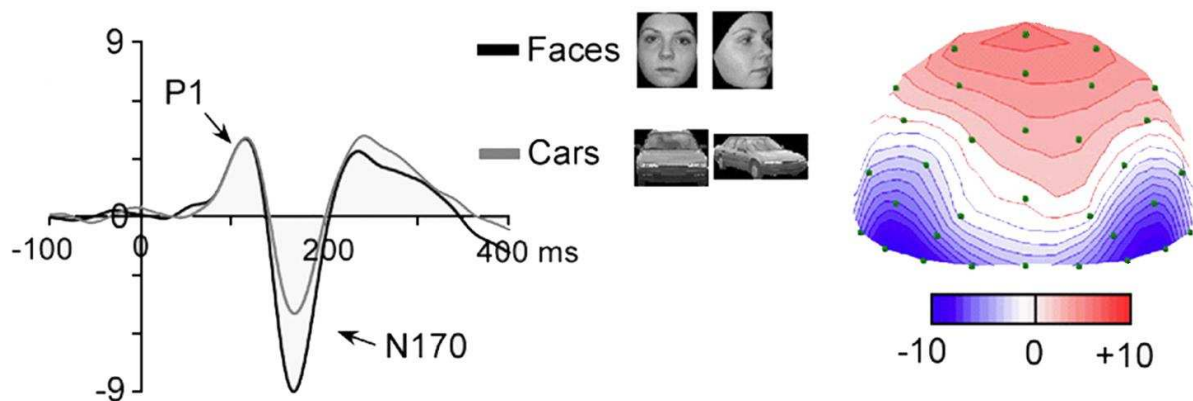
Figure 1.3: An example waveform (left) showing a representation of the P1 and face sensitive N170 event related potentials, with a topographical representation (right) of the activity at the exact moment of the N170 peak (image adapted from Rossion & Jacques, 2008).

The first face sensitive ERP, and one which has received particular attention, is the N170. Discovery of the component is often cited to be Bentin and colleagues (Bentin, Allison, Puce, Perez, & McCarthy, 1996) but the discovery was made in other papers at a similar time (Bötzel, Schulze & Stodieck, 1995; George, Evans, Fiori, Davidoff, & Renault, 1996). Twenty years of research on the N170 have produced a substantial understanding of

the N170 component (for a review see Eimer, 2011). The N170 is a negative deflecting peak occurring approximately 130-200 ms after stimulus onset recorded over occipitotemporal regions (for an example see Figure 1.3). This component can also be detected with magnetoencephalography (MEG) and is termed the M170 (Halgren, Raij, Marinkovic, Jousmäki, & Hari, 2000; Harris & Nakayam, 2008). The N170 is a type of N1 component and is therefore present after seeing most visual stimuli. However, the N170 is significantly larger after seeing face stimuli. The component is affected by low-level image characteristics such as size, contrast and luminance, and some suggest the N170 is merely a consequence of this (Thierry, Martin, Downing, & Pegna, 2007), though this evidence is strongly disputed based on methodological flaws (discussed in Rossion & Jacques, 2008). On the opposite side of the scalp to the N170 is the vertex positive potential (VPP) which is now considered to reflect the same underling processes (Itier and Taylor, 2002; Jemel et al., 2003; Joyce & Rossion, 2005).

The N170 shows large inversion effects, with inverted face stimuli producing a larger (Bentin et al., 1996 ;Eimer, 2000a; Rossion et al., 2000; Sagiv and Bentin, 2001) and later (Bentin et al., 1996 ; Eimer, 2000a; Itier, Latinus, & Taylor, 2006; Itier, Alain, Sedore, & McIntosh, 2007; Rossion et al., 2000; Sagiv and Bentin, 2001) peak compared to upright faces. The inversion effect in behavioural responses (slower and less accurate with inverted faces) is hypothesised to be due to holistic processing, possibly because of expertise with upright faces. These findings therefore indicate that the N170 is associated with holistic or configural processing, a perceptual stage in face recognition. In further support that the N170 reflects perceptual processing, one study shows the N170 to be affected by face viewpoint (Caharel, d'Arripe, Ramon, Jacques, & Rossion, 2009) which requires changes in perceptual processing but not necessarily memorial processing, and other studies have found the N170 to be unaffected by familiarity with the faces (Bentin and Deouell, 2000; Eimer 2000c).

The absence or presence of an N170 recorded from participants with developmental prosopagnosia could suggest perceptual (apperceptive) or post-perceptual (associative) impairments, respectively. A small selection of developmental prosopagnosia case studies have shown reduced sensitivity to faces in the N170 (Bentin et al., 1999; Bentin et al., 2007; Kress & Daum, 2003), while others show no difference compared to controls (Eimer et al., 2012; Harris, Duchaine, & Nakayama, 2005; Minnebusch, Suchan, Ramon, & Daum, 2007; Towler, Gosling, Duchaine, & Eimer, 2012). For example, Righart and De Gelder (2007) found a reduced N170 in 50 percent of their sample, and Minnebusch et al. (2007) found this in 75% of their sample. Interestingly, one larger group study (n = 16) found no group differences for the N170 to upright faces, but discovered that nearly the entire group failed to show the N170 inversion effect (Towler et al., 2012). This finding appears to show some degree of homogeneity within developmental prosopagnosia, though the effect has yet to be replicated and their criteria for the prosopagnosia category was poorly defined.

In sum, evidence for N170 differences in developmental prosopagnosia compared to controls is weak. The evidence is largely based on case studies or inconsistent effects in group studies, that if anything show an impaired inversion effect indicative of atypical holistic processing. Consequently, it is difficult to draw conclusions about N170 differences but the findings seem to suggest heterogeneity in impaired N170 for upright faces and possibly impaired N170 for inverted faces.

Following the onset of the N170, the next prominent and relevant component is the N250r (Begleiter, Porjesz, & Wang, 1995; Schweinberger, Pfütze, & Sommer, 1995; Schweinberger, Pickering, Jentzsch, Burton, & Kaufmann, 2002) although originally called the early repetition effect (ERE). This component is a negative deflection with a peak occurring between approximately 230-330 ms after stimulus onset (for an example see figure

1.4) and is typically recorded over ventral temporal regions. The component represents a

negative deflection to the target face when it has been preceded by a face of the same identity

compared to a different identity. The component has an MEG monologue, the M250r

(Schweinberger, Kaufmann, Moratti, Keil, & Burton, 2007). Of particular interest, the N250r

is present for repetition of individual identity, not just repetition of image (Bindemann,

Burton, Leuthold, & Schweinberger, 2008), though some have observed a larger response to

identical image repetition suggesting residual perceptual processing (Schweinberger et al.,

2002). The N250r is elicited but considerably smaller for immediate repetition of two

unfamiliar compared to familiar faces (Begleiter et al., 1995; Pfütze, Sommer, &

Schweinberger, 2002; Schweinberger et al., 1995). The component is hypothesised to be

linked to re-activation of a face representation in working memory (Schweinberger & Burton,

2003), that is, when a memory of a face stored in working memory is reactivated.
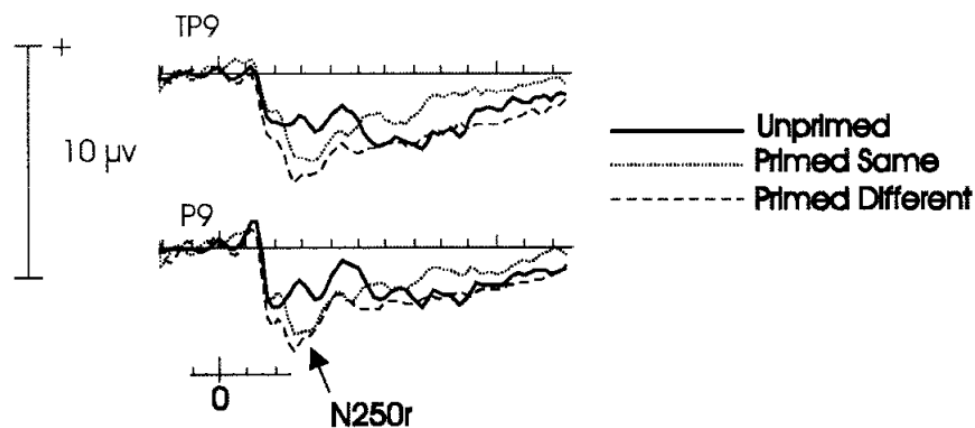
Figure 1.4: Data from Pickering and Schweinberger (2003) used as an example of the N250r
event related potential.

No published study has explored the N250r in prosopagnosia. Instead, the allied

component N250 is employed, which is assumed to represent a match between the percept of

the face stimulus and the stored visual memory of the face. Note a key difference here, in

that the N250r reflects access to the face representation from working memory and the N250 reflects access to the face representation from long-term visual memory. Caution should be taken in applying previous results from one ERP to the other as they may not be reflecting the same underlying processes. However, an insight into what the N250 studies have revealed so far may still be valuable. Studies, all from Martin Eimer's lab, have shown that the N250 for recognised faces in developmental prosopagnosia is no different to those from unimpaired individuals (Towler & Eimer, 2012) but is later for recently learned faces (Parketny, Towler, & Eimer, 2015) and may even be present for explicitly unknown faces, suggesting covert recognition (Eimer, Gosling, & Duchaine, 2012). These findings may help guide predictions for N250r effects in future studies.

The next well published ERP following the N250r, is the N400. This component is a negative deflection with a peak occurring between approximately 300-500 ms after stimulus onset (for an example see figure 1.5) and is typically recorded over central-parietal regions. Initially demonstrated following semantic incongruency during sentence reading (Kutas & Hillyard, 1980), the component was suggested to represent semantic processing. An N400 effect is also observed when a word is preceded by another semantically associated word (e.g. bread and butter) compared to an unrelated words pair (Holcomb, 1993; Rugg, 1987), and also with word-picture pairs (Connolly, Byrne, & Dyman, 1995). For face stimuli, the N400 for semantically related pairs is smaller than for unrelated pairs (Barrett, Rugg, & Perrett, 1988; Barrett and Rugg, 1989; Olivares, Saavedra, Trujillo-Barreto, & Iglesias, 2013; Schweinberger et al., 1995; Bobes, Valdes-Sosa, & Olivares, 1994). The N400 effect for face pairs represents access to semantic activity. This therefore requires prior recognition of the individuals but may not reflect face recognition in itself. Indeed, the effect is still present when names instead of faces are used (Schweinberger, 1996).
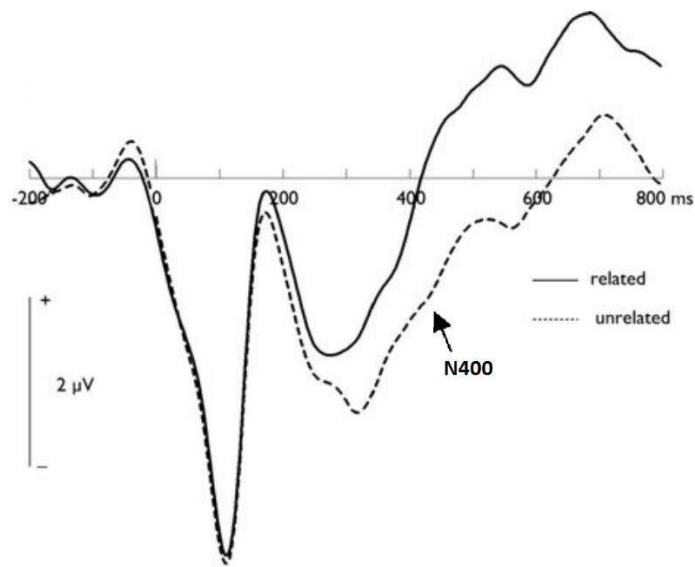
Figure 1.5: Image adapted from Wiese (2011) as an example of the N400 event related potential.

No published study has explored the N400 in developmental prosopagnosia. One paper (Eimer, 2000b) has shown the N400 to be entirely absent in acquired prosopagnosia, and another testing developmental prosopagnosia has found different topography and longer latency for an ERP within a similar catchment window (300-500 ms) as the N400 (Burns, Tree, & Weidermann, 2014). Despite the N400 being a widely researched component (Kutas & Federmeier, 2011), there is surprisingly little research on its presentation in prosopagnosia.

The N400 is sometimes referred to as the early aspect of a late positive component (LPC). Similarly, the late aspect of the LPC is sometimes referred to when observing the P600. This component is a positive deflection with a peak occurring between approximately 400-700 ms after stimulus onset and is broadly distributed over central electrode regions. The P600 is restricted to familiar faces only (Gosling & Eimer, 2011) but different speculations have been made regarding what the P600 reflects, such as name recollection (Díaz, Lindín, Galdo-Alvarez, Facal, & Juncos-Rabadán, 2007; Gosling & Eimer, 2011) or episodic memory (Gosling & Eimer, 2011; Olichney et al., 2008), but it is also suggested to

reflect similar underlying mechanisms to the N400 (Eimer, 2000b). Evidence of the P600 in developmental prosopagnosia, shows that the component is later and reduced compared to controls (Parketny et al., 2015) and is only present for overt recognition (Eimer et al., 2012). Not only is there minimal research for the P600 in developmental prosopagnosia, there is also a lack of understanding of what process the component reflects, other than being somehow related to memory recall.

6. Consistency of Impairment

While the literature on prosopagnosia grows, very few re-administer their tests overtime. Consequently, there is limited, if any, knowledge regarding the permanency or variability of the condition. Impairments observed in one testing session may no longer be present in a second testing session. Test-retest reliability is high for other developmental disorders such as dyslexia (Berninger, Nielson, Abbott, Wijsman, & Raskind, 2008; Cotton, Crewther, & Crewther, 2005; Lefly & Pennington, 2000) and autism (Matson, Gonzalez, & Rivet, 2008; Silverman, Saavedra, & Pina, 2001), but is unreported for many others including dyscalculia, amusia, color agnosia, and attention deficit disorder. For developmental prosopagnosia, there are no published estimates of test-retest reliability.

Existing measures of reliability are related to the tasks used to assess developmental prosopagnosia. Calculating the internal reliability of tests for faces is critical to understanding whether they are consistent measures of the construct they are claiming to measure. This is typically measured using Cronbach's alpha (an average of all possible split-half correlations). Given the frequent use of the CFMT to 'diagnose' developmental prosopagnosia, a number of tests have proceeded to show high internal reliability in various studies (Bowles et al., 2009; Herzmann, Danthiir, Schacht, Sommer, & Wilhelm, 2008;

Wilmer et al., 2010). Surprisingly, only recently has one paper reported the internal

reliability of this test in a group of participants with developmental prosopagnosia (Esins et

al., 2016), and they found the Cronbach's alpha was significantly lower than that of the

control group (0.26 compared to 0.82). Of all 14 experimental conditions contained within

the 7 tests, over a quarter showed significantly lower internal reliability in the prosopagnosia

group compared to the control group. The authors suggested this may be due to the

inconsistent use of compensatory strategies or inconsistent internal processing of any one

chosen strategy.

The report of low internal consistency in developmental prosopagnosia beckons the

question of external consistency, that is, the extent to which scores are consistent over time.

External consistency is typically not reported because measures of internal reliability, at least

for measures of memory, are preferred (e.g. Dennett et al., 2012; McKone et al., 2011) over

test-retest scores because of practise effects affecting performance both between testing

sessions and for some participants differently to others (Wilmer et al., 2010). However, test-

retest scores are still documented for face tests (e.g. Logan, Wilkinson, Wilson, Gordon, &

Loffler, 2016; Thomas, Lawler, Olson, & Agguirre, 2008) but again do not separate the

scores for a developmental prosopagnosia group from control participants. A literature

search fails to find any evidence of test-retest reliability in developmental prosopagnosia, and

only one report in acquired prosopagnosia (De Haan, Young, & Newcombe, 1991).

The importance of test-retest reliability should not be underestimated. The variability

and stability of prosopagnosia, both long-term and short-term, is unknown. This means that

the participant's score in a test may only represent their impairment at that specific time, and

testing a day, or even an hour, later could produce an entirely different set of outcomes.

Indeed this could be the source of the wide spread heterogeneity in prosopagnosia that is

currently unexplained. In addition, current approaches to 'diagnosis' do not account for

changes over time and are therefore invalid without confirmation of consistency of impairment. Lastly, the assessment of attempts to ameliorate developmental prosopagnosia that use pre- and post-therapy differences may be confounded with inconsistent impairment.

7. Rehabilitation Approaches

Current therapies for prosopagnosia can be divided into three different types. Firstly, behavioural compensation methods that do not attempt to alter the face recognition processes in the individual, but circumvent them by teaching alternative behavioural techniques that utilise intact processes, such as feature comparisons. Secondly, behavioural remedial methods that attempt to improve the impaired face recognition processes of the individual to encourage face processing methods akin to those in typically developed face recognition. Thirdly, neurostimulation or pharmacological interventions that do not require (though can be combined with) behavioural training and seek to restore lost function via synaptic plastic change and reactivation of damaged or neighbouring circuitry.

Reported attempts to improve face recognition abilities in developmental prosopagnosia are restricted to just six studies, four of which are case studies. The first of these involved teaching the child to identify key facial features and encouraging him to focus on these for recognition (Brunsdon, Coltheart, Nickels & Joy, 2006). This strategy significantly improved performance on the training faces, but despite subjective reports of improvements in real life, objective measures showed the improvements did not generalise to novel faces. A similar strategy, also taught over one month, met with similar success that persisted at the four week retest, but again the improvements did not generalise (Schmalzl, Palermo, Green, Brunsdon & Coltheart, 2008). Remedial therapy (i.e. ones which encourage typical face processing and do not try to circumvent them with alternative recognition

strategies) has met with more success, with two of the three studies showing generalised improvements while the fourth, despite 10 months practising with one face, showed no significant changes (Dalrymple, Corrow, Yonas, & Duchaine, 2012). The two successful programmes attempted to increase sensitivity to Euclidean distances between features (DeGutis, 2007; 2014). The earlier of these studies trained one individual for approximately 14 months, with the result that improvements in face perception and recognition generalised and were even accompanied with an increased N170 response. When the technique was repeated in the later group study, three weeks of training provided significant benefits. However, even after 14 months of training, the benefits were not maintained unless training was continued, but heavy demands on the participant is likely to reduce compliance. Lastly, Bate and colleagues (2014) reported generalised improvements following inhalation of the hormone oxytocin. The mechanism of effect is not clear but is suggested to be a consequence of increased blood flow to the FFA and amygdala during oxytocin conditions. Unfortunately, the report did not examine whether the benefits were maintained.

By contrast, twelve studies have explored the therapy possibilities in acquired prosopagnosia (for a detailed discussion of these see Bate & Bennetts, 2014; DeGutis, Chiu, Grosso, & Cohan, 2014). Many of these studies have met with a certain degree of success, but only one has shown generalisable benefits. This study used a method called vestibular stimulation, and after just four hours of treatment the patient showed significantly improved face matching ability (Wilkinson, Ko, Kilduff, McGlinchey, & Milberg, 2005).

8. Vestibular stimulation

The vestibular organs, made up of three semicircular canals and two otolith organs (utricle and saccule) positioned in the inner ear, provide information about gravity, acceleration

and velocity of the head. The electrical information generated from the organ's hair cells travels down the vestibulocochlear nerve to the brain stem, then primarily to the thalamus before dispersing across dense anatomical projections to areas mostly concerned with spatial perception and memory (Fitzpatrick & Day, 2004). The vestibular system has attracted increasing interest as a therapeutic pathway because it can be easily manipulated using, for example, thermal or galvanic current.

The vestibular receptors form connections with all levels of the central nervous system (for reviews, see Khan & Chang, 2013, and Tascioglu, 2005). Signals from the five vestibular organs in each ear are innervated by the vestibular ganglion, which form the vestibular nerve. Together with the cochlear nerve, these extend through the internal auditory meatus and across the cerebellopontine angle to enter the pons and terminate at the vestibular nuclei complex. Projections from the vestibular nuclei extend to the cerebellum as well as the thalamus and then the hippocampus via the parietal cortex. In primates, the parietoinsular vestibular cortex (PIVC), area 2v of the intraparietal sulcus, and area 3av in the central sulcus (Grüsser, Pause, & Schreiter, 1990) are widely considered the main cortical areas for receiving and integrating vestibular input. However, the human homologue of the PIVC is yet to be fully understood, with data suggesting a lateral cortical temporoparietal area (Kahane, Hoffmann, Minotti, & Berthoz, 2003), the right hemispheric parietal opercular area (Zu Eulenburg, Caspers, Roski, & Eickhoff, 2012) or the retroinsular cortex and parietal operculum (Lopez, Blanke, & Mast, 2012).

Relevant to face recognition, Bense and colleagues (Bense, Stephan, Yousrey, Brandt & Dieterich, 2001) report decreases in blood flow (therefore reduced activation is assumed) to Brodmann Areas (BA) 18 and 19 (but sparing BA 17, the striate cortex) which form part of the extra-striate cortex which includes face processing regions, during galvanic vestibular stimulation (GVS). In addition, Lobel and colleagues (Lobel, Kleine, Bihan, Leroy-Willig &

Berthoz, 1998) report bilateral increases in blood flow of the central sulcus, but also of the temporal-parietal junction and intraparietal sulcus. Furthermore, GVS has been shown to also increase activation in the hippocampal area (Vitte, Derosier, Caitu, Berthoz, Hasboun, & Soulié, 1996). These areas of activation, particularly the hippocampus and other temporal regions, are strongly associated with explicit long-term visual memory and memory retrieval (e.g. Ishai, Haxby & Ungerleider, 2002; Smith, Geddes, Baek, Darlington & Zheng, 2010).

There are a variety of methods to induce vestibular stimulation, including physically rotating the participant, electrical current through the skin, auditory vibrations, and changing the ear canal temperature by means of cold water irrigation, cold air, infrared, and electronic Peltier units. Galvanic and caloric by water irrigation are the most commonly adopted methods. Galvanic vestibular stimulation is the transcutaneous administration of small electrical currents to the nerve through electrodes placed behind the ears on the mastoid processes (Coats, 1972). This current modulates the firing rate of the vestibular neuroepithelial hairs which are designed to detect changes in motion within each of the vestibular organs (Fitzpatrick & Day, 2004). Specificity of what vestibular activity the stimulation is modulating is not clear. During GVS, excitation is recorded in a wide range of neurons, including those associated with the semicircular canals and also the otilith organs (Peterson, Fukushima, Hirai, Schor, & Wilson, 1980; Wilson, Peterson, Fukushima, Hirai, & Uchino, 1979). However, while sensations associated with otilith system activation (e.g. rocking) are experienced (Bent, Bolton, & Macefield, 2006), this is not true of sensations regarding semicircular canal activation. One suggestion (Cohen, Yakushin, & Holstein, 2012) is that activity in areas associated with semicircular canal activity habituate quickly to the GVS (Courjon, Precht, & Sirkin, 1987) whereas those for otilith activity keep firing.

Alternatively, caloric vestibular stimulation (CVS, see Figure 1.6) traditionally involves the instillation of cold or warm water into the ear canal using a syringe and thin tubing

(Miller & Ngo, 2007). The temperature change causes convection currents in the fluid (endolymph) of the horizontal semicircular canal (Aw, Haslwanter, Fetter, & Dichgans, 2000) which affect the signals transmitted by the neuroepithelium, leading to relative increased activity in the contralateral hemisphere (Bächtold, Baumann, Sándor, Kritos, Regard & Brugger, 2001).
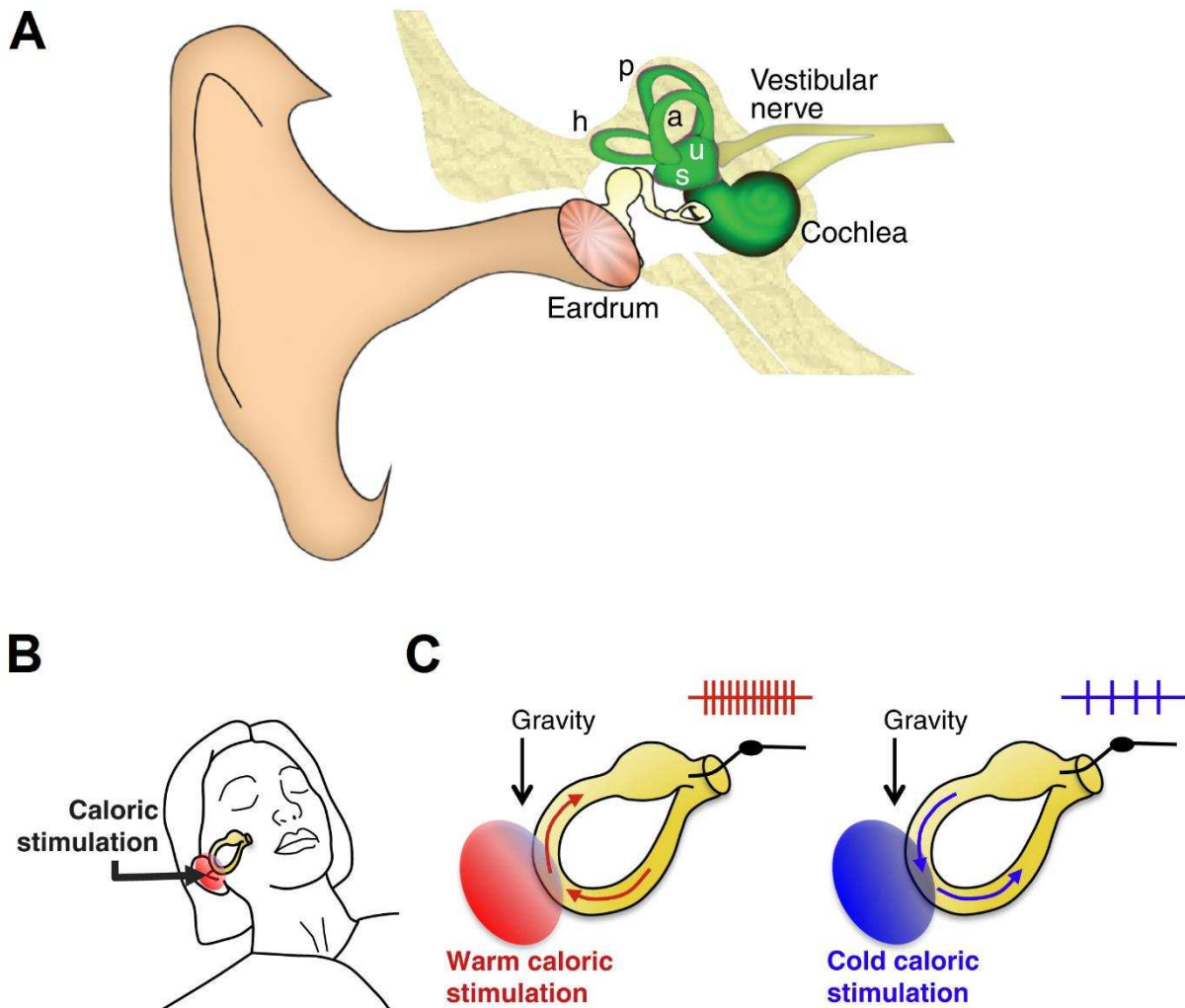
Figure 1.6: (A) Anatomy of the inner ear located in the temporal bone, showing the utricle (u) and saccule (s) otilith organs which encode horizontal and vertical accelerations, repectively, and the horizontal (h), anterior (a), and posterior (p) semicircular canals encoding angular accelerations. (B) Example administration of caloric vestibular stimulation with the head angled at 30° from horizontal. (C) Induced convection currents in the horizontal semicircular canal and subsequent modulation of firing rates in the vestibular afferents. Image adapted from Lopez and Blanke (2014).

These two methods of stimulation modulate the signals from the vestibular organs in different ways and consequently have similar but not identical activation patterns. Lopez, Blanke, and Mast (2012) conducted a meta-analysis of neuro-imaging studies identifying the

areas of activation from CVS compared to GVS. Areas of overlap, which they considered to form the vestibular cortex, were in seven main regions, included the temporoparieto-insular and retroinsular cortex, parietal cortex, frontal cortex, cingulate cortex, thalamus, basal ganglia, and cerebellum. Considering the physical size of the vestibular organs, their stimulation, regardless of method, causes massively widespread activation across both cortical and subcortical structures. Interestingly, despite only modulating one semicircular canal, CVS was shown to induce more wide-spread BOLD activation than GVS which activates regions associated with all five structures.

In comparison to CVS and GVS, alternative neurostimulatory or neuromodulatory techniques, such as transcranial magnetic stimulation (TMS) or transcranial direct current stimulation (tDCS) induce relatively focal areas of stimulation and, without more invasive techniques, are currently unable to reach subcortical areas. Furthermore, the equipment needed to administer CVS and GVS is inexpensive, portable, and requires minimal training to operate independently.

9. Cognitive effects of CVS and GVS

The first use of CVS as a medical tool occurred more than a century ago (Bárány, 1906) and is commonly used now to diagnose balance disorders and brainstem function. The effects of CVS on mood and cognition are surprisingly widespread, including alterations in the mental imagery of bodies and objects (Mast Merfeld, & Kosslyn, 2006), modulation of tactile acuity (Ferre, Sedda, Gandola, & Bottini, 2011), induction of depersonalisation (Sang, Jauregui-Renaud, Green, Bronstein, & Gresty, 2006), and improved treatment-resistant mania (Dodson, 2004), post-stroke central pain (McGeoch & Ramachandran, 2008), spatial neglect, hemianesthesia, anosognosia and somatoparaphrenia (Bottini et al., 2005). With particular

relevance to this thesis, CVS has been shown to improve memory. In one study (Bächtold et al., 2001), object locations and word meanings were learned during CVS and then tested 15 minutes later. Recall was faster in both tasks when performed during stimulation compared to a sham. It is pertinent to note that for both tasks, the participant needed to associate a memory with an image on the screen, much like the process involved in recognising face stimuli.

Three studies have investigated the use of GVS with face recognition. In one study (Wilkinson, Nicholls, Pattenden, Kilduff & Milberg, 2008), participants were asked to memorise 24 faces and their names, and were tested 10 minutes later with questions about the memorised faces. Correct responses from those who received GVS were made 0.5 seconds earlier than those from people in the control condition (with no loss in overall accuracy). As mentioned, GVS has also been shown to have beneficial effects on a patient with acquired prosopagnosia (Wilkinson, Ko, Kilduff, McGlinchey & Milberg, 2005). Patient R.C., a 61 year old male, suffered a stroke 22 years prior to study enrolment. The lesion affected the entire right temporal lobe, inferior frontal gyrus, right basal ganglia, right thalamus, and right superior parietal lobe. In the task, R.C. was instructed to match a target face presented simultaneously with one of two sample faces. Pre-stimulation, R.C. performed at chance (50%), but after two consecutive periods of GVS performance reliably increased to 70%. Unfortunately no reaction time data was provided in the article and conclusions must therefore be drawn with care. Wilkinson and colleagues (Wilkinson, Ferguson & Worley, 2012) also investigated the effect of GVS on the N170. The experimental task was simply to distinguish between famous and non-famous faces, and therefore, perhaps unsurprisingly given the simplicity of the task, there was no effect on behavioural performance. However, relative to the no stimulation condition, stimulation significantly increased the amplitude of the N170 in the group of 16 healthy

students which may be relevant in developmental prosopagnosia where it is sometimes reduced (Bentin et al., 1999; Bentin et al., 2007; Kress & Daum, 2003).

As introduced above, the effects of CVS have not been tested with face recognition, only GVS has been used. There is reason to suggest that larger effects may be observed with CVS because of its widespread cortical activation and larger evidence base for clinical effect. The use of water irrigation is messy, uncomfortable, and often induces side effects of nausea and vertigo. However, a device has been produced that administers vestibular stimulation by means of electronically controlled temperature changes in the ear canal that obviates these symptoms. This opens up new therapeutic opportunities that will be explored in the current thesis.

10. Opening study questions

The inability to recognise faces can have profound consequences on an individual's daily living, from slight embarrassment to social anxiety disorder. In acquired prosopagnosia, two subtypes emerge; apperceptive and associate, a distinction that has been modelled at cognitive and physiological levels. However, the distinction is less clear in DP. Only a small number of studies have administered a wide battery of tests to participants with developmental prosopagnosia, and all have methodological problems such as small sample size (typically n < 5) and not reporting reaction times or allied physiological data. It is therefore unclear how much the condition is perceptual and how much it is memorial. For example, someone with developmental prosopagnosia may show subtle perceptual deficits that, without administration of memorial tests that would have unveiled a more dramatic associative deficit, are wrongly attributed as the main source of the problem. This thesis will employ a more comprehensive test battery on a large sample of participants with

developmental prosopagnosia to better determine perceptual and memorial contributions to the condition.

Numerous papers exploring the functional activity underlying face recognition have found clear correlates for apperceptive and associative subtypes of acquired prosopagnosia, yet the evidence for functional impairment for developmental prosopagnosia is far less clear. Whereas some studies identify differential activation patterns between prosopagnosia and controls, others claim no difference exists or it differs between participants. This confusion has provided very little in the way of a physiological explanation for the impairments. The same can also be said for the early perceptual ERP, the N170. Some participants appear to have a reduced or absent N170, while for others it is unaffected. Moving beyond the N170 to ERPs related to identity recognition and semantic recall, little research exists. This thesis therefore aims to contribute new and much needed data to this area. The N250r and N400 components provide correlates of different stages of face recognition, but there is too little known about the P600 to be able to offer insight into impaired face recognition. The tasks that induce the N250r and the N400 also have the advantage of providing both direct (forced choice) and indirect (priming) measures. This thesis will measure ERP responses to corroborate the behavioural data and will present data from post-perceptual ERPs previously unreported in developmental prosopagnosia to better understand the underlying physiology of the condition.

The consistency of impairments overtime is unknown in prosopagnosia. The widespread heterogeneity reported in the condition could be a consequence of highly variable symptoms or test insensitivity. Temporal consistency can be established be employing a test-retest design with sufficient separation to help prevent against practise effects. Therefore, this thesis will estimate consistency of impairment over time to better understand the variability of impairments in developmental prosopagnosia.

Current therapies for developmental prosopagnosia have been largely unsuccessful. Attempts commonly lead to benefits that do not extend beyond the training set, or do not persist, even after 14 months of training. In contrast, vestibular stimulation has induced behavioural improvements in both acquired prosopagnosia and typically developing participants. This thesis will test the effects of CVS on the behaviour and physiology in developmental prosopagnosia as a possible new source of therapy.

11. The Thesis Plan

Chapter Two of this thesis sought to understand the perceptual and memorial contributions to developmental prosopagnosia. A large battery of tests, sensitive to subjective and objective features of the condition, was administered to a large group of participants categorised as having developmental prosopagnosia. Critically, the test battery contained tests that assess all of the most commonly reported perceptual and memorial impairments.

Chapter 3 assessed the underlying physiology of the impairments in developmental prosopagnosia. Face sensitive ERPs were measured that correlate with face perception (N170) and two different levels of face recognition, familiarity (N250r) and semantic recall (N400). These post-perceptual ERPs have never before been reported for developmental prosopagnosia. Simultaneously, the effects of CVS of these processes were assessed. All participants were tested twice, once without stimulation and once with stimulation. All of the tasks were compared to an age matched control group.

Lastly, Chapter 4 assessed the consistency of the impairments observed in the ERP experiments by conducting a complete retest of the developmental prosopagnosia group. The retest was conducted one year after the previous testing to eliminate any practise effects.

This is the first evidence of test-retest scores in developmental prosopagnosia and provides qualitative data for the variability in impairments over time.

Overall, the thesis provides new and compelling evidence for the underlying impairments, the ERP response pattern, and the consistency of impairments in developmental prosopagnosia. Furthermore, the thesis presents data for a promising new form of therapy for the condition.

CHAPTER II

Introduction II: Perceptual and Memorial Contributions to Developmental

Prosopagnosia

Prosopagnosia is characterised by a difficulty in recognising people by their facial appearance in the absence of visual, sensory or general intellectual impairment (see Bate, 2012). Initial case reports featured individuals with acquired prosopoagnosia, but there have been increasing reports of cases that appear to be developmental in origin, showing no structural lesion (though see Garrido et al., 2009 for evidence of subtle changes in grey matter volume) and seemingly evident from the early years of life (e.g., Temple, 1992; Kracke, 1994; Ariel & Sadeh, 1996; Bentin, Deouell, & Soroker, 1999; Grueter et al., 2007; Avidan, Tanzer, & Behrmann, 2011; Rivolta, Palermo, Schmalzl, & Coltheart, 2012). Both acquired and developmental prosopagnosia (DP) show a relatively high co-occurrence with difficulties in object identification (Behrmann, Avidan, Marotta, & Kimchi, 2005; Gauthier, Behrmann, & Tarr, 1999) and poor navigational skill (De Haan & Campbell, 1991; Duchaine, Parker, & Nakayama, 2003; Jones & Tranel, 2001). However, while there is evidence to suggest that the symptoms of acquired prosopagnosia can stem primarily from an apperceptive deficit (a difficulty in face perception), or an associative deficit (a difficulty in linking the face percept to semantic information; (Dalrymple et al., 2011; De Renzi, 1986; De Renzi, Faglioni, Grossi, & Nichelli, 1991; Tippett, Miller, & Farah, 2000), the distinction is not always clear. This is even less so in DP as most reported cases are accompanied by some type of perceptual impairment. Here I report the results of a group study in which no single perceptual deficit was either necessary or sufficient for DP to occur. In the majority of cases, no perceptual

deficit was apparent at all. These observations support the idea that there is an isolable memorial component in DP.

Most studies of DP report allied perceptual impairment. A deficit is most commonly identified in recovering configural or holistic information. This is typically inferred by an unusual inversion effect or composite effect (Avidan et al., 2011; Behrmann et al., 2005; Lee, Duchaine, Wilson, & Nakayama, 2009; Le Grand et al., 2006; Nunn, Postma, & Pearson, 2001; Palermo et al., 2011; Schmalzl, Palermo, & Coltheart, 2008), and deficits associated with the recovery of first-order relations (Garrido, Duchaine, & Nakayama, 2008), and the apprehension of global-local hierarchical levels (as tested with Navon compound letters; Navon, 1977) (Avidan et al., 2011; Behrmann et al., 2005: Schmalzl et al., 2008).

Perhaps unsurprisingly, more general tests of face perception, such as the Cambridge Face Perception Test (Avidan et al., 2011; Chatterjee & Nakayama, 2012; Dingle, Duchaine, & Nakayama, 2005; Duchaine, Germine, & Nakayama, 2007), face matching paradigms (Ariel & Sadeh, 1996; Behrmann et al., 2005; Humphreys, Avidan, & Behrmann, 2007; Lee et al., 2009; Nunn et al., 2001) and matching across viewpoint (Behrmann et al., 2005; Duchaine, 2000; Duchaine et al., 2006; Lee et al., 2009; Schmalzl et al., 2008), that are not designed to isolate a particular type of processing (such as the recovery of first- or second-order relations), also unveil impairment. Problems such as apprehending emotional expression, age, and gender (Ariel & Sadeh, 1996, De Haan & Campbell, 1991) accompany these perceptual deficits, but it is unclear whether these latter problems are purely perceptual in nature.

At first glance, these findings might be taken as evidence that DP is perceptual in origin, with impaired recall arising because faces are not adequately encoded at the level of the structural

description. However, close examination of the published data indicates that there is considerable variability in the perceptual deficits reported (Avidan et al., 2011; Stollhoff et al., 2011). For example, despite frequent reports of a deficit in configural processing, this is not always present (Duchaine, 2000; Le Grand et al, 2006). Even family members with congenital prosopagnosia do not show consistent patterns of perceptual deficits (Schmalzl et al., 2008). This raises the question as to whether, on one hand, DP has multiple perceptual origins or, on the other hand, some (or all) of the perceptual deficits co-occur with a memory impairment but do not cause difficulties in face recognition per se. Given the relatively small number of individuals with DP who have been investigated with this aim and the lack of uniform assessment, it is still difficult to decide between these alternatives.

Evidence for the idea that memory deficits can occur independently of perceptual impairments is limited to only a few studies. Two of these (Dalrymple, Garrido, & Duchaine, 2014; McKone et al., 2011) reported a memory impairment with intact face perception. Dalrymple and colleagues showed this dissociation in five of sixteen adult participants while McKone and colleagues showed it in four of the six tested. However, face perception was only assessed via the Cambridge Face Perception Test (CFPT; Duchaine, Germine, & Nakayama, 2007) which, despite its widespread use (Bowles et al., 2009), may not fully capture and differentiate subtle but relevant perceptual deficits (see e.g., Chatterjee & Nakayama, 2012; and DPs F30a and M29 in Garrido, Duchaine, & Nakayama, 2008). More compelling evidence was reported by Lee et al. (2009), who administered a test battery to three family members classified as DP. The father showed characteristically low scores on the most widely used tests for prosopagnosia, the Cambridge Face Memory Test (CFMT; Duchaine & Nakayama, 2006) and a famous faces recognition task, both of which are predominantly memory based. However, he showed normal performance on all five perception tests, including the CFPT (Duchaine,

Germine, & Nakayama, 2007), face detection, emotional expression recognition (from eyes alone), and matching across viewpoint. Although suggestive of a memory impairment, the study did not report, except in one experiment (the face detection task), whether his normal accuracy scores were accompanied by normal reaction times. This is problematic because some cases of prosopagnosia are better characterised by slowed rather than inaccurate responding (Delvenne, Seron, Coyette, & Rossion, 2004; Gauthier, Behrmann, & Tarr, 1999).

The aim of the present study was to cast further light on the role of perceptual and memory factors in DP. Specifically, I wanted to investigate whether some cases of DP are more likely memorial than perceptual in origin. Of course, the value of any such investigation rests on the detail and diversity of the test battery administered. To assess face memory, the two most commonly used tests were employed: (1) the CFMT which probes the ability to learn new faces, and (2) a famous faces task to assess longer-term memory. The commonly used CFPT was administered, which probes the ability to make fine-grained perceptual distinctions between unfamiliar faces. These measures were supplemented with the comprehensive battery devised by Schmalzl et al. (2008), which has proven sensitivity to the main types of perceptual deficit reported in this population. The battery comprises tasks that together estimate the ability to both detect and individuate faces, including measures that are sensitive to the detection of first- and second-order spatial relations, global/local processing, holistic processing, the detection of feature and contour changes, viewpoint matching, and judgements of facial expression. Note that the Famous Faces task (Exp. 1) and the CFMT (Exp. 2) are reported first, as these were employed to classify DP, followed by the CFPT (Exp. 3) and the test battery (Exp. 4 to 10). However, during data collection, these tasks were administered in a different order (Exp. 4 to 10, Exp. 3, Exp. 2, and Exp. 1). Below I describe the ability of 11 individuals

with face recognition difficulties to perform these tests, relative to a group of age- and gender-matched controls.

General Method

Participants

All participants reported normal or corrected-to-normal vision and no documented history of brain injury. To eliminate influence of the own-race recognition bias (see Brigham & Malpass, 1985), participants were all Caucasian and had lived in the UK for the last three years. Ethical approval for this study was granted by the School of Psychology Ethics Board at the University of Kent, and written informed consent was obtained from each participant at enrolment.

Developmental prosopagnosics

Thirty individuals responded to adverts placed in the local newspaper and on the University departmental website which encouraged people to get in touch if they experienced face recognition difficulties. Both advertisements referred to "face blindness" rather than "developmental prosopagnosia" and featured the following statements, "Do you have difficulty recognising friends when not expecting to see them?" and "Do you have difficulty keeping up with characters in a film?". These questions were not intended to confirm the presence of prosopagnosia, but to confirm a degree of face recognition difficulty that was worthy of further assessment. An interview with each self-referred participant was subsequently conducted in a research laboratory. This lasted approximately one hour and sought confirmation that the recognition difficulties were symptomatic of DP. To be considered for study, each individual had to (1) answer "yes" to all eight questions about their activities of daily living (ADL) using a DP questionnaire devised by DeGutis, DeNicola, Zink, McGlinchey and Milberg (2011), (2)

confirm that the recognition problem had been evident since childhood, (3) confirm that the difficulty had not followed from a traumatic or other prominent neurological event, and (4) confirm that it was not accompanied by a developmental (e.g., autism or Asperger's syndrome) or psychiatric disorder. Unprompted, many individuals also mentioned the presence of similar symptoms in a family relative. To confirm that a participant has prosopagnosia, each individual then had to score below two standard deviations of the control group's accuracy on the Famous Faces Task and below the established 58.4% cut-off (i.e., an overall score of 42 or below; see e.g., Bennetts, Butcher, Lander, Udale, & Bate, 2015; Duchaine & Nakayama, 2006) on the CFMT (see Experiments 1 and 2 below). Of the 30 self-referred individuals, 11 (3 male, 2 left-handed) aged 25-62 years (mean = 46.1, SD = 14.8) met all of these criteria and were therefore deemed eligible for enrolment.

Control group

Eleven (4 male, 2 left-handed) participants aged 25-62 years old (mean = 46.2, SD = 14.4) were recruited to closely match the age, gender, and handedness of the DP group. Typically the difference in age between each control and their corresponding DP was less than one year, but up to two years maximum. All confirmed that they were unaware of a problem recognising faces, answered no to all questions on the ADL questionnaire, and did not report a significant neurological or psychiatric history.

General Procedure

Testing was conducted in a psychology laboratory at the University of Kent. The computerised tasks were administered on a Microsoft Windows® computer and stimuli were presented on a 20.1 inch Dell™ monitor at a resolution of 1024 × 768 pixels. Participants were seated at a distance of 1 meter from the computer screen, with their eyes level with the top of the monitor.

Responses were given using a Dell™ USB keyboard. All participants took part in all experiments described below.

Statistical Approach

The responses of each group to the various experimental manipulations were analysed using either univariate or mixed-effects ANCOVAs (with significant interactions explored using Bonferroni-corrected pairwise comparisons). For those tasks that required a speeded response, inferential analyses were only performed on mean correct reaction times that fell within two standard deviations of each participant's mean and that had been log transformed to base power 10. MANCOVAs were also performed on the accuracy and RT data to determine whether a linear combination of perceptual sub-test scores, accounting for age, was more associated with the DP than control group. To provide a handle on individual variability within the DP group, individual test scores were also compared to the control group using Crawford and Howell's (1998) modified t-tests, which are less vulnerable to the inflated Type 1 error rate that can occur when z-scores are calculated from relatively small control group samples. To increase the likelihood of uncovering perceptual impairment (and thereby refute the prediction that some cases of DP are more memorial than perceptual in nature), these t-tests were initially uncorrected for multiple comparisons (alpha = 0.05). As can be seen in Figure 2.2, this liberal criterion seems appropriate given that the prevalence of perceptual impairment turned-out to be low. Cronbach's alpha was calculated (based on the 22 control and DP participants) for all experiments in which performance in both groups was below ceiling, except the CFMT and CFPT for which internal consistency reliability has been widely reported (Bowles et al., 2009). Although the controls and DPs were closely age-matched, the range was relatively large (37 years). Given evidence that age can affect performance on key measures such as the CFMT and CFPT (Bowles et al., 2009), age was therefore included as a co-variate in all statistical tests

that follow. As described below, there were only two instances (in the configural and holistic tasks) in which experimental performance deteriorated as age increased but importantly these effects did not interact with Group and only influenced variables that are widely seen as less diagnostic of face perception impairment than other variables measured here. Mindful of the finding by Bowles et al. (2009) that age-related norms become especially important in studies that recruit prosopagnosics over the age of 50, an additional statistical procedure was performed in which age was added as a binomial variable (< 50 years [n=12] vs. > 50 years [n=10]) to each experimental analysis. This factor again failed to reach significance in any of the experiments (alpha = 0.05) except in several sub-conditions of the Composite task but here it neither interacted with Group nor with the inversion effect (i.e. the outcome of most interest). Given the modest and limited nature of the age effects that emerge when age is treated as a binomial variable, no further mention is made of them in this report.

Experiment 1

Famous Faces Task

To assess long-term retention and retrieval of faces, participants were presented with pictures of well-known celebrities to name. Face images were gathered from the internet and all depicted a frontal view with minimal hair occlusion and no visible accessories such as hats and earrings (though glasses were permissible if usually worn). Images were cropped to include only the face and external features (hair, ears, jaw line) and presented on a black background (see Figure 2.1). Cropped stimuli that were less than 320 pixels high were rejected, and faces were proportionally resized to subtend no more than $3.9° \times 5.3°$ of visual angle with a resolution of 72 pixels per inch. Seventy-seven images were presented in a random order, with each trial beginning with a central fixation cross for 500ms, ending with a blank screen

for 500ms. The face remained on-screen until an unspeeded response was made. Participants were asked to name out loud the celebrity or, if the name did not come to mind but they knew the identity of the person, to provide a distinguishing semantic fact. Either response was marked as "correct" if the individual was identified. Task duration was approximately eight minutes.
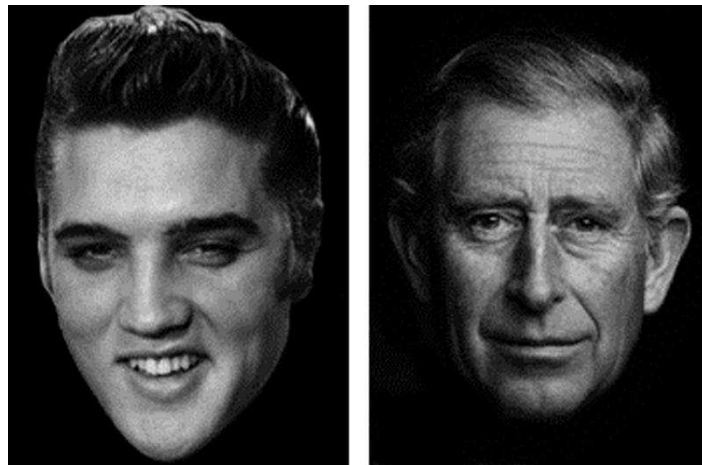
Figure 2.1: Example stimuli from the Famous Faces Task (FFT).

Internal consistency, as assessed using Cronbach's alpha, was .970. A univariate ANCOVA showed that the DP group (mean = 40.7%, SD = 15.1) performed the task less accurately than the control group (mean = 82.3%, SD = 9.9), $F(1,19) = 59.85$, $p < .001$. There was no main effect of Age, $F(1,19) = 1.56$, $p = .08$. The individual test scores of the DPs were compared to the controls using Crawford and Howell's (1998) modified t-distribution and, as with all other experiments, are presented in Figure 2.2.

Accuracy

	1	2	3	4	5	6	7	8	9	10	11
FFT	-2.69	-6.22	-6.35	-4.84	-2.95	-2.69	-5.47	-2.57	-2.95	-3.32	-4.33
CFMT	-2.82	-4.61	-3.71	-4.61	-2.82	-2.97	-3.71	-2.52	-2.82	-2.82	-2.97
CFPT Upright	-0.12	3.02	1.91	1.73	0.81	-0.67	1.18	0.07	2.10	1.55	-0.86
CFPT Inversion Index	0.52	1.83	1.63	0.16	0.36	-0.78	1.49	-0.02	1.37	0.60	-0.46
Configural	0.19	0.19	-0.52	0.90	-	0.19	-0.52	0.90	-	-	0.19
Mooney	0.56	-0.81	-0.81	0.56	0.56	0.56	-0.81	0.56	-0.12	0.56	0.56
Composite	0.44	0.29	1.63	0.59	1.63	-0.31	-0.01	0.44	-1.06	-0.16	1.18
Composite Effect	0.84	0.15	-0.78	-0.78	-0.78	-0.78	-1.01	0.15	0.38	1.07	-0.54
Jane Upright	-1.26	-1.01	-0.63	0.00	-0.50	-0.50	-1.77	-0.25	-1.51	-0.50	-1.14
Jane Inversion Effect	-0.21	1.24	0.41	-0.83	0.21	1.86	-1.24	-1.86	-0.41	1.24	-1.24
Viewpoint	-1.44	-0.48	-1.44	-0.48	-0.96	-0.48	-2.39	0.00	-1.91	-4.31	0.00
Expressions	1.21	0.45	-0.31	-1.07	-1.07	1.59	-3.35	0.83	-0.69	-2.21	1.21

Response times

	1	2	3	4	5	6	7	8	9	10	11
Configural	0.23	3.46	1.13	-0.35	-	1.81	-0.16	-1.19	-	-	-0.43
Mooney	-1.15	1.43	1.66	1.95	0.56	0.78	-0.10	-0.29	1.61	0.96	1.68
Composite	0.41	1.41	1.21	0.47	0.32	0.88	1.12	-0.48	1.21	2.14	0.05
Composite Effect	-0.13	-0.19	1.25	-1.02	-0.85	-0.41	0.14	0.01	-0.01	0.55	-0.28
Jane Upright	-0.08	2.62	1.78	1.35	1.74	0.28	0.77	0.25	-0.31	2.63	-0.37
Jane Inversion Effect	0.74	3.22	2.33	0.61	1.73	1.52	0.45	-1.78	3.57	0.17	0.67
Viewpoint	1.04	1.79	0.08	0.49	-0.44	0.32	1.28	0.86	0.80	2.27	-0.43

Figure 2.2: Mean accuracy scores and reaction times (RTs; where appropriate) for each developmental prosopagnosia (DP) participant (1 to 11) expressed as modified t-scores, derived by comparison to the control group mean and standard deviation. Scores that are significantly below the performance of the control group (i.e., >2.23, corresponding to alpha = .05, df = 10) are shaded in black. FFT = Famous Faces Task; CFMT = Cambridge Face Memory Test; CFPT = Cambridge Face Perception Test.

Experiment 2

Cambridge Face Memory Test (CFMT)

The CFMT was administered to assess participants' ability to learn and retrieve the identities of new faces. In contrast to the other face perception tests described below which rely only on detection or matching, the CFMT incorporates both a perceptual and memorial component (for details see Duchaine & Nakayama, 2006). In short, participants are first asked

to learn six faces, each from three viewpoints. The recall phase is then divided into three blocks. In Block 1 (introduction) participants are briefly shown a target face and then immediately shown 3 more faces and asked which one they just saw. Given that the study and test images are the same, the task can be correctly performed using image matching alone. In Block 2 (novel images), participants are first shown all six target faces in frontal view and given 20 seconds to review them. Across 30 trials, they are then required to pick out each of these faces from a three-face line-up consisting of a new image of the target identity and two foils (see Figure 1.2). Block 3 (noise) is a 24 trial variation of Block 2 (including the same 20 second review screen) but with Gaussian noise added to the stimuli to keep performance below ceiling and place greater reliance on mechanisms believed central to face recognition. Responses were unspeeded and task duration was approximately eight minutes.

Accuracy was, as is typical, highest for the introduction condition (mean = 90.4%, SD = 17.2), followed by the novel images condition (mean = 64.4%, SD = 21.4) and lowest for the noise condition (mean = 50.9%, SD = 21.4). A one-way ANCOVA revealed a main effect of Group, $F(1,19) = 77.84$, $p < .001$, indicating that controls (mean = 81.8%, SD = 8.9) were more accurate than the DPs (mean = 51.0%, SD = 6.9). There was no main effect of Age ($F(1,19) = 0.01$, $p = .91$). Individual performance is described in Figure 2.2.

Experiment 3

Cambridge Face Perception Test (CFPT)

The CFPT was administered to assess the perception of facial similarity (for details of the original task see Duchaine, Germine, & Nakayama, 2007). In short, participants are given 60 seconds to arrange six frontal facial images according to their similarity to a ¾ view target face that appears directly above (see Figure 2.3). Eight different upright arrangements are then

duplicated, inverted, and intermixed to create 16 pseudorandomised trials. Scores reflect how many deviations the participants' order of the faces falls from the correct order. Task duration was approximately 18 minutes.
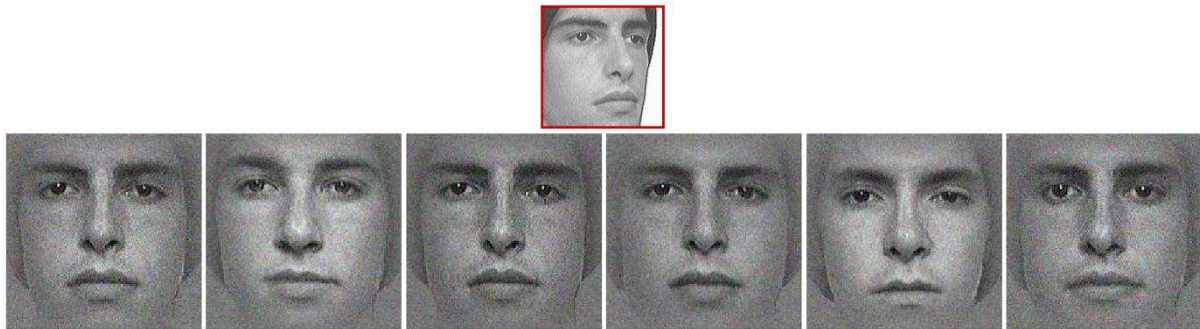
Figure 2.3: Example stimuli from the Cambridge Face Perception Test (CFPT).

The mean correct responses for all participants are shown in Table 1. A 2 (Group) $\times$ 2 (Orientation) ANCOVA failed to find significant effects involving Orientation (all $Fs(1,19) \leq$ 3.66, ps $\geq$ .07), Group or Age (both $Fs(1,19) \leq$ 2.48, ps $\geq$ .13). An inversion index was calculated using the formula: (upright - inverted) / (upright + inverted) (see Wilkinson, Ko, Wiriadjaja, Kilduff, McGlinchey, & Milberg, 2009). A normal inversion effect (negative index value) was demonstrated for each group and when interrogated with ANCOVA was found to be comparable across Group and Age (all $Fs(1,19) \leq$ 2.31, ps $\geq$ .15). Individual analysis indicated that only one of the 11 DPs demonstrated significant impairment on the CFPT in comparison with the control group. This impairment was only evident in the upright CFPT condition but did not eliminate the inversion effect.

Table 2.1: Means and standard deviations across participant group and task

Task	Condition	Accuracy				RT			
		Controls		DPs		Controls		DPs	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
FFT	Overall	.82	.10	.41	.15				
CFMT	Intro	.99	.02	.82	.21				
	Novel	.83	.10	.46	.11				
	Noise	.68	.16	.34	.09				
	Overall	.82	.09	.51	.07				
CFPT (Inaccuracy)	Upright	4.41	1.30	5.73	1.68				
	Inversion Index	-0.32	0.17	-0.21	0.15				
Configural	GlobalCon	.99	.02	.99	.02	2.74	0.06	2.82	0.14
	GlobalIncon	.98	.03	.98	.03	2.77	0.08	2.84	0.13
	LocalCon	1.00	.00	.99	.02	2.79	0.11	2.80	0.12
	LocalIncon	.97	.04	1.00	.00	2.81	0.10	2.85	0.13
	Overall	.98	.02	.99	.01	2.78	0.08	2.83	0.12
Mooney	Overall	.98	.04	.98	.02	2.92	0.07	2.98	0.07
Composite	IntactDiff	.93	.08	.93	.06	3.24	0.14	3.36	0.10
	IntactSame	.81	.18	.87	.12	3.30	0.17	3.39	0.08
	MisalignedDiff	.86	.16	.89	.10	3.20	0.11	3.31	0.10
	MisalignedSame	.95	.04	.97	.04	3.16	0.14	3.24	0.12
	Overall	.89	.07	.92	.06	3.22	0.12	3.32	0.09
CFE	Overall	0.14	0.17	0.11	0.13	-0.14	0.12	-0.15	0.08
Jane Upright	Spacing	.85	.08	.78	.10	3.32	0.17	3.52	0.23
	Feature	.95	.06	.95	.05	3.21	0.12	3.34	0.12
	Contour	.83	.15	.69	.13	3.32	0.15	3.43	0.20
	Overall	.88	.08	.81	.05	3.28	0.14	3.42	0.16
Jane Inversion Effect	Spacing	0.16	0.14	0.18	0.13	-0.13	0.14	0.10	0.22
	Feature	0.07	0.10	0.05	0.08	-0.11	0.14	-0.05	0.13
	Contour	0.14	0.13	0.11	0.13	-0.08	0.17	-0.01	0.17
	Overall	0.12	0.07	0.11	0.06	-0.10	0.11	0.01	0.15
Viewpoint	Front	.99	.02	.99	.02	3.26	0.11	3.33	0.14
	45 Degrees	.95	.05	.90	.06	3.40	0.14	3.51	0.11
	90 Degrees	.85	.08	.79	.07	3.49	0.14	3.63	0.11
	Overall	.93	.03	.89	.04	3.37	0.13	3.47	0.11
Expressions	Overall	.85	.05	.83	.09				

Note: DP = developmental prosopagnosia group; RT = reaction time; FFT = Famous Faces

Task; CFMT = Cambridge Face Memory Test; CFPT = Cambridge Face Perception Test;

CFE = Composite Face Effect.

Experiment 4

Basic Configural Processing

A global-local Navon task (Navon, 1977) was administered to assess basic configural processing, that is the perception of both global/local structure and global precedence. This task does not utilise face stimuli and it remains unclear whether it relates to DP (Duchaine, Yovel, & Nakayama, 2007). However, it has been used repeatedly as an analogue to assess the bias for processing faces holistically compared to on an individual featural level (Behrmann et al., 2005; Bentin, DeGutis, D'Esposito, & Robertson, 2007; Macrae & Lewis, 2002; Perfect, Weston, Dennis, & Snell, 2008; Weston & Perfect, 2005). Stimuli consisted of the outline of either a circle or square (global shapes), which was made up from either small circles or small squares (local shapes) (see Figure 2.4). Stimuli subtended $6.8° \times 6.8°$ of visual angle. On congruent trials, local and global elements matched (i.e. the big square was composed of small squares), while on inconsistent trials they had a different appearance (i.e. the big square was composed of small circles). The first block consisted of 40 randomised trials (20 congruent and 20 incongruent) with each trial beginning with a central fixation cross for 500 milliseconds (ms), followed by the stimulus which remained on screen until a response was made, ending with a blank screen for 500ms. Participants were asked to identify the large shape by pressing "c" or "s" on a standard computer keyboard quickly and accurately (as was the case for all tests described below, the mapping between response button and stimulus selection was counterbalanced across participants). The second block was identical to the first but participants were asked to identify the smaller shape. Task duration was approximately four minutes.
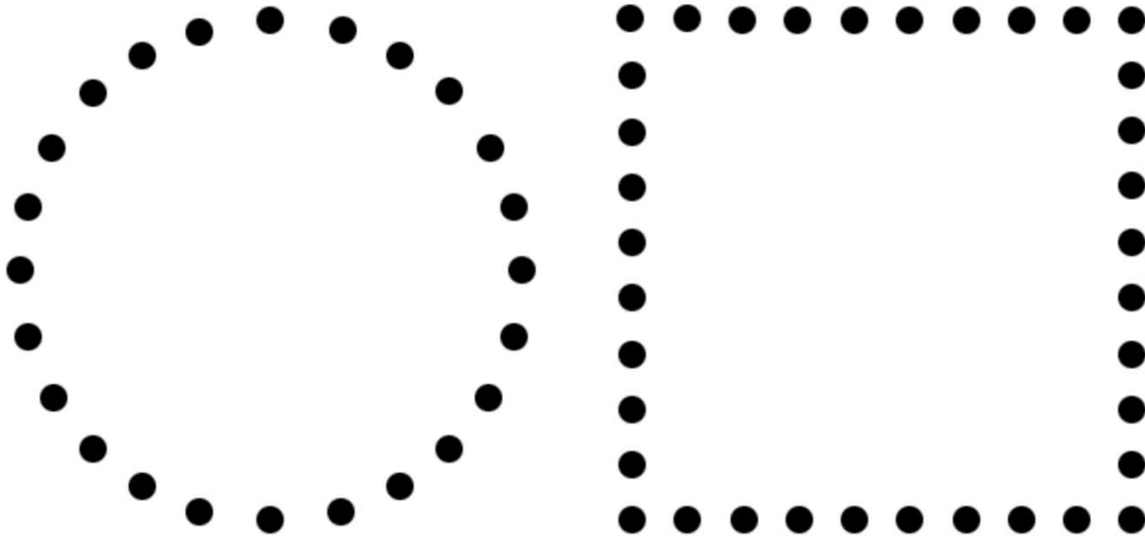
Figure 2.4: Example stimuli from the configural task.

The responses of three DPs were removed prior to analysis because they confused the response key mappings. Cronbach's alpha for the RT data was .839 but was not calculated for the accuracy data as the mean accuracy for both groups was near ceiling (98%). The mean correct responses and reaction times for all other participants are shown in Table 1. The correct responses were then subjected to a 2 (Group) $\times$ 2 (Global Target/Local Target) $\times$ 2 (Consistent/Inconsistent) repeated-measures ANCOVA controlling for age as a covariate. Local judgements were more accurate than global judgments, $F(1,16) = 5.62$, $p < .05$, and were moderated by age, $F(1,16) = 4.66$, $p < .05$, such that they became easier as age decreased. Neither the main effect of Consistency nor the Consistency $\times$ Target interaction term reached significance (both Fs $(1,16) \leq 0.93$, ps $\geq .35$), though this may have arisen because of the observed ceiling effects in mean accuracy. A three way interaction of Group, Target and Consistency was also observed, $F(1,16) = 7.55$, $p < .05$, and was driven by more accurate responses to local targets in the control group when they were consistent (mean = 100.0, SD = 0.0) compared to inconsistent (mean = 96.8, SD = 4.0) with the global target ($p < .01$). The main effects of Group and Age were not reliable (both Fs$(1,16) < 0.27$, ps $\geq .61$).

The same ANCOVA was applied to the RT data and indicated that reaction times were shorter for consistent versus inconsistent trials, $F(1,16) = 4.94$, $p < .05$. The main effect of Target was not significant, $F(1,16) = 1.18$, $p = .30$, but a significant Target $\times$ Consistency interaction emerged, $F(1,16) = 5.60$, $p < .05$, which was driven by an effect of Consistency only at the Local level ($p < .001$). However, this interaction effect reduced as age increased, $F(1,16) = 5.01$, $p < .05$. There was no main effect of Age, $F(1,16) = 1.11$, $p = .31$, and no main effect, $F(1,16) = 1.11$, $p = .31$, or interactions involving the factor Group (all $Fs(1,16) \leq 3.43$, $ps \geq .08$). Finally, performance was once again assessed at the level of the individual (see Figure 2.2). Only one (DP 2) of the 11 DPs demonstrated significant impairment in the Configural task, showing slower RTs than the controls.

Experiment 5

Face Detection

The Mooney task (Mooney, 1957, 1960) was administered to assess the detection of first-order relations (e.g., the basic configuration of a pair of eyes above a nose and mouth), which are believed to be important for recognising a face as a face. Stimuli consisted of degraded facial images in which all colour was transformed to black or white. Individual pixels of contrasting luminance were removed to form smooth blocks of colour so that shadows and highlights were made salient. These stimuli were then duplicated and individually rearranged to form images with no discernible form, thus creating the non-face images. Stimuli subtended $6.9° \times 6.9°$ of visual angle. The task consisted of 40 trials, each containing a pair of stimuli comprising one face and one non-face (see Figure 2.5). Trials began with a central fixation for 500 ms, followed by the stimulus which remained on screen until a response was made, ending with a blank screen for 500 ms. Participants were instructed to choose quickly and accurately

which image depicted a face by pressing "1" or "2". Task duration was approximately two minutes.
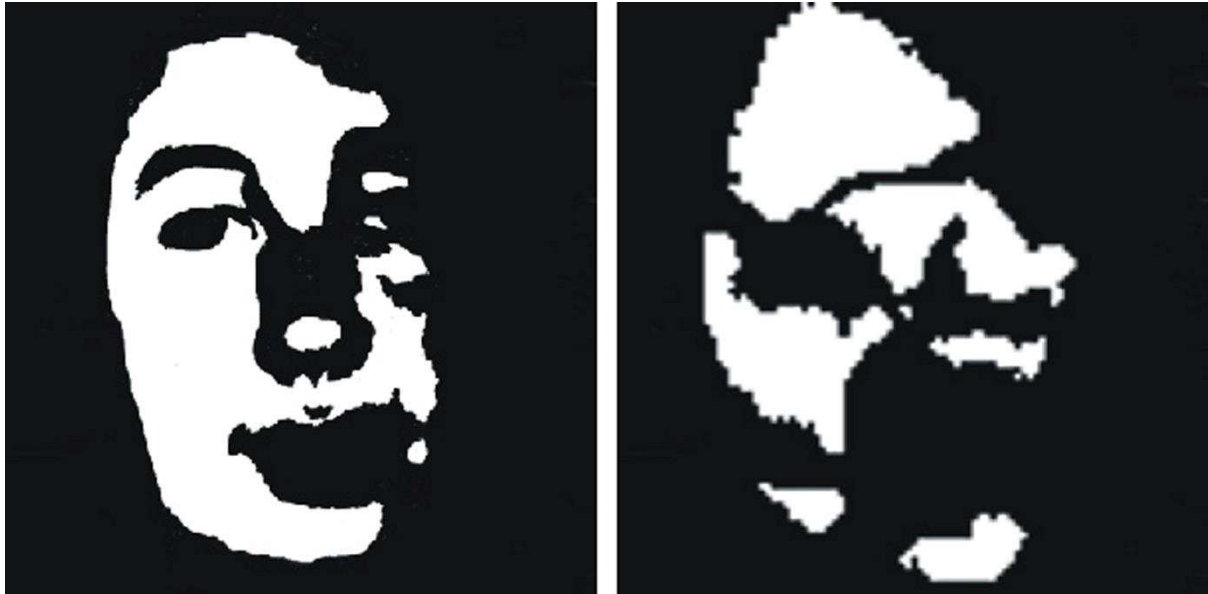
Figure 2.5: Example stimuli from the Mooney task.

Cronbach's alpha for the RT data was .864 but was again not calculated for the accuracy data as the mean accuracy for both groups was near ceiling (98%). The mean correct responses and reaction times are shown in Table 1. Accuracy and reaction times were analysed in separate univariate ANCOVAs as a function of Group and Age and produced no statistically significant differences (Age in accuracy $F(1,19) = 2.86$, $p = .11$; Group in RTs $F(1,19) = 3.88$, $p = .06$; all other $Fs(1,19) \leq 0.47$, $ps \geq .50$). Individual analysis indicated that none of the 11 DPs demonstrated impairment in this task (see Figure 2.2).

Experiment 6

Holistic Processing

Holistic processing refers to the simultaneous encoding of multiple features and their integration into a coherent whole. This capacity is taken as a core requirement for normal face processing (see Rossion, 2008). The most widely recognised measure of holistic face processing is the Composite Faces Task (Le Grand, Mondloch, Maurer, & Brent, 2004). Participants are presented with pairs of faces in which the tops and bottoms of each face are either aligned or misaligned (see Figure 2.6). While the bottom halves of the faces are always different, the top halves can match and participants are asked to determine as quickly and accurately as possible whether the top halves of the faces are the same or different. In the current experiment, responses were registered by pressing "s" or "d" on the keyboard. The composite face effect relies on the assumption that misaligned faces disrupt holistic processing and force recognition to be based on features instead. This should therefore lead to better featural discrimination performance, of the top halves of the faces, in the misaligned compared to the aligned condition (Maurer et al., 2002).

The first block (aligned) consisted of 48 randomised trials evenly divided into "same" or "different" conditions. In the second block (misaligned), the bottom half of the face stimuli was shifted half-way to the right. Each trial began with a central fixation cross for 500 ms, followed by the stimulus which remained on screen until a response was made, ending with a blank screen for 500 ms. Aligned face stimuli subtended $6.6° \times 9.8°$ of visual angle and misaligned face stimuli subtended $9.8° \times 9.8°$ of visual angle. Task duration was approximately six minutes.
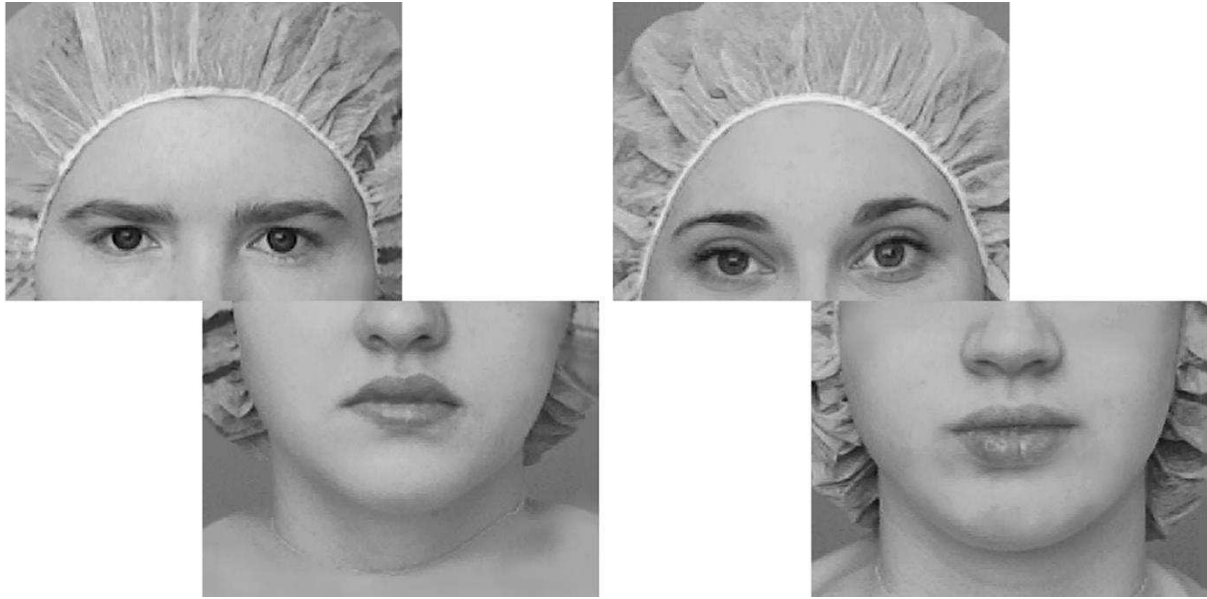
Figure 2.6: *Example stimuli ("different–misaligned" condition) from the composite task.*

Cronbach's alpha was .794 for the accuracy data and .958 for the RT data. The mean correct responses and reaction times are shown in Table 1. A 2 (Group) × 2 (Alignment) ANCOVA (with age again entered as a covariate) of the accuracy scores revealed a main effect of Age, $F(1,19) = 9.22$, $p < .01$, whereby accuracy reduced as age increased, but importantly this did not interact with Group. Analysis of the RT data showed that the control group generated shorter reactions overall, $F(1,19) = 4.87$, $p < .05$, but no other effects reached significance (all $Fs(1,19) \leq 1.70$, $ps \geq .21$).

There are two commonly applied methods to calculate a face composite effect (that is, the illusion that the top halves of two faces are different when aligned with two different bottom halves of faces). The traditional measure is obtained by subtracting performance on same-misaligned trials from that on same-aligned trials. Using this method, reliability was .56 for the accuracy data and .70 for the RT data (calculated using the subtraction method described by DeGutis, Wilmer, Mercado, & Cohan, 2013). Both groups generated negative face composite effect scores in accuracy and positive scores in RT, which is suggestive of holistic processing. Allied univariate ANCOVAs revealed no effect of Group or Age in either the

accuracy or RT data (Age in accuracy $F_{(1,19)} = 4.07$, $p = .06$; all other $Fs_{(1,19)} \leq 0.47$, $ps \geq$ .50). The second method of assessing the composite effect is to calculate an inversion index (misaligned - aligned) / (misaligned + aligned) (see Avidan et al., 2011). Again, both groups showed positive indices in accuracy and negative indices in RT, indicative of holistic processing. The ANCOVAs showed no significant Group or Age differences (all $Fs_{(1,19)} \leq$ 0.92, $ps \geq .35$). No individual DPs showed impairment in this task (see Figure 2.2).

Experiments 7 & 8

Detection of Spacing, Feature and Contour Changes

In Experiment 7, the Jane Task (Le Grand Mondloch, Maurer, & Brent, 2001) was used to estimate participants' sensitivity to subtle changes in either the identity or spacing of individual features. Several reports exist of impairment in this task in DP (Schmalzl et al., 2008; Le Grand et al., 2006; Rivolta et al., 2012), justifying its inclusion here. All stimuli were derived from just one face (Jane's) that had been altered from the original in one of three ways; (1) in the spacing condition the eyes were moved in/out (see Figure 2.7) or the eyes and mouth were moved up/down, (2) in the feature condition, the eyes and mouth were replaced with those from another face, and (3) in the contour condition, the internal part of the face was combined with the contour from another face. Each of the three conditions consisted of 30 randomised pairs of faces (presented alongside each other), with one of the faces altered on 50% of the trials. Each face stimulus subtended $6.6° \times 9.8°$ of visual angle. Each trial began with a central fixation cross for 500ms, followed by the stimulus until a response was made, and ending with a blank screen for 500ms. Participants were asked to indicate quickly and accurately whether the two faces presented were the same or different by pressing "s" or "d".

In Experiment 8, the entire procedure was repeated but with the faces inverted. Sensitivity to second-order relations is strongly affected by inversion (Freire, Lee, & Symons, 2000; Leder & Bruce, 2000) so individuals who normally make use of this information to identify faces should be adversely affected by the manipulation. Task duration was approximately 11 minutes for each of the two experiments.

Figure 2.7: *Example stimuli ("different" condition) from the Jane task.*

Cronbach's alpha in Experiment 7 was .794 for the accuracy data and .976 for the RT data, and in Experiment 8 was .881 for the accuracy data and .979 for the RT data. The mean correct responses and reaction times are shown in Table 1. In Experiment 7, a 2 (Group) $\times$ 3 (Feature change) ANCOVA indicated that the controls were generally more accurate than the DPs, $F(1,19) = 5.87$, $p < .05$. Group also interacted with Feature change, $F(2,38) = 3.72$, $p <$

.05; while DPs were more accurate at detecting feature compared to contour changes, the controls showed no such sensitivity (p < .001). There was no main effect or interaction involving Age (both $Fs \leq 0.15$, ps $\geq$ .78). Despite the higher group accuracy for controls, none of the individual DPs showed evidence of impairment (see Figure 2.2).

The RT analysis indicated that controls were generally faster to respond than the DPs, F(1,19) = 4.77, p < .05, and that, together, differences in features (mean = 3.275, SD = 0.13) were more quickly detected than differences in spacing (mean = 3.421, SD = 0.22) or contour (mean = 3.373, SD = 0.18) (F(2,38) = 3.32, p < .05) (pairwise p-values < .005). Group and Feature did not interact with each other, F(2,38) = 1.33, p = .28. There was no main effect or interaction involving Age (both $Fs \leq 0.69$, ps $\geq$ .51). Again, however, the overall impairment in the DPs was not robust at an individual level, with only two participants (DP 2 and DP 10) showing significantly slower RTs to the controls (see Figure 2.2).

The effect of inversion in the two groups was explored by subtracting the performance in the inverted experiment from that in the upright experiment. This was calculated separately for the accuracy and RT data. Reliability was .31 for the accuracy data and .94 for the RT data (DeGutis et al., 2013). Both groups showed normal inversion effects, producing accuracy and RT inversion effect scores that were above and below zero respectively for all Feature conditions except the spacing change condition in which the DP group did not show a negative RT inversion effect score. A 2 (Group) × 3 (Feature change) ANCOVA was conducted separately on the inversion effect scores of correct responses and reaction times. No main effects or interaction terms reached significance in the accuracy inversion analysis (Feature change F(2,38) = 2.08, p = .14; Feature × Age F(2,38) = 2.49, p = .10; all other Fs $\leq$ 0.90, ps $\geq$ .35), but in the RT data controls (mean = -.102, SD = 0.11) produced a more negative inversion effect score than the DPs (mean = .007, SD = 0.15) (F(1,19) = 4.27, p = .05). The Group × Feature change interaction was also significant, F(2,38) = 3.62, p < .05, and was driven

by a lower (i.e. negative) inversion effect score in the spacing change condition in the controls vs. DPs (p < .01). There was no main effect or interaction involving Age (both Fs ≤ 1.37, ps ≥ .27). At the individual level, and when collapsing across spacing and contour conditions (inversion effects are not expected in feature change detection; reliability rises to .49 and .94 for accuracy and RT respectively [DeGutis et al., 2013]), only DP 2, DP 3 and DP 9 produced a smaller inversion effect score (RT data only) than the control group (see Figure 2.2). An alternative inversion index calculated again using the formula: (upright - inverted) / (upright + inverted) (see Wilkinson, Ko, Wiriadjaja, Kilduff, McGlinchey, & Milberg, 2009) yielded the same pattern of statistically significant and non-significant effects.

Experiment 9

Viewpoint Matching

To assess the ability to form viewpoint-independent representations of faces, participants were asked to quickly and accurately match unfamiliar faces presented at different horizontal viewpoints, which were full frontal, 45-degree mid-profile, and 90-degree side profile view. Each trial consisted of a full frontal view of an unfamiliar face simultaneously presented below three other faces, one of which was the same person and the remaining two were foils. All three were presented at either the same or a different viewpoint to the target. Participants had to indicate, via button press, which of the three test faces was the same person depicted below in full frontal view (see Figure 2.8). Each face subtended, on average, 3.5° × 7.9° of visual angle and was positioned within a frame subtending 16.3° × 12.3° of visual angle. Of the 60 trials, 20 showed front views of the three faces, in another 20 trials the faces were at 45° (10 left, 10 right), and in the remaining 20 the faces were at 90° (10 left, 10 right). All trials were randomised and began with a central fixation cross for 500ms, followed by the

stimulus which remained on screen until a response was made, ending with a blank screen for 500ms. Task duration was approximately five minutes.

Figure 2.8: Example stimuli presentation from the viewpoint task.

Cronbach's alpha was .426 for the accuracy data and .937 for the RT data. The low Cronbach's alpha for the accuracy data is due to highly consistent participant performance. Mean accuracy across the three conditions was 93.3% (SD = 3.3) for controls and 88.9% (SD = 4.4) for the DP group. As expected, accuracy in both groups was highest when the 3 test faces were full-frontal (control mean = 99.1%, SD = 2.0; DP mean = 98.6%, SD = 2.3) followed by the 45 degree side views (control mean = 95.5%, SD = 4.7; DP mean = 89.5%, SD = 6.1), and then the 90 degree profiles (control mean = 85.5%, SD = 7.9; DP mean = 78.6%, SD = 7.4).

A 2 (Group) × 2 (Angle) ANCOVA conducted on the accuracy data revealed a main effect of Angle (Front > 45 > 90), $F(2,38) = 4.06$, $p < .05$, and a significant main effect of Group, $F(1,19) = 6.72$, $p < .05$, which reflected higher accuracy for the control group (mean = .93, SD = .03; though note that accuracy was still high in the DP group [mean = .89, SD = .04]). No other main effect or interaction terms reached significance (Group × Angle interaction F = 2.66, $p = .08$; all other Fs ≤ 0.40, ps ≥ .54). However, despite the higher control group accuracy only two individuals (DP 7 and DP 10) showed evidence of significant impairment (see Figure 2.2).

The same ANCOVA conducted on the RT data also showed a main effect of Angle, $F(2,38) = 23.11$, $p < .001$. The main effect of Group, $F(1,19) = 4.31$, $p = .052$, and the Group by Angle interaction did not reach significance, $F(2,38) = 2.96$, $p = .06$ (all other Fs ≤ 1.56, ps ≥ .22). At the individual level, only one participant (DP 10) performed significantly below the group mean.

Experiment 10

Judgments of Emotional Expression

Judgements of emotional expression are generally spared in prosopagnosia (Humphreys et al., 2007; Nunn et al., 2001) but have been shown in some cases of DP (De Haan & Campbell, 1991; Duchaine et al., 2006; Minnebusch et al., 2007). Consequently, emotion recognition was also assessed here. Stimuli were made up of 48 faces depicting the six emotions of happiness, sadness, anger, fear, surprise or disgust (Ekman & Friesen, 1976) (see Figure 2.9), each subtending 5.8° × 8.8° of visual angle. Each trial began with a central fixation cross for 500 ms, followed by a face stimulus, which remained on screen until a response was made, and ending with a blank screen for 500 ms. Participants were asked to

quickly and accurately judge the emotional expression of the face by pressing one of six buttons (a reference key showing which button should be pressed for each emotion was placed below the monitor during the experiment). Task duration was approximately 5 minutes.
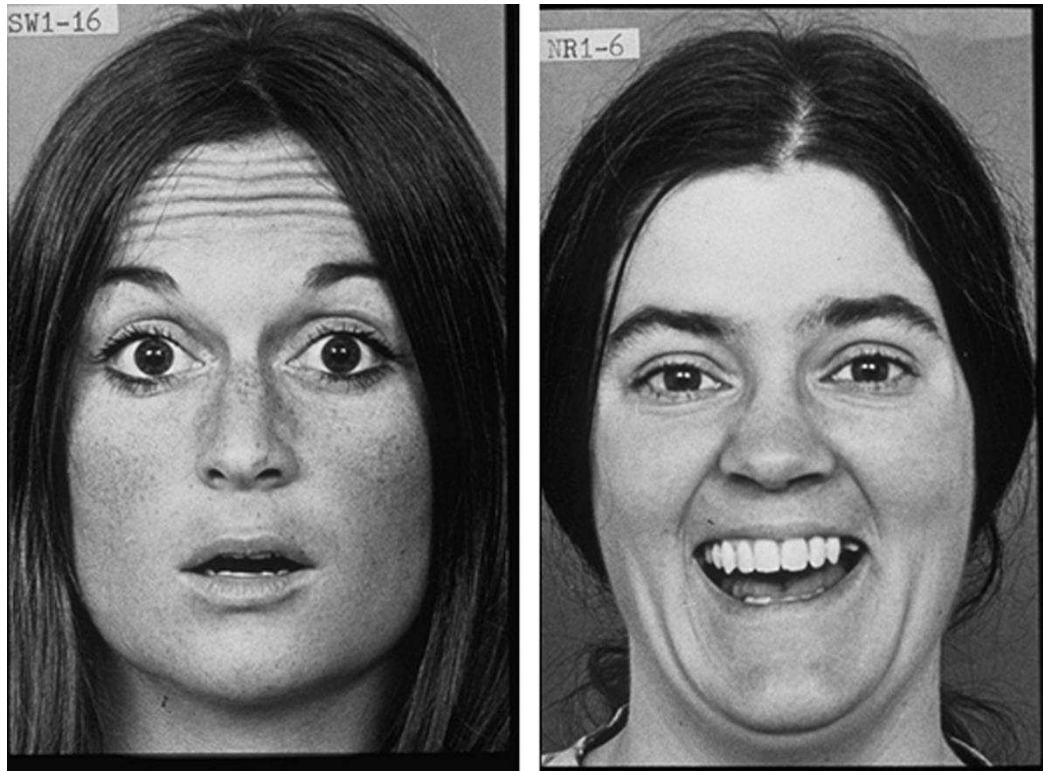
Figure 2.9: Example stimuli (surprise and happiness) from the emotional expression task.

Cronbach's alpha was .535 and the following results should consequently be interpreted with caution. Mean accuracy was 85.0% for the controls (SD = 5.3) and 83.3% for the DP group (SD = 8.5). A univariate ANCOVA failed to find reliable differences involving Group or Age (both $Fs(1,19) \leq 0.83$, $ps \geq .37$). One individual (DP 7) produced a mean score that was significantly below the control group mean (see Figure 2.2). Given the need to select one of six possible response buttons, RTs were not analysed.

MANCOVA Results

Four separate MANCOVAs were performed on the accuracy and reaction time data to assess whether a more general perceptual difference existed between the DP and control group, as characterised by performance across multiple tests rather than on any one test in particular. Age was controlled for as a covariate and Group was the fixed factor.

(1) MANCOVA of Accuracy data:

The following 16 dependent variables were included: the upright and inversion index scores from Experiment 3, the four conditions and face composite effect scores from Experiment 6, the three conditions from Experiment 7, the spacing and contour inversion effect scores from Experiments 7 and 8, the three conditions from Experiment 9, and the overall score from Experiment 10. Note that 'total' summary scores were not included because of their inter-dependence with the condition-specific scores from which they were derived. Due to collinearity ($r > .8$), scores in the upright condition from Experiment 3 and the intact-same condition from Experiment 6 were also removed. The MANCOVA failed to reach statistical significance, $F(14,6) = 2.77$, $p = .11$, Wilk's $\Lambda = 0.134$, partial $\eta^2 = .87$, and Age was not a significant factor, $F(14,6) = 2.5$, $p = .14$.

(2) MANCOVA of RT data:

The following 14 dependent variables were included: the overall score from Experiment 5, the four conditions and face composite effect scores from Experiment 6, the three conditions from Experiment 7, the spacing and contour inversion effect scores from Experiments 7 and 8, and the three conditions from Experiment 9. Due to collinearity, data from the intact-diff condition from Experiment 6, the spacing condition from Experiment 7, and the 45 degree condition from Experiment 9 were removed. The MANCOVA failed to reach statistical significance, F(10,

10) = 1.78, p = .19, Wilk's Λ = 0.360, partial η^2 = .64, and Age was not a significant factor, F(10,10) = 0.8, p = .67.

(3) MANCOVA of RT and Accuracy data combined:

All accuracy and RT variables described in the above MANCOVAs were included but again failed to produce a reliable effect, F(19,1) = 3.81, p = .39, Wilk's Λ = 0.014, partial η^2 = .99. Age was not a significant factor, F(19,1) = 0.50, p = .83.

(4) MANCOVA of those six measures that might be considered most sensitive to face perception deficits; the face composite effect accuracy, viewpoint total accuracy, CFPT accuracy, face detection RT, face composite effect RT, and the Jane inversion effect RT. Again, the test did not reach significance, F(6,14) = 2.55, p = .07, Wilk's Λ = 0.478, partial η^2 = .52) and Age was not a significant factor, F(6,14) = 0.81, p = .58.

General Discussion

We present data from 11 individuals with DP as defined by self-reported difficulties in face recognition and relative impairment on the famous faces test (DP group mean = 41%; control group mean = 82%), and the CFMT (DP group mean = 51%; control group mean = 82%). Group analysis revealed some evidence of perceptual impairment compared to the controls; accuracy was lower for one of the experimental conditions sensitive to global-local interference (Experiment 4), for detecting Jane Upright changes (Experiment 7) and viewpoint matching (Experiment 9), while RTs were slightly longer on subtests of the Jane Upright (Experiment 7) and composite processing tasks (Experiment 6), and there was a reduced inversion effect in one condition of the Jane task (Experiment 7 & 8). However, when set against the severe recognition failure shown by the DPs on the Famous Faces Test and the CFMT, these perceptual impairments were subtle; accuracy was always well above chance, less than 15% of

the corresponding control mean, and accompanied by normal inversion effects. Additionally, unlike other instances of prosopagnosia (e.g., Avidan et al., 2011; Wilkinson et al., 2009), the observed reaction time differences only ever differed from the control means by milliseconds rather than seconds making it difficult to see how these could explain the profound identification problems. The failure to clearly distinguish between the perceptual capacities of the DPs and controls could not be easily attributed to differences in age. Compellingly, the patterns observed at the group level were also not evident at the individual level: six of the 11 DPs showed no evidence of perceptual impairment, while the remainder showed heterogeneous impairment profiles. In line with this, the MANCOVAs failed to identify a combination of perceptual test scores that differentiated the DPs from the controls. Together, these data suggest that no single perceptual impairment, or combination of perceptual impairments, is necessary or sufficient for face recognition failure (as defined by the ADL scale, Famous Faces Test and CFPT) to occur in DP.

Examination of individual performance in Figure 2.2 shows that the most prevalent perceptual impairments were observed on the Jane Upright RT, Jane RT inversion effect, and the viewpoint matching task. For each of these tasks, two or more individuals performed significantly worse than the control group mean. The fact that so few participants encountered difficulty with the viewpoint task is noteworthy because, along with the CFPT, Mooney test, and composite inversion effect, this task is seen as particularly sensitive to face identification impairment (by contrast, the configural task does not involve faces, the emotional expression task does not involve face identification and the Jane Upright task may lack sensitivity to DP – see Yovell and Duchaine, 2006). Given that performance was so much more impaired on the memory compared to perceptual tests, I suggest it unlikely that the identification problems were perceptual in origin, and propose that many of the isolated perceptual deficits were merely co-

morbid to DP. That is, they were simply allied rather than directly causal. This alludes to the idea that DP, like its acquired counterpart, has a dissociable memorial element (Barton, 2008; Damasio, Tranel, & Damasio, 1990; De Renzi et al., 1991). If true then perhaps some of the perceptual deficits seen in DP are motivational in origin and reflect a reluctance to maximise perceptual face processing during laboratory testing. For example, if the individual knows that he/she is unable to retain facial information then the attentional focus during visual processing may move to non-defining facial attributes or other aspects of appearance such as voice, body shape or posture. Alternatively, and admittedly speculatively, it remains possible that a higher-level memory impairment compromises lower-level perceptual processing via disruptive back-propagated messages.

The non-uniform profile of perceptual impairment seen in the current sample is similar to that reported by Schmalzl et al. (2008) who administered the same battery to members of one family, of whom seven were categorised as DP. All showed a different combination of perceptual impairments but, unlike in the present sample, six of those individuals also failed to show a face inversion effect on the Jane task. Interestingly, one of the seven individuals who showed evidence of DP at screening (i.e. reported that they could not identify photographs of family members), showed no evidence of impairment on any of the perceptual tests. This led the authors to conclude that although a subtle perceptual deficit could not be ruled out, the individual's impairment "seems to lie in the ability to associate a visual percept of a face with an individual identity" (p.113). The current data suggest that the divergent pattern seen in this family member may not be uncommon.

For future research, I suggest the need for additional, larger group studies (as opposed to the single-case or small-group approaches that currently predominate) that are more amenable to

robust inferential statistical procedures and that provide a better measure of individual heterogeneity that, as was the case here, belies averaged group effects. While the sample size recruited in the present study is larger than that in most other developmental prosopagnosia studies, it is most likely still too small to speak to the behaviour of the wider population. And although the conjoint administration of the CFMT and Famous Faces Test allows one to separately probe, on one hand, the learning and short-term retention of unfamiliar faces and, on the other, the recollection of faces that have been seen many times over, neither is sensitive to other important determinants of memory ability. A related concern is that while some measures, such as the CMFT, seem able to detect individual differences in performance, others may not have such sensitivity and, accordingly, the profile of individual data reported here must be seen as more suggestive than definitive. This problem of measurement is compounded by the fact that no experimental test of DP has received diagnostic validation. Consequently, it remains unclear what percentage change in these tests constitutes a minimal clinically important difference.

CHAPTER III

Introduction III: Electroencephalographic explorations of the effects of CVS on face recognition

In the previous chapter, a sample of 11 individuals with DP completed a battery of 10 experiments. All DPs were impaired at the FFT and CFMT, which are both memory based tasks, but six showed no impairment in perceptual tasks, while the remaining five showed only minimal impairment in the perceptual tasks. The implication is that DP is more memory based then perception based. While the FFT and CFMT tests may be highly sensitive to the presence of DP, they do not provide any specificity in terms of what memory processes might be affected. The current chapter will address this shortcoming.

As recognition can be divided into perception and memory, the act of remembering can be further characterised into two domains; remembering the episodic information that occurred during encoding of the event, or merely having a looser sense of familiarity. The element that separates the two is one of conscious experience. This division of memory is based on original work by Tulving (1985) who suggested a difference between remembering, and simply knowing, which he claimed to be a difference of autonoetic consciousness (the ability to mentalise a situation). The dual-process model of memory recognition (Jacoby, 1991; Joordens & Hockley, 2000; Mandler, 1980; Reder, Nhouyvanisvong, Schunn, Ayers, Angstadt, & Hiraki, 2000; Yonelinas, 1994) develops this idea, and redefines the two processes as a feeling of familiarity and explicit recollection. While familiarity is considered to provide a continuous value about the test item, a quantitative change happens involving episodic retrieval which enables recollection. Though there is support for a single-process model of recognition memory (Dunn, 2004; McClelland & Chappell, 1998; Shiffrin & Steyvers, 1997), which more

parsimoniously denotes one continuum of strength of familiarity, there is generally more support for the dual-process model (Diana, Reder, Arndt, & Park, 2006; Yonelinas, 2002). Furthermore, double dissociations of neural correlates have been identified for familiarity and recollection processes (Vilberg & Rugg, 2007; Yonelinas, 2002), with the parahippocampal gyrus associated with familiarity but not recollection, and the hippocampus associated with recollection but not familiarity. The present chapter will therefore employ three experiments that separately examine face perception, familiarity and recollection. In the case of DP, the experiments will test if the two domains of remembering are differentially impaired, as a fault in either of these could explain the memory deficits seen in the previous chapter.

There is already some evidence to suggest that familiarity and recollection are dissociable in DP. In an early paper using two cases of acquired prosopagnosia (Tranel & Damasio, 1985), both patients showed large galvanic skin responses to familiar faces that they did not explicitly recognise, and yet they showed no such autonomic response to unfamiliar faces. The first article to offer firm evidence that implicit recognition existed in developmental (congenital) prosopagnosia used an identity priming task to show that explicitly unrecognised celebrities primed other famous faces compared to unfamiliar faces (Avidan & Behrmann, 2008). These papers demonstrated familiarity with a face without recollection. However, what these papers also demonstrate is that the familiarity is not always overt. Any experiment designed to test for familiarity in DP therefore requires measures of covert recognition, preferably in addition to ones for overt recognition. In the present chapter, this will be assessed using repetition priming which can provide both overt and covert measures of familiarity.

Electroencephalography

To gain a neuro-physiological understanding of the impairments in the DP sample, a further stream of data will be collected from electroencephalography (EEG). This high

temporal resolution technique allows patterns of brain activity can be time-locked to stimuli presentation, allowing event related potentials (ERPs) to be examined across experimental conditions and participant groups. This technique has also previously been used in DP to differentiate between familiarity and recollection (Burns, Tree, & Weidemann, 2014) as well as detecting covert responses to familiarity (Eimer, Gosling, & Duchaine, 2012).

One of the most researched perceptual ERPs in the face literature is the N170, a negative voltage deflection occurring 170 ms after stimulus onset, with well over 200 published articles measuring this component (see Eimer, 2011). The N170 is significantly more negative in response to faces and is the first face-specific electrophysiological response. This early face perception process is thought to reflect structural encoding of faces (Eimer & McCarthy, 1999).

Several ERP studies have investigated the N170 in people with DP, but often there is large variability between patients whereby some show a reliable N170, others show it reduced to that seen with non-face stimuli (Kress & Daum, 2003), and one other (Towler, Gosling, Duchaine & Eimer, 2012) shows a typical N170 to upright faces but remained the same amplitude with inverted faces (known to disrupt configural processing [Tanaka & Farah, 1993]) which is classically enlarged in adults. The impairment in the present sample of DPs appears to be memorial in nature, with limited perceptual impairment. It is therefore predicted that the N170, a response to facial stimuli with no memory element required, will not be different to the control group's N170, and thereby bolster the conclusions of Chapter Two with corroborative, physiological evidence.

Whereas the N170 is present regardless of familiarity, the first component sensitive to identity is the N250 (a negative amplitude deflection 250 ms post stimulus onset; Tanaka, Curran, Porterfield & Collins, 2006). This component has also been observed as a response to immediate face repetition (e.g. Schweinberger, Pfütze & Sommer, 1995), termed the N250r. The N250r is thought to reflect access to stored facial representations (Bindemann, Burton,

Leuthold & Schweinberger, 2008). The repetition effect is larger for familiar face repetitions and has been reported as twice the size when the probe stimulus is of the same image as the prime, compared to when it is simply of the same identity (Cooper, Harvey, Lavidor, & Schweinberger, 2007). The N250r component has not been investigated in prosopagnosia, but recent investigations with the N250 have lead to some interesting hypotheses. Eimer, Gosling and Duchaine (2012) looked at the N170, N250 and P600f (a late ERP reflecting explicit recognition) in 12 patients with DP. In responses to faces, half of the group showed an N250 comparable to controls, but no P600f, on 'face unknown' trials. If the N250 reflects access to stored face representations, then the data suggests that these patients have covert recognition of the stimuli that does not become explicit. This could also reflect a sub-category of prosopagnosia. Should the present DP group provide a high number of incorrect trials, the ERP data will provide a measure of whether familiarity is absent or merely covert.

Finally, the N400 (a negative amplitude peaking near 400 ms post stimulus onset) is a late ERP associated with familiar faces. Initially identified as a marker for priming in word recognition (Kutas & Federmeier, 2000), it has been found to be modulated by familiarity with faces, and indeed can be used as an index of familiarity with a specific face (Eimer, 2000). Originating from central-parietal regions, Bentin and Deouell (2000) noticed significant negativity in later epochs of the ERP waveforms only when participants had viewed familiar faces. They concluded that this N400 reflected semantic activity involved in the identification of familiar faces, and more specifically, the activation of person identity nodes (PINs; Bruce & Young, 1986). The N400 is also present when people are recognised by their names not their faces (Schweinberger, 1996) which suggests that the N400 reflects post-perceptual processing stages of face recognition. The ERP has also been investigated in priming; Barrett and Rugg (1989) found the N400 to be larger when the probe face was of the same occupational category (and so similar semantic information) as the prime face. Only one study has

specifically investigated the presence of an N400 in a patient with AP (Eimer, 2000) and they found, similar to the patients N170, that the N400 was not present. As the N400 reflects recollection of specific details after seeing a face, a relatively late stage in memory retrieval, it is predicted that the present DP group will show a significant difference in the N400 compared to the control group.

In one recent study using a remember/know paradigm following training with unfamiliar faces (Burns, Tree, & Weidemann, 2014), DPs showed a typical EEG response (measured in the ERP window 300-500ms) to familiar faces that they could not recollect, but showed an atypical EEG response to faces they could recollect. The data suggest that the DPs show no impairments in tasks assessing familiarity, and that the impairments instead lie in processes associated with recollection. However, the authors discuss the possibility that the chosen time window (300-500ms) may reflect familiarity effects (based on MacKenzie & Donaldson, 2007), thereby suggesting that recollection impairments maybe a consequence of earlier (or parallel) familiarity impairments. The N400 used in the present study to index recollection is most similar to the 300-500ms window measurement used by Burns and colleagues. However, to delineate recollection and familiarity, the present study is adopting the view that the N250r is the earliest point at which familiarity is indexed and will be dissociated from the electrophysiology involved with recollection.

Chapter Two showed consistent impairment in both the CFMT and FFT tasks, both of which require just familiarity and not necessarily recollection to successfully perform the task. This leads me to predict impairment in both the N250r and N400 tasks. If DP can indeed exist in the absence of perceptual impairment, then no or smaller differences are predicted in the N170 task as this indexes face perception. Should this be the case then this will further clarify, with excellent temporal resolution, the nature of impairment in this DP sample.

Galvanic and Caloric Vestibular Stimulation

Despite the high prevalence of prosopagnosia, and the negative social impact of having the condition, surprisingly few attempts have been made to develop a therapy to increase face recognition abilities. Even fewer have been successful. As discussed in Chapter One, anatomical, clinical, normative and ERP data all indicate the possible modulatory effects of vestibular stimulation on face recognition. In the present chapter, the effects of vestibular stimulation will be tested in controls and participants with DP, in the N170, N250r, and N400 tasks.

Two studies that investigated the use of GVS with face recognition both found beneficial effects, one with shorter reaction times in healthy adults (Wilkinson, Nicholls, Pattenden, Kilduff & Milberg, 2008), and the other with higher accuracy in face matching in one case of acquired prosopagnosia (Wilkinson, Ko, Kilduff, McGlinchey & Milberg, 2005). Wilkinson and colleagues (Wilkinson, Ferguson & Worley, 2012) also found that (in a task with no meaningful behavioural measure) stimulation significantly increased the amplitude of the N170 in a group of 16 healthy students.

The effects of vestibular stimulation on face recognition described above are all based on GVS, the delivery of small electrical currents over the mastoid processes. One significant disadvantage with this method is the strong artifact that is introduced to the EEG signal when the two are used simultaneously. By contrast, the temperature changes induced to the inner ear canal during CVS cause no such artifact. There are, however, several other reasons to choose CVS over GVS in this study. First, mechanistic differences between CVS and GVS drive similar but not identical patterns of activation in the cortex, and it is CVS that causes more widespread activation (Lopez, Blanke, & Mast, 2012). There is also greater evidence of a link between CVS and memory than GVS and memory (Bächtold, Baumann, Sándor, Kritos,

Regard & Brugger, 2001). Third, my access to a CVS device and the accompanying support from the device suppliers, Scion, means that the study will be completed more efficiently.

To date, the effects of CVS on face recognition in DP have not been tested, and while the effects of GVS are encouraging, they may not be replicated using CVS. If the effects of CVS on the N170 are similar to those of GVS, then an enhanced amplitude in both groups will emerge. Although there is only one paper to support this prediction, there are none to support predictions for the memory based N250r and N400 tasks. That said, those brain regions associated with visual memory and memory retrieval such as the hippocampus and temporal regions receive increased blood flow during GVS (Lobel, Kleine, Bihan, Leroy-Willig & Berthoz, 1998; Vitte, Derosier, Caitu, Berthoz, Hasboun, & Soulié, 1996) and it may be that this increase in blood flow enhances the processing capabilities of these regions, resulting in improved behavioural performance. This reasoning would stand for both the DP and control group. The link between increased blood flow and scalp ERP recording is very weak, but one might predict that with increased cortical activity, an enhanced negativity of the N250r and N400 might follow.

General Method

Participants

All participants reported normal or corrected-to-normal vision and no history of brain trauma. Informed consent was obtained from all participants prior to the onset of the experiment. Due to own race bias effects in face recognition (see Brigham & Malpass, 1985), all participants were required to be Caucasian, brought up in the UK, and have lived in the UK for the last three years. Ethical approval for this study was granted by the School of Psychology

Ethics Board at the University of Kent. Participants were financially compensated for their time. Participant names have been replaced with numbers.

Controls

Twenty neuro-typical adults (seven male, one left-handed) aged 18-66 years old (mean = 49.45, SD = 14.57), independent from the previous experiments, volunteered to take part in the experiments. Recruitment was through the Kent Adult Research Unit (KARU) database of individuals who are interested in volunteering for scientific research. All controls completed a famous faces task and the CFMT (Duchaine & Nakayama, 2006) to assess their face processing skills. No scores on these two tests reached the threshold for a classification of impaired, as defined in Chapter Two.

Prosopagnosics

Of the 11 DP individuals deemed eligible for enrolment in the previous chapter (based on impaired activities of daily living, face recognition problems evident since birth, absence of known trauma or neurological events, absence of developmental or psychiatric disorders, and impairment on both the CFMT and FFT tasks), eight (three males, one left-handed) were willing and able to return for further assessment (DP5, DP9, and DP11 did not return). This group were aged 27-61 years old (mean = 49.63, SD = 13.45) which was not significantly different to the control group, $t(26) = 0.03$, $p = .98$. No DPs reported any subjective changes in their face recognition abilities and none had received any form of training or therapy since completing the test battery.

Apparatus and Procedure

Participants were contacted via email to arrange testing times and were given information about the study prior to their arrival. All experiments were conducted in a noise and temperature controlled laboratory in the School of Psychology at the University of Kent. The testing session outline was verbally described to the participant before they were asked to read the information sheet and sign a consent form (appendix A). The participants head was measured to inform of the correct EEG cap size and friendly conversation was encouraged during the application of the electrodes (details below). Once the EEG was set up, participants were shown the real-time output of their EEG signal and made aware of ocular and muscular artifacts, as well as alpha activity.

The participant's chair was then reclined backwards such that their head was 30° from horizontal. This moves the horizontal semicircular canal into a truly horizontal position and maximises the effects of stimulation (Black, Rogers, Ade, Nicoletto, Adkins & Laskowitz, in press; Saladin, 2012). This position was also retained through the no stimulation conditions. A 17 inch flat-screen computer monitor was positioned 120 cm away from their eyes and positioned such that it could be seen when the participant looked straight ahead. Screen resolution was set to 1024 pixels × 768 pixels. Images were tightly framed in a canvas 170 pixels wide × 230 pixels tall. This resulted in images approximately 76 mm tall which formed a viewing angle of 3.63°. To reduce glare from overhead lighting, the laboratory was illuminated from a desk lamp behind the monitor. Two closed-circuit ceiling cameras were located in the room, one directed towards the participant and one directed toward the CVS control unit, allowing observations from the control room and for the participant to be alone during testing. Experiments were programmed in E-Prime 2 (Psychology Software Tools) and responses were collected using a keyboard positioned on the participant's lap.

The three tasks (odd-ball eliciting the N170, familiarity judgements eliciting the N250r, and nationality judgements eliciting the N400) were conducted in sequence with a break of 3 minutes between each one. The order of the experiments was counterbalanced across each group of participants. All participants completed each experiment twice, once with stimulation and once without. So not to exceed the recommended dose of 30 minutes of stimulation as set by the device manufacturers, Scion, stimulation was administered in either Experiment 2, or Experiments 1 and 3. During the second testing session for each participant, experiments previously in the stimulation condition were now in the no-stim condition, and vice-versa. To prevent any short-term carry-over effects of CVS, experiments in the stimulation condition were always conducted in the second half of each session. Furthermore, the second testing session was held a minimum of 7 days after the first session so that any medium-term carry-over effects could dissipate. The order of experiments and whether stimulation was present was also counterbalanced across each group of participants, though the order of Experiments 1 and 3 was kept consistent between first and second testing sessions. Figure 3.1 shows this design graphically. As an example protocol, one participant might experience Experiment 1 and 3 without stimulation followed by Experiment 2 with stimulation, then a week later they would experience Experiment 2 without stimulation followed by Experiments 1 and 3 with stimulation.
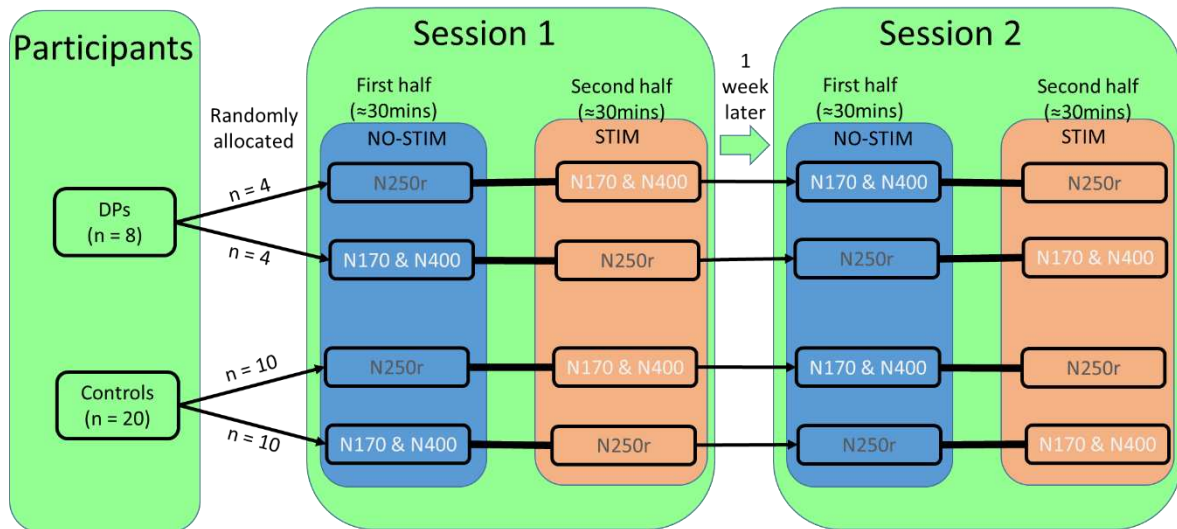
Figure 3.1: Illustration of the experimental design, showing the order in which the experiments were conducted and when stimulation was administered. Note that the order of N170 and N400 was counterbalanced.

Electrophysiological recording and preparation

Continuous EEG was recorded using a 64 channel Acticap system and Brain Vision recording software (Brain Products GmbH, Germany), The active electrodes were configured on the scalp according to the International 10-20 System (Jasper, 1958) using midline electrodes Fz, Cz, CPz, Pz, POz, and Oz, left hemisphere electrodes Fp1, Af3, Af7, F1, F3, F5, F7, FC1, FC3, FC5, FT7, C1, C3, C5, T7, CP1, CP3, CP5, TP7, P1, P3, P5, P7, PO3, PO7, PO9, and O1, and the homologous electrodes for the right hemisphere. Site FCz was used as on-line reference and site AFz was used for ground. Electrooculography (EOG) was recorded from below the participant's right eye (vertical EOG) and from the lateral canthus of the participant's left eye (horizontal EOG). SuperVisc-Gel® (Brain Products GmbH, Germany) was syringed under each electrode for signal transduction. The analogue signal from all electrodes was sampled at 500 Hz and impedances were kept below 5 kΩ for each channel.

Off-line analysis was performed using BrainVision Analyser 2 software (BrainProducts GmbH, Germany). The VEOG channel was re-referenced to the Fp2 channel, and all EEG

channels were recalculated to an average reference. Any channels showing zero signal (due to hardware failure) were excluded from analysis on a participant by participant basis before the average re-reference. All EEG channels were band-pass filtered with a low cut-off set to 0.1 Hz and a high cut-off set to 40Hz, both with a roll off of 12 dB/oct.

Independent component analysis (ICA) was performed separately for each participant and each experiment using a restricted infomax algorithm, before horizontal and vertical eye-movement components were manually identified and removed. Artifact rejection was performed semi-automatically with maximum voltage step thresholds set to 50 µV/ms, maximum difference thresholds set to 200 µV per 200ms, and maximum amplitude thresholds of ±200 µV. Remaining trials were then visually inspected for any remaining artifacts. Trials with visibly discernable alpha activity were not rejected as these fluctuations were not time locked and therefore add only minimal noise to the average waveform compared to removal (Luck, 2005).

Caloric vestibular stimulation

Stimulation was delivered using the Generation 2.5 CVS device developed by Scion NeuroStim. The device consists of an aluminium earpiece attached to a Peltier array (to exchange the heat from the external ear canal to a water cooled heat sink) contained on the left side of a headset (see Figure 3.2). Temperature changes to the earpiece are powered and controlled with the attached control unit.
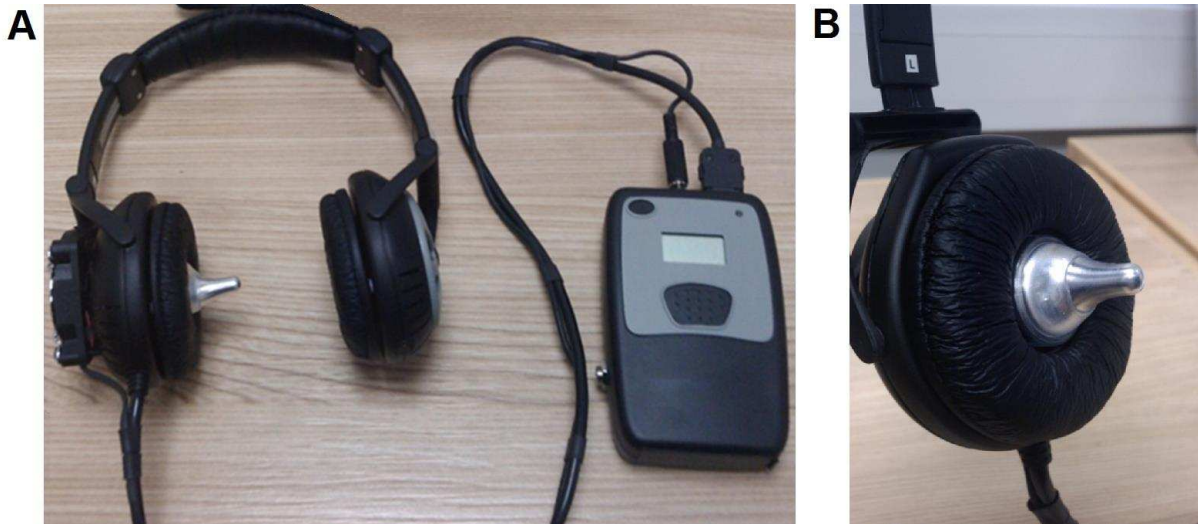
Figure 3.2: (A) The Gen 2.5 CVS device headset and control unit. (B) Closer view of the aluminium earpiece.

In all experiments the protocol was for the temperature of the earpiece to follow a saw-tooth waveform with a frequency of one cycle per three minutes, cooling to 17 degrees Celsius and restarting the cycle at 35 degrees Celsius. The cycling is required to maintain convection currents in the semicircular canal and prevent habituation, and also to help reduce unpleasant side effects associated with the rapid introduction of cold temperatures into the external ear canal. After each dose of CVS, a log of activity was retrieved confirming that the desired temperature changes were reached. In the no-stimulation condition, the headset was not worn.

Planned comparisons

As described in Chapter One, the experiments of the present chapter are to be replicated in Chapter Four to assess test-retest reliability. As is conventional in neuropsychological research, the same baseline control group data will be used for comparison in both chapters. To prevent inflated type I error that results from the multiple comparisons that follow from comparing a different dataset to a common dataset (i.e. the control group), a correction was made by dividing the alpha region by two (significant $p < .025$) for most of the contrasts (not

including post-hoc pairwise comparisons). Mindful of the need to catch effects that just missed such strict significance, but which were strongly predicted a priori, an alpha of .05 was retained for a small subset of contrasts. While there exists a self-fulfilling argument that the effects we predict have a higher chance of being significant because of this adjustment, equally there is a risk that real group differences will be missed. Furthermore, the replication chapter will provide additional support that any differences discovered in this chapter are genuine and not due to random error.

EXPERIMENT 1

Does the N170 differ between DPs and controls and is it modifiable with CVS?

The N170 was induced via an oddball paradigm similar to the one used in some of the original N170 experiments (Bentin, Allison, Puce, Perez & McCarthy, 1996) which yielded a robust response.

Multiple studies have investigated the N170 in DP, but with mixed conclusions; some studies find a reduced N170 compared to controls (Kress & Daum, 2003) whereas other studies find no difference (Towler et al., 2012). The present experiment will further explore these findings with a large group of individuals with DP. As our sample showed memorial based impairments, it is hypothesised that the N170 of the DP sample will not be different from the N170 of the control group. As outlined above it is also hypothesised that the CVS stimulation will enlarge the N170 in both the DPs and controls.

Stimuli

Four different categories of stimuli were used: Flowers, butterflies, houses and unfamiliar faces. Face stimuli were balanced for gender and images were trimmed to include the head only, removing the neck but including hair. House stimuli were front view images of real-life houses. Butterfly stimuli were top-down images. All images were monochrome, tightly framed within a 170 pixels wide × 230 pixels tall canvas, centred on a white background and with surrounding details removed (see Figure 3.3). Twenty stimuli for each category were gathered. Face and house stimuli were duplicated and inverted, thus creating six categories and 120 stimuli in total.
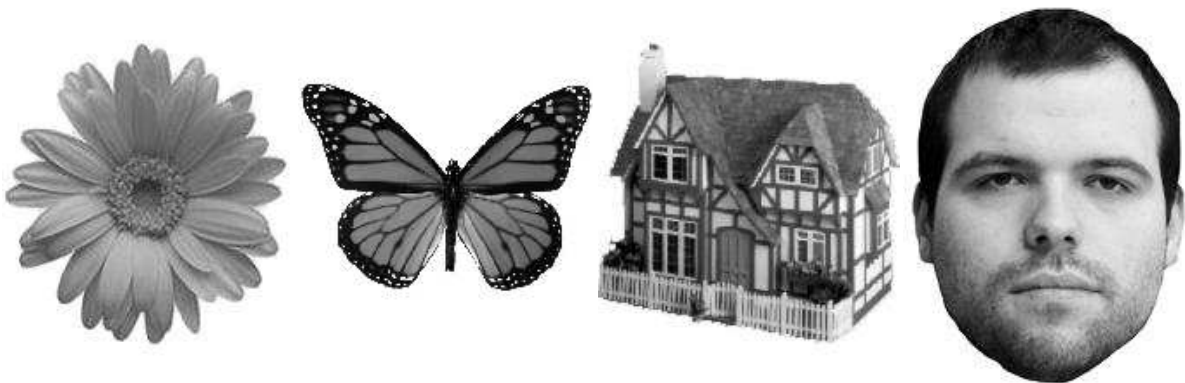
Figure 3.3: Example images from each stimuli category in Experiment 1, Chapter Three.

Design and Procedure

The procedure was identical to that reported in the General Method with the following exceptions. Participants were first shown an instruction screen which asked them to watch the presentation of images and to press the space bar whenever they saw a butterfly (the oddball). Each stimulus was presented in a random order for 250 ms, followed by a blank white screen for 1250 ms. After all 120 stimuli were shown, there was a break, the length of which was

defined by the participant but did not exceed 30 seconds. After this break the same block of stimuli was presented again, in random order. In total there were three blocks containing 360 trials, separated by two break periods.

This entire procedure was repeated on separate days, once with CVS and once without, the order of which was counterbalanced. In the stimulation condition, the CVS device protocol was initiated 30 seconds prior to the onset of the first stimulus and remained on during between-block breaks.

The target stimuli were set to butterfly to keep EEG responses to the stimuli of interest free from artifacts such as motor planning and muscle movement. The task of identifying butterflies was intended for engagement purposes only so the behavioural measures were irrelevant to our hypotheses.

The dependent variable of interest was the EEG record; specifically the amplitude and latency of the N170 ERP. The independent variables were object (face and house), inversion (upright and inverted), sample group (prosopagnosic and control), and CVS (with and without stimulation).

ERP analysis

EEG preparation was as described in the general method. Specific to this experiment, epochs lasting 1100 ms and starting 100 ms before the onset of each stimulus were generated off-line from the continuous EEG recording. These ERP waveforms were then corrected to a 100 ms baseline before being averaged separately for each condition. In total, 24 ERP waveforms were available for each electrode: One for each of the six stimulus categories, in both prosopagnosia and control groups, for both stimulation and no-stimulation conditions.

N170 amplitude and latency were defined as the maximal negative voltage within the parameters of 120 ms and 220 ms post-stimulus onset. These peaks were identified and

exported using BrainVision Analyser 2 software. Only peak data from electrodes sites at which the N170 was most prominent both in previous research (Guillaume et al., 2009; Morgan, Klein, Boehm, Shapiro, & Linden, 2008; Bauser, Schriewer, & Suchan, 2015; Zheng, Mondloch, & Segalowitz, 2012) and in the present study were used for analysis (P7, P8, PO7, PO8, PO9 and PO10). The three channels from the left hemisphere were pooled together as one, and the calculation repeated for the right hemisphere.

Results (1)

Electrophysiological data

N170

Latency (ms) and amplitude (μV) of N170 peaks were analysed using separate mixed measures ANOVAs using the design 2(Group: DP vs. control) x 2(CVS: with vs. without) x 2(Object: face vs. house) x 2(Inversion: upright vs. inverted) x 2(Hemisphere: left vs. right). Interactions were explored using Bonferroni corrected pairwise comparisons.

i) Latency

There was a main effect of Inversion, $F(1, 26) = 70.87$, $p < .001$, which was qualified by an Inversion $\times$ Object interaction, $F(1, 26) = 40.41$, $p < .001$, whereby an inversion effect was seen in Faces (upright mean = 169.89, SD = 14.88; inverted mean = 179.96, SD = 14.33), $p < .001$, but not in Houses (upright mean = 173.10, SD = 17.21; inverted mean = 173.75, SD = 16.48), $p = .43$.

As predicted, there was no main effect of Group (control mean = 173.08, SD = 16.08; DP mean = 176.90, SD = 16.01), $F(1, 26) < 1$, $p = .49$, and no Group $\times$ Object interaction, $F(1, 26) = 1.98$, $p = .17$.

There was no main effect of CVS (no-stim mean = 173.64, SD = 15.52; stim mean = 174.71, SD = 16.75), $F(1, 26) < 1$, $p = .34$. The only significant interaction involving CVS was a CVS $\times$ Hemisphere $\times$ Object $\times$ Group four way interaction, $F(1, 26) = 14.29$, $p < .001$. The means for the individual conditions involved in this interaction are plotted in Figure 3.4. Breaking down the interaction by the CVS variable, the key manipulation, the Hemisphere $\times$ Object $\times$ Group interaction was only significant in the stimulation condition, $F(1, 26) = 9.81$, $p < .004$, and not during the no-stimulation condition, $F(1, 26) = .03$ $p = .87$. Exploring the two way interactions within this, and using Bonferroni adjustment, an Object $\times$ Group interaction was found only within the right hemisphere, $F(1, 26) = 10.92$, $p < .005$, (driven by a higher latency for faces [mean = 176.48, SE = 3.17] compared to houses [mean 169.73, SE = 4.27] in the control group, $p < .005$), and a Object $\times$ Hemisphere interaction found only in the control group, $F(1, 19) = 10.69$, $p < .005$ (which was driven by higher latency for faces [mean = 176.48, SE = 3.19] compared to houses [mean = 169.73, SE = 4.48] only in the right hemisphere, $p < .01$). In summary, stimulation induced a longer latency for faces compared to houses in the right hemisphere of the control group.

All other main effects and interactions were not significant, all $Fs(1, 26) \leq 3.87$, $p \geq .06$.
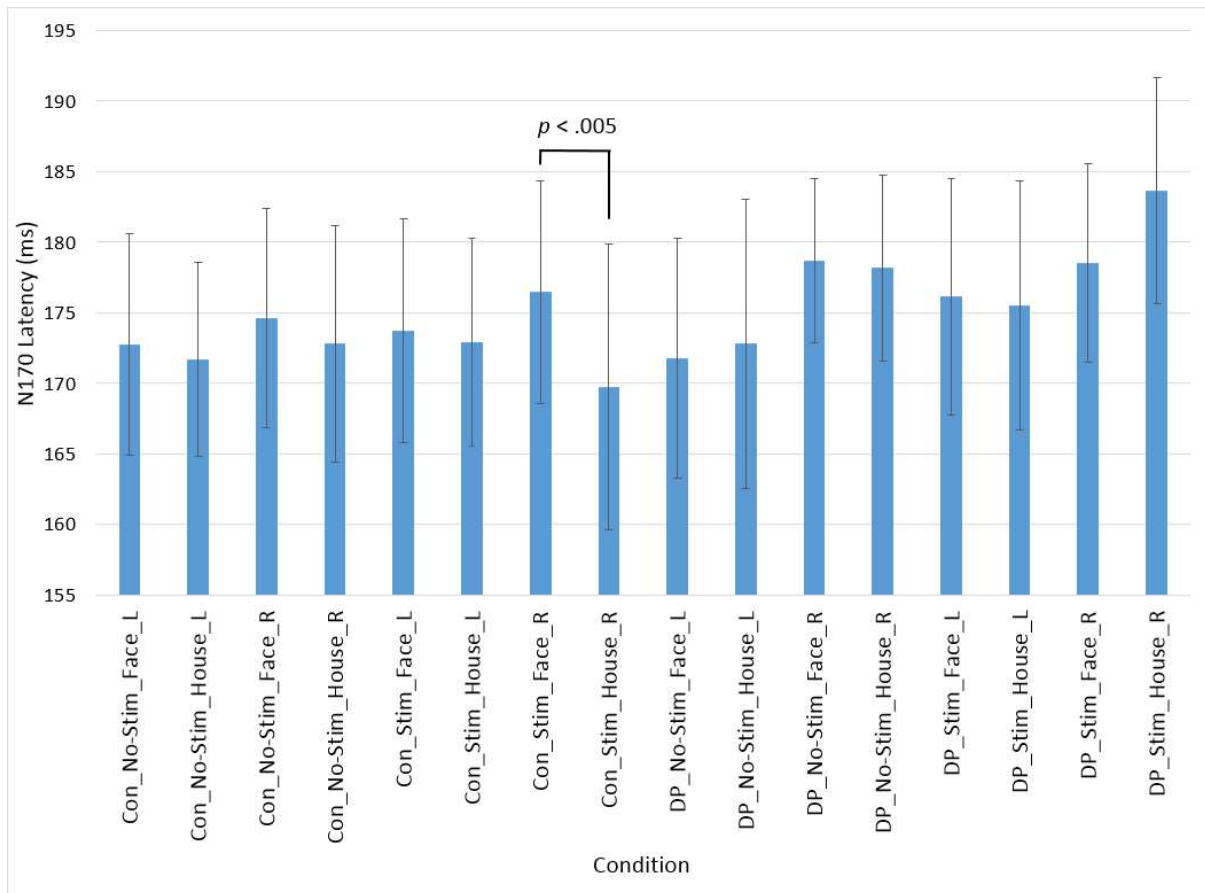
Figure 3.4: Means with standard error bars of Group × CVS × Object × Hemisphere interaction in N170 latency.

ii) Amplitude

There was a main effect of Object (face mean = -5.48, SD = 4.30; house mean = -4.01, SD = 3.90), $F(1, 26) = 7.12$, $p < .05$ but the main effect of Inversion was not significant using the adjusted p-value of .025, $F(1, 26) = 4.31$, $p = .048$. The Inversion × Object interaction term was not significant, $F(1, 26) = 4.11$, $p = .053$.

As predicted, there was no main effect of Group (control mean = -4.85, SD = 4.39; DP mean = -4.48, SD = 3.55), $F(1, 26) < 1$, $p = .80$, and no Group × Object interaction, $F(1, 26) < 1$, $p = .51$.

There was no main effect of CVS (no-stim mean = -4.99, SD = 4.27; stim mean = -4.49, SD = 4.05), $F_{(1, 26)} = 2.28$, $p = .14$. All other main effects and interactions were not significant, all $Fs_{(1, 26)} \leq 3.85$, $p \geq .06$.
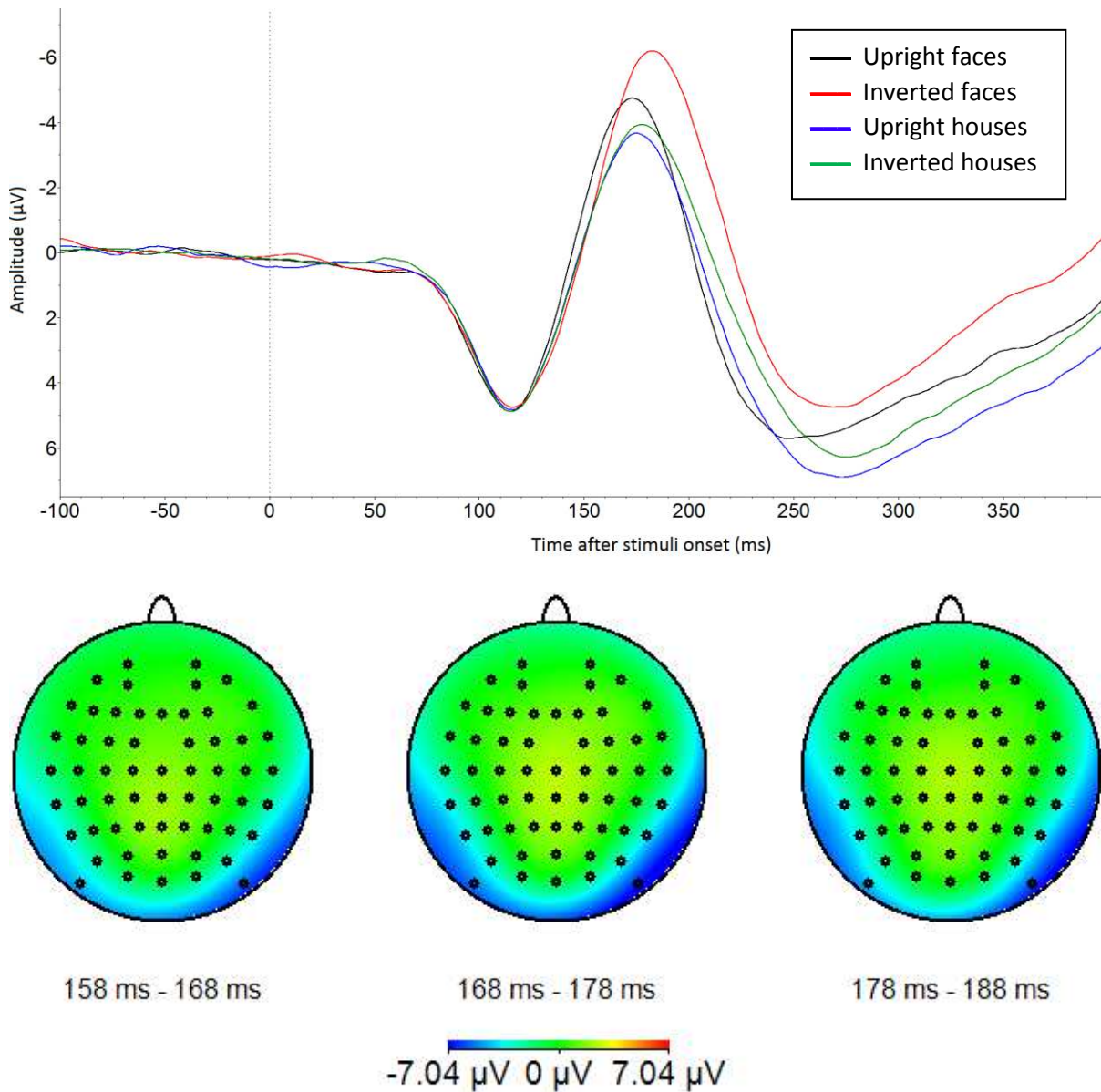
Figure 3.5: Pooled right hemisphere N170 peaks for faces (black and red) and houses (blue and green) in upright and inverted orientations, collapsed across Group and CVS variables. Topographical plots show the location of the N170 activity, and surrounding time-frames, for the upright face condition.
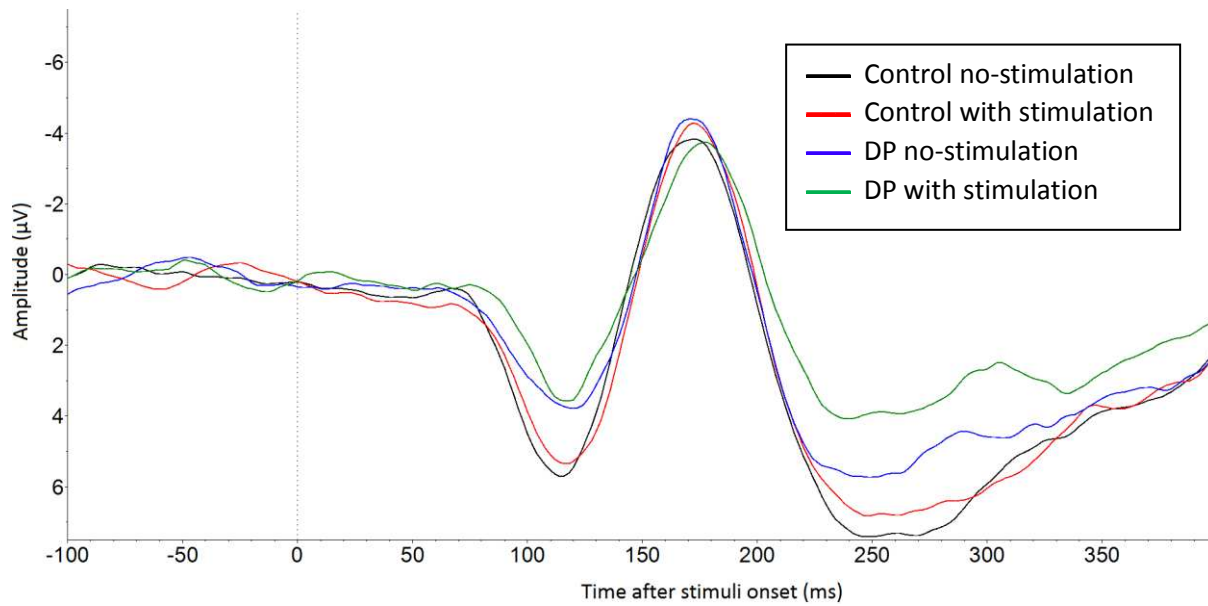
Figure 3.6: Pooled right hemisphere N170 peaks for upright faces in prosopagnosia (blue and green) and control participants (black and red) during no-stimulation and stimulation.

Discussion (1)

Figures 3.5 and 3.6 show a prominent negative peak and imply that the experiment successfully elicited the N170. Consistent with previous literature, face stimuli elicited a larger N170 than non-face stimuli. The data showed typical inversion effects with later N170s for inverted face stimuli, but this was not coupled with the typical increase in N170 amplitude (an effect that did not withstand family wise error adjustments related to the planned replication in Chapter IV). These effects occurred in both controls and DPs. It should be noted that the data offered one more unusual finding. The data lacked a main effect of hemispheric laterality, another established feature of the N170.

The N170 latency and amplitude are in the direction of being later and larger for the DP group, but variance around these values is considerably high. There was also no evidence to suggest the inversion effect varied between the groups.

In contrast with previous GVS literature, we failed to find evidence that CVS has an overall modulatory effect on the amplitude or latency of the N170, for either controls or DPs. However, the likely source of the four way interaction suggests that CVS has a quite specific effect in the control group, modulating the N170 latency differently for faces compared to houses. We will return to this complicated interaction in the next chapter if replicated.

N170 aside, Figure 3.6 shows a visibly discernible group difference in the amplitude of the P1 peak. The P1, a marker of visual perception, is now generally considered to be not face sensitive (Rossion & Caharel, 2011) so it is not clear why participants with face specific deficits would exhibit differences at this level. However, if group differences are detected in the ERPs measured in the remaining experiments, quantitative changes in P1 should be assessed as a possible predecessor.

EXPERIMENT 2

Does the N250r differ between DPs and controls and is it modifiable with CVS?

The design of the present experiment was inspired by the paradigm of Pickering and Schweinberger (2003) but uses faces instead of names for both the prime and probe stimuli. The task is designed to elicit the N250r, a marker of familiarity, but not explicit knowledge, following presentation of a face.

This experiment provides the first investigation into the N250r of individuals with DP. Similar to the heterogeneity in N170 differences in DPs, previous studies have found that a similar ERP, the N250 (which differs from the N250r in that it reflects access to long-term visual memory instead of a visual representation in working memory) is present in some, but not all DPs. As the key impairment of the present DP group appears to be memorial, it is possible that an N250r component will be absent. The presence of the N250 also correlates

with performance in a famous faces task (Rivolta, Palermo, Schmalzl, & Coltheart, 2012). If the findings from N250 studies can be applied to the N250r, then impairment in the FFT observed in Chapter Two predicts an atypical N250r in the present task.

Contrary to the N170 task above, the current task also produces a behavioural measure. This measure is an explicit judgement of whether the celebrity face is familiar. Both the N250r task and the CFMT tasks can be completed successfully if feelings of familiarity are intact. As all DPs showed impairment in the CFMT task, it is predicted that the DPs will perform significantly worse than the control group in the N250r task. However, while overall scores maybe impaired, an effect of priming (which is independent from absolute scores) may still be present.

The effects of vestibular stimulation on face sensitive ERPs beyond the N170 have never been tested. However, all face sensitive ERPs beyond 200ms are thought to reflect memorial processes and, as discussed, there is evidence of beneficial effects of CVS on memory (Bächtold et al., 2001). Studies have also shown that damage to the vestibular system causes impairments in (spatial) memory and hippocampal (a region typically associated with memory storage and consolidation) atrophy (Besnard et al., 2012; Brandt et al., 2005). Additionally, CVS has been shown to affect blood flow to the temporal cortex (Takeda, 1996), an area associated with maximal N250r peaks (Pickering & Schweinberger, 2003). This leads to the prediction that changes induced by CVS may interfere with memory for faces.

Method (2)

Stimuli

Stimuli consisted of 60 unfamiliar faces and 60 familiar faces. For each familiar face there were two different images, resulting in 180 stimuli in total. Familiar faces were of UK

or USA celebrities. Face images were at, or near, front-on view and were of straight or smiling faces. Stimuli were balanced for gender and images were trimmed to include the head only, removing the neck but including hair. All images were greyscale (to minimise confounding effects of low-level characteristics), tightly framed within a 170 pixels wide × 230 pixels tall canvas, centred on a black background and with surrounding details removed (see Figure 3.7).
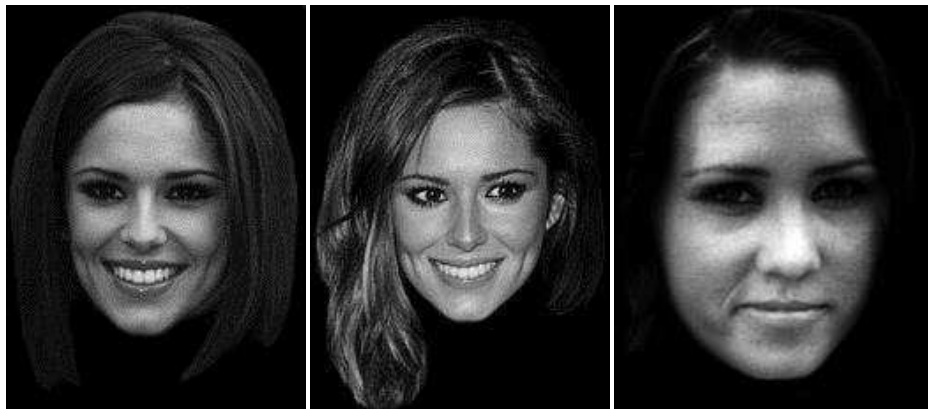
Figure 3.7: Example images from each stimuli category in experiment 2, Chapter Three.

Design and Procedure

The procedure was as reported in the General Method with the following exceptions. Participants were first shown an instruction screen which instructed them to decide whether the second face in each pair is famous or not. Responses were indicated by pressing the 'z' key for famous and the 'm' key for not famous. For each trial, a white fixation cross was presented in the centre of the screen for 500 ms before the prime face was shown for 500 ms. A red fixation cross was then presented for 1300 ms before the target face was shown for 1500 ms. Responses were only accepted during the 1500 ms of the target face presentation. The trial concluded with a blank black screen displayed for 2500 ms. All stimuli were presented in the centre of the screen.

There were four priming conditions: SameImage was when the target stimulus was an identical image of the familiar face in the prime stimuli; SameID was when the target stimulus was of a different image of the familiar face in the prime stimuli; DiffID was when the target stimulus was of a different familiar face to the familiar face in the prime stimulus; and NonFamous was when the target stimulus was an unfamiliar face and the prime stimulus was of a familiar face. The same set of 60 familiar faces was used as the prime stimuli for each condition. In the DiffID and NonFamous conditions, the choice of target image was randomly selected from the corresponding category (i.e., familiar or unfamiliar). The pairs of stimuli were presented in a random order forming 240 trials separated by three breaks. Each of the 60 famous faces was shown only once for each priming condition, therefore not all possible target-prime combinations for DiffID and Nonfamous were shown.

This entire procedure was repeated on two separate days, once with CVS and once without. Due to the length of the experiment, the CVS device needed to be reset (including replacing the ice-cold water in the cooling baffle) during each break so that the maximum cycle of 10 minutes was not exceeded. The CVS device protocol was initiated 30 seconds prior to the start of each block.

The dependent variables were accuracy, reaction time, and the EEG record, specifically the mean amplitude of the N250r ERP. The independent variables were priming condition (same image, same ID, different ID and nonfamous), sample group (prosopagnosic and control), and CVS (with and without stimulation). Additionally, the ERP analyses included a hemisphere independent variable (left and right).

ERP analysis

EEG preparation was as described in the General Method. Specific to this experiment, epochs lasting 1700 ms and starting 200 ms before the onset of each target stimulus were

generated off-line from the continuous EEG recording (note that the prime triggers were not used). These ERP waveforms were then corrected to a 200 ms baseline before being averaged separately for each condition. In total, 16 ERP waveforms were available for each electrode; one for each of the four stimulus categories, in both sample groups, for both stimulation and no-stimulation conditions.

N250r mean amplitude was defined as the average voltage within the parameters of 250 ms and 400 ms post-stimulus onset. The data for correct responses only were exported using BrainVision Analyser 2 software. Only mean amplitude data from electrodes sites at which the N250r was most prominent both in previous research (Caharel, Collet, & Rossion, 2015; Faerber, Kaufmann & Schweinberger, 2015; Kaufmann, Schulz, & Schweinberger, 2013; Neumann & Schweinberger, 2009; Neumann, Mohamed, & Schweinberger, 2011) and in the present study were used for analysis (TP7, TP8, P7, P8, PO7, PO8, PO9 and PO10). The four channels from the left hemisphere were pooled together as one, and the calculation repeated for the right hemisphere.

Results (2)

Behavioural data

Accuracy scores were coded as 0 for incorrect and 1 for correct. Averages and standard deviations were then calculated from trial reaction times individually for each participant. Participant's reaction times for incorrect trials, and also those beyond 2.5 standard deviations of the participant's overall mean, were removed. Analysis was conducted using SPSS statistical software. Effect sizes are given in terms of partial eta squared (η_p^2).

Mean accuracy and mean correct reaction times were interrogated using separate 4 (Prime: SameImage vs. SameID vs. DiffID vs. NonFamous) × 2 (Group: DP vs. control) × 2

(CVS: with vs. without) mixed measures, GLMs. Greenhouse-Geisser correction was used when assumed sphericity was violated. Interactions were explored using Bonferroni corrected pairwise comparisons.

Accuracy

The task elicited a main effect of Prime, $F(3, 78) = 15.10$, MSe $= 0.38$, $p < .001$, $\eta_p^2 = .37$, whereby the highest accuracy was achieved in the NonFamous condition (mean $= 0.94$, SD $= 0.14$) which was significantly higher ($p < .05$) than the SameImage condition (mean $= 0.80$, SD $= 0.20$), which in turn was significantly higher than both the SameID (mean $= 0.77$, SD $= 0.22$; $p < .005$) and the DiffID conditions (mean $= 0.77$, SD $= 0.21$; $p < .005$) which were not significantly different from each other ($p = 1.00$).

There was a significant main effect of Group, $F(1,26) = 8.65$, MSe $= 1.29$, $p < .01$, $\eta_p^2 = .25$, whereby the prosopagnosic group (mean $= 0.70$, SD $= 0.22$) were significantly less accurate than the control group (mean $= 0.87$, SD $= 0.18$). All other interactions with Group were not significant, all Fs ≤ 1.28, ps $\geq .29$.

There was no main effect of CVS, $F(1,26) = 2.61$, MSe $= 0.04$, $p = .12$, $\eta_p^2 = .09$, but there was a significant Prime $\times$ CVS interaction, $F(1.69,43.92) = 6.20$, MSe $= 0.03$, $p < .01$, $\eta_p^2 = .19$. Pairwise comparisons did not show any differences between stim and no-stim for the four conditions, and the interaction (see Figure 3.8) was instead driven by significant differences between SameImage and DiffID ($p < .005$), SameImage and NonFamous ($p < .001$), SameID and NonFamous ($p < .001$), and DiffID and NonFamous ($p < .001$) in the no-stim condition, but only a difference between SameImage and SameID in the stim condition ($p < .005$).
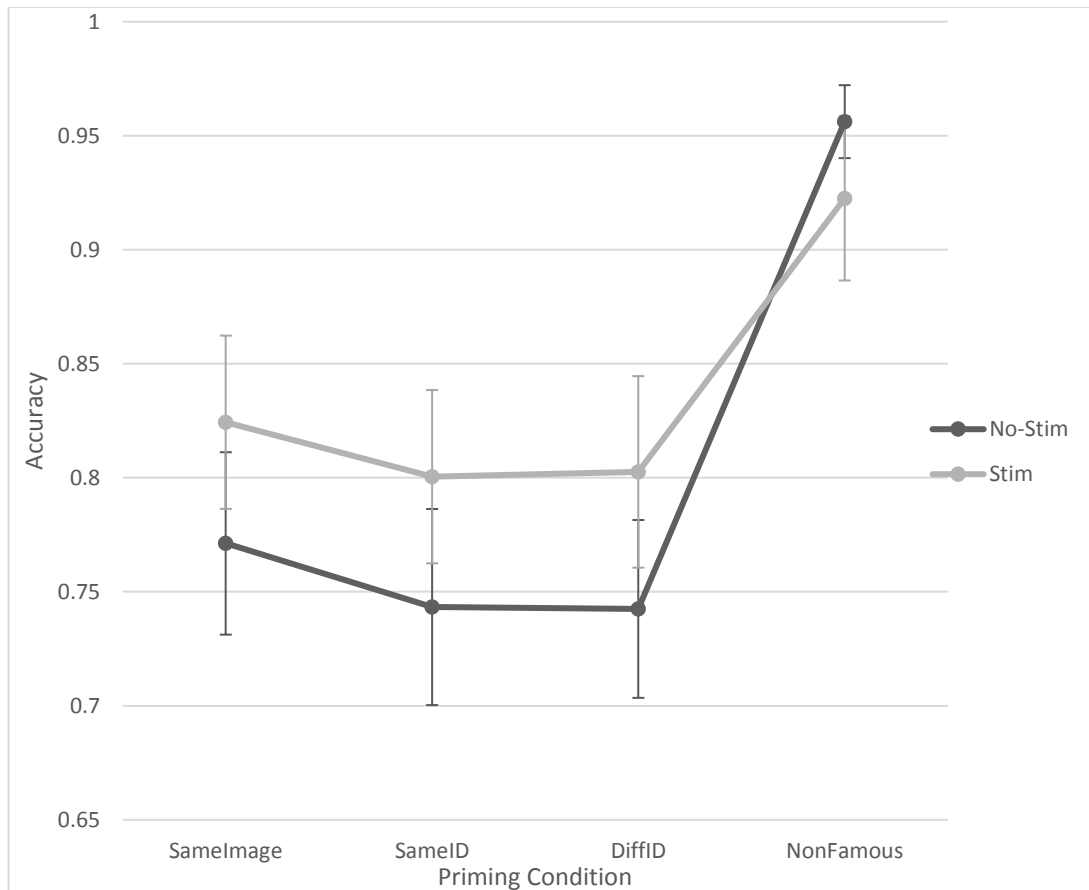
Figure 3.8: Means with standard error bars for a Prime × CVS interaction in the accuracy scores for the N250r task.

Reaction time

The task elicited a main effect of Prime, $F(3, 78) = 52.76$, $MSe = 567284.00$, $p < .001$, $\eta_p^2 = .67$, whereby the shortest reaction times were achieved in the SameImage condition (mean $= 768.85$, $SD = 178.03$), followed by the SameID (mean $= 875.63$, $SD = 214.83$) condition, followed by DiffID (mean $= 934.41$, $SD = 182.37$), with NonFamous having the longest RTs (mean $= 1044.09$, $SD = 176.39$), all of which were significantly different from each other, all $ps < .005$.

The prosopagnosia group were slower (mean $= 992.81$, $SD = 221.71$) than the control group (mean $= 870.92$, $SD = 198.49$) but the difference was not reliable, $F(1,26) = 3.23$, MSe

= 679178.51, p = .08, η_{p2} = .11. All other interactions with Group were not significant, all Fs ≤ 1.92, ps ≥ .13.

There was no main effect of CVS, F(1,26) = 0.76, MSe = 17482.75, p = .39, η_p^2 = .03, but there was a significant Prime × CVS interaction, F(3,78) = 4.26, MSe = 8761.27, p < .01, η_p^2 = .14. Pairwise comparisons did not show any differences between stim and no-stim for the four conditions, and the interaction (see Figure 3.9) was instead driven by significant differences between all priming conditions except DiffID and NonFamous (the two conditions that do not prime individual identity) in the no-stim condition (all other ps < .001).
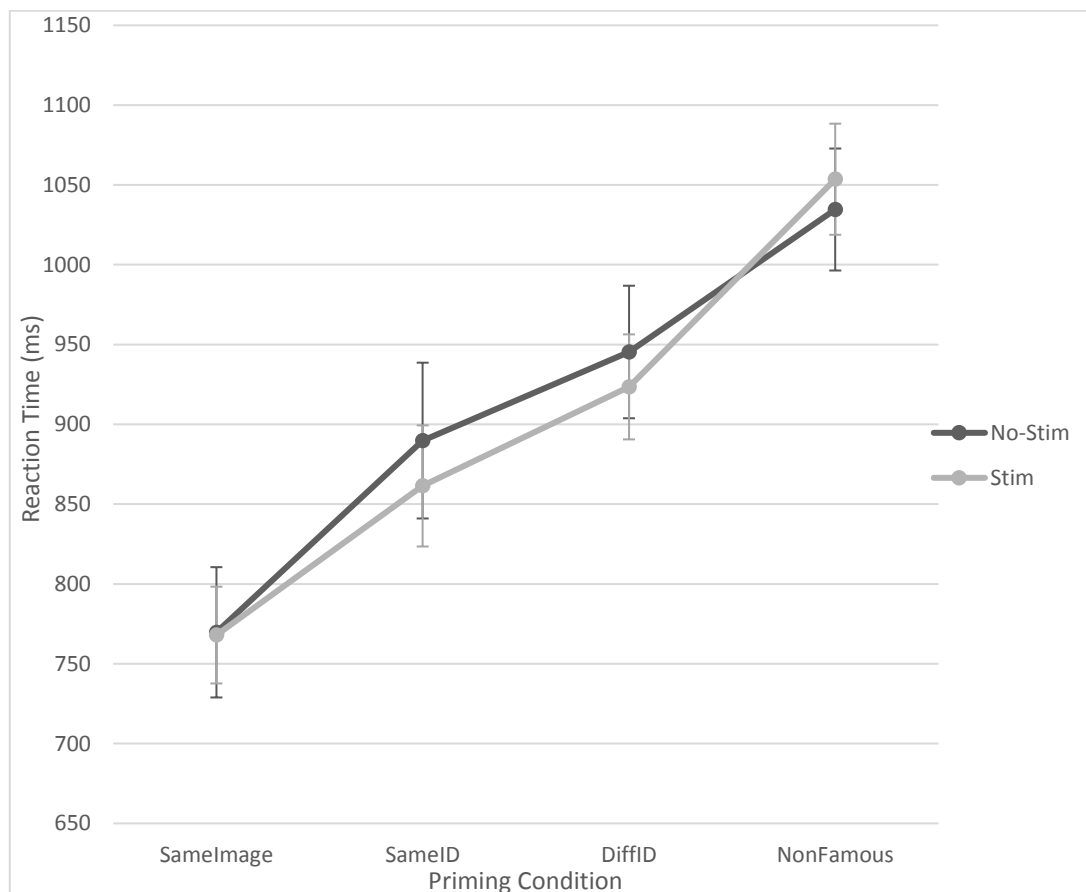
Figure 3.9: Means with standard error bars for a Prime × CVS interaction in the reaction times scores for the N250r task.

Electrophysiological data

N250r

Mean activity values (µV) of N250r peaks were analysed using ANOVA with the design 2(Group: DP vs. control) x 2(CVS: with vs. without) x 4(Prime: SameImage vs. SameID vs. DiffID vs. NonFamous) x 2(Hemisphere: left vs. right). Interactions were explored using Bonferroni corrected pairwise comparisons.

The task elicited a main effect of Prime, $F(3, 78) = 6.55$, MSe = 14.24, $p < .001$, $\eta_p^2 = .20$, whereby the largest N250r (in the present data, this corresponds to the least positive) was in the SameImage condition (mean = 1.83, SD = 2.49), which was not significantly different from in the SameID (mean = 2.19, SD = 2.84), both of which were significantly larger than in the DiffID condition (mean = 2.73, SD = 2.89), with the smallest N250r in the NonFamous condition (mean = 2.67, SD = 2.50) which was significantly different from the SameImage condition only.

The prosopagnosia group had a larger N250r (mean = 1.72, SD = 2.70) than the control group (mean = 2.61, SD = 2.66) but this was not reliable, $F(1,26) = 0.93$, MSe = 72.36, $p = .34$, $\eta_p^2 = .03$.

There was no main effect of CVS, $F(1,26) = 0.23$, MSe = 1.18, $p = .64$, $\eta_p^2 = .01$, but there was a significant Group $\times$ CVS interaction, $F(1,26) = 6.30$, MSe = 32.94, $p < .05$, $\eta_p^2 = .20$ (see Figure 3.10). Pairwise comparisons reveal no significant differences between conditions, but the pairs with the most contrasting mean differences and p-values suggest the difference between groups is reduced under stim conditions.

All other main effects and interactions were not significant, all Fs ≤ 1.56, ps $\geq .21$.
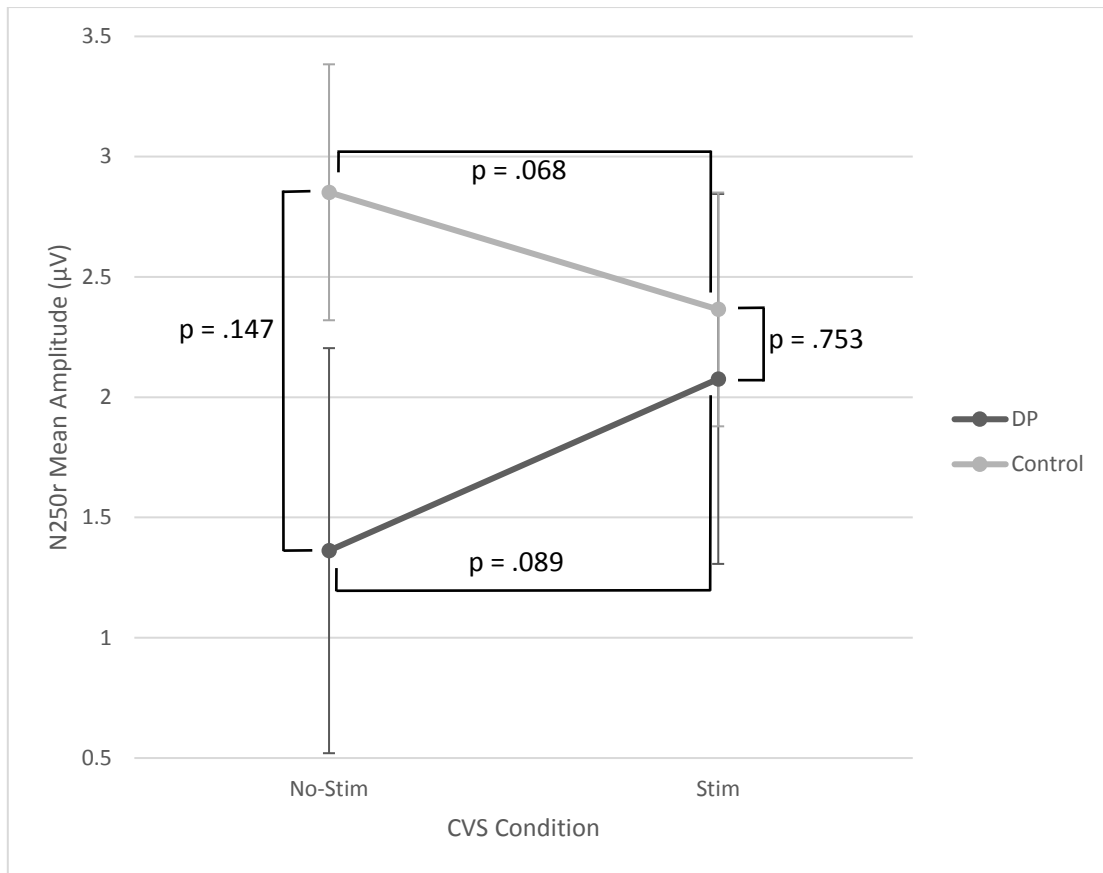
Figure 3.10: Means with standard error bars for a Group × CVS interaction in the N250r

mean activity values (pairwise p-values are not adjusted for multiple comparisons).
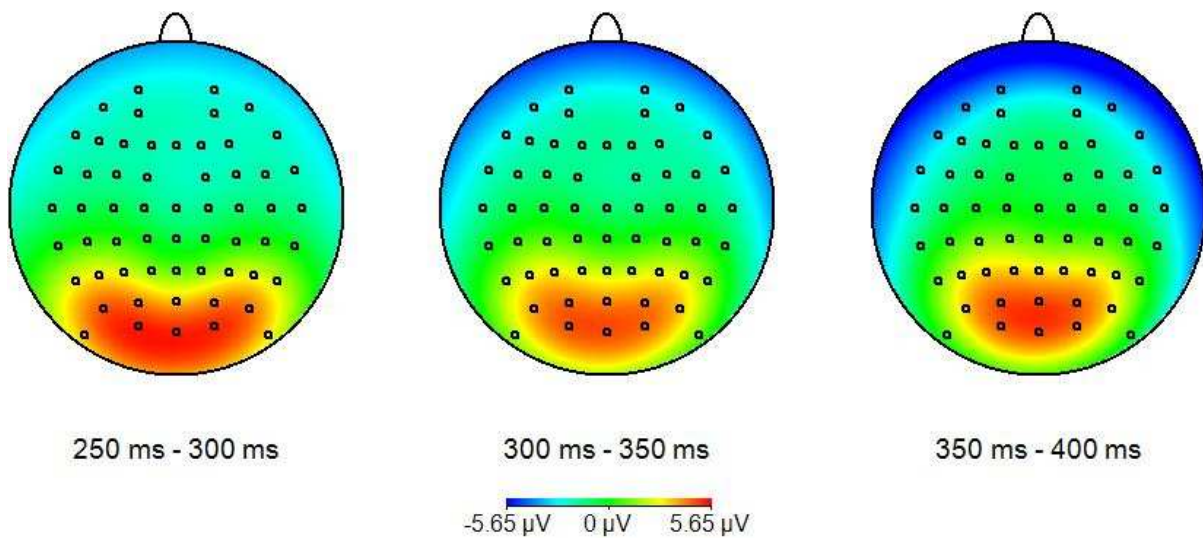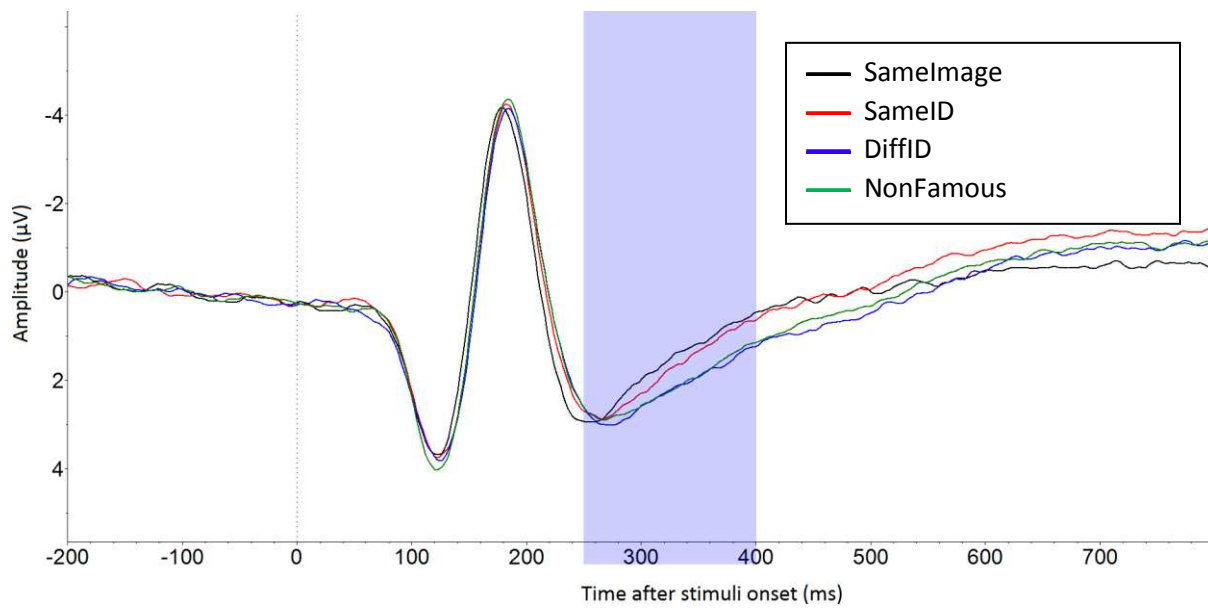
Figure 3.11: N250r waveform for SameImage (black), SameID (red), DiffID (blue), and NonFamous (green) conditions, collapsed across Group and CVS variables. Topographical plots show the location of the N250r activity, and surrounding time-frames, for the SameImage condition.
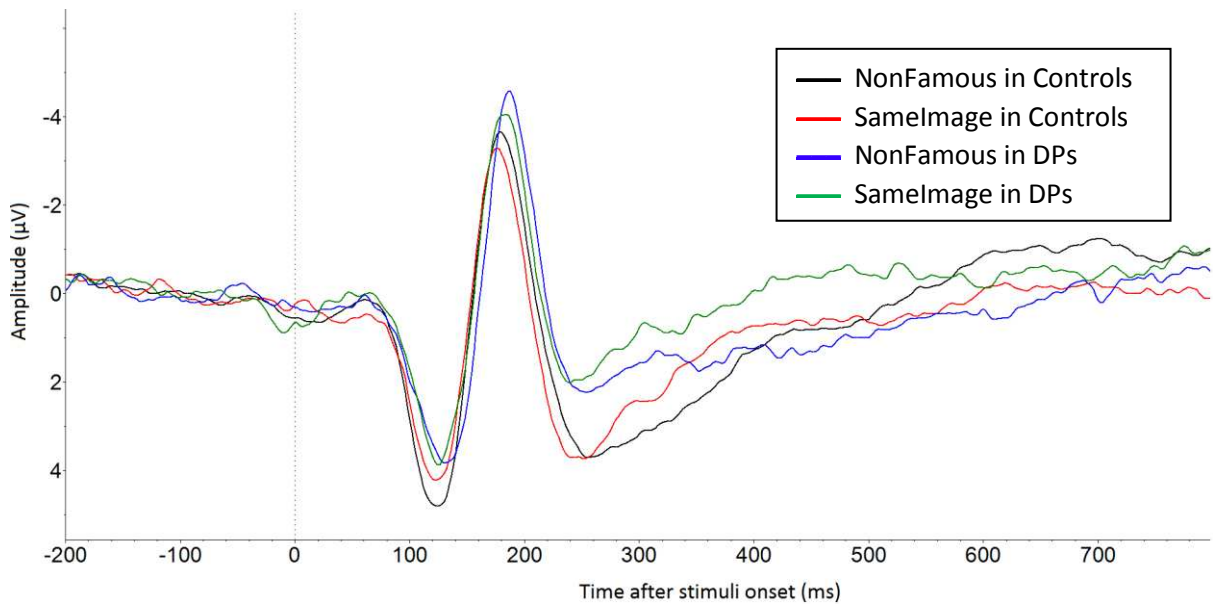
Figure 3.12: Maximal N250r effects (NonFamous compared to SameImage) in controls and prosopagnosia groups in the no-stimulation condition only, pooled across all eight electrodes.
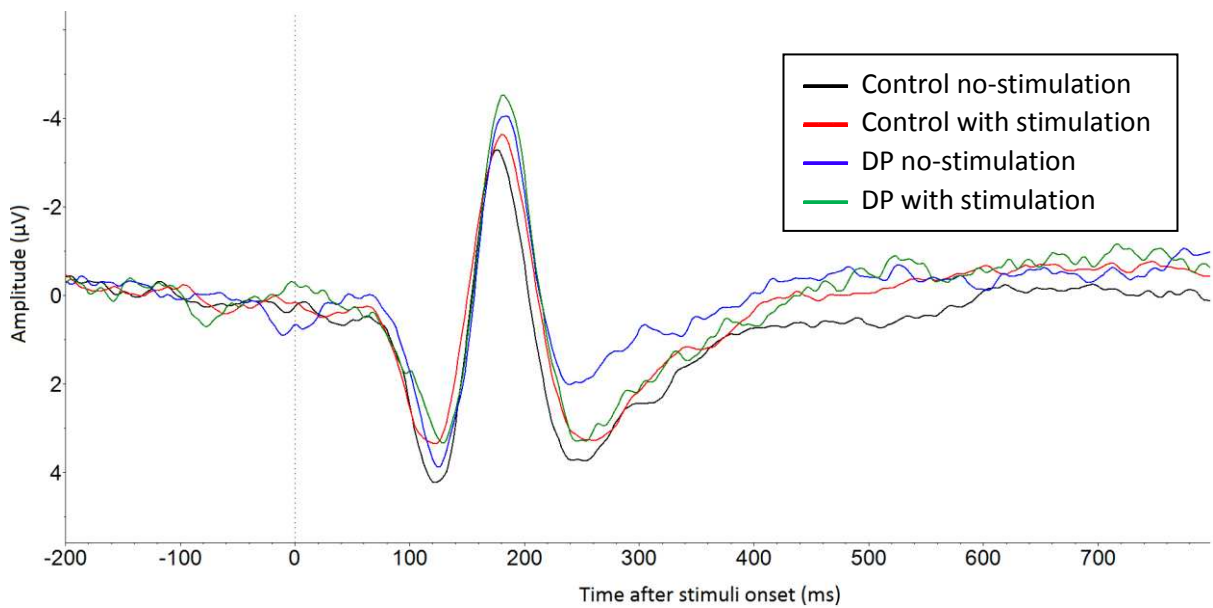
Figure 3.13: N250r waveform for the SameImage condition in controls (black and red) and prosopagnosia (blue and green) groups during no-stim and stim conditions.

Discussion (2)

The paradigm successfully elicited the N250r ERP component. The sizes of the ERPs were in the expected order, with the largest N250r for famous faces primed with the same image, followed by those primed with a different picture of the same person, followed by those primed with another celebrity, and lastly the smallest N250r for non-famous faces, the control condition. However, perhaps due to substantial variance, only the identical image condition was significantly different from the non-famous condition. The response time data provide support that the different priming conditions manipulated different performance outcomes, but a robust N250r effect was not consistently observed.

The accuracy data did not map linearly to the RT data, with non-famous faces having the longest RTs but the highest accuracy. It may be easier to determine that a face is not famous compared to famous, but the task requires an exhaustive 'cognitive search' for a matching memory of the face, and therefore takes longer. The process of recognising a famous face can halt once a match has been found, but this positive match appears to be more prone to error.

Also in the accuracy data, scores for the SameID and DiffID conditions were not different from each other, but were both significantly lower than the SameImage condition. This may suggest that it was only the image information, and not the facial information itself, that primed participants' accuracy to the second image. Though this account may fit appropriately with the DP group given the difficulty in extracting information from faces, the effect did not interact with the group variable, and perhaps, therefore, suggests that this lack of difference between SameID and DiffID should instead be viewed more sceptically, such as reflecting a type II error or a flaw in the experimental design such as not demanding sufficient attention to the prime image.

In line with predictions, the prosopagnosia group were significantly less accurate at recognising the face as famous compared to the control group. This is a critical finding to the study and identifies the first point of impairment in face recognition. This implies that the source of developmental prosopagnosia is after perceptual processing and at the earliest point at which face memory is required.

Alongside impaired accuracy, the developmental prosopagnosia group were no slower at making their judgement, and showed no difference in the N250r ERP response. One explanation for the typical RTs and N250r with impaired accuracy derives from the fact that these are based on correct trials only. Although the RTs and N250r are typical when the response is accurate, the data does not tell us about their patterns when incorrect response is given. While it is a valuable observation that the N250r is not permanently impaired in DP, the possibility remains that the component may show transient irregularities, and that an atypical N250r is associated with an incorrect response. A repetition of the analysis using only the incorrect trials could provide useful information to support or reject this hypothesis, but the comparatively low number of incorrect trials here does not permit such analysis. It would also be appropriate at this point to repeat that the DP group consists of only eight participants. While this may be a good size for a clinical sample of this nature, it is small in terms of EEG research and negatively impacts the already high variance seen in grand average waveforms. This issue would be considerably compounded if trial numbers were also low, as would be the case if using only incorrect trials.

There were no main effects of CVS in the N250r task, but there was an interaction with priming condition in the accuracy scores (Figure 3.8) and RTs (Figure 3.9). An explanation of these interactions is not straightforward but it would appear that something interesting is occurring. Analyses of the interaction between CVS and priming conditions showed there was

no effect of stimulation in any one priming condition. However, the complex pattern of differences between each of the priming conditions was changed during stimulation, and this appears to relate to CVS having a differential effect on trials where face memory is recalled, as will be explained below.

In the accuracy data without stimulation, means were higher for unfamiliar faces compared to any of the familiar faces. With stimulation, these differences were no longer reliable. This implies that CVS is acting differently for faces that have a memory trace. It is yet to be determined whether this is because of a benefit to familiar faces or a detriment to unfamiliar faces, but inspection of the means would suggest that it is perhaps both.

In the RT data without stimulation, famous faces primed with the same identity (identical and non-identical images) were recognised as familiar quicker than unfamiliar faces, yet with stimulation, even faces primed with a different celebrity were recognised as familiar faster than unfamiliar faces. Again, CVS is generating a difference between the unfamiliar faces and familiar faces trials. It should be noted, however, that while this difference was not reliable without stimulation, the p-value of .009 was rendered non-significant following arguably conservative Bonferroni adjustments.

Together, the accuracy scores and RTs provide compelling evidence that CVS differentially affects trials with a memory component, but it is not clear which memory process CVS may be affecting. The benefit could be related to the face representation in memory being matched for the target face, or to the memory of the familiarity status of the prime face matched to the familiarity of the target face, or both.

However, 'famousness' is not the only difference between the familiar and unfamiliar trials as the NonFamous condition also requires a different response compared to the other 75% of trials which require a 'famous' response. Perhaps CVS is interacting with their choice of button presses, for example increasing alertness to a change in the usual response? However,

the argument that CVS is merely increasing concentration does not explain the diminished accuracy in the non-famous condition, unless the increased concentration is stealing resources from face processing.

An additional interesting effect of CVS emerged in an interaction with group for the N250r mean amplitude (Figure 3.10). The interaction is difficult to interpret due to the pairwise comparisons identifying no clear source. Without stimulation, the N250r is smaller (though not significantly) for the DPs compared to the controls, but there does appear to be a pattern that the difference is reduced under stimulation. Stimulation appears to be having a 'normalising' effect but statistically the two groups did not differ even without stimulation, and the effect of stimulation was not reliable in either group. Given such uncertainty, the best approach may be to defer further investigation until the replication study in the next chapter is complete.

It is possible that levels of familiarity with the celebrities chosen for the study vary, and that some stimuli may be poor photographical representations of that celebrity. If recognition of these images was degraded in some way, their ability to prime, and be primed by, another image would be reduced. The implication is that any underlying priming effect, and its subsequent moderation by CVS, would be reduced. To explore this possibility, a by-item assessment was conducted on all familiar face stimuli presented. Given that trials with familiar faces were reacted to quicker than trials with unfamiliar faces, the overall mean RT for each individual familiar face stimulus was compared to the overall mean RT for all unfamiliar face stimuli (mean = 1038.70, SD = 248.94). Only four of the 180 stimuli yielded a mean RT above that for unfamiliar faces. Consequently, all trials where those faces were included (646 trials) were removed from the data set which was then re-analysed. The pattern of significant effects (main effects of Prime, Group and a CVS $\times$ Prime interaction in accuracy; main effects of Prime and a CVS $\times$ Prime interaction in RT) remained the same. Importantly, the F values for

the CVS $\times$ Prime interaction were very similar but marginally lower in both the accuracy data, $F(1.71, 44.54) = 5.99$, MSe $= 0.008$, $p < .01$, $\eta_p^2 = .19$, and in the RT data, $F(3, 78) = 4.02$, MSe $= 2078.77$, $p < .05$, $\eta_p^2 = .13$. This preserved pattern of effects suggests that the effects of CVS were not under-estimated by the inclusion of familiar faces that might in fact have been relatively unfamiliar to the DPs.

The extent to which this particular paradigm can robustly elicit the N250r has not been established in this experiment, indeed where previously published articles have often documented a lateral bias to the right hemisphere (Olivares, Iglesias, Saavedra, Trujillo-Barreto, & Valdés-Sosa, 2015), an effect was not observed in the present study. However, one particular observation that may be fruitful to investigate further is the time of offset of the N250r (whether this be defined as when two conditions regain equivalent amplitude or when they are merely no longer significantly different). It is unusual for the catchment window of the N250r to extend beyond around 400ms in previously published literature, yet in the present data, and as can be seen in Figure 3.12, the N250r ERP in the control group appears to continue until approximately 500ms. Given the substantial variation in N250r latencies with different experimental designs (Schweinberger & Neumann, 2016) this may not be of substantial concern, but the N250r offset in the DPs extends, on average, to around 700ms. An evaluation of whether this late offset correlates with behavioural performance would be an interesting hypothesis to test.

In sum, the key findings from this experiment are that the DPs were less accurate and no slower than the controls at categorising famous from non-famous faces. Significant interactions with CVS suggest subtle differences in effect between famous and non-famous conditions, and possibly a normalising effect on the N250r. Lastly, an N250r ERP effect was observed with image repetition but failed to reach significance during identity or 'famousness' repetition.

EXPERIMENT 3

Does the N400 differ between DPs and controls and is it modifiable with CVS?

In the last experiment the N250r was used as a correlate for face familiarity. In the present experiment the N400 will be used as a correlate for face recollection. The N400 component is thought to correlate with semantic activity involved in the identification of familiar faces (Bentin & Deouell, 2000; Engst, Martín-Loeches, & Sommer, 2006). The N400 is also present when semantic information such as occupation is primed (Barrett and Rugg, 1989), confirming that the N400 is post-perceptual. To capture this component, a repetition paradigm very similar to that used to elicit the N250r in the previous experiment, is typically used. Instead of manipulating familiarity, semantic information is instead used as a prime. Nationality (British and American) was chosen as the variable because this has previously been shown to successfully prime responses (McNeill & Burton, 2002).

Recollection of semantic information after seeing a face is a late process in the chain of successful face recognition. While a failure to identify an individual can still be accompanied with successful face perception and feelings of familiarity, recall of semantic information cannot be achieved. Therefore, it is in this semantic recall task that group differences are most likely to be seen. As observed in the N250r task, impairments may be confined to accuracy and not necessarily RTs as well.

There is very little evidence to guide predictions of the N400 in DP. Eimer (2000) found the N400 to be absent in one case of acquired prosopagnosia, and an alternative and later ERP correlate of semantic processing (P600f, where 'f' denotes it is associated with face stimuli) has been shown to be absent in DP, even for those with covert face familiarity (Eimer, Gosling, Duchaine, 2012). Nevertheless, if memorial processes are impaired in DP, then it is predicted that the corresponding N400 component will be smaller or absent.

Explicit judgements of nationality as a whole may be poor for the DPs, but an effect of priming remains possible. This was observed for the previous measure of familiarity whereby the group performed worse but an effect of priming was present, an effect that did not interact with group. As the group did not perform at floor for any of the tests they have completed, it is possible and likely that a priming effect will emerge based on the faces that they do recollect.

Similar to the N250r, the effects of CVS on the N400 have never been tested. The N400 is normally recorded over central and parietal midline regions (Wiese 2011). Research has identified a number of different sources of the N400, but the component is likely to reflect a wave of activity from posterior superior temporal gyrus to the anterior temporal lobe, (Kutas & Federmeier, 2011) affecting other areas such as the frontal lobe and the hippocampal system (Grunwald & Kurthen, 2006). Vestibular stimulation has been shown to stimulate overlapping regions including the hippocampus (Vitte et al., 1996), middle and superior temporal gyri (Bense et al., 2001; Friberg, Olsen, Roland, Paulson, & Lassen, 1985), and the temporal parietal junction (Lobel et al., 1998). Using the same logic described in the previous experiment, it is hypothesised that the changes in blood flow from CVS will modulate the activity in those regions. Consequently, it is predicted that accuracy and RTs for judgements requiring recollection of semantic information will be improved. As the N400 reflects access to semantic information, and this access is hypothesised to be facilitated in some way with CVS, it is therefore predicted that the N400 priming effect will be larger during active stimulation.

Method (3)

Stimuli

Stimuli consisted of 60 familiar faces (different from Experiment 2) each in two different images, resulting in 120 stimuli in total. Familiar faces were of 30 UK and 30 US

celebrities. Face images were at, or near, face-on view and were of straight or smiling faces. Stimuli were balanced for gender and images were trimmed to include the head only, removing the neck but including hair. All images were greyscale, tightly framed within a 170 pixels wide × 230 pixels tall canvas, centred on a black background and with surrounding details removed (see Figure 3.14).
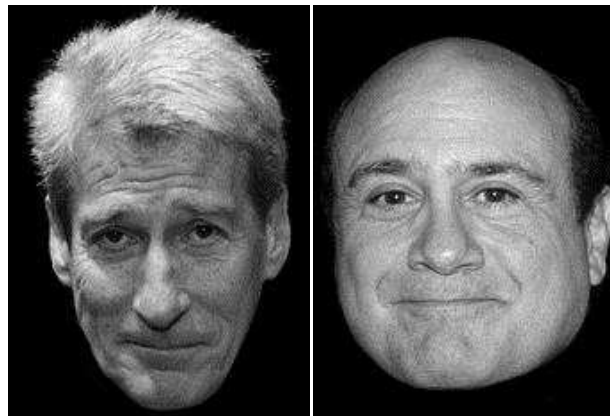
Figure 3.14: Example images from each stimuli category in Experiment 3, chapter III.

Design and Procedure

The procedure was as reported in the General Method with the following exceptions. Participants were first shown an instruction screen which instructed them to decide whether the second face in each pair is British or not. Responses were given by pressing the 'z' key for British and the 'm' key for not British. For each trial, a white fixation cross was presented in the centre of the screen for 500 ms before the prime face was shown for 500 ms. A green fixation circle was then presented for 1000 ms before the target face was shown for 2000 ms. Responses were only accepted during the 2000 ms of the target face presentation. All stimuli were presented in the centre of the screen.

There were two conditions: Consistent is when the target stimulus is of the same nationality as the prime stimulus; and Inconsistent is when the target stimulus is of different

nationality as the prime stimulus. The first set of 60 familiar faces was presented as primes with a randomly selected familiar face of the same nationality. The first set of 60 familiar faces was presented as primes again but this time paired with a randomly selected familiar face of a different nationality. The pairs of stimuli were presented in a random order forming 120 trials separated by two breaks.

This entire procedure was repeated on separate days, once with CVS and once without. In the stimulation condition, the CVS device protocol was initiated 30 seconds prior to the onset of the first stimulus and remained on during between-block breaks.

The dependent variables were accuracy, reaction time, and the EEG record, specifically the mean amplitude of the N400 ERP. The independent variables were Consistency (consistent and inconsistent), Group (prosopagnosia and control), and CVS (with and without stimulation).

ERP analysis

EEG preparation was as described in the general method. Specific to this experiment, epochs lasting 1700 ms and starting 200 ms before the onset of each target stimulus were generated off-line from the continuous EEG recording (note that the prime triggers were not used). These ERP waveforms were then corrected to a 200 ms baseline before being averaged separately for each condition. In total, 8 ERP waveforms were available for each electrode/channel: One for each of the two stimulus categories, in both groups, for both stimulation and no-stimulation conditions.

N400 mean amplitude was defined as the average voltage within the parameters of 250 ms and 450 ms post-stimulus onset. The data for correct responses only were exported using BrainVision Analyser 2 software. Only mean amplitude data from the Cz electrodes site at which the N400 was most prominent both in previous research and in the present study were used for analysis.

Results (3)

Behavioural data

Data preparation was as described for Experiment 2, Chapter Three. Correct answers were coded as 1 and incorrect as 0, and reaction times for incorrect trials and those 2.5 standard deviations from the mean were excluded. Mean accuracy and mean correct reaction times were interrogated using separate 2 (Prime: Consistent vs. Inconsistent) × 2 (Group: DP vs. control) × 2 (CVS: with vs. without) mixed measures, GLMs. Greenhouse-Geisser correction was used when assumed sphericity was violated.

Accuracy

Accuracy was above chance (mean = .79, SD = .15) but there was no main effect of Prime, $F(1,26) = 0.21$, MSe = 0.01, $p = .65$, $\eta_p^2 = .01$.

The DP group's accuracy (DP mean = .69, SD = .14) was significantly lower than that of the control group (mean = .83, SD = .13), $F(1,26) = 8.59$, MSe = 0.47, $p < .01$, $\eta_p^2 = .25$.

There was no main effect of CVS, $F(1,26) = 0.63$, MSe = 0.01, $p = .44$, $\eta_p^2 = .02$. All interactions failed to reach significance, $Fs(1,26) \leq 1.99$, $ps \geq .17$.

Reaction time

There was no main effect of Prime, $F(1,26) = 0.51$, MSe = 845.03, $p = .48$, $\eta_p^2 = .02$, or Group, $F(1,26) = 0.65$, MSe = 77219.58, $p = .43$, $\eta_p^2 = .02$.

Reaction times were shorter under stimulation (mean = 1064.67, SD = 172.57) compared to no-stim (mean = 1110.58, SD = 195.21) but this was not a reliable difference, $F(1,26) = 3.22$, MSe = 59734.83, $p = .09$, $\eta_p^2 = .11$. No interactions were significant, $Fs(1,26) \leq 2.62$, $ps \geq .12$.

Electrophysiological data

N400

The main effect of Prime was not significant using the p-value adjusted for multiple comparisons, $F(1,26) = 4.51$, $MSe = 1.49$, $p < .05$, $\eta_p^2 = .15$, whereby the N400 was not reliably different for the consistent condition (mean = -2.70, SD = 1.94) compared to the inconsistent condition (mean = -3.00, SD = 1.81).

There was, however, a main effect of Group, $F(1,26) = 5.06$, $MSe = 53.06$, $p < .05$, $\eta_p^2 = .16$, with the control group having a significantly larger N400 (mean = -3.28, SD = 1.75) than the DPs (mean = -1.76, SD = 1.75).

There was no main effect of CVS, $F(1,26) = 1.52$, $MSe = 2.17$, $p = .23$, $\eta_p^2 = .06$, and all interactions failed to reach significance, $F(1,26) \leq 3.28$, $p \geq .08$ (with the closest being CVS $\times$ Prime $\times$ Group).
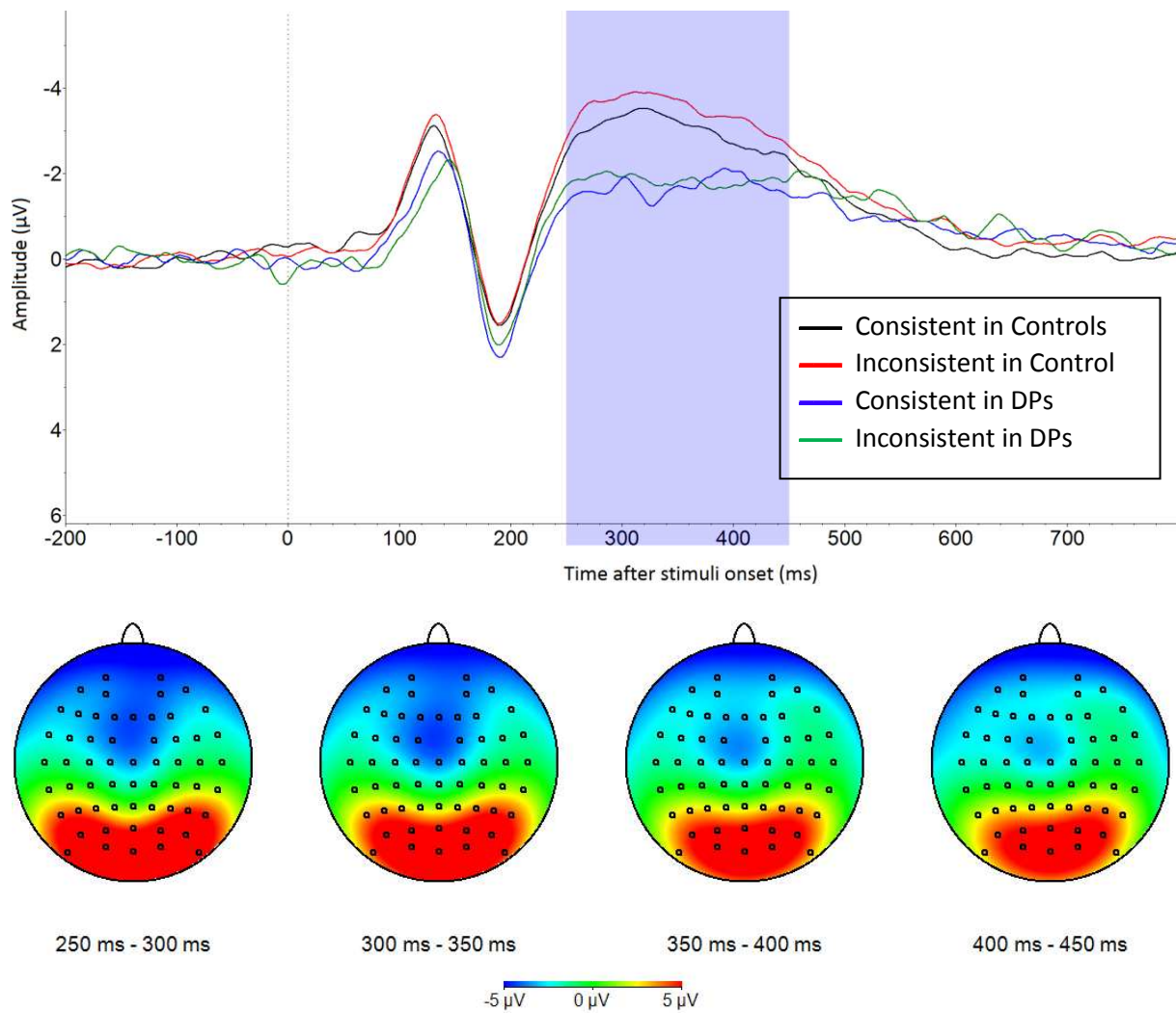
Figure 3.15: N400 waveform for the control (black and red) and prosopagnosia (blue and green) groups in the consistent and inconsistent conditions, collapsed across the CVS variable. Topographical plots show the location of the N400 activity, and surrounding time-frames, for the inconsistent condition in the control group.
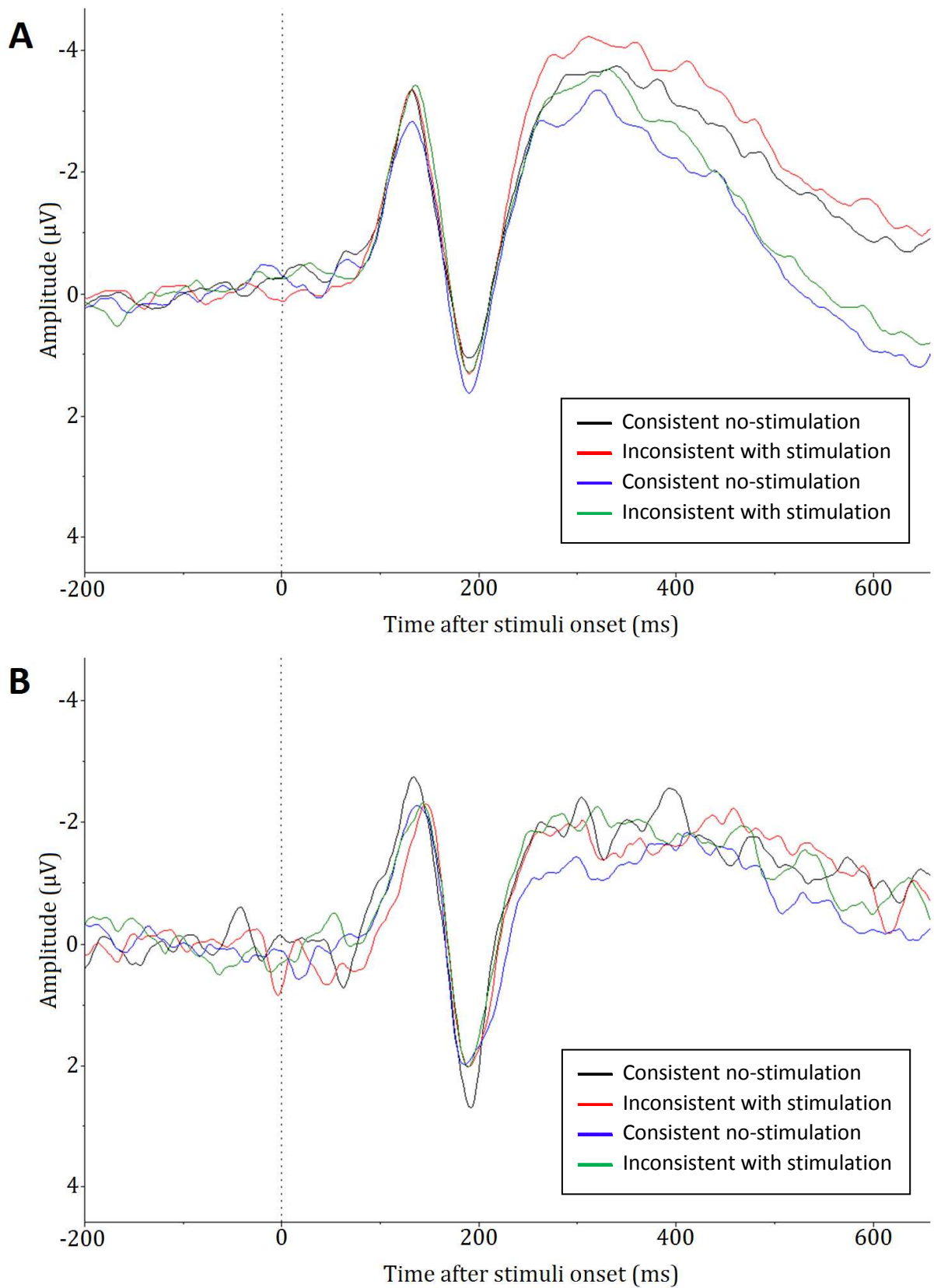
Figure 3.16: N400 waveform for each consistent and inconsistent conditions during no-stimulation (black and red) and stimulation (blue and green), in (A) controls and (B) DPs.

The experiment failed to elicit a significantly reduced N400 in response to the repetition of nationality. While the p-value of this test was below .05, the test did not withstand adjustments for multiple comparison necessitated by the later experimental replication in Chapter Four. The lack of priming effect did not interact with group, which suggests there was no evidence for an N400 effect (i.e. the difference between consistent and inconsistent conditions) in either the controls or DPs. When considering Figure 3.15, it is difficult to distinguish a traditional N400 effect for the DP group, but a conspicuous difference between consistent and inconsistent appears visible for the control group, so the possibility arises that high variance may be masking ERP effects. Topographical plots (Figure 3.15) suggest that focused use of the Cz electrode was suitable, but also that the more anterior FCz (used as on-line reference) may have been more appropriate had it been available. In concordance with the ERP data is the absence of effect of nationality priming on behavioural response.

Importantly, the N400 was significantly smaller for the DPs compared to the controls. This is the first ERP evidence for the source of impairment in this DP group. This effect was corroborated with the behavioural finding that the DPs were significantly less accurate than the controls at determining the nationality of the celebrities shown to them. This implies that DPs are significantly impaired at associating semantic information with a face.

There were no significant changes in either accuracy scores, RTs or N400 mean amplitudes as a consequence of using CVS. This implies that the effects of CVS are not associated with memory for semantic information. Although a trend emerged in the reduction of RTs during stimulation, the role of chance cannot be confidently discounted.

In summary, an overall smaller N400 evident in the DP group corresponded with lower accuracy scores. A priming effect between the consistency conditions was not present in any

measure in either group, suggesting either high variance or that Nationality did not prime. While reaction times were subtly shorter during stimulation, CVS was considered to have no effect in this experiment.

General Discussion

In Chapter Two, a group of individuals with developmental prosopagnosia completed a battery of tests to provide a handle on which core aspects of face processing they were impaired. With minimal variation, the sample showed most difficulty in memorial tasks and little impairment in perceptual tasks. To further understand the nature of this impairment profile, Chapter Three deployed experiments to assess two separate divisions of memory; 'familiarity' (Experiment 2) and 'recollection' (Experiment 3), while seeking to confirm an absence of perceptual impairment (Experiment 1).

To assess perceptual differences between the DP and control groups, an odd ball task was designed to elicit the N170 (with no relevant behavioural measure), an ERP marker for face perception. In this task, the group of eight DPs showed no differences in latency or amplitude in comparison to the control group. This pattern is in line with the perceptual performance seen in Chapter Two and indicates the absence of perceptual impairment. Despite thorough investigation into behavioural and physiological responses to face perception tasks, I have failed to find any substantial evidence of impairment.

To assess familiarity, a repetition priming paradigm was designed to elicit the N250r component, a marker for face familiarity, which is accompanied by a behavioural response of explicit familiarity. Consistent with a memory deficit, the DP group were significantly less accurate at identifying celebrities (but responded at an equivalent speed) compared to the

control group. This implies that the origin of impairments in DP is at the earliest stage that face memory is required.

The impairments in judging familiarity identified in the accuracy scores were not accompanied by differences in the N250r. This is a likely consequence of the ERP waveform being composed of only correct trials. While the state of the N250r for incorrect trials is unknown (due to insufficient incorrect trials), this does indicate that the N250r is not permanently atypical and that correct responses are associated with a typical N250r response.

In Experiment 3, recollection was assessed, again in a repetition priming paradigm designed to elicit the N400 component, a marker for semantic access. In line with the hypothesis, the DP group were significantly less accurate at identifying the nationality of celebrities (and responded at an equivalent speed) compared to the control group. Compellingly, this effect was accompanied by a reduced N400 mean amplitude for the DPs in comparison to the control group. This implies that DPs are impaired at associating semantic information with a face.

There are two additional points of interest from Experiment Three. Firstly, and unlike for the N250r task, poor accuracy was corroborated with a different ERP signature. As in Experiment Two, the grand average waveform entered into the analysis represented only the trials with a correct response. While this resulted in a typical ERP component in the N250r task, the suggestion for the N400 is that even when the behavioural response is correct, the ERP is atypical. Second, no effect of priming was observed in either group for either behavioural or ERP responses. Accuracy was well above chance indicating that the chosen stimuli were known, and the priming effects of British/American nationality have been observed before in the behavioural measures of other studies (McNeill & Burton, 2002), so it is unclear why they were not captured in this particular paradigm.

In sum, the results of Chapter Three are consistent with Chapter Two, in that the DPs show no differences in the ERPs of the perceptual task, but do show poor accuracy in both the familiarity and recollection memory tasks. This provides strong evidence that the source of impairments in this sample of DP is memory based. Differences in electrophysiological responses were only found in the N400, a marker for semantic access used in recollection.

A second goal of this chapter was to determine if vestibular stimulation improved face recognition. Only one confined effect of CVS was found in the N170 task, which was on the N170 latency for face stimuli in the right hemisphere of the control group. Although novel and face specific, the effect awaits replication in the next chapter given its complexity and unexpected nature. Contrary to previous findings from the only other available data on vestibular stimulation with EEG (Wilkinson, Ferguson, & Worley, 2012), no effects were observed on the amplitude of the N170. Previously, vestibular stimulation has been shown to increase the N170 amplitude. However, this was observed under different conditions: not only was this with a different sample group, it was also a different task, and importantly, it was a different method of stimulation, GVS. While CVS acts directly on the organs themselves (mostly the semicircular canals), GVS permeates to the labyrinth of the inner ear affecting both the semicircular canals and the otoliths. Consequently, the two methods elicit different cortical effects (Lopez, Blanke, & Mast, 2012).

There were no main effects of CVS in either the behavioural or ERP measures in the N250r task, however, there were two interactions of particular interest. First, it appears that CVS is affecting accuracy scores (and to some extent also RTs) in trials with familiar faces differently to trials with unfamiliar faces. Because a memory trace exists for familiar faces but not unfamiliar, this implies that CVS maybe having an effect on the memory of faces. This is in-line with existing evidence for the link between CVS and memory improvements (Bächtold

et al., 2001) and is particularly interesting given that the source of impairments in this sample of DP appears to be memory based. However, the omnibus interaction effect is subtle and is not supported with reliable CVS differences in any individual priming condition. For this reason, more confident inferences should await replication.

Second, and while non-significant pairwise comparisons muddy the picture, there is a sense that CVS may be reducing the N250r mean amplitude difference between the DP and control groups. Without stimulation, the difference in ERP sizes between groups is not reliable, but with stimulation this difference becomes much smaller. It cannot yet be claimed that CVS is normalizing the N250r, but this will be a key source of interest in the replication because it would suggest that CVS is modulating the underlying electrophysiology associated with face memory.

Lastly, there was no effect of CVS in the N400 task. As it was in this task that poor performance was related to an altered ERP, an effect of CVS would have been of special interest. Responses were noted as being faster during stimulation, but this effect did not reach significance.

In terms of study limitation, there are multiple established methods that could have been employed for analysing the ERP data. One that is likely to be rewarding in the future is single trial analysis which provides information about ERP variation within each participant. This may be particularly relevant given the heterogeneity of DP. However, though this technique may be achievable with prominent components such as the N170, the task faces significant technical challenges for later and noisier components of N250r, N400 and beyond where the component may not be visible in any one trial.

In conclusion, the three experiments reported in the present chapter provide insight into the impairments of the DP sample. Given the substantial individual heterogeneity in responses

in previously published DP papers, it is pleasing to find significant group effects in both behavioural and ERP measures that reaffirm and develop the conclusion of Chapter Two that the condition is primarily memorial in nature. The chapter also provides novel information about the effects of CVS in face recognition. These effects are limited and subtle, and it is not yet clear whether they are robust and reliable findings, an issue to which I next turn.

CHAPTER IV

Introduction IV: Study replication, test reliability and DP consistency

The previous chapter had two purposes. Firstly, it sought to identify the nature of the memory impairment that is the cause of difficulty in this sample of DPs. This was achieved by employing three tasks that individually tapped ascending levels of the face information processing hierarchy, beginning with perception, followed by familiarity recognition, and finally recollection of semantic information. Secondly, the ameliorative potential of CVS was tested in each of these experiments with the aim of improving behavioural performance and mapping this to changes in underlying ERPs.

As predicted, and consistent with intact perception, the DP group showed no differences in the N170 compared to controls. For the memory based experiments, the DP group were less accurate in both the familiarity and recollection tasks. This impaired behavioural performance was paired with an atypical N400 ERP correlate (less negative deflection compared to controls) but a typical N250r ERP. Evidence across Chapters Two and Three identify the source of impairments in DP to be memory based.

A face specific effect of CVS was observed in the N170 task, whereby the latency of the N170 was longer for faces compared to houses during stimulation, but this effect was confined to the right hemisphere of the control group. Perhaps of more interest, effects of CVS were seen in the N250r task and appear to relate to memory. Of the four priming conditions in this task, CVS affected the accuracy scores on trials where memory of the face was required differently to unfamiliar faces. The evidence that CVS acts on face memory is of particular importance given the memorial basis of impairments in this sample of DP. It remains unclear whether CVS was of benefit to familiar face recognition, detriment to unfamiliar face rejection,

or both. There was also a differential effect of CVS on the N250r ERP depending on group. The pattern of this effect suggests that differences between the group's N250r components are smaller with CVS. In this way, stimulation was normalizing the physiology of the DPs. However, this CVS effect were subtle and the source of the interaction was not statistically established. There were no effects of CVS in the N400 task.

The present chapter will assess whether these impairments are consistent over time. There is evidence to suggest that internal reliability of face recognition tests, even for the CFMT which is typically used to categorise DP, is high for unimpaired participants (Bowles et al., 2009; Herzmann et al., 2008; Wilmer et al., 2010). Indeed the Cronbach's alpha values from Chapter Two are based on the DPs and controls combined, but the large majority were high. However, for DPs analysed separately, internal reliability maybe significantly lower (Esins et al., 2016), perhaps as a consequence of a continually changing face recognition strategy. There are no published attempts to measure external reliability in DP. Test-retest measures are available for many face recognition tests (e.g. Bird, Papadopoulou, Ricciardelli, Rossor, & Cipolotti, 2003; McKone 2011; Wilmer et al., 2010), but a separate value for DPs is not reported.

Obtaining data for test-retest reliability will inform current understanding of the construct of DP, revealing whether impairments are permanent or relatively transient. This has considerable implications for clinical diagnosis because it would depend on the pattern of impairments specific only to the time of diagnosis and prevent any meaningful rehabilitation plan from being drawn up. A second consequence, and one relevant to this thesis, is one of evaluating potential interventions using pre and post measurements. Without consistency over time, fluctuations in performance between testing sessions may be incorrectly attributed to an intervention and not the underlying condition. Furthermore, from a theoretical perspective,

there would be a real risk of over-determining the data if it were so variable over time. While this may not be the case, test-retest reliability is not questioned.

A literature search into test-retest reliability for developmental disorders reveals that in many cases it is not assessed. Of the more heavily researched developmental disorders, test-retest reliability values are high for dyslexia (Berninger, Nielson, Abbott, Wijsman, & Raskind, 2008; Cotton, Crewther, & Crewther, 2005; Lefly & Pennington, 2000) and autism (Matson, Gonzalez, & Rivet, 2008; Silverman, Saavedra, & Pina, 2001). However, for many other developmental disorders, such as dyscalculia, amusia, color agnosia, developmental coordination disorder, and attention deficit disorder, test-retest reliability is generally not assessed. In the rare incidences where reliability is reported, it is often for the tests themselves using control data, and therefore does not speak to the consistency of scores from the atypical group. Even when the test-retest reliability scores for the atypical comparison group are reported, these are not then compared to those of a control group, so the source of the reliability cannot be adequately attributed to either the measure or the participant sample.

The Current Study

The previous chapter consisted of two testing sessions spread seven days apart, thereby providing data for a measure of test-retest reliability across a relatively short time period in both controls and DPs. The present chapter will build on the previous one by assessing the long term consistency of the DPs impairments; the two sessions comprising of all three experiments will be repeated 12 months later. This will provide two measures of short-term consistency and one of long-term consistency.

Reliability over time will be assessed in two different ways. First, the same ANOVAs run on the data from Chapter Three will be re-run to determine whether the patterns of effects

are still present one year later. To supplement these ANOVAs, correlations will examine the relationship between behavioural scores from one testing session to the next. While the ANOVA's show differences within sessions, they do not speak in quantitative terms to differences between sessions. Additionally, ANOVA is less sensitive to the consistency of individual participant's scores overtime. For example, correlations but not ANOVA would detect whether one half of the group are impaired at time 1, and only the other half impaired at time 2.

With respect to the effects of CVS, the impact on behavioural and electrophysiological responses appears to be localised and subtle. A normalising effect of CVS on the N400 amplitude in the DPs could not be confirmed with post-hoc analyses, but there was a clear trend that merits further investigation. Also, an understanding of the significant interactions between CVS and priming condition in the N250r behavioural data is not straightforward, and one would be more compelled to undertake it if the effect could be reproduced. At present, it appears the stimulation is having a subtle benefit for only the familiar face priming conditions, though this was not confirmed by post-hoc tests. It is not clear whether the subtle effects may intimate type I errors or whether these changes do truly exist. The replication design proposed in the current chapter can go some way to test this.

In sum, the present study will replicate the DP testing of Chapter Three. The key aim is to assess short- and long-term consistency of deficits in the DP group. This will be achieved by first assessing whether the group differences observed in Chapter Three are observed again, and secondly to assess consistency across the four testing sessions (sessions 1 and 2 from Chapter Three and sessions 3 and 4 from the present chapter). Lastly, the present chapter will test if the previously observed CVS effects are replicable.

General Method

Participants

All eight DP participants from Chapter Three were asked to return for repeat testing. Seven (three males, one left-handed) were willing and able to return for further assessment but one was no longer in the country (DP 3). This group were now aged 32-62 years old (mean = 53.86, SD = 10.65). No DPs reported any subjective changes in their face recognition abilities and none had received any form of training or therapy since their previous visit. Informed consent was obtained again from all participants prior to the onset of the experiment. Participants were financially compensated for their time.

Apparatus and Procedure

At least 12 months, and no more than 15 months, after each DP's participation in Chapter Three (sessions 1 and 2), they were contacted via email to arrange testing times and were given information about the study prior to their arrival. All experiments were conducted in exactly the same environment as sessions 1 and 2 and the setup procedure was identical to Chapter Three. After the EEG was set up, participants were once again shown the real-time output of their EEG signal and made aware of ocular and muscular artifacts, as well as alpha activity. All experiment scripts and apparatus (except the CVS unit, as described below) were identical to Chapter Three.

For each participant, the procedure protocol that they were assigned in session 2 was repeated for session 3, and the protocol for session 1 was repeated for session 4. In doing this, any possible effects of practice between sessions 1 and 2 were counterbalanced in sessions 3 and 4. Sessions 3 and 4 were again separated by 7 days so any short-term carry-over effects of CVS could dissipate. Figure 4.1 shows the graphical representation of the design, extending

from the original design in Figure 3.1. As can be seen, the protocol for an individual in sessions 1 and 2 was repeated for sessions 4 and 3 respectively. Responses from DPs were compared to the baseline data from the control group collected in Chapter Three, in line with convention adopted in other neuropsychology studies that compare patients to controls over time (e.g. Wilkinson, Morris, Milberg, & Sakel, 2013).
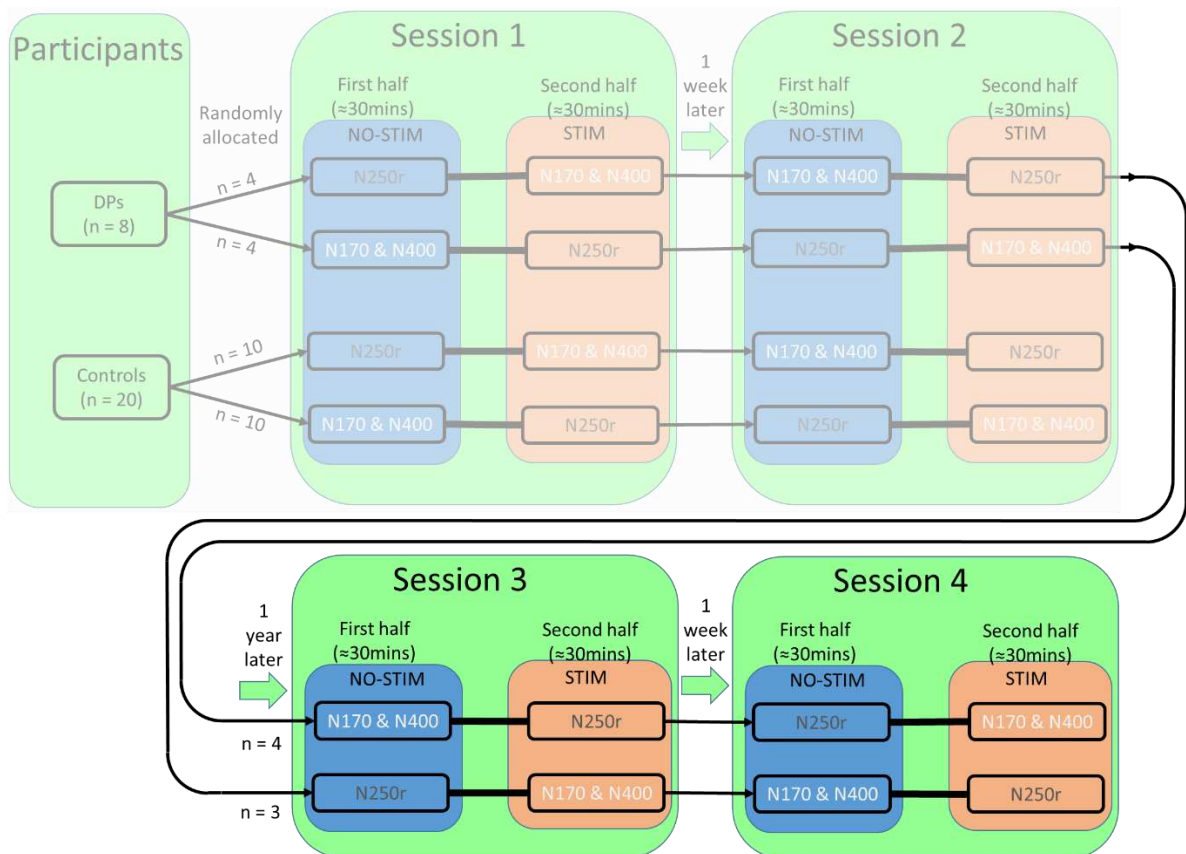
Figure 4.1: Illustration of the experimental design, showing the order in which the experiments were conducted and when stimulation was administered. The lightened top area is a replication of Figure 3.1 and represents Chapter Three; the bottom area represents the replication in the present chapter. Note that the order counterbalancing of the N170 and N400 experiments was kept consistent for each individual throughout all sessions.

Electrophysiological recording and preparation

EEG hardware, software and setup was identical to that of Chapter Three. Off-line analysis, including ICA, followed an identical procedure to Chapter Three except for a different approach to channels with zero signal. In the case of zero signal, contiguous electrodes (minimum of three) were averaged together and this manufactured signal was used to replace the zero signal. On no occasion did a key electrode (i.e. one named here as being used for creating an ERP) produce zero signal, therefore the recreated electrode signals were used only in the creation of topography maps.

Caloric vestibular stimulation

Vestibular stimulation was administered using a device very similar to that used in Chapter Three. Following technological advances in the design of the CVS hardware, the Generation 2.5 unit was replaced with the Generation 3.0 unit (as described in Black et al., 2016). Although the delivery technology has improved, most importantly the waveform itself remained unchanged. Aside from being more durable and easier to use, the newer unit does not require the replacement of water in the reservoir and therefore avoids associated interruption in its use. To maintain efficient heat exchange, the unit instead employs small internal fans attached to the heatsink which produce a quiet sound in both ears. The second key difference is the use of an earpiece in both ears and not just the left. In the present study, the right earpiece sat inside the ear but was not active and covered by a soft plastic cone. The newer control unit also displayed a real-time plot of the temperature change waveform, but this was turned away from the participant.

The protocol of the left earpiece was identical to that used in Chapter Three, a saw-tooth waveform cycling from 35 degrees Celsius to 17 degrees Celsius and back every three minutes. After each dose of CVS a log of activity was retrieved confirming the desired

temperature changes. As before in the no-stimulation condition, the headset was not worn but the chair was reclined.

Planned comparisons

As described, the experiments of the present chapter are a replication of the previous chapter. The same baseline control group data were used for both comparisons, so to prevent inflated type I error from making multiple comparisons that draw on the same data set (i.e. the control group) twice, the analysis in Chapter Three applied a correction in which the alpha region was divided by two (significant $p < .025$) for all effects except for the main effects of Group. This was because the main effects of group were strongly predicted and a very stringent alpha would run the risk of a type II error. Along similar lines, effects that were previously found to be significant were assessed using a one-tail test in this replication because the direction of effect was clearly predicted.

EXPERIMENT 1

A replication of Experiment 1, Chapter Three: 'Does the N170 differ between DPs and controls and is it modifiable with CVS?'

The N170 was induced with the same oddball paradigm script described in Chapter Three. The previous results showed no differences between the N170 of the control group and the DP group, suggesting a typical face perception ERP. This is consistent with the key finding from Chapter Two that the DP group presented little or no difficulties in face perception tasks. No main effects of CVS were found, except one interaction that was confined to the control group.

Method (1)

The same script programmed for Chapter Three was used for this experiment. Consequently, the stimuli and design were identical. The procedure was also identical with the exception of only testing the DP group. The dependent variables were the amplitude and latency of the N170 ERP. The independent variables were stimuli type (houses and faces), inversion (upright and inverted) and CVS (with and without stimulation). Finally, the ERP analysis followed an identical process to that described in Chapter Three, including the same 120 ms to 220 ms peak window and the same six parietal occipital electrodes.

Results (1)

Electrophysiological data

N170

Latency (ms) and amplitude (µV) of N170 peaks were analysed using separate mixed measures ANOVAs with the same design as previously used: 2(Group: DP vs. control) x 2(CVS: with vs. without) x 2(Object: face vs. house) x 2(Inversion: upright vs. inverted) x 2(Hemisphere: left vs. right). Interactions were explored using Bonferroni corrected pairwise comparisons.

i) Latency

As before, there was a main effect of Inversion, $F(1, 25) = 80.25$, $MSe = 27.53$, $p < .001$, $\eta_p^2 = .76$, whereby latency was longer for inverted (mean = 175.95, SD = 15.16) compared to upright stimuli (mean = 170.76, SD = 15.34). Also as before, the Object × Inversion

interaction also reached significance, $F(1, 25) = 26.32$, MSe = 77.38, $p < .001$, $\eta_p^2 = .51$, with pairwise comparisons showing an inversion effect (i.e. longer latency for inverted compared to upright) for face stimuli ($p < .001$), and a longer latency for inverted faces compared to inverted houses ($p < .001$).

As before, there was no main effect of CVS (no-stim mean = 173.49, SD = 15.01; stim mean = 173.23, SD = 15.92), $F(1, 25) = 0.28$, MSe = 155.74, $p = .60$, $\eta_p^2 = .01$. The 4-way interaction observed previously did not reach significance. All other main effects and interactions were not significant, all $Fs(1, 6) \leq 2.86$, $ps \geq .10$.

ii) Amplitude

As before, there was a main effect of Object, $F(1, 25) = 5.51$, MSe = 20.26, $p < .05$, $\eta_p^2 = .18$, whereby faces (mean = -5.58, SD = 4.06) elicited a larger amplitude (more negative) than houses (mean = -4.18, SD = 3.96). However, unlike before, this effect was moderated by Inversion, $F(1, 25) = 8.21$, MSe = 2.84, $p < .01$, $\eta_p^2 = .25$, with pairwise comparisons showing amplitude was larger (more negative) for inverted faces compared to inverted houses ($p < .005$).

As before, there was no main effect of CVS (no-stim mean = -5.10, SD = 4.10; stim mean = -4.66, SD = 4.03), $F(1, 25) = 1.07$, MSe = 9.15, $p = .31$, $\eta_p^2 = .04$. All other main effects and interactions failed to reach significance, all $Fs(1, 6) \leq 2.96$, $ps \geq .10$.
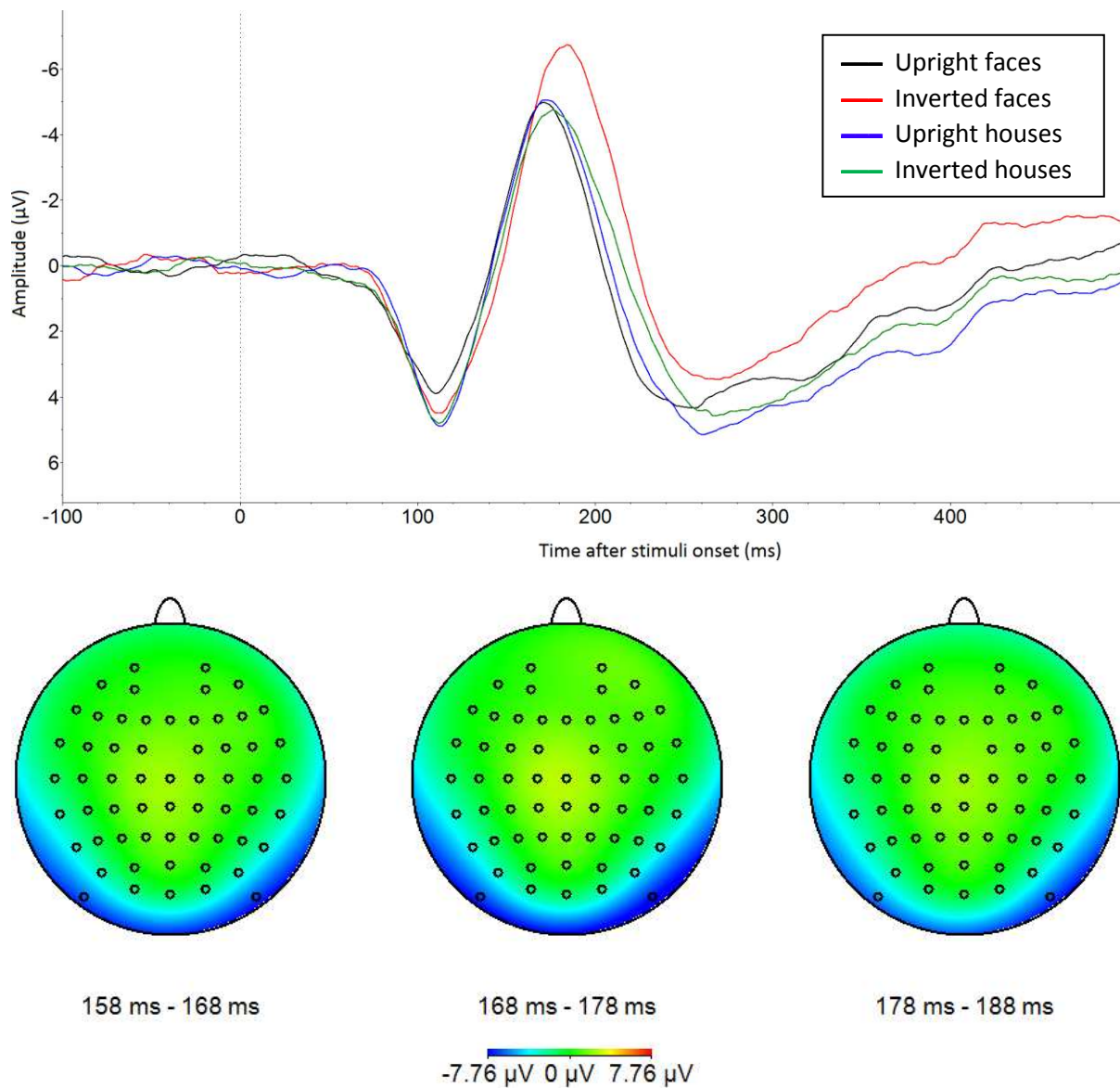
Figure 4.2: Pooled right hemisphere N170 peaks for faces (black and red) and houses (blue and green) in upright and inverted orientations, collapsed across CVS variable for the DP group only. Topographical plots show the location of the N170 activity, and surrounding time-frames, for the upright face condition for the DP group.
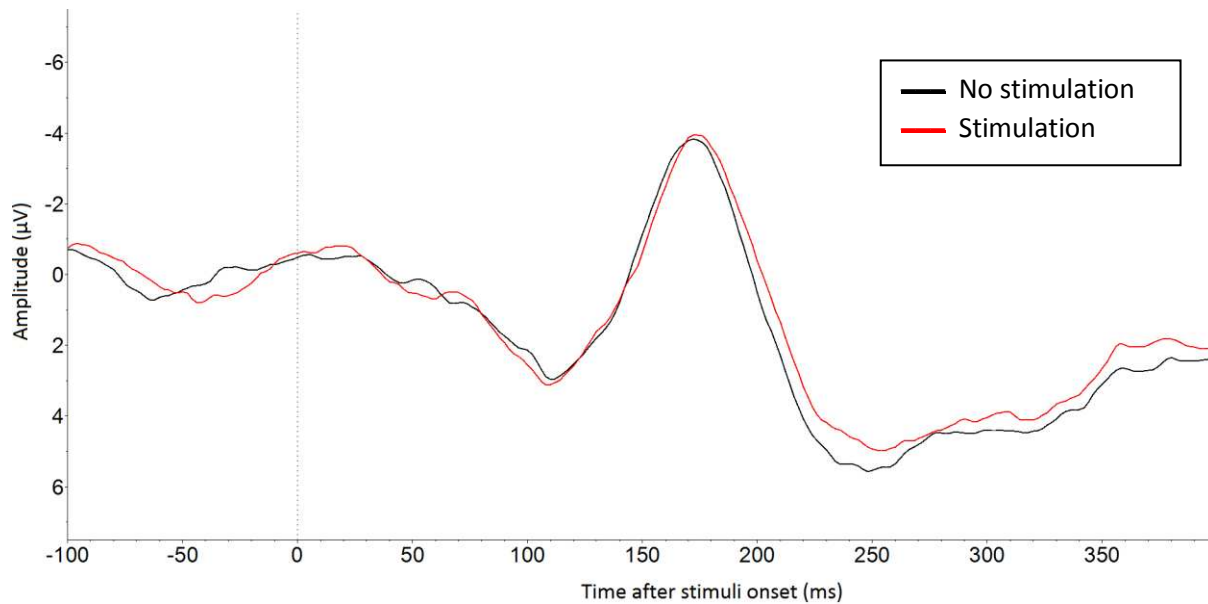
Figure 4.3: Pooled right hemisphere N170 peaks for upright faces in DP group during no-stimulation (black) and stimulation (red).

Discussion (1)

The present experiment sought to replicate the N170 data reported in Chapter Three. Previously, there were no overall differences between DPs and controls, and the only effect involving group (and CVS) was within a 4-way interaction, characterised by a longer latency for face compared to houses in the right hemisphere of the control group during stim. In this replication, there was still no evidence of impairment in the DP group, and the 4-way interaction was not reliable. There were still no effects of CVS.

As predicted, the expected effects of inversion and its interaction with object type that were significant in the latency and amplitude data of the previous chapter, were again all significant in this replication. However, a new Inversion × Object interaction also emerged in the amplitude data. This interaction was driven by inverted faces having a larger N170 than inverted houses, a standard and expected effect for this ERP. This tells us that a failure to find

a difference between the DPs and controls was not because the paradigm had been poorly implemented.

It is relevant to note that pairwise comparisons of the Inversion × Object interaction in the amplitude data, significant here but not in in Chapter Three, did not show a difference between faces and objects when presented upright. This is a standard effect found in N170 studies, and indeed what defines the ERP as face sensitive. It is not clear why this effect did not emerge here, but this does mean that if DPs, unlike typical face recognisers, do happen to process upright faces in a similar way to objects (in terms of N170 amplitude at least), the present paradigm may not have been able to detect it. Furthermore, and similarly to Chapter Three, the data lacked a main effect of hemispheric laterality, an established, though not always present, feature of the N170 (the possibility remains that these participants did not have a hemispheric bias).

Similarly to Chapter Three and as predicted, there were no effects of CVS. The two sources of evidence that vestibular stimulation modulates the N170, (Experiment 1 Chapter Three; Wilkinson, Ferguson, & Worley, 2012) are both in unimpaired participants. Whilst there is no evidence that the N170 is different for DPs, it is an interesting observation that vestibular stimulation has never been observed in that group, whereas it has in controls.

EXPERIMENT 2

A replication of Experiment 2, Chapter Three: 'Does the N250r differ between DPs and controls and is it modifiable with CVS?'

The N250r, a marker of face familiarity but not explicit knowledge, was induced using exactly the same repetition paradigm script described in Chapter Three. The previous results showed the DP group to be less accurate, but no slower at categorising faces as familiar. The

behavioural difference was not accompanied by changes in the mean activity values of the N250r.

There were no main effects of stimulation, but there were interactions between CVS and familiarity priming conditions in the behavioural data. The effects were subtle and related to differences between conditions that changed during stimulation and were not a primary effect of stimulation. In the accuracy data without stimulation, means were higher for unfamiliar faces compared to any of the familiar faces, and accuracy was higher for SameImage primes compared to DiffID primes. With stimulation, the only significant difference was between SameImage and SameID primes (see Figure 3.8). In the RT data without stimulation, all priming conditions were different to each other, yet with stimulation RTs for DiffID primes became no different to those for unfamiliar faces (see Figure 3.9). This replication study reported here will provide evidence of whether these effects are replicable, or whether they may represent type I error or inconsistency of the stimulation itself.

Lastly, the previous chapter showed an interaction between stimulation and group, whereby CVS tended to normalize the N250r ERP between the groups. Again, this was a subtle effect of stimulation that could not withstand post-hoc family wise error rate corrections, so pairwise comparisons were not significant. If the group ERPs and effects of CVS are consistent, then the trend will once again appear here.

Method (2)

The same script programmed for Chapter Three was employed for this experiment. The stimuli and design were identical except only the DPs were retested. The dependent variables were accuracy, RT and the mean amplitude of the N250r ERP. The independent variables were priming condition (SameImage, SameID, DiffID, NonFamous), stimulation (with and without

CVS), and where applicable, hemisphere (left and right). Finally, the ERP analysis followed an identical process to that described in Chapter Three, including the N250r parameters of a time window from 250 ms to 400 ms and two electrode pools using the eight temporal, occipital, and parietal electrodes.

Results (2)

Behavioural data

Data was prepared and trimmed in the same way as for Chapter Three. Mean accuracy and mean correct reaction times were interrogated using separate 4 (Prime: SameImage vs. SameID vs. DiffID vs. NonFamous) × 2 (Group: DP vs. control) × 2 (CVS: with vs. without) mixed measures, GLMs. Greenhouse-Geisser correction was used when assumed sphericity was violated. Interactions were explored using Bonferroni corrected pairwise comparisons.

Accuracy

As before, a main effect of Prime was observed, $F(1.15, 28.67) = 19.60$, MSe = 0.06, p < .001, $\eta_p^2 = .44$, with pairwise comparisons revealing accuracy for the NonFamous condition (mean = 0.96, SD = .09) was higher than SameImage (mean = 0.80, SD = .19), SameID (mean = 0.79, SD = .20), and DiffID (mean = 0.79, SD = .20) conditions (all ps < .001) (same pattern of means but a slightly different pattern of which comparisons reached significance compared to Chapter Three).

As before, there was a significant main effect of Group, $F(1,25) = 4.91$, MSe = 0.14, p < .05, $\eta_p^2 = .16$, whereby the DP group (mean = 0.74, SD = 0.20) were significantly less accurate than the control group (mean = 0.87, SD = 0.18). All other interactions with Group were not significant, all Fs ≤ 2.42, ps ≥ .07.

An effect not seen previously, was a main effect of CVS, $F(1, 25) = 10.53$, $MSe = 0.01$, $p < .005$, $\eta_p^2 = .30$, with accuracy significantly higher under stimulation (mean $= .86$, SD $= .17$) compared to no-stimulation (mean $= .81$, SD $= .20$). However, this was moderated by an interaction with Prime, as before, $F(1.70, 42.61) = 8.42$, $MSe = 0.05$, $p < .001$, $\eta_p^2 = .25$ (see Figure 4.4). They revealed that without stimulation, accuracy for the NonFamous condition was higher than all other conditions ($ps < .001$) but with stimulation these differences were not reliable ($ps \geq .006$). A Bonferroni correction for this post-hoc analysis is based on 16 comparisons and is therefore very strict, however, a more liberal correction does not affect the general pattern of results. The comparisons also revealed a simple main effect of CVS and increase in accuracy in the SameID and DiffID conditions ($ps < .005$), but not in the SameImage ($p = .005$) and NonFamous ($p = .36$) conditions. Note that a slightly more liberal correction would reveal a simple main effect of CVS in all priming conditions except NonFamous. Note also the very similar appearance in the plots of the means (Figure 3.8 and Figure 4.4).
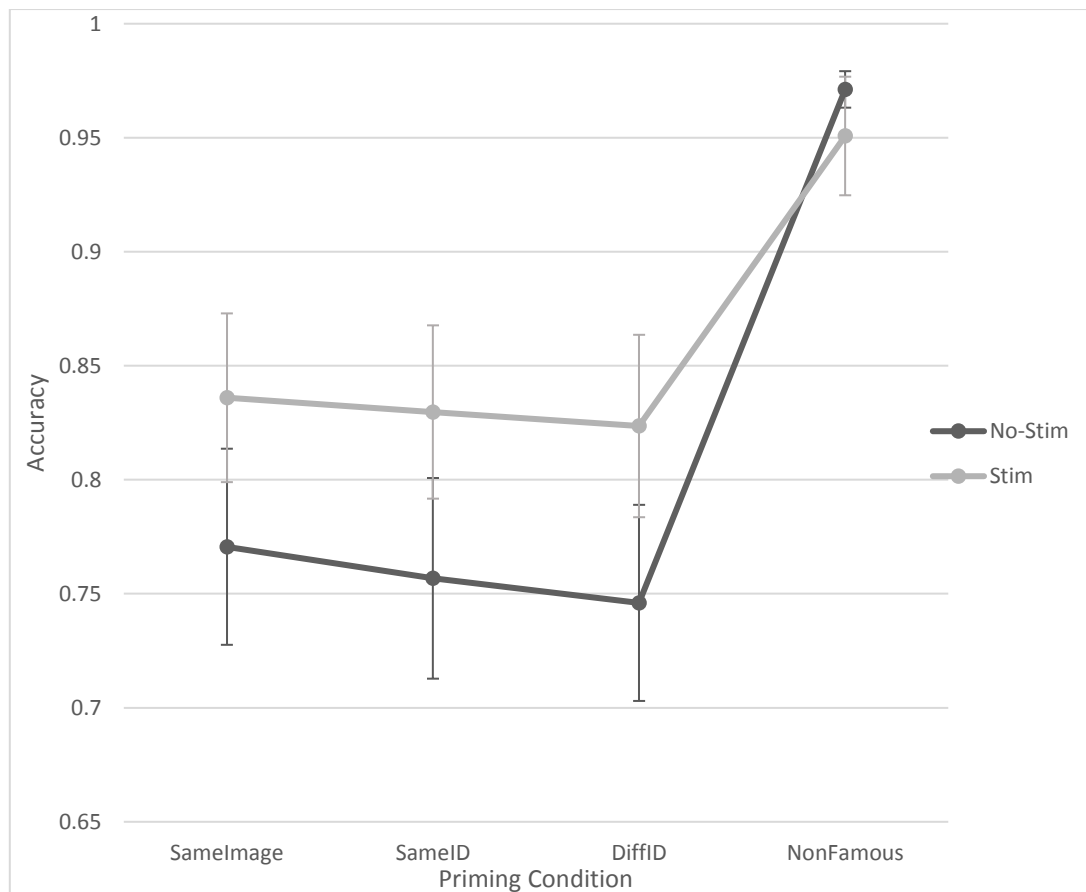
Figure 4.4: Means with standard error bars of a Prime × CVS interaction in the accuracy scores for the N250r task.

Reaction time

As before, a main effect of Prime was observed, $F(1.44, 36.04) = 38.78$, MSe $= 26927.74$, $p < .001$, $\eta_p^2 = .61$, whereby RTs were shortest for SameImage (mean = 764.35, SD $= 192.06$), followed by SameID (mean $= 867.36$, SD $= 216.56$) then DiffID (mean $= 936.91$, SD $= 193.11$) and the longest RTs for NonFamous (mean $= 1036.22$, SD $= 178.30$). Similar but not identical to before, pairwise comparisons revealed only DiffID and NonFamous were not reliably different from each other.

As before, there was no main effect of Group, $F(1, 25) = 2.56$, MSe $= 221544.95$, $p = .12$, $\eta_p^2 = .09$.

Reaction times were shorter under active stimulation (mean = 898.21, SD = 203.87) compared to no stimulation (mean = 904.21, SD = 232.41), but the difference was, as before, not reliable, $F(1, 25) = 0.45$, MSe = 29296.80, $p = .51$, $\eta_p^2 = .02$. However, as before, there was a CVS × Prime interaction, $F(2.17, 54.22) = 7.55$, MSe = 3463.44, $p < .001$, $\eta_p^2 = .23$ (see Figure 4.5). Corrected pairwise comparisons show RTs for NonFamous were only significantly longer than DiffID and SameID conditions under stimulation. Note that a slightly more liberal correction would reveal that just NonFamous and DiffID are only significantly different with stimulation, the exact pattern observed in Chapter Three. Note again the very similar appearance in the plots of the means (Figure 3.9 and Figure 4.5).
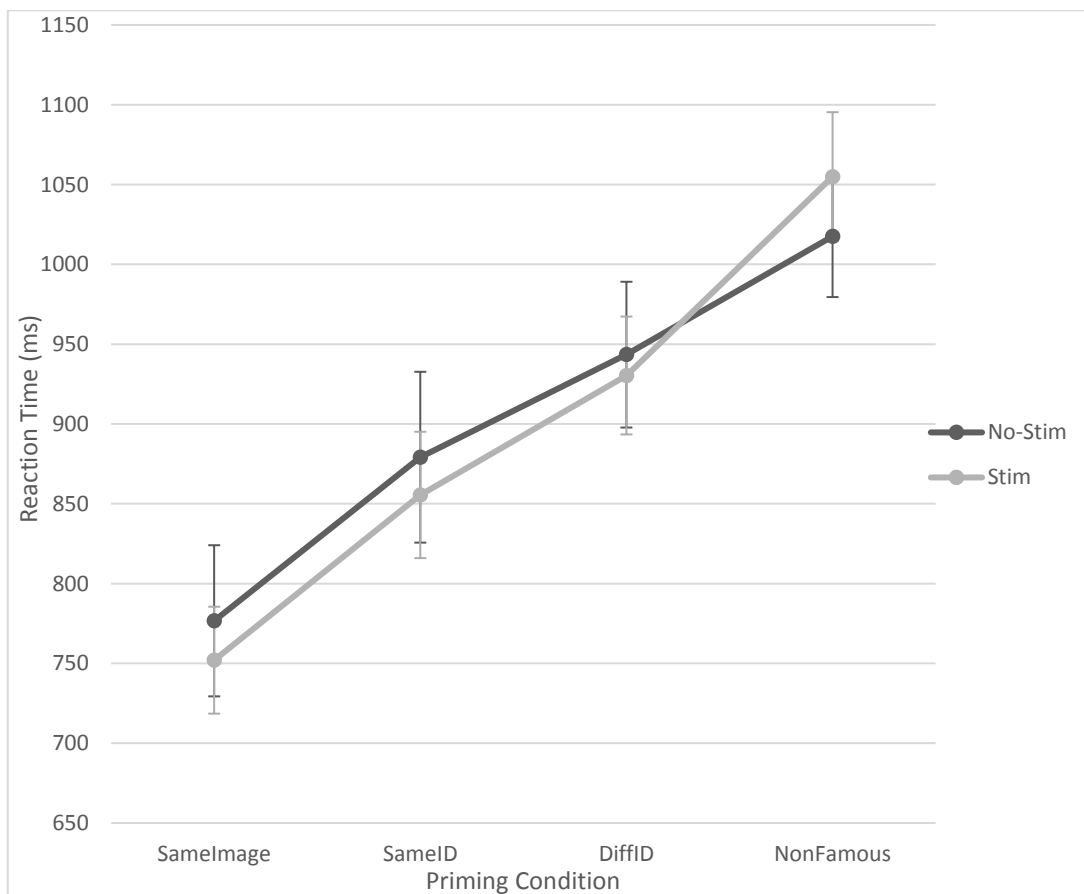
Figure 4.5: Means with standard error bars of a Prime × CVS interaction in the reaction time scores for the N250r task.

Electrophysiological data

N250r

Mean activity values (μV) of N250r peaks were analysed using ANOVA with the design 2(Group: DP vs. control) x 2(CVS: with vs. without) x 4(Prime: SameImage vs. SameID vs. DiffID vs. NonFamous) x 2(Hemisphere: left vs. right). Interactions were explored using Bonferroni corrected pairwise comparisons.

As before, there was a main effect of Prime in the N250r mean activity values, $F(2.14, 53.56) = 3.70$, MSe $= 2.89$, $p < .05$, $\eta_p^2 = .13$. No pairwise comparisons were significant, though mean values were in the expected direction with SameImage (mean $= 2.04$, SD $= 2.63$) as the largest (least positive), followed by SameID (mean $= 2.31$, SD $= 2.92$), and with DiffID (mean $= 2.85$, SD $= 2.92$) and NonFamous (mean $= 2.77$, SD $= 2.44$) being the two smallest.

As before, there was no main effect of Group, $F(1, 25) = 0.20$, MSe $= 82.69$, $p = .66$, $\eta_p^2 = .01$.

There was no main effect of CVS, $F(1, 25) = 0.03$, MSe $= 6.26$, $p = .88$, $\eta_p^2 = .001$, with the N250r being no different with stimulation (mean $= 2.36$, SD $= 2.55$) compared to no-stimulation (mean $= 2.62$, SD $= 2.92$). The Group $\times$ CVS interaction reported before did not reach significance. All other main effects and interactions were also not significant, Fs ≤ 2.61, ps $\geq .12$.
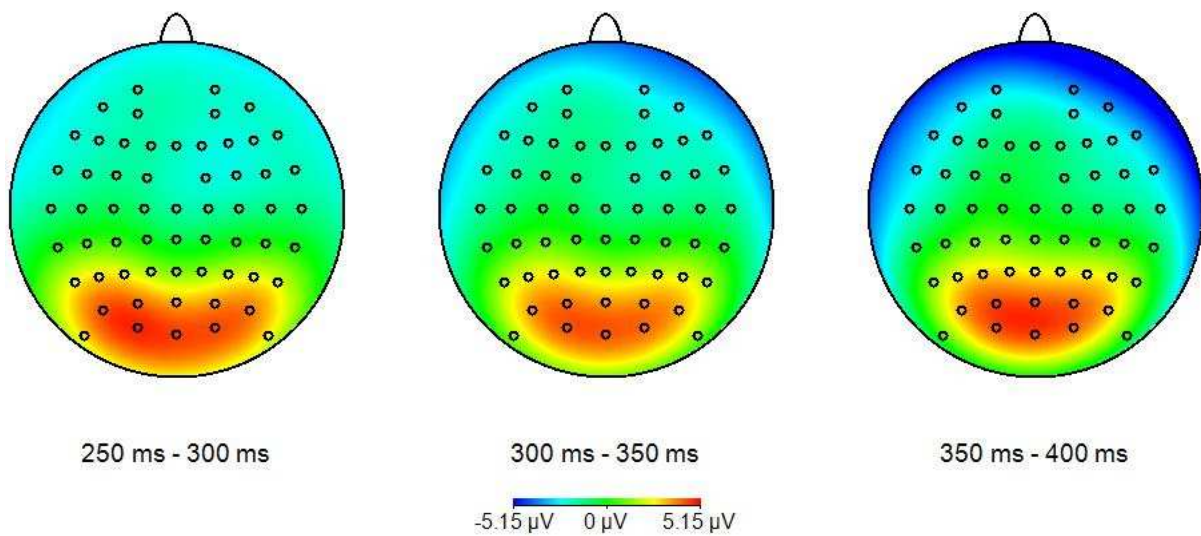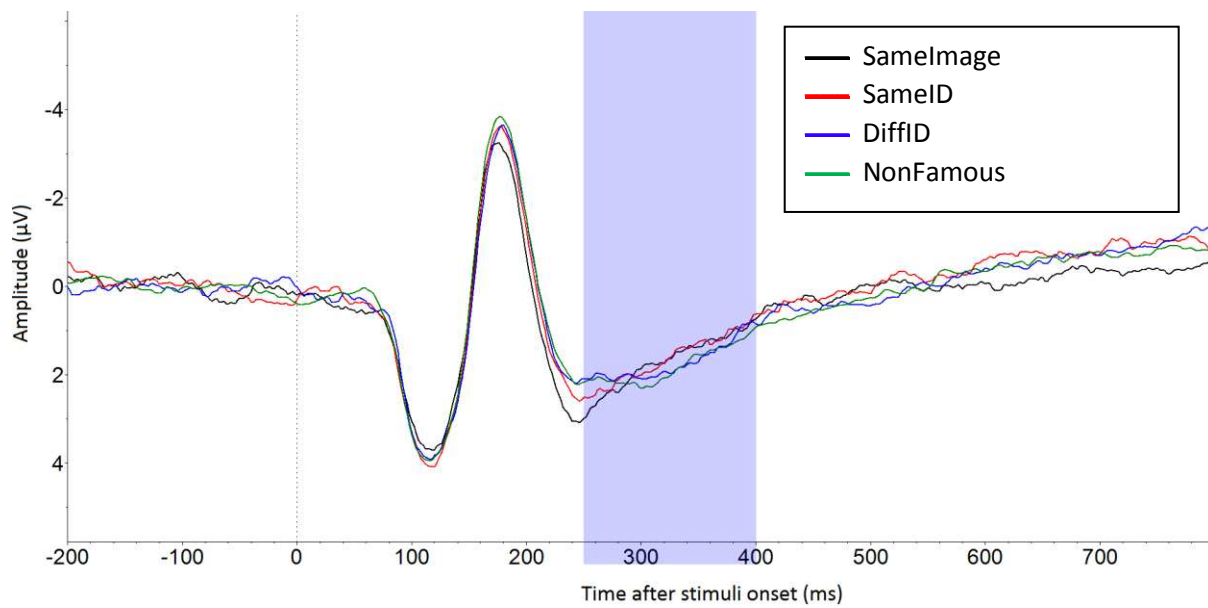
Figure 4.6: N250r waveform for SameImage (black), SameID (red), DiffID (blue), and NonFamous (green) priming conditions for the DP group, collapsed across Hemisphere and CVS variables. Topographical plots show the location of the N250r activity, and surrounding time-frames, for the SameImage condition for the DP group.
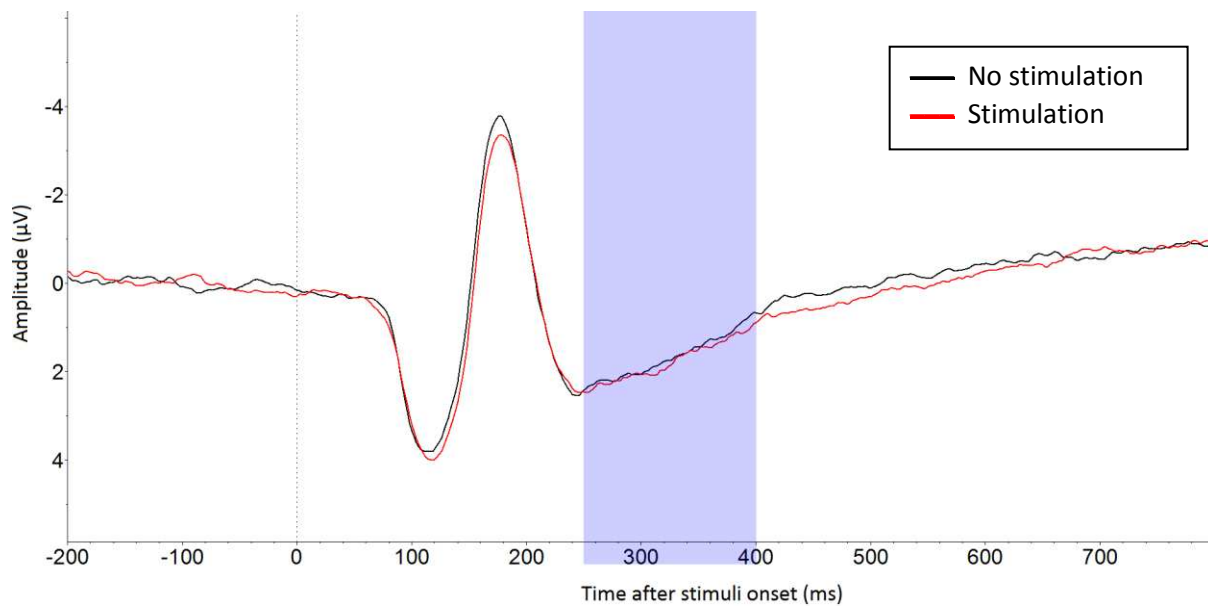
Figure 4.7: N250r waveform for CVS stimulation (red) and no-stimulation (black) for the DP group collapsed across Hemisphere and Prime variables.

Discussion (2)

The present N250r experiment was a replication from the previous chapter to assess the consistency of DP impairments in a task involving familiarity memory. Previously, the DP group showed lower accuracy compared to controls with no difference in reaction times. In this replicated experiment, the same pattern was observed; lower accuracy with typical reaction times. This shows consistency of impairment and also adds to the evidence from Chapters Two and Three that the source of impairment in the DPs is memory based. Chapter Three also described how CVS affected behavioural performance in trials with familiar faces differently to trials with unfamiliar faces. This effect was also replicated.

Similar to Chapter Three, effective priming of face familiarity was revealed in all measures. The interaction between each of the priming conditions, that is, which ones were reliably different from each other, differed between chapters. For example, in the accuracy data, the present chapter found the NonFamous condition to be significantly higher than all

others, yet for the previous chapter the pattern of which conditions were different from the others was far less simple. While fluctuations in means and SDs should be expected, the pattern of means between the original and this replication were very similar and demonstrate reliability in the paradigm.

In contrast to the original, this replication discovered a main effect of CVS, whereby accuracy was significantly improved during active stimulation by an average of nearly 5 percent. This effect was moderated by priming condition, an effect present in accuracy and RT in both chapters. For the accuracy data, the key similarity between the testing sessions is that during stimulation, accuracy for unfamiliar faces is no longer higher than the other conditions. The present study now reveals that this is because stimulation is increasing accuracy only in the conditions that require familiarity with the famous faces. From this, it can be determined that CVS is acting selectively on faces with which the observer has familiarity, and this infers that the locus of effect is with facial memory.

The interaction between stimulation and priming condition was also present in the RT data. The pattern of effects that is consistent between the original and the replication is that the RTs for the NonFamous condition are only slower than the DiffID condition during active stimulation. Pairwise comparisons do not reveal why this is, but the means suggest that stimulation may have a mildly beneficial effect in the DiffID condition, and a mildly detrimental effect in the NonFamous condition; effects that are too small for significant simple main effects of CVS, but ones which come together to form a significant interaction. This implies, as is the case in the accuracy scores, that CVS is affecting trials that require memory differently to those trials that do not, which further supports the inference that the locus of effect of CVS is with facial memory.

It could be hypothesised that face memory networks are receiving an extra 'boost' from CVS that enables them to somehow function in a way that improves their ability, perhaps by

improving search efficiency for example, and this consequently increases the chances of quickly and correctly categorising a face as familiar. In parallel, the time allowed for an exhaustive search before rejecting the face as unfamiliar may not be affected. CVS is affecting processes associated with the memory representation of a face; trials with unfamiliar faces do not have a memory trace and this is why they are not responding to stimulation.

An alternative explanation might be one regarding a placebo effect. The participants are aware of the hypothesis that CVS may effect face recognition, and they may naively believe that the benefits are more likely to be positive, despite never being tested before. In this situation, they may feel that they should be recognising more celebrities, and therefore increase their rate of 'famous' responses. However, while a random increase in 'famous' responses across trials will increase accuracy and reduce RTs for famous trials, one should also observe an accompanying reduction in accuracy for unfamiliar faces, for which there is no evidence.

Lastly, in the previous chapter a Group × CVS interaction was observed, whereby stimulation appeared to be bring the N250r mean amplitudes of the DPs and controls closer together, though pairwise comparisons could not verify this reliably. In the present chapter's replication, there was no evidence of this effect. We are therefore unable to further endorse this potentially normalising effect.

EXPERIMENT 3

A replication of Experiment 3, Chapter Three: 'Does the N400 differ between DPs and controls and is it modifiable with CVS?'

The N400, thought to reflect activation of semantic information involved in familiar face identification, was induced using the same repetition paradigm script described in Chapter Three. The previous results showed the DP group were less accurate but no slower at

determining the nationality of celebrities. This behavioural difference was paired with a reduced N400. There was also no effect of priming in behavioural or ERP responses. Lastly, there was no main effect of CVS or interactions involving this variable, so any such effect in the present experiment is unexpected.

Method (3)

The same script programmed for Chapter Three was employed for this experiment. The stimuli, design, and procedure were identical except only the DP group were retested. The dependent variables were accuracy, RT and the mean amplitude of the N400 ERP. The independent variables were priming condition (consistent and inconsistent) and CVS (with and without stimulation). Finally, the ERP analysis followed an identical process to that described in Chapter Three, including using the N400 parameters of a time window from 250 ms to 450 ms and using electrode Cz.

Results (3)

Behavioural data

Data was prepared and trimmed in the same way as for Chapter Three. Mean accuracy and mean correct reaction times were interrogated using separate 2 (Prime: Consistent vs. Inconsistent) $\times$ 2 (Group: DP vs. control) $\times$ 2 (CVS: with vs. without) mixed measures, GLMs. Greenhouse-Geisser correction was used when assumed sphericity was violated.

Accuracy

As before, there was no main effect of Prime, $F(1, 25) = 0.16$, MSe = 0.004, p = .69, $\eta_p^2 = .01$, with accuracy for Consistent trials (mean = .81, SD = .13) not reliably different to Inconsistent trials (mean = .80, SD = .14).

As before, there was a main effect of Group, $F(1, 25) = 6.16$, MSe = 0.05, p < .05, $\eta_p^2 = .20$, whereby the DP group (mean = .71, SD = .12) were less accurate than the control group (mean = .83, SD = .13).

As before, there was no effect of CVS, $F(1, 25) = 0.12$, MSe = 0.01, p = .73, $\eta_p^2 = .005$ (no-stim mean = .79, SD = .13; stim mean = 0.82, SD = .14). All other interactions were not significant, all $Fs(1, 25) \leq 2.96$, ps $\geq$ .10.

Reaction time

Unlike before, there was a significant main effect of Prime, $F(1, 25) = 6.32$, MSe = 1436.17, p < .05, $\eta_p^2 = .20$, whereby RTs were 20ms shorter for Inconsistent trials (mean = 1080.20, SD = 188.07) compared to Consistent trials (mean = 1100.61, SD = 188.26), the reverse of what was expected.

As before, there was no main effect of Group, $F(1, 25) = 0.96$, MSe = 121159.93, p = .34, $\eta_p^2 = .04$, with RTs for the Control group (mean = 1071.02, SD = 191.91) were not reliably shorter than those for the DP group (mean = 1145.79, SD = 165.41). As before, there was also no effect of CVS, $F(1, 25) = 1.11$, MSe = 20080.51, p = .30, $\eta_p^2 = .04$ (no-stim mean = 1108.26, SD = 197.01; stim mean = 1072.55, SD = 177.65). All other interactions were not significant, all $Fs(1, 25) \leq 2.32$, ps $\geq$ .14.

Electrophysiological data

N400

Unlike before, there was a main effect of Prime, $F(1, 25) = 9.14$, MSe $= 0.35$, $p < .01$, $\eta_p^2 = .27$, with the N400 mean amplitude values smaller (less negative) for Consistent condition (mean $= -2.89$, SD $= 1.82$) than Inconsistent condition (mean $= -3.27$, SD $= 1.77$).

Unlike before, there was no main effect of Group, $F(1, 25) = 1.25$, MSe $= 10.62$, $p = .27$, $\eta_p^2 = .05$, with the DP group (mean $= -2.48$, SD $= 1.84$) not reliably different from the control group (mean $= -3.28$, SD $= 1.75$). As before, the main effect of CVS was not reliable, $F(1, 25) = 3.19$, MSe $= 1.57$, $p = .09$, $\eta_p^2 = .11$, but was tending in the direction of a smaller (less negative) N400 with active stimulation (mean $= -2.85$, SD $= 1.94$) compared to no stimulation (mean $= -3.30$, SD $= 1.63$). All other interactions were not significant, all $Fs(1, 25) \leq 0.59$, ps $\geq .45$.
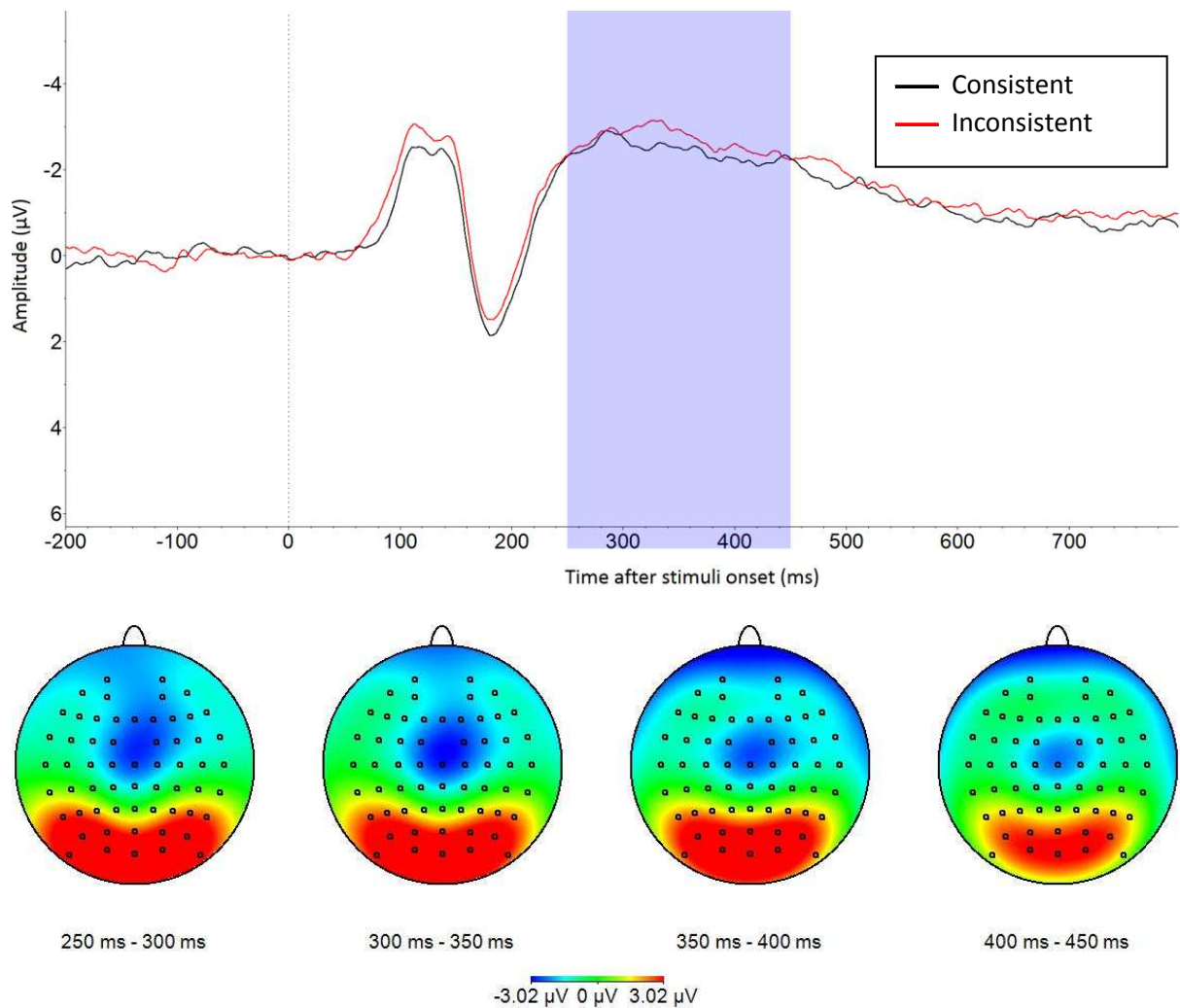
Figure 4.8: N400 waveform for the consistent (black) and inconsistent (red) conditions for the DP group, collapsed across the CVS variable. Topographical plots show the location of the N400 activity, and surrounding time-frames, for the inconsistent condition in the DP group.

Discussion (3)

The present N400 experiment was again a replication of the same experiment conducted in the previous chapter. Previously, the DP group were significantly less accurate at determining the famous face's nationality with no difference in reaction times. This

behavioural change mapped onto a smaller N400 component. In the present experiment, the behavioural differences were replicated, thereby showing consistency of impairment. However, the N400 was no longer reliably different from the controls'.

Whereas priming had no effect in the previous chapter, an effect was observed in the ERP data, with a larger N400 when a face was preceded by face of the same nationality. A priming effect was also observed in the RTs, but in the opposite direction, longer RTs in the consistent condition. A lack of priming in the previous chapter, combined with a small effect size of 20ms, subjective reports from the DPs (and indeed a majority of the controls) that the task was extremely difficult, and evidence of nationality priming being successful in previous literature, all suggest this result may be type I error.

Similar to the previous chapter, no reliable evidence emerged for effects related to CVS. It may be worth noting that in the RTs of the original study, and in the ERP of the replication study, CVS effects were approaching significance (shorter RTs, $p = .085$ and smaller N400, $p = .086$ respectively). While the role of chance cannot therefore be discounted, a regular pattern of this nature may indicate that a more sensitive design could reveal an effect that must be dismissed here.

General Reliability Results

Test-retest reliability

The previous section revealed that the majority of effects in the N170 and N250r tasks, but not in the N400 task, were replicated from Chapter Three. Importantly, all behavioural impairments were replicated which provides novel insight into the reliability of impairments in DP overtime.

Comparing ANOVAs is limited in how much it can reveal about reliability. To resolve this, correlations will be calculated to examine the relationship between sessions. Whereas ANOVAs compare differences between groups, correlations can explore the relationship between sessions, thereby providing a measure for reliability from one testing session to the next. In addition, while the ANOVAs revealed that group differences exist in both Chapters, the effect is defined in a binary way and lacks the quantitative sensitivity provided by correlations. For example, there was a consistent main effect of Group in both Chapters, yet accuracy for DPs in the familiarity task was 0.70 in the original and 0.74 in the replication and the ANOVA is unable to highlight this. Lastly, the ANOVA uses group mean scores and fails to capture individual differences. Hypothetically, if at retest half of the DP group performed worse and the other half performed better, the group mean remains the same despite the inconsistency of the group's impairments.

To assess the reliability of the experiments, correlations were first calculated between session 1 and session 2 (those used for Chapter Three) for the 20 controls only. This was repeated for each priming condition, and the average then collapsed across conditions. It should be noted that in doing this, the CVS variable was ignored. This is mostly because the variable is well counterbalanced but partly because of a lack of CVS main effects or simple main effects in all statistical tests except the accuracy data for the N250r experiment in sessions 3 and 4. It was therefore assumed that the effects of CVS would not impact the correlation of scores between sessions. Correlations were calculated for accuracy and RT data (see Table 4.1), and not ERP data. Consequently, the N170 task is excluded from this analysis because there was no behavioural measure.

Note that the Cronbach's alpha is not an appropriate statistic in these methodological designs because it assesses a correlation of scores across each specific trial, yet in these paradigms individual trials could not be aligned between participants. This was because not

all possible combinations were presented to each participant, but instead each particular trial was randomly constructed using the pool of available faces.

Table 4.1: Correlations in behavioural responses for the control group between Session 1 and Session 2.

Experiment	Measure	Priming Condition	r	p
Exp 2 (N250r)	Accuracy	SameImage	0.86	< .001
		SameID	0.78	< .001
		DiffID	0.84	< .001
		NonFamous	0.72	< .001
		Average	0.84	< .001
	Reaction time	SameImage	0.84	< .001
		SameID	0.90	< .001
		DiffID	0.83	< .001
		NonFamous	0.75	< .001
		Average	0.77	< .001
Exp 3 (N400)	Accuracy	Consistent	0.82	< .001
		Inconsistent	0.76	< .001
		Average	0.84	< .001
	Reaction time	Consistent	0.88	< .001
		Inconsistent	0.91	< .001
		Average	0.91	< .001

In the sample of 20 control participants, all of the correlations were highly significant, with the majority being very strong correlations (r > .8) and the rest as strong (r > .7; definitions

suggested by Evans, 1996). This implies that any weak correlations in the DP group are not because of intrinsic variability in the measure.

Building on the reliability of the experiments, it is now possible to explore whether the same reliability exists in the DP data. This can be achieved by first calculating the correlation between session 1 and session 2 and then comparing this to the same correlations obtained from the control group. The retest of the DP group also allows a correlation between session 2 and session 3 (a delay of one year) and a correlation between session 3 and session 4 (a delay of one week again). These correlations are presented in Table 4.2. Because there are three tests for each condition, the alpha region is divided by three, therefore only values where $p < .017$ should be considered significantly correlated (as marked in black font colour).

Table 4.2: Correlations in behavioural responses for the developmental prosopagnosia group between Session 1 and Session 2, between Session 2 and Session 3, and between Session 3 and Session 4. Grey font represents non-significant *at the Bonferonni adjusted α value for the three tests* (p > 0.17).

Experiment	Measure	Priming Condition	S1-S2 Correlation (n = 8)		S2-S3 Correlation (n = 7)		S3-S4 Correlation (n = 7)	
			r	p	r	p	r	p
Exp 2 (N250r)	Accuracy	SameImage	0.82	0.01	0.91	0.004	0.78	0.04
		SameID	0.85	0.01	0.80	0.03	0.74	0.06
		DiffID	0.71	0.05	0.68	0.09	0.47	0.29
		NonFamous	0.90	0.002	0.16	0.73	0.82	0.02
		Average	0.88	0.004	0.90	0.01	0.71	0.07
	Reaction time	SameImage	0.82	0.01	0.78	0.04	0.82	0.03
		SameID	0.78	0.02	0.89	0.01	0.82	0.03
		DiffID	0.87	0.01	0.84	0.02	0.95	0.001
		NonFamous	0.88	0.004	0.87	0.01	0.82	0.02
		Average	0.87	0.01	0.87	0.01	0.90	0.01
Exp 3 (N400)	Accuracy	Consistent	0.44	0.28	0.71	0.07	0.26	0.57
		Inconsistent	0.72	0.04	0.80	0.03	0.68	0.09
		Average	0.69	0.06	0.93	0.003	0.46	0.30
	Reaction time	Consistent	0.37	0.36	0.64	0.12	0.68	0.10
		Inconsistent	0.56	0.15	0.80	0.03	0.59	0.16
		Average	0.48	0.22	0.72	0.07	0.67	0.10

Starting with the N250r data, the test-retest reliability over one week from session 1 to session 2 is very strong (r > .8) for the average scores, and also comparatively high when split by each priming condition. Test-retest reliability over one year from session 2 to session 3 is also very strong in the average scores. However, the one year reliability in the accuracy data is considerable weaker when separating the data by each priming condition, and especially low for the non-famous condition (r = .16). While this may indicate differences in processing in direct (average scores) and indirect (priming) measures, the reliabilities in the RT data for priming conditions remain strong. The second short-term reliability assessment from session 3 to session 4 mostly shows correlations that did not reach significance, though the average RT score is very strong. The general message from the data is that reliability is typically very strong in the average scores, even over one year, with performance measured by RTs being the most reliable.

The N400 test-retest reliability tells a different story. In both short-term measures and the long-term measure, nearly all correlations fail to reach significance. The only exception is for the average accuracy score over one year.

It should be noted that the sample sizes of 7 and 8 lower the power of the design. This means that with n = 8, correlations of r < .71 will not reach significance, yet generally a correlation of this magnitude is considered strong. Furthermore, the family wise error rate correction means only correlations of r ≥ .80 will reach significance. While this should not change the message of the data, it does imply that a higher power study may be informative.

To summarise, the control group data demonstrate that the N250r and N400 tasks have good reliability. In comparison to these values, the DPs show very strong test-retest reliability in the N250r familiarity task over both one week and one year, but weaker reliability between sessions 3 and 4. In contrast, all but one correlation for the N400 semantic recall task failed to

reach significance except for average RT. While this demonstrates weaker reliability in this task, the small DP group may have led to an underestimation of the true reliability.

General Discussion

The aim of the present chapter was to assess the consistency of impairment in a group of participants with DP. This was achieved by replicating the experiments in Chapter Three one year later. Consistency of impairment was assessed by two methods: comparing ANOVA effects between chapters and correlating performance between sessions. In the original study, accuracy was impaired in the DP group when making judgements involving familiarity (N250r task) and semantic recollection (N400 task), but no group differences were observed in the face perception (N170) task. These effects were fully replicated in the present study, showing high consistency of impairment when tested one year later. The additional measure of test-retest reliability between each testing session was calculated to assess both short term (one-week) and long term (one-year) reliability. The control group demonstrated strong test-retest reliability confirming the reliability of the test instruments. Test-retest reliability for the DP group was strong in the N250r task but considerably weaker in the N400 task, though lack of power through low sample size may be underestimating these values.

A second aim was to re-assess the effects of CVS in the DP group. Subtle interactions were observed in Chapter Three whereby CVS had no direct effect but moderated the differences between priming conditions in the N250r task, and had a normalising effect on the N250r component between groups. Importantly, the present chapter reproduced the CVS with priming interaction in the N250r task and also found a main effect of CVS whereby stimulation increased the accuracy of familiarity with famous faces. Post-hoc comparisons of the interaction this time showed direct effects of stimulation in two of the face familiarity priming

conditions (SameID and DiffID). The present chapter found no further evidence of the normalising N250r ERP effect. A summary showing which statistical effects were replicated can be seen in Figure 4.9.

Measure	Effect	T1 & T2	T3 & T4
N170 Latency	Inversion	*	*
	Inversion×Object	*	*
	CVS×Hemisphere×Object×Group	*	ns
N170 Amplitude	Object	*	*
	Inversion×Object	ns	*
N250r Acc	Prime	*	*
	Group	*	*
	CVS	ns	*
	Prime×CVS	*	*
N250r RT	Prime	*	*
	Prime×CVS	*	*
N250r MeanAmp	Prime	*	*
	Group×CVS	*	ns
N400 Acc	Group	*	*
N400 RT	Prime	ns	*
N400 MeanAmp	Prime	ns	*
	Group	*	ns

Note: * = Statistically significant; ns = non-significant; T1 = Session one; T2 = Session two;

T4 = Session three; T4 = Session four

Figure 4.9: A summary of significant statistical effects replicated across Chapters Three and Four. All significant effects from the ANOVAs of Chapters Three and Four are included.

The temperature changes induced by the CVS device are easily discernible to the participants which makes it difficult to sham. When questioned, many of the participants reported being conscious of these temperature changes although they did also mention becoming accustomed to these overtime. The study design counterbalanced the order of stim/no-stim across participants, but one uncontrolled factor was participants' familiarity with the device which was greater during the replicatory study. The question therefore arises as to

whether the 'cleaner' pattern of effects seen in the replicatory study partly reflect this increased familiarity. I am unable to dismiss this possibility although note that the key effect involving stimulation was present in both experiments so is unaffected by this variable. In the future, one way in which to test for familiarity effects would be to administer the study twice in close succession and then probe the participant's sense of familiarity and tolerability with the device. If familiarity significantly constrains the effects of stimulation then any observed effects should correlate with these subjective ratings.

All of the DPs reported show significant memorial impairment, but Chapter 2 also observed perceptual impairments in one or two cases. If CVS is acting on only memorial processes then the presence of an unresolved perceptual deficit may obscure any beneficial effect. To explore this possibility, the DP (DP 2) showing the most severe impairment in the perceptual tasks of Chapter 2 was removed from the data set which was then re-analysed. This alteration removed the majority of reliable Group effects in the accuracy data of both N250r and N400 experiments in both Chapters 3 and 4, an effect that could as much reflect reduced statistical power than any underlying task by group interaction. Importantly, the removal of this participant did not diminish the stimulation effect and in fact gave rise to an additional effect, $F(1, 25) = 7.63$, MSe $= 14760.84$, $p < .05$, $\eta_p^2 = .23$, whereby RTs on the nationality semantic recall task became shorter during CVS (mean $= 1054.68$, SD $= 167.17$) compared to no stim (mean $= 1111.65$, SD $= 198.66$). This effect may suggest that the effect of CVS is stronger in those DPs who present without a perceptual deficit. The removal of this participant for explorative purposes had minimal effect on the CVS $\times$ Prime interaction, and therefore failed to reveal any insight regarding the mechanism of CVS effects. It is also worth noting that the removal of DP10 (the next most severe perceptual impairments) further amplified this effect, $F(1, 24) = 8.91$, MSe $= 14627.13$, $p < .01$, $\eta_p^2 = .27$. This emerging patterns warrants

further investigation in which the performance of a relatively equal sample of DPs with and without perceptual impairment are compared.

Replicability of Behavioural and ERP ANOVA Effects

When observing the patterns across the N170 experiment of both Chapters Three and Four, there is little compelling evidence to suggest the N170 is abnormal in the DP group. This directly supports the study hypothesis that the group show no evidence of face perception impairment. There is perhaps some reason to think that the design could have missed specific effects, such as not showing the typical effect of a larger N170 for upright faces compared to houses, but this needs to be compared to the large body of evidence in Chapter Two demonstrating typical face perception in this group.

In Experiment 2 which was a face familiarity task, low accuracy scores in the DP group compared to the control group were replicated. That said, there was no evidence that the group show atypical RTs, which was also replicated in both studies. There was also no evidence of atypical N250r mean amplitudes. The confirmation of low accuracy replicates original findings and shows consistency of this impairment over at least one year.

In Experiment 3 which was a semantic recollection task, the DP group produced lower accuracy scores but typical RTs compared to controls. The present chapter replicated both findings which not only supports the hypothesis that face memory is impaired in the DP group, but also shows that the impairment is consistent even after one year. However, the DP group also showed a significantly smaller N400 compared to the controls in the original study which was not replicated. While the data suggests the atypical N400 in the DP group is not consistent and fluctuates over time, this failure to replicate could instead be a consequence of type I error in the previous chapter, or type II error on the present chapter. Unfortunately, previous

literature cannot help here because no one has recorded the N400 from a DP participant at multiple long delays. So while this study provides the first evidence that the ERP maybe inconsistent in DP, the conclusion requires further support from additional testing.

The present chapter also discovers an effect of nationality priming in both the RTs and ERP that was not present in Chapter Three. However, while the N400 was smaller for trials with a prime of the same nationality (an expected effect), RTs were longer (the reverse of what is expected). Such varied effects of priming across different measures and chapters suggests the paradigm was unable to adequately prime nationality in either group.

Replicability of CVS ANOVA Effects

An additional study hypothesis was that CVS may improve face memory abilities. CVS had limited and un-replicated effects on the N170 and no reliable effects in the N400 experiment. The strongest evidence that improvement occurred was in the N250r experiment. In the original study, CVS modulated the differences between familiarity priming conditions. The pattern of effects was complicated and did not yield easily to exploration but in general suggested a beneficial effect when memory was needed. In this replication study, the pattern was again evident.

In the N250r accuracy data, the replicated pattern of effects was that CVS affected trials with familiar faces differently to trials with unfamiliar faces. As discovered in the present chapter, this was due to CVS increasing accuracy in trials with familiar faces and having no effect for unfamiliar faces. Likewise in the N250r RT data, the replicated effect was that CVS was affecting trials with familiar faces differently to unfamiliar faces. These patterns show remarkable consistency between test and retest (see Figure 4.10). Interestingly, these beneficial effects of CVS are not dependent on face recognition ability but benefit performance in both

the DPs and control group. There is now considerable evidence from the present thesis that CVS is having a differential effect with familiar faces. Because the difference between familiar and unfamiliar faces is that of a memory of a face representation, this implies that the locus of effect of CVS is face memory based. One alternative suggestion described earlier is that the CVS effect is merely a placebo and causing an increased bias to respond 'famous'. However, no reliable reduction in accuracy or increase in RT is seen in the NonFamous condition.
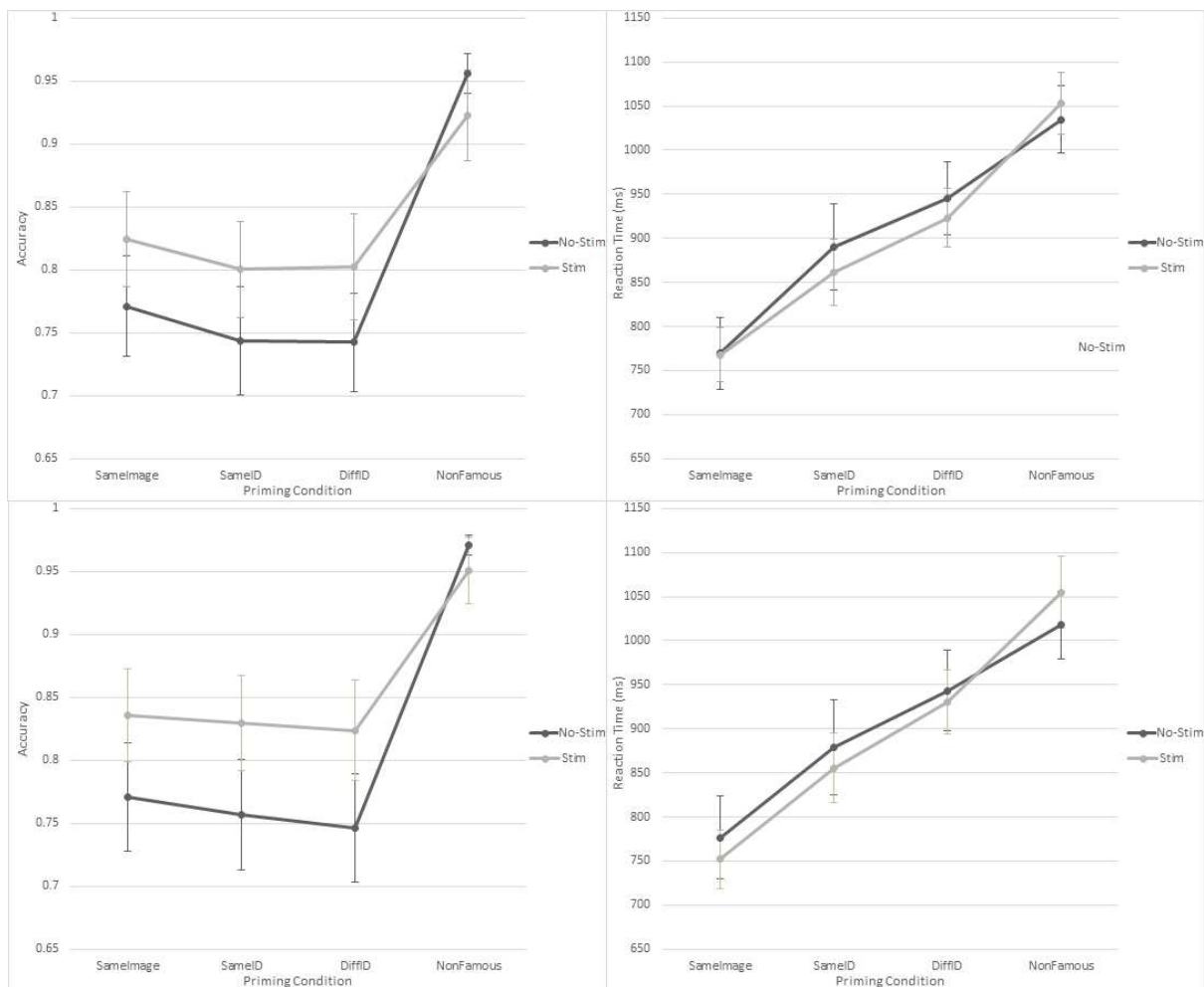
Figure 4.10: Replication of Figures 3.8, 3.9, 4.4 and 4.5 for ease of reference. These represent the means scores and standard error bars for the CVS × Prime interactions from the N250r task in accuracy scores (left panels) and RTs (right panels) in Chapter Three (top panels) and the present chapter (bottom panels).

In addition to the interaction between CVS and prime conditions, other evidence that CVS affects face recognition comes from an interaction with group in Chapter Three. Pairwise comparisons did not reveal the source of the interaction and the effect was not replicated in the present study. There is little evidence to explain why the effect was not replicable, but one suggestion might be that the effects of CVS are not highly consistent. The reason for this suggestion is because other CVS effects were not replicated, for example the beneficial effect of CVS on accuracy scores discovered in the present chapter ($\eta_p^2 = .30$) but not in Chapter Three ($\eta_p^2 = .09$).

Between Sessions Reliability in Accuracy Scores and RTs

To begin, correlations using only the control data indicated high external reliability in both the accuracy scores and RTs of the N250r and N400 tasks. Of the 16 correlations for each priming condition and overall average, all showed strong relationships (r > .7) and the majority can be described as very strong (r > .8). These values from a group of 20 unimpaired participants show that scores between two different testing sessions (separated by one week) were highly related, and this demonstrates high consistency of the tasks. The implication is that when the data from a group of impaired participants is then subjected to the same analysis, low correlations would therefore demonstrate low consistency associated with the participants and not the tasks themselves.

Test-retest correlations for the DPs in the N250r task reveal high reliability between sessions 1 and 2 (short-term), and also high reliability between sessions 2 and 3 (long-term), though the long-term reliability is surprisingly low for the non-famous condition. This same reliability is not observed for the N400 task, with only the correlation for average RTs between sessions 2 and 3 reaching significance. These data tell us that consistency of responses in DP

is high for decisions of face familiarity (i.e. those measured in the N250r familiarity task), but low when asked to explicitly recall information about a familiar face (i.e. during the N400 recollection task).

Between sessions 3 and 4 (also short-term) very few correlations reach significance for either task. This is a distinct difference to the other short-term reliability measure between sessions 1 and 2. It is not clear why the reliability is lower here, especially considering the high reliability demonstrated over one year. Interestingly, one might expect practise effects to emerge by the fourth testing session, yet this would cause an increase in reliability, the reverse of what was found. However, it should be noted that the non-significant r values are typically above .7 for the N250r task which is typically considered a strong correlation. This in turn suggests that a more sensitive design (i.e. more participants) would reveal more reliability than is reported in this data set.

One notion relating to reliability that the present study was unable to capture, is the possibility that not all DPs have consistent impairments. There is little reason to think that reliability is stable across DPs; perhaps some have more consistent responses than others. Future studies may therefore choose to employ a design whereby reliability can be assessed at an individual level.

In sum, the present chapter has replicated the three experiments from Chapter Three to assess whether the impairments of the DP group are consistent over time, and to assess whether their responses are reliable. Typical performance in the N170 task continues to support the hypothesis that impairments in this sample of DP are not perception based. Impaired accuracy with intact RTs in both the familiarity and recollection based memory tasks was replicated. Compared to baseline, the impairments are consistent over at least a one year delay. Group differences in the N400 component were not replicated and suggest transient changes in the

underlying EEG of the DP group. In addition to ANOVA comparisons, test-retest reliability measures from the control group demonstrate strong reliability in the experimental instruments. The test-retest correlations then reported for the DP group demonstrated strong reliability in the N250r task after delays of one week and even of one year. Reliability was considerably lower for the N400 task whereby only correlations for long-term accuracy scores were significant. Replicable effects of CVS were confined to the N250r familiarity task and affect familiar face targets differently to unfamiliar faces. There is now stronger evidence that CVS maybe affecting face familiarity memory processes.

CHAPTER V

General Discussion

Thesis summary

Developmental prosopagnosia can have significant negative consequences on an individual's daily life, yet up to 2.9 percent of the population may be affected (Bowles et al., 2009). The present thesis had four core study questions. First, it is unclear from existing data whether the source of impairment in developmental prosopagnosia is perceptual or memorial in origin. Four previous studies reported memorial impairments with intact perception in developmental prosopagnosia but each of these were methodologically flawed. Chapter Two addressed this shortcoming by administering a large battery of ten tests to 11 participants with developmental prosopagnosia, as defined by rigorous criteria based on the CFMT, a FFT and interview. Each task assessed a different aspect of face recognition and was thereby capable of detecting a wide range of impairments. While less than half of the participants showed impairment in a minority of perceptual tests, the majority showed no impairment in any of the eight perceptual tests, implying that the source of their impairment was memorial in nature.

Second, the thesis sought to better understand the underlying physiology of the condition and to corroborate impairments found at the behavioural level. Despite a growing body of evidence, it is not clear whether the N170, an ERP reflecting face perception, is atypical in developmental prosopagnosia. Post-perceptual ERPs N250r and N400 have not been reported in developmental prosopagnosia but both reflect different aspects of face memory (familiarity and recollection, respectively) that could be impaired. Chapter Three

administered three separate tasks which are known to induce the N170, N250r and N400. Corroborating the findings of Chapter Two, no differences were found between the N170 of participants with developmental prosopagnosia and an age-matched control group. In the N250r task (sensitive to face familiarity), accuracy scores for the developmental prosopagnosia group were significantly lower compared to the controls. In the N400 task (reflecting semantic recollection) the developmental prosopagnosia group again showed reduced accuracy scores, and this was corroborated with a significantly reduced N400 component.

Third, there are no current reports of the persistency of symptoms in developmental prosopagnosia. Without knowing the variability and permanence of the condition, data from any one testing session cannot be generalised to other testing sessions or other participants. From a clinical angle, 'diagnoses' of developmental prosopagnosia beyond the time of testing remains open to question. Chapter Four is the first report into the stability of impairment in developmental prosopagnosia. Importantly, high correlations in the control data indicated reliable experimental tasks. The majority of effects in the N170 ERP were replicated, indicating good consistency overtime. Effects in the N250r, including impaired accuracy in the DP group, were also replicated. Correlations of behavioural performance in the N250r task showed high reliability, over both one week and one year durations. Impaired accuracy in the N400 task was replicated, but an atypical N400 and nationality priming effects were not. The majority of correlations in performance between sessions in the N400 task were not significant, suggesting low reliability for this task in developmental prosopagnosia.

Fourth, current therapies for developmental prosopagnosia are particularly unsuccessful. Despite sometimes months of training, the benefits are either short lived, do not generalise to faces beyond the training set, or both (Bate & Bennetts, 2014; DeGutis et al., 2014). One method that has shown successful in allied groups is vestibular stimulation.

During the experiments of Chapters Three and Four, CVS was administered to all participants and was compared to performance with no stimulation. In the N250r task, CVS interacted with familiarity priming condition, in that increased accuracy and reduced RTs were observed only when the target face was famous. A normalising effect of CVS in the N250r component was found but not replicated. Lastly, CVS had no effect in the N400 task.

In sum, the present thesis shows that in this sample of participants with developmental prosopagnosia, the source of the impairment is memorial in nature. Every participant was impaired at both tasks involving memory in the test battery, and only negligibly impaired on perceptual tasks. In support of the source being memorial, consistent group differences were found in the memory based familiarity and semantic recollection tasks, though external reliability was particularly poor for the semantic recollection task for the developmental prosopagnosia group. Reliable effects of CVS were confined to the N250r task, whereby CVS increased accuracy and reduced RTs only in conditions where the target face was famous. These findings support previous reports of the benefits of vestibular stimulation in face recognition and pave the way for a promising form of relief from the condition.

Chapter Two Summary

A division of impairments commonly applied to acquired prosopagnosia is between apperceptive, relating to impairments in perception, and associative, relating to post-perceptual impairments in linking the face percept to previously encoded semantic information. This distinction is less in clear in developmental prosopagnosia. While a difficulty or inability to recognise a face is a failure in recalling a memory, it is not clear whether this is necessarily preceded by a perceptual failure. Most studies do report perceptual impairments, and most commonly this is an impairment in configural processing.

While this may suggest that developmental prosopagnosia is perceptual in origin, there are reports of developmental prosopagnosia in the absence of perceptual impairments. These reports, however, have significant weaknesses in detecting perceptual impairments. Two of the studies only used one measure of face perception (Dalrymple et al., 2014; McKone et al., 2011) and one did not report RTs (Lee et al., 2009). Chapter Two sought to address these weaknesses using a large sample of participants with developmental prosopagnosia.

Significant care was taken to confirm the presence of developmental prosopagnosia in each participant using two commonly employed objective tests (CFMT and FFT) and interview. The control group was selected based on matching age, gender and handedness. In total, participants were administered ten tasks, eight of which were designed to assess different aspects of face perception. As a whole, it is not clear which aspects of face recognition the test battery could not capture. Reliability (based on all 22 participants) was generally high, with lower values acknowledged for the expression recognition task and for the accuracy in the composite task and the inversion condition of the Jane task.

Following thorough statistical interrogation, some indication of perceptual impairments emerged. Reduced accuracy was in three experiments (global-local, Jane Upright, and viewpoint tasks) and increased RTs observed in subtests of two experiments (Jane Upright and Composite tasks) and the Jane inversion effects. However, these effects were subtle in comparison to the substantial impairments in memory based tasks. This is endorsed by the individual differences table (Figure 2.2) revealing 6 of the 11 participants with developmental prosopagnosia showed no evidence of impairment in any of the face perception tasks. Furthermore, the MANCOVA results showed that when face perception tasks were combined, a group effect still did not manifest.

The data from Chapter Two directly respond to the first study question. Although subtle perceptual impairments were observed, the majority of evidence suggests that the

significant face recognition impairments observed in the group are memorial in nature. This implies that face perception and face memory may be largely separable components comparable to those reported in acquired prosopagnosia. This dataset also speaks to the suggestions of heterogeneity in the condition. The sample sizes in the study are not conducive with strong conclusions of individual differences, but the data are suggesting that in those that do show perceptual impairment, the patterns of effects may be different for each person. Similar diversity in memory based face recognition tasks may also reveal heterogeneity in the memory based impairments.

Chapter Three Summary

The data from the behavioural test battery suggested that the source of face recognition impairments in this sample of participants with developmental prosopagnosia is memorial. To lend mechanistic specificity to this finding, tasks were chosen that assessed two divisions of memory; recollection and mere familiarity. To better understand the underlying physiology of the memory impairments, corroborative ERPs were also measured.

To confirm the absence of impairment at the perceptual stage, a task was administered to measure the N170, a face sensitive response understood to reflect structural encoding of a face (Eimer & McCarthy, 1999). Previous literature on this component in prosopagnosia is inconsistent, with papers showing it to be absent, reduced or intact. One group study (Towler et al., 2012) showed a typical N170 in developmental prosopagnosia to upright faces, but the same group failed to show an inversion effect.

In the N170 task, no evidence was found that the latency or amplitude differed between the developmental prosopagnosia group and the control group. Furthermore, an inversion effect in the latency of the N170 was not dependent on group. No evidence was

found that the underlying electrophysiology of face perception in developmental prosopagnosia is different to that of controls.

Following the N170, the first component sensitive to identity, and thereby showing a memorial element, is the N250. This component reflects access to stored face representations, such as when recognising a famous face. An allied component, the N250r, is present when a face stimulus is immediately repeated, but is larger when that face is of a celebrity. The component is present when a face is recognised, but is not associated with semantic information. Whereas the N250 is associated with long-term visual memory, the N250r is associated with working memory (but also long-term when the repeated face is of a celebrity). A small number of papers have investigated the N250 in developmental prosopagnosia, but no reports have been published regarding the N250r. It is not clear how closely the properties of the N250r match those of the N250, but reports have currently identified the N250 as being no different in developmental prosopagnosia compared to controls. Interestingly, the N250 component is sometimes still present when the identity is known to the individual but the face is not explicitly recognised, and may therefore represent covert recognition (Eimer et al., 2012), but again this effect may not be true of the N250r. To elicit the N250r, a repetition priming paradigm was employed. This has the additional benefit of providing direct (forced choice) and indirect (priming) measures. The task was to respond famous or not-famous to target faces, thereby providing data about identity familiarity without requiring access to semantic information.

In the N250r task, the developmental prosopagnosia group were significantly less accurate at detecting famous faces and with no differences in RTs. This implies that impairments in this sample of developmental prosopagnosia are detectable at the earliest stage that memory is required. The impaired accuracy was not corroborated with differences

in the N250r but this may be because the ERP is formed only from corrects responses. An effect of priming condition was found in every measure (accuracy scores, RTs and N250r mean amplitude) but because this did not interact with group this means the developmental prosopagnosia group showed typical priming. Therefore, while the developmental prosopagnosia group showed impairment in the accuracy scores of the forced choice task (a direct measure), no impairment was detected in face familiarity priming (an indirect measure).

Following the N250r, the first component sensitive to semantic information, an additional memory component, is the N400. This component is understood to reflect person identification and can be induced with faces as well as written names. The component can be enlarged (increased negativity) with semantic priming, such as preceding the target celebrity face with a different celebrity of the same occupation (Barrett & Rugg, 1989). This component, like the N250r, has not been reported in developmental prosopagnosia. One study has reported the N400 in acquired prosopagnosia and found it to be absent (as was the N170 in this patient). To elicit the N400, a repetition priming paradigm was used again. The prime was nationality because this semantic information has been successfully primed before (McNeill & Burton, 2002). The participants' task was to respond British or not-British to the target face.

In the N400 task, the developmental prosopagnosia group were significantly less accurate at determining the nationality of the celebrity face than the control group, with no differences in the RTs. This behavioural effect was corroborated with a smaller N400 component than the controls. This implies that as well as impairments in judging familiarity, this sample of developmental prosopagnosia also have impairments in associating semantic information with faces. It should be noted that no effect of priming was observed in accuracy

scores, RTs or ERP for this task, for both the developmental prosopagnosia and control groups, questioning the validity of the priming measure.

In sum, these three tasks provided physiological data for face perception, face familiarity and semantic recollection. The results of Chapter Three show no impairment in the N170 task, impaired accuracy in the familiarity task and impaired accuracy with corroborating ERP in the recollection task. Consistent with the findings of Chapter Two, these data suggest that the source if impairment in the developmental prosopagnosia group is not perceptual, but at the first point at which face memory is required. Because the N250r generated from famous faces reflects both working memory and long-term memory, it is impossible to suggest which of these is impaired in this sample of developmental prosopagnosia. The data suggest that the impairment observed in the N400 task may be merely a consequence of the preceding face memory failure. Without a match between stimuli and face memory, there is no signal to process for associated information.

Chapter Three also introduced a novel form of therapy. Current therapies for prosopagnosia have been largely unsuccessful. Six papers (four case studies) have attempted to improve face recognition skills in developmental prosopagnosia; one showed no benefits (Dalrymple et al., 2012), two did not generalise beyond the training set (Brunsdon et al., 2006; Schmalzl et al., 2008), and the remaining three did not report maintained benefits (Bate et al., 2014; DeGutis 2007, 2014). The twelve attempts of therapy in acquired prosopagnosia have met with more success, but only one has reported benefits beyond the training set, and this involved GVS (Wilkinson et al., 2005).

There are multiple reasons why vestibular stimulation might be beneficial in developmental prosopagnosia, especially if the origin is of a memorial nature: Associative

memory has been improved with CVS (Bächtold et al., 2001), memory (and hippocampal volume) is reduced in patients with bilateral vestibular loss (Brandt et al., 2005) and vestibular stimulation increases blood flow to brain areas associated with memory (Lobel et al., 1998; Vitte et al., 1996). However, there is also evidence that GVS directly improves face recognition in typically developing adults (Wilkinson et al., 2008), including increasing the N170 (Wilkinson et al., 2012), and, as mentioned above, in one profound case of acquired prosopagnosia (Wilkinson et al., 2005). Based on widespread cortical activation, clinical use, and to some extent, available infrastructure, CVS was chosen over GVS for this thesis.

The effect of CVS in the N170 task was limited to a subtle 4-way interaction whereby the N170 latency was longer for faces compared to houses in the right hemisphere of the control group during stimulation. Interpretation was withheld depending on its replication, which was not achieved. The largest effects of CVS were in the N250r task. An interaction between CVS and priming condition was significant in both the accuracy scores and RTs. Post-hoc comparisons revealed, in both accuracy and RTs, that the interaction was not driven by effects of CVS in individual priming conditions, but was instead related to the pattern of differences between priming conditions that differed with stimulation. A parsimonious explanation was not straightforward, but the suggestion was made that stimulation was differentially affecting behaviour in the non-famous condition compared to the others. Lastly, CVS also interacted with group in affecting the mean amplitude of the N250r. Post-hoc analysis did not provide useful information, but the means suggests a normalising effect, whereby the N250r of each group was more similar during stimulation. In the N400 task, no effect of CVS was detected. In sum, the effects of stimulation were mostly in the N250r task, but the exact locus of its effects was not clear.

Chapter Four Summary

The stability overtime of impairments in developmental prosopagnosia has not been previously established. Despite the growing literature base, only one case study of acquired prosopagnosia has employed a test-retest design. Consequently, it is not known whether the performance in face recognition in developmental prosopagnosia observed in any one testing session is stable or specific to that point in time. If the pattern of impairments varies across time, this has considerable consequences. Current approaches to 'diagnosis' of developmental prosopagnosia (impairment in both CFMT and FFT) do not accommodate variation over time. In addition, attempts to characterise the condition might be dependent on time of testing; impairments may be linked to failures in retrieving configural information at time 1, but, for example, linked to failures in retrieving associative memories at time 2. Indeed, the apparent heterogeneity of the condition in the literature may be a consequence of this. It is perhaps surprising, therefore, that tests of impairment consistency have not been reported.

To rectify the gap in the literature, Chapter Four reported a complete replication of the tasks used in Chapter Three, including the administration of CVS. Consistency was assessed by comparing the pattern of effects from the ANOVAs and using correlations between sessions. While ANOVAs provided a measure of within experiment differences, correlations provided a quantitative measure of between testing session differences, and are more sensitive to changes in each individuals' scores.

An illustration of the consistency of impairments overtime shown in the ANOVA results can be found in Figure 4.9. As in the original study reported in Chapter Three, no group differences in the N170 were detected in the replication. This further supports the results from Chapter Two that this sample of developmental prosopagnosia show little or no

indication of perceptual impairment. Even with an argument that the study was

underpowered, and as described in Chapter Two, it would be difficult to ascribe the profound

impairments in face recognition observed in these participants to such minor differences at

the perceptual level. Lastly, a 4-way interaction involving CVS and localised to the control

group was not replicated. It is not clear what can be unveiled from the interaction but it was

not a predicted effect and may suggest a type I error.

All effects in the N250r task that did not interact with CVS were replicated, including

the group effect of impaired accuracy. This provides confidence that the task was reliable

and, importantly, that the impairment in the developmental prosopagnosia group is consistent.

This is the first evidence of consistent impairment in developmental prosopagnosia. The

normalising effect of CVS on the N250r mean amplitude was not replicated, but interesting

CVS effects on the behavioural data were observed. A new main effect of CVS was

observed in the accuracy scores of the replication, but as in the original, this was qualified by

the priming condition. The Prime $\times$ CVS interaction in the RTs was also replicated. The

patterns of means in these interactions are remarkably similar across testing sessions (see

Figure 4.10 for direct comparisons). A suggestion in Chapter Three postulated that CVS may

be having a differential effect in trials with familiar faces compared to trials with unfamiliar

faces. Not only does the replication data continue to support this, but the post-hoc analyses

of the accuracy data confirm the suggestion that CVS improves accuracy scores in all priming

conditions that require memory of the familiarity of the prime face.

In the N400 task, the finding of impaired accuracy scores in the developmental

prosopagnosia group was replicated. This shows again that the impairment in developmental

prosopagnosia has some temporal stability. However, the reduced N400 in the

developmental prosopagnosia group observed in the original was not evident in the

replication. An effect of priming was observed in the RTs and corroborated by an ERP

effect, but only in the replication. It is not clear why priming was unsuccessful in the original study. Importantly, the effect did not interact with group, and therefore does not speak to the reliability of impairments in developmental prosopagnosia but may instead reflect poor reliability of the task itself. This was assessed using the correlation data summarised below.

Additionally to the comparisons of ANOVAs described above, correlations were calculated from the data of one session to the next. The control group's data was separated from the developmental prosopagnosia group's data to establish the test-retest reliability of the tasks themselves. Table 4.1 shows that correlations were either strong or very strong for all conditions (and the average scores) of both memory based tasks. This implies that low correlations in the developmental prosopagnosia group are a consequence of poor test-retest reliability and not the tasks themselves.

In the N250r task, the developmental prosopagnosia group showed very high test-retest reliability across a one week duration. The average scores are also very high over the one year duration, though this is weaker for some of the individual priming conditions, in particular the non-famous condition. Generally, there is high reliability in the N250r task across both short and long term.

In contrast, the N400 showed a considerably different pattern. Except for the accuracy scores across a one year duration, all other correlations were not significant. This implies that performance in the N400 at time 1 does not significantly predict the score at time 2, and therefore points to poor temporal stability of impairment in this task.

The final pattern to describe is the correlations between sessions 3 and 4. These sessions were separated by one week, identical to the duration between sessions 1 and 2. However, the strong correlations between sessions 1 and 2 revealed in the N250r task, was not repeated between sessions 3 and 4. It is likely that the participants are benefitting from

practise by session four, and therefore one might expect to find the highest correlations between sessions 3 and 4, yet the opposite pattern is observed. One other difference between these sets of correlations is the reduced sample size in sessions 3 and 4 from n = 8 to n = 7. While this is not a large change, it may have been enough for the correlations to not reach significance. Indeed the correlation values are mostly above .7 (the same cannot be said for the N400 equivalent which are all below .7) which is ordinarily categorised as a strong correlation.

In sum, the comparison of ANOVA effects between the original (Chapter Three) and the replication (Chapter Four) show high consistency in the N170 and N250r tasks, but less consistency in the N400 tasks. That said, impairments in accuracy scores for the developmental prosopagnosia group were replicated in both memory based tasks. The correlations substantiate this by showing high reliability in the N250r across both the short- and long-term, but the majority of correlations for the N400 failed to reach significance. Lastly, the increased accuracy and reduced RTs induced by CVS in trials with familiar faces in the N250r task was replicated, and bolsters the suggestion that CVS affects face representations in memory.

Main Theoretical Insights

One of the core aims of this thesis was to uncover whether the source of the impairments in developmental prosopagnosia was perceptual or memorial in nature. As the above summaries highlight, a consistent theme has emerged. First identified in the test battery of Chapter Two, the group have significant impairments in memory based tasks such as the CFMT and the FFT. Critically to the study question, there was minimal evidence of

impairment in perceptual tasks. The test battery included a thorough examination of face perception, and impairments were largely confined to a small number of sub-conditions. When exploring individual differences in Chapter Two, it emerged that less than half of the group showed signs of perceptual impairment, and even then the impairments were confined to not more than three specific conditions (which varied between those that did show perceptual impairment).

The same narrative of intact face perception and impaired face memory appeared in the two other empirical chapters. No impairments were detected in the N170 task, while accuracy scores in the two memory based tasks were significantly less compared to the controls. These effects were replicated. Corroborating evidence of a group difference in the N400 mean amplitude was present in Chapter Three but this was not replicated. This is the first report of the N250r and N400 ERPs in developmental prosopagnosia.

The present thesis does not argue that developmental prosopagnosia cannot arise from face perception impairments. Instead, it argues that impairments in face perception are not necessary for the condition to manifest. It also implies that the division of apperceptive and associative types, as seen in acquired prosopagnosia, may also be appropriate for developmental prosopagnosia. That said, the possibility exists for simultaneous but unrelated impairments in both face perception, and face memory domains. In which case, a future challenge would be in separating whether memorial deficits in associative developmental prosopagnosia are a consequence of impaired face perception or whether they are two unrelated and co-occurring impairments.

Chapters Three and Four used separate paradigms to explore two divisions of memory: familiarity and recollection (Tulving, 1985). The purpose of this was to assess what

different aspects of memory are impaired in developmental prosopagnosia. It emerged that

accuracy scores in both of these tasks were impaired compared to controls.

According to cognitive models of face recognition (Bruce & Young, 1986; Burton et

al., 1990; Ellis & Young, 1990), face familiarity is achieved with FRU activation. This stage

immediately follows structural encoding, but is before semantic information is retrieved at

the PIN stage. This suggests that the origin of impairment in this sample of developmental

prosopagnosia is at the earliest point at which memory is required, face familiarity. That is,

the encoded percept of the face is not being successfully matched with a FRU in memory. As

predicted by the models, the signal cannot progress to the PIN where semantic information is

accessed. The model therefore suggests that in the present thesis, either there are

simultaneous but unrelated irregularities impacting the N250r and N400, or that the

impairment in the N400 task is a consequence of a lost signal at the N250r stage.

Physiological models of face recognition (e.g. Haxby, Hoffman, & Gobbini, 2000)

propose that following processing of face features in the OFA, the signal then feeds to the

FFA for processing of invariant aspects if the face. It is during this stage that individual

identity is represented but before the extended system where semantic information is

accessed. Without successful processing in this core system, the signal does not progress to

the extended system where biographical information is accessed. The data in Chapters Three

and Four again fit this model and suggest that failed processing in the core system (as

demonstrated with impaired familiarity judgements) leads to failure to access the extended

system (as demonstrated by impaired nationality judgements). Again, the origin of

impairment in this sample of developmental prosopagnosia is at the first stage in which

memory is required.

Both the cognitive and physiological models claim that impairment in the N400 task

will result as a direct consequence of impairment in the N250r task. This is not to say that

impairments cannot exist at the N400 stage independently of performance in the N250r stage, but that impairment in the N250r task should always lead to impairment in the N400 task.

The stability of impairments overtime in developmental prosopagnosia has not been previously reported. Consequently, the validity of published findings was rested on the assumption that impairments do not vary over time. By comparing and correlating the results from separate testing sessions, the present thesis provides the first report of rest-retest reliability in developmental prosopagnosia. Firstly, the absence of impairment in the N170 task was replicated. For the N250r task, impairments in accuracy were replicated and high reliability was seen between sessions separated by both one week and one year. In the N400 task, impairments in accuracy scores (but not the atypical ERP response) were replicated, but the correlations of performance across different sessions suggested poor test-retest reliability across any duration.

While the test-retest reliability data suggest that impairments in tasks of face familiarity are relatively consistent, the proposal is less clear for judgements requiring access to semantic information. The primary impairment of reduced accuracy in the N400 was replicated, but the correlations did not corroborate this. One data model that would replicate the N400 reliability pattern is if the performance of half the developmental prosopagnosia group increased at time 2, and the other half decreased at time 2; the group effect remains, but the correlation will be low. This implies that the accuracy and RTs collected from a developmental prosopagnosia group in a semantic association task are less stable over time than controls. However, this variability in individual performance maybe averaged out by using a group. In contrast, individual consistency appears to be much higher for judgements of mere familiarity with a face.

There is little basis on which to explain why impairments in the N250r task are consistent, but impairments in the N400 task are inconsistent. Perhaps the consistent impairment in the N250r task suggests a stable underlying deficiency where its processes are persistently compromised. In contrast, perhaps the variance surrounding the impairments in the N400 task reflects the nature of the information it receives. That is, the information reaching the processes involved in semantic access is noisy or incomplete, and not represented in a consistent manner.

Despite consistent group impairment in the both the N250r and N400 memory based tasks across two Chapters, this was only corroborated with an atypical N400 ERP response in Chapter Three. The question arises as to where in the electrophysiological response lies the source of the impairment. If face recognition is impaired but the associated ERP is not, perhaps this is suggestive of performance being inhibited by processes beyond face recognition, such as those involved in decision making. However, while some argue that impaired performance need not necessarily be paired with a physiological difference (for a review see Wilkinson & Halligan, 2004), one explanation to the lack of corroborating evidence is that the ERPs are based only on the trials when responses were correct. This suggests that the ERP response may be transiently affected, whereby typical ERPs are associated with correct responses, but atypical ERPs are associated with incorrect responses. The present studies did not have enough incorrect trials to produce a waveform that could adequately test this theory. It does, however, show that typical ERPs are achievable in developmental prosopagnosia, and occurred relatively commonly in this sample.

Alternatively, perhaps key impairments do not manifest in ERP measures, and approaches such as phase synchrony (the phenomenon whereby neurons of common purpose, such as face perception, synchronise firing activity) would be more revealing (e.g. Dobel,

Junghöfer, & Gruber, 2011; Németh, Zimmer, Schweinberger, Vakli, & Kovács, 2014). Perhaps EEG is not the best approach for detecting the physiological differences in developmental prosopagnosia, and that a hemodynamic response, such as measured by functional magnetic resonance imaging (fMRI) or positron emission tomography (PET), may be a better metric for capturing differences in developmental prosopagnosia.

The effects of CVS during the face processing tasks of this thesis were most convincing in the N250r task. In Chapter Three an effect emerged that CVS may be reducing the difference in N250r mean amplitudes between the groups, though this was not supported in statistical tests, and was not replicated. However, the headline effect was observed in its interaction with priming condition. Post-hoc comparisons revealed (though only statistically significant in Chapter Four) that CVS was increasing accuracy scores and reducing RTs only in the conditions where the target face was famous. The possibility was raised that CVS was having a placebo effect making participants more likely to respond 'famous'. Alternatively, CVS may have affected the processing of the prime face, which was always famous, and again make the participants more likely to respond 'famous'. However, both of these suggestions must be rejected because of no significant decrease in accuracy in the non-famous condition. A more compelling suggestion is that CVS is affecting processing of the target face, perhaps by means of optimising the match between the percept of the face and the memory of the face representation. This suggestion is consistent with the literature highlighting the link between CVS and memory (Bächtold et al., 2001).

Interestingly, the interaction is present in both the developmental prosopagnosia and control group. This implies that CVS is exerting an ameliorative rather than a remediating effect; it is improving rather than restoring. This has significant consequences for its

application. If CVS, as suggested, improves the matching of a face percept with memory, it has potential applications for the police force, customs, or eye-witnesses.

Limitations and Future Directions

One key limitation, and one which in theory should be simple to rectify, is the sample size. It is common for studies in developmental prosopagnosia to use sample sizes of less than five participants, though larger studies have recently been published. Of the 30 individuals that responded to my adverts for participants with face recognition difficulties, 11 of them were 'diagnosed' as having developmental prosopagnosia. This constitutes a large group size for this condition, but would be considered grossly under powered in studies using typically developing participants, especially when including EEG measures. One of the dangers of using a small sample is that the results are less generalizable, and, as suggested from the N400 data in this thesis, inconsistent impairments can reduce reliability in small samples. A larger sample size always adds power to a design, but there are specific examples in this thesis where effects were subtle or where strong correlations were not significant. The data has provided important new findings and insight into developmental prosopagnosia, but smaller effects may have been missed.

The test battery of Chapter Two provided a wide range of perceptual tasks, but it was not exhaustive. Measures of face attractiveness or age may identify different impairments, and more sensitive measures of configural processing may be developed. The more inclusive a test battery can be, the more detailed the profile of impairments for each individual will be. A deeper understanding of the differences between individuals may provide explanations for the apparent heterogeneity of the condition. Measures not included in the present studies but advised for future studies, are ones of general object recognition skills and memory performance. While prosopagnosia is still present alongside visual agnosia and memory

disorders, the impairments of the group could not be ascribed solely to face stimuli. The present thesis was also unable to present data on covert recognition because not enough responses were incorrect. As the covert route to recognition diverges from the overt route that was assessed here, there was a stream of evidence that was missed.

The present thesis chose to divide memory into familiarity and recollection, yet many more divisions are possible. For example, it is not clear whether memory impairments (for those participants where perceptual impairments are absent) are related to the encoding or storage of a new face, or the recall of a face memory. Furthermore, whether the impaired memory for faces is limited to long-term memory, or whether it also persists in working memory. Indeed, future studies may wish to test whether these memory effects are confined to faces, or whether the stimulation poses clinical promise in conditions such as amnesia.

The effects of CVS on face memory are particularly exciting and prove to be a rewarding route to continue; the possibility of inducing super-recognition in participants has immediate applied value. Following evidence of its effect in face recognition, one possible future study would be to assess the longevity if these benefits. Stimulation was limited to 30 minutes in these studies, but a program of regular stimulation may provide larger and longer effects.

A future study may also wish to replicate the benefits of CVS, perhaps adopting an ABAB/BABA design in a control group to control for placebo effects and assess maintenance of the benefits. However, currently sham stimulation options, such as reduced temperature changes or smaller earpiece, either still cause significant stimulation or are immediately obvious to the participant. GVS has a significant advantage in this respect due to sub-sensory stimulation.

Perhaps the most appropriate route for the CVS research to continue, would be for randomised control trials (RCT) with a larger sample of developmental prosopagnosics. This

approach is the most rigorous method of assessing the cause and effect relationships in potential treatments, the core features of which are random and blinded allocation to active stimulation or sham groups. I would propose multiple measures that each assess subtly different aspects of face familiarity or other routes of access to face representations, as well as more general measures of memory.

Lastly, the present studies provide valuable information to guide 'diagnosis' of developmental prosopagnosia. Diagnosis must include a measure of face memory, and this should be based on forced choice decisions and not priming measures. Tasks should focus on testing familiarity judgements of faces because a diagnosis based on this should still be valid at least one year later. There is less durability overtime for tests that rely on semantic recollection and measures of ERP responses do not appear to provide reliable diagnostics of developmental prosopagnosia. Impaired performance should be based on low accuracy scores compared to a control baseline. One suggestion therefore, is to simply present the individual with a collection of images of celebrities and appropriate unfamiliar face stimuli and ask them to judge whether they are famous. Alternatively, ask the individual to categorise the collection into piles of famous and not famous. While these suggested tasks lack clinical validation, they provide quick and simple methods for making a recommendation of further assessment.

Conclusion

The source of impairment in developmental prosopagnosia can be memorial in nature. These deficits appear at the first point in which face memory is required and deny access to semantic information, yet the N250r and N400 ERP responses are not reliably different to controls. Impairments with judging face familiarity are consistent over time, but there is

more variation in judgements that require access to semantic information. This implies that diagnosis, pre- and post-therapy measures and study inferences should be based on tasks assessing familiarity and not semantic access. Caloric vestibular stimulation reliably improved behavioural performance, possibly by optimising face memory recall processes, but robust trials are needed. The present thesis contributed to knowledge by providing the first reported evidence confirming the N250r and N400 memory processes, and test-retest reliability in developmental prosopagnosia. Future studies may wish to capitalise on the demonstrated and wide spread possibilities of CVS.

REFERENCES

Ariel, R., & Sadeh, M. (1996). Congenital visual agnosia and prosopagnosia in a child: A case report. Cortex, 32(2), 221-240. doi:10.1016/S0010-9452(96)80048-7

Avidan, G., & Behrmann, M. (2008). Implicit familiarity processing in congenital prosopagnosia. Journal of Neuropsychology, 2(1), 141-164. doi:10.1348/174866407X260180

Avidan, G., & Behrmann, M. (2009). Functional MRI reveals compromised neural integrity of the face processing network in congenital prosopagnosia. Current biology, 19(13), 1146-1150. doi:10.1016/j.cub.2009.04.060

Avidan, G., & Behrmann, M. (2014). Impairment of the face processing network in congenital prosopagnosia. Frontiers in Bioscience, E6, 236-257. Retrieved from https://www.bioscience.org/

Avidan, G., Hasson, U., Malach, R., & Behrmann, M. (2005). Detailed exploration of face-related processing in congenital prosopagnosia: 2. Functional neuroimaging findings. Journal of Cognitive Neuroscience, 17(7), 1150-1167. doi:10.1162/0898929054475145

Avidan, G., Tanzer, M., & Behrmann, M. (2011). Impaired holistic processing in congenital prosopagnosia. Neuropsychologia, 49(9), 2541-2552. doi:10.1016/j.neuropsychologia.2011.05.002

Aw, S. T., Haslwanter, T., Fetter, M., & Dichgans, J. (2000). Three-dimensional spatial characteristics of caloric nystagmus. Experimental brain research, 134(3), 289-294. doi:10.1007/s002210000460

Bächtold, D., Baumann, T., Sándor, P.S., Kritos, M., Regard, M., & Brugger, P. (2001). Spatial- and verbal-memory improvement by cold-water caloric stimulation in healthy subjects. Experimental Brain Research, 136, 128-132. doi:10.1007/s002210000588

Bárány, R. (1906). Untersuchungen über den vom Vestibularapparat des Ohres reflektorisch

   ausgelösten rhythmischen Nystagmus und seine Begleiterscheinungen. Monatsschrift fuer

   Ohrenheilkunde und Laryngo-Rhinologie, 40, 193-297.

Barrett, S. E., & Rugg, M. D. (1989). Event-related potentials and the semantic matching of faces.

   Neuropsycholgia, 27(7), 913-922. doi:10.1016/0028-3932(89)90067-5

Barrett, S. E., Rugg, M. D., & Perrett, D. I. (1988). Event-related potentials and the matching of

   familiar and unfamiliar faces. Neuropsychologia, 26(1), 105-117. doi:10.1016/0028-

   3932(88)90034-6

Barrett, S.E. and Rugg, M.D. (1989) Event-related potentials and the semantic matching of faces.

   Neuropsychologia, 27(7), 913-922. doi:10.1016/0028-3932(89)90067-5

Barton, J. J. (2008). Structure and function in acquired prosopagnosia: lessons from a series of 10

   patients with brain damage. Journal of Neuropsychology, 2(1), 197-225.

   doi:10.1348/174866407X214172

Barton, J. J. S, & Cherkasova, M. (2003). Face imagery and its relation to perception and covert

   recognition in prosopagnosia. Neurology, 61(2), 220-225.

   doi:10.1212/01.WNL.0000071229.11658.F8

Barton, J. J. S. (2001). Covert recognition in acquired and developmental prosopagnosia. Neurology,

   57(7), 1161-1168. doi:10.1212/WNL.57.7.1161

Barton, J. J. S. (2008). Structure and function in acquired prosopagnosia: Lessons from a series of 10

   patients with brain damage. Journal of Neuropsychology, 2(1), 197-225.

   doi:10.1348/174866407X214172

Barton, J. J. S., Cherkasova, M. V., Press, D. Z., Intriligator, J. M., & O'Connor, M. (2003).

   Developmental prosopagnosia: A study of three patients. Brain and Cognition, 51(1), 12-30.

   doi:10.1016/S0278-2626(02)00516-X

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

Barton, J. J., & Cherkasova, M. (2003). Face imagery and its relation to perception and covert recognition in prosopagnosia. Neurology, 61(2), 220-225. doi:10.1212/01.WNL.0000071229.11658.F8

Barton, J. J., Cherkasova, M. V., & Hefter, R. (2004). The covert priming effect of faces in prosopagnosia. Neurology, 63(11), 2062-2068. doi:10.1212/01.WNL.0000145772.77040.24

Barton, J. J., Cherkasova, M., & O'Connor, M. (2001). Covert recognition in acquired and developmental prosopagnosia. Neurology, 57(7), 1161-1168. doi:10.1212/WNL.57.7.1161

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

Barton, J. J., Zhao, J., & Keenan, J. P. (2003). Perception of global facial geometry in the inversion effect and prosopagnosia. Neuropsychologia, 41(12), 1703-1711. doi:10.1016/S0028-3932(03)00115-5

Bate, S. (2012). Face Recognition and its Disorders. Hampshire, UK: Palgrave Macmillan.

Bate, S., & Bennetts, R. J. (2014). The rehabilitation of face recognition impairments: a critical review and future directions. Frontiers in Human Neuroscience, 8(491), 1-17. doi:10.3389/fnhum.2014.00491

Bate, S., & Tree, J. J. (2016). The Definition and Diagnosis of Developmental Prosopagnosia. The Quarterly Journal of Experimental Psychology, 1-8. doi:10.1080/17470218.2016.1195414

Bate, S., Bennetts, R., Parris, B. A., Bindemann, M., Udale, R., & Bussunt, A. (2015). Oxytocin increases bias, but not accuracy, in face recognition line-ups. Social cognitive and affective neuroscience, 10(7), 1010-1014. doi:10.1093/scan/nsu150

Bate, S., Cook, S. J., Duchaine, B., Tree, J. J., Burns, E. J., & Hodgson, T. L. (2014). Intranasal

    inhalation of oxytocin improves face processing in developmental prosopagnosia. Cortex, 50,

    55-63. doi:10.1016/j.cortex.2013.08.006

Bate, S., Haslam, C., Jansari, A., & Hodgson, T. L. (2009). Covert face recognition relies on

    affective valence in congenital prosopagnosia. Cognitive neuropsychology, 26(4), 391-411.

    doi:10.1080/02643290903175004

Bate, S., Haslam, C., Tree, J. J., & Hodgson, T. L. (2008). Evidence of an eye movement-based

    memory effect in congenital prosopagnosia. Cortex, 44(7), 806-819.

    doi:10.1016/j.cortex.2007.02.004

Bauer, R. M. (1984). Autonomic recognition of names and faces in prosopagnosia: A

    neuropsychological application of the guilty knowledge test. Neuropsychologia, 22(4), 457-

    469. doi:10.1016/0028-3932(84)90040-X

Bauer, R. M. (1986). The cognitive psychophysiology of prosopagnosia. In H. D. Ellis, M. A. Jeeves,

    F. Newcombe & A. Young (Eds.), Aspects of face processing (pp. 253-267). Netherlands:

    Martinus Nijhoff Publishers.

Begleiter, H., Porjesz, B., & Wang, W. (1995). Event-related brain potentials differentiate priming

    and recognition to familiar and unfamiliar faces. Electroencephalography and clinical

    neurophysiology, 94(1), 41-49. doi:10.1016/0013-4694(94)00240-L

Behrman, M., & Avidan, G. (2005). Congenital prosopagnosia: face-blind from birth. Trends in

    Cognitive Sciences, 9(4), 180-187. doi:10.1016/j.tics.2005.02.011

Behrmann, M., Avidan, G., Marotta, J. J., & Kimchi, R. (2005). Detailed exploration of face-related

    processing in congenital prosopagnosia: 1. Behavioral Findings. Journal of Cognitive

    Neuroscience, 17(7), 1130-1149. doi:10.1162/0898929054475154

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

Bense, S., Stephan, T., Yousry, T. A., Brandt, T., & Dieterich, M. (2001). Multisensory Cortical Signal Increases During Vestibular Galvanic Stimulation (fMRI). Journal of Neurophysiology, 85, 886-899. Retrieved from http://jn.physiology.org/

Bent, L. R., Bolton, P. S., & Macefield, V. G. (2006). Modulation of muscle sympathetic bursts by sinusoidal galvanic vestibular stimulation in human subjects. Experimental brain research, 174(4), 701-711. doi:10.1007/s00221-006-0515-6

Bentin, S., & Deouell, L. Y. (2000). Structural encoding and identification in face processing: ERP evidence for separate mechanisms. Cognitive Neuropsychology, 17(1-3), 35-55. doi:10.1080/026432900380472

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological Studies of Face Perception in Humans. Journal of Cognitive Neuroscience, 8(6), 551-565. doi:10.1162/jocn.1996.8.6.551

Bentin, S., DeGutis, J. M., D'Esposito, M., & Robertson, L. C. (2007). Too many trees to see the forest: Performance, event-related potential, and functional magnetic resonance imaging manifestations of integrative congenital prosopagnosia. Journal of Cognitive Neuroscience, 19(1), 132-146. doi:10.1162/jocn.2007.19.1.132

Bentin, S., Deouell, L. Y., & Soroker, N. (1999). Selective visual streaming in face recognition: Evidence from developmental prosopagnosia. Neuroreport, 10(4), 823-827. doi:10.1097/00001756-199903170-00029

Berninger, V. W., Nielson, K. H., Abbott, R. D., Wijsman, E., & Raskind, W. (2008). Writing problems in developmental dyslexia: Under-recognized and under-treated. Journal of School Psychology, 46(1), 1-21. doi:10.1016/j.jsp.2006.11.008

Besnard, S., Machado, M. L., Vignaux, G., Boulouard, M., Coquerel, A., Bouet, V., ... & Lelong-Boulouard, V. (2012). Influence of vestibular input on spatial and nonspatial memory and on hippocampal NMDA receptors. Hippocampus, 22(4), 814-826. doi:10.1002/hipo.20942

Bindemann, M., Attard, J., & Johnston, R. A. (2014). Perceived ability and actual recognition accuracy for unfamiliar and famous faces. Cogent Psychology, 1(1), 986903. doi:10.1080/23311908.2014.986903

Bindemann, M., Burton, A. M., Leuthold, H., & Schweinberger, S. R. (2008). Brain potential correlates of face recognition: Geometric distortions and the N250r brain response to stimulus repetitions. Psychophysiology, 45(4), 535-544. doi:10.1111/j.1469-8986.2008.00663.x

Bird, C. M., Papadopoulou, K., Ricciardelli, P., Rossor, M. N., & 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. doi:10.1348/014466503322528946

Black, R. D., Rogers, L. L., Ade, K. K., Nicoletto, H. A., Adkins, H. D, & Laskowitz, D. T. (in press). Non-Invasive Neuromodulation using Time-Verying Caloric Vestibular Stimulation. IEEE.

Bobes, M.A., Valdes-Sosa, M. and Olivares, E (1994) An ERP study of expectancy violation in face perception. Brain and Cognition, 26(1), 1-22. doi:10.1006/brcg.1994.1039

Bodamer, J. (1947). Die Prosop-Agnosie. Archiv für Psychiatrie und Nervenkrankheiten, 179(1-2), 6-53. doi:10.1007/BF00352849

Bottini, G., Paulesu, E., Gandola, M., Loffredo, S., Scarpa, P., Sterzi, R., ... & Vallar, G. (2005). Left caloric vestibular stimulation ameliorates right hemianesthesia. Neurology, 65(8), 1278-1283. doi:10.1212/01.wnl.0000182398.14088.e8

Bötzel, K., Schulze, S., & Stodieck, S. R. G. (1995). Scalp topography and analysis of intracranial sources of face-evoked potentials. Experimental Brain Research, 104(1), 135-143. doi:10.1007/BF00229863

Bowles, D. C., McKone, E., Dawel, A., Duchaine, B., Palermo, R., Schmalzl, L., ... & 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. doi:10.1080/02643290903343149

Brandt, T., Schautzer, F., Hamilton, D. A., Brüning, R., Markowitsch, H. J., Kalla, R., ... & Strupp,
M. (2005). Vestibular loss causes hippocampal atrophy and impaired spatial memory in
humans. Brain, 128(11), 2732-2741. doi:10.1093/brain/awh617

Brigham, J. C., & Malpass, R. S. (1985). The role of experience and contact in the recognition of
faces of own- and other-race persons. Journal of Social Issues, 41(3), 139-155.
doi:10.1111/j.1540-4560.1985.tb01133.x

Bruce, V. (1979). Searching for politicians: An information-processing approach to face recognition.
The Quarterly Journal of Experimental Psychology, 31(3), 373-395.
doi:10.1080/14640747908400734

Bruce, V., & Le Voi, M. E. (1983). Recognizing faces [and discussion]. Philosophical Transactions
of the Royal Society of London B: Biological Sciences, 302(1110), 423-436.
doi:10.1098/rstb.1983.0065

Bruce, V., & Young, A. (1986). Understanding face recognition. British Journal of Psychology,
77(3), 305-327. doi: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.
doi:10.1080/02643290500441841

Burgess, N. (2008). Spatial cognition and the brain. Annals of the New York Academy of Sciences,
1124(1), 77-97. doi:10.1196/annals.1440.002

Burns, E. J., Tree, J. J., & Weidemann, C. T. (2014). Recognition memory in developmental

prosopagnosia: electrophysiological evidence for abnormal routes to face recognition.

Frontiers in human neuroscience, 8(622), 1-15. doi:10.3389/fnhum.2014.00622

Burton, A. M., Bruce, V., & Johnston, R. A. (1990). Understanding face recognition with an

interactive activation model. British Journal of Psychology, 81(3), 361-380.

doi:10.1111/j.2044-8295.1990.tb02367.x

Burton, A. M., Young, A. W., Bruce, V., Johnston, R. A., & Ellis, A. W. (1991). Understanding

covert recognition. Cognition, 39(2), 129-166. doi:10.1016/0010-0277(91)90041-2

Caharel, S., Collet, K., & Rossion, B. (2015). The early visual encoding of a face (N170) is

viewpoint-dependent: A parametric ERP-adaptation study. Biological psychology, 106, 18-

27. doi:10.1016/j.biopsycho.2015.01.010

Caharel, S., d'Arripe, O., Ramon, M., Jacques, C., & Rossion, B. (2009). Early adaptation to

repeated unfamiliar faces across viewpoint changes in the right hemisphere: evidence from

the N170 ERP component. Neuropsychologia, 47(3), 639-643.

doi:10.1016/j.neuropsychologia.2008.11.016

Capgras, J. (1923). L'illusion des" sosies" dans un delire systematise chronique. Bulletin de la

Société clinique de médecine mentale, 11, 6-16. Retrieved from

http://www.biusante.parisdescartes.fr/

Capgras, J., & Reboul-Lachaux, J. (1994). L'Illusion des' sosies' dans un délire systématisé

chronique. [The illusion of doubles in chronic systematized delusions] History of Psychiatry,

5(17), 119-133. doi:10.1177/0957154X9400501709

Carey, S., De Schonen, S., & Ellis, H. D. (1992). Becoming a Face Expert [and Discussion].

Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences,

335(1273), 95-103. doi:10.1098/rstb.1992.0012

Chatterjee, G., & Nakayama, K. (2012). Normal facial age and gender perception in developmental

    prosopagnosia. Cognitive Neuropsychology, 29(5-6), 482-502.

    doi:10.1080/02643294.2012.756809

Coats, A. C. (1972). The sinusoidal galvanic body-sway response. ActaOto-Laryngologica, 74(3),

    155-162. doi:10.3109/00016487209128436

Cohen, B., Yakushin, S. B., & Holstein, G. R. (2012). What does galvanic vestibular stimulation

    actually activate? Frontiers in Neurology, 2, 90. doi:10.3389/fneur.2011.00090

Connolly, J.F., Byrne, J.M. and Dyman, C.A. (1995) Assessing adult receptive vocabulary with

    event-related potentials: An investigation of cross-modal and cross-form priming. Journal of

    Clinical and Experimental Neuropsychology, 17(4), 548-565.

    doi:10.1080/01688639508405145

Cooper, T. J., Harvey, M., Lavidor, M., & Schweinberger, S. R. (2007). Hemispheric asymmetries in

    image-specific and abstractive priming of famous faces: Evidence from reaction times and

    event-related brain potentials. Neuropsychologia, 45(13), 2910-2921.

    doi:10.1016/j.neuropsychologia.2007.06.005

Cotton, S. M., Crewther, D. P., & Crewther, S. G. (2005). Measurement error: implications for

    diagnosis and discrepancy models of developmental dyslexia. Dyslexia, 11(3), 186-202.

    doi:10.1002/dys.298

Courjon, J. H., Precht, W., & Sirkin, D. W. (1987). Vestibular nerve and nuclei unit responses and

    eye movement responses to repetitive galvanic stimulation of the labyrinth in the rat.

    Experimental Brain Research, 66(1), 41-48. doi:10.1007/BF00236200

Crawford, J. R., & Howell, D. C. (1998). Comparing an individual's test score against norms derived

    from small samples. The Clinical Neuropsychologist, 12(4), 482-486.

    doi:10.1076/clin.12.4.482.7241

Dalrymple, K. A., Corrow, S., Yonas, A., & Duchaine, B. (2012). Developmental prosopagnosia in

    childhood. Cognitive neuropsychology, 29(5-6), 393-418.

    doi:10.1080/02643294.2012.722547

Dalrymple, K. A., Garrido, L., & Duchaine, B. (2014). Dissociation between face perception and

    face memory in adults, but not children, with developmental prosopagnosia. Developmental

    Cognitive Neuroscience, 10, 10-20. doi:10.1016/j.dcn.2014.07.003

Dalrymple, K. A., Oruç, I., Duchaine, B., Pancaroglu, R., Fox, C. J., Iaria, G., . . . Barton, J. J. S.

    (2011). The anatomic basis of the right face-selective N170 in acquired prosopagnosia: A

    combined ERP/fMRI study. Neuropsychologia, 49(9), 2553-2563.

    doi:10.1016/j.neuropsychologia.2011.05.003

Damasio, A. R., Tranel, D., & Damasio, H. (1990). Face agnosia and the neural substrates of

    memory. Annual review of neuroscience, 13(1), 89-109.

    doi:10.1146/annurev.ne.13.030190.000513

Damasio, A., Damasio, H., & van Hoessen, G. (1982). Prosopagnosia: Anatomical basis and

    behavioral mechanisms. Neurology, 32(4), 331-341. doi:10.1212/WNL.32.4.331

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

    and processing impairments. Frontiers in Bioscience, Elite, 6, 159-174. doi:10.2741/E699

De Gutis, J. M., Bentin, S., Robertson, L. C., & D'Esposito, M. (2007). Functional plasticity in

    ventral temporal cortex following configural training with faces in a congenital

    prosopagnosic. Journal of Cognitive Neuroscience, 9(11), 1790-1820.

    doi:10.1162/jocn.2007.19.11.1790

De Haan, E. H. F., & Campbell, R. (1991). A fifteen year follow-up of a case of developmental

    prosopagnosia. Cortex, 27(4), 489-509. doi:10.1016/S0010-9452(13)80001-9

De Haan, E. H., Bauer, R. M., & Greve, K. W. (1992). Behavioural and physiological evidence for covert face recognition in a prosopagnosic patient. Cortex, 28(1), 77-95. doi:10.1016/S0010-9452(13)80167-0

De Haan, E. H., Young, A. W., & Newcombe, F. (1991). Covert And Overt recognition in prosopagnosia. Brain, 114(6), 2575-2591. doi:10.1093/brain/114.6.2575

De Renzi, E. (1986). Current issues on prosopagnosia. In H. D. Ellis (Ed.), Aspects of Face Processing (pp. 243-252). Dordrecht, Netherlands: Nijhoff.

De Renzi, E., Faglioni, P., Grossi, D., & Nichelli, P. (1991). Apperceptive and associative forms of prosopagnosia. Cortex, 27(2), 213-221. doi:10.1016/S0010-9452(13)80125-6

De Renzi, E., Perani, D., Carlesimo, G. A., Silveri, M. C., & Fazio, F. (1994). Prosopagnosia can be associated with damage confined to the right hemisphere: An MRI and PET study and a review of the literature. Neuropsychologia, 32(8), 893-902. doi:10.1016/0028-3932(94)90041-8

DeGutis, J. M., Bentin, S., Robertson, L. C., & D'Esposito, M. (2007). Functional plasticity in ventral temporal cortex following cognitive rehabilitation of a congenital prosopagnosic. Journal of Cognitive Neuroscience, 19(11), 1790-1802. doi:10.1162/jocn.2007.19.11.1790

DeGutis, J. M., Chiu, C., Grosso, M. E., & Cohan, S. (2014). Face processing improvements in prosopagnosia: successes and failures over the last 50 years. Frontiers in human neuroscience, 8(561), 1-14. doi:10.3389/fnhum.2014.00561

DeGutis, J., Cohan, S., & Nakayama, K. (2014). Holistic face training enhances face processing in developmental prosopagnosia. Brain, 137, 1781-1798. doi:10.1093/brain/awu062

DeGutis, J., DeNicola, C., Zink, T., McGlinchey, R., & Milberg, W. (2011). Training with own-race faces can improve processing of other-race faces: Evidence from developmental prosopagnosia. Neuropsychologia, 49(9), 2505-2513. doi:10.1016/j.neuropsychologia.2011.04.031

DeGutis, J., Wilmer, J., Mercado, R. J., & Cohan, S. (2013). Using regression to measure holistic face processing reveals a strong link with face recognition ability. Cognition, 126, 87-100. doi:10.1016/j.cognition.2012.09.004

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(5), 597-612. doi:10.1016/j.neuropsychologia.2003.10.008

Dennett, H. W., McKone, E., Tavashmi, R., Hall, A., Pidcock, M., Edwards, M., & Duchaine, B. (2012). 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. doi:10.3758/s13428-011-0160-2

Diamond, R., & Carey, S. (1986). Why faces are and are not special: an effect of expertise. Journal of Experimental Psychology: General, 115(2), 107-117. doi:10.1037/0096-3445.115.2.107

Diana, R. A., Reder, L. M. (2006). Models of recognition: A review of arguments in favor of a dual-process account. Psychonomic Bulletin & Review, 13(1), 1-21. doi: 10.3758/BF03193807

Díaz, F., Lindín, M., Galdo-Alvarez, S., Facal, D., & Juncos-Rabadán, O. (2007). An event-related potentials study of face identification and naming: The tip-of-the-tongue state. Psychophysiology, 44(1), 50-68. doi:10.1111/j.1469-8986.2006.00483.x

Dingle, K. J., Duchaine, B. C., & Nakayama, K. (2005). A new test for face perception. Journal of Vision, 5(8), 40. doi:10.1167/5.8.40

Dinkelacker, V., Grüter, M., Klaver, P., Grüter, T., Specht, K., Weis, S., ... & Fernandez, G. (2011). Congenital prosopagnosia: multistage anatomical and functional deficits in face processing circuitry. Journal of neurology, 258(5), 770-782. doi:10.1007/s00415-010-5828-5

Dobel, C., Junghöfer, M., & Gruber, T. (2011). The role of gamma-band activity in the representation of faces: reduced activity in the fusiform face area in congenital prosopagnosia. PloS one, 6(5), e19550. doi:10.1371/journal.pone.0019550

Dodson, M. J. (2004). Vestibular stimulation in mania: a case report. Journal of Neurology, Neurosurgery & Psychiatry, 75(1), 168-169. Retrieved from http://jnnp.bmj.com/

Drouler, O., & Adil, S. (2015). Could Face Presence In Print Ads Influence Memorization? The Journal of Applied Business Research, 31(4), 1403–1408. doi:10.19030/jabr.v31i4.9326

Duchaine, B. C. (2000). Developmental prosopagnosia with normal configural processing. Cognitive Neuroscience and Neuropsychology, 11(1), 79-83. doi:10.1097/00001756-200001170-00016

Duchaine, B. C., Parker, H., & Nakayama, K. (2003). Normal recognition of emotion in a Prosopagnosic. Perception, 32, 827-838. doi:10.1068/p5067

Duchaine, B. C., Yovel, G., Butterworth, E .J., & Nakayama, K. (2006). Prosopagnosia as an impairment to face-specific mechanisms: Elimination of the alternative hypotheses in a developmental case. Cognitive Neuropsychology, 23(5), 714-747. doi:10.1080/02643290500441296

Duchaine, B., & Nakayama, K. (2005). Dissociations of Face and Object Recognition in Developmental Prosopagnosia. Journal of Cognitive Neuroscience, 17(2), 1-13. doi: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. doi:10.1016/j.neuropsychologia.2005.07.001

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. doi:10.1080/02643290701380491

Duchaine, B., Yovel, G., & Nakayama, K. (2007). No global processing deficit in the Navon task in 14 developmental prosopagnosics. Social Cognitive and Affective Neuroscience, 2(2), 104-113. doi:10.1093/scan/nsm003

Dunn, J. C. (2004). Remember–know: A matter of confidence. Psychological Review, 111(2), 524-542. doi:10.1037/0033-295X.111.2.524

Eimer, M. (2000a). Effects of face inversion on the structural encoding and recognition of faces: Evidence from event-related brain potentials. Cognitive Brain Research, 10(1), 145-158. doi:10.1016/S0926-6410(00)00038-0

Eimer, M. (2000b). Event-related brain potentials distinguish processing stages involved in face perception and recognition. Clinical neurophysiology, 111(4), 694-705. doi:10.1016/S1388-2457(99)00285-0

Eimer, M. (2011). The Face-Sensitive N170 Component of the Event-Related Brain Potential. In G. Rhodes, A. Calder, M. Johnson & J. V. Haxby (Eds.), Oxford Handbook of Face Perception (pp. 329-344). UK: Oxford University Press.

Eimer, M., & McCarthy, R. A. (1999). Prosopagnosia and structural encoding of faces: Evidence from event-related potentials. Cognitive Neuroscience, 10(2), 255-259. doi:10.1097/00001756-199902050-00010

Eimer, M., Gosling, A., & Duchaine, B. (2012). Electrophysiological markers of covert face recognition in developmental prosopagnosia. Brain, 135(2), 542-554. doi:10.1093/brain/awr347

Ekman, P., & Friesen, W. V. (1976). Pictures of facial affect. Palo Alto, CA: Consulting Psychology Press.

Ellis, H. D. (1981). Theoretical aspects of face recognition. In G. Davies, H. Ellis & J. Shepherd (Eds.), Perceiving and remembering faces (pp. 171-197). London: Academic Press.

Ellis, H. D. (1983). The role of the right hemisphere in face perception. In A. Young (Ed.), Functions of the right cerebral hemisphere (pp. 33-64). London: Academic Press.

Ellis, H. D., & Florence, M. (1990). Bodamer's (1947) paper on prosopagnosia. Cognitive Neuropsychology, 7(2), 81-105. doi:10.1080/02643299008253437

Ellis, H. D., & Lewis, M. B. (2001). Capgras delusion: a window on face recognition. Trends in cognitive sciences, 5(4), 149-156. doi:10.1016/S1364-6613(00)01620-X

Ellis, H. D., Young, A. W., Quayle, A. H., & De Pauw, K. W. (1997). Reduced autonomic responses to faces in Capgras delusion. Proceedings of the Royal Society of London B: Biological Sciences, 264(1384), 1085-1092. doi:10.1098/rspb.1997.0150

Ellis, H., & Young, A. (1988). Training in face-processing skills for a child with acquired prosopagnosia. Developmental Neuropsychology, 4(4), 283-294. doi:10.1080/87565648809540412

Engst, F. E., Martín0Loeches, M., & Sommer, W. (2006). Memory systems for structural and semantic knowledge of faces and buildings. Brain Research, 1124(1), 70-80. doi:10.1016/j.brainres.2006.09.038

Esins, J., Schultz, J., Stemper, C., Kennerknecht, I., & Bülthoff, I. (2016). Face Perception and Test Reliabilities in Congenital Prosopagnosia in Seven Tests. i-Perception, 7(1), 1-37. doi:10.1177/2041669515625797

Evans, J. D. (1996). Straightforward statistics for the behavioral sciences. Pacific Grove, CA: Brooks/Cole Publishing.

Evans, J. J., Heggs, A. J., Antoun, N., & Hodges, J. R. (1995). Progressive prosopagnosia associated with selective right temporal lobe atrophy: A new syndrome? Brain, 118(1), 1-13. doi:10.1093/brain/118.1.1

Faerber, S. J., Kaufmann, J. M., & Schweinberger, S. R. (2015). Early temporal negativity is sensitive to perceived (rather than physical) facial identity. Neuropsychologia, 75, 132-142. doi:10.1016/j.neuropsychologia.2015.05.023

Fasold, O., von Brevern, M., Kuhberg, M., Ploner, C. J., Villringer, A., Lempert, T., & Wenzel, R. (2002). Human Vestibular Cortex as Identified with Caloric Stimulation in Functional Magnetic Resonance Imaging. NeuroImage, 17(3), 1384-1393. doi:10.1006/nimg.2002.1241

Ferre, E. R., Sedda, A., Gandola, M., & Bottini, G. (2011). How the vestibular system modulates tactile perception in normal subjects: a behavioural and physiological study. Experimental brain research, 208(1), 29-38. doi:10.1007/s00221-010-2450-9

Fisher, R. A. (1921). On the "Probable Error" of a Coefficient of Correlation Deduced from a Small Sample. Metron, 1, 3-32. Retrieved from http://link.springer.com/journal/40300

Fitzpatrick, R. C., & Day, B. L. (2004). Probing the human vestibular system with galvanic stimulation. Journal of Applied Physiology, 96, 2301-2316. doi:10.1152/japplphysiol.00008.2004

Francis, D., Riddoch, M. J., & Humphreys, G. W. (2002). "Who's that girl?" Prosopagnosia, person-based semantic disorder, and the reacquisition of face identification ability. Neuropsychological Rehabilitation, 12(1), 1-26. doi:10.1080/09602010143000158

Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. Perception, 29, 159-170. doi:10.1068/p3012

Friberg, L., Olsen, T. S., Roland, P. E., Paulson, O. B., & Lassen, N. A. (1985). Focal increase of blood flow in the cerebral cortex of main during vestibular stimulation. Brain, 108(3), 609-623. doi:10.1093/brain/108.3.609

Furl, N., Garrido, L., Dolan, R. J., Driver, J., & Duchaine, B. (2011). Fusiform gyrus face selectivity relates to individual differences in facial recognition ability. Journal of Cognitive Neuroscience, 23(7), 1723-1740. doi:10.1162/jocn.2010.21545

Gainotti, G., & Marra, C. (2011). Differential contribution of right and left temporo-occipital and anterior temporal lesions to face recognition disorders. Frontiers in Human Neuroscience, 5(55), 1-11. doi:10.3389/fnhum.2011.00055

Garrido, L., Duchaine, B., Nakayama, K. (2008). Face detection in normal and prosopagnosic individuals. Journal of Neuropsychology, 2(1), 119-140. doi:10.1348/174866407X246843

Garrido, L., Furl, N., Draganski, B., Weiskopf, N., Stevens, N., Tan, G. C.-Y., . . . Duchaine, B. (2009). Voxel-based morphometry reveals reduced grey matter volume in the temporal cortex of developmental prosopagnosics. Brain, 132(12), 3443-3455. doi:10.1093/brain/awp271

Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: Exploring mechanisms for face recognition. Vision research, 37(12), 1673-1682. doi:10.1016/S0042-6989(96)00286-6

Gauthier, I., Behrmann, M., & Tarr, M. J. (1999). Can face recognition really be dissociated from object recognition? Journal of Cognitive Neuroscience, 11(4), 349-370. doi:10.1162/089892999563472

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. Nature Neuroscience, 3(2), 191-197. doi:10.1038/72140

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. Journal of Cognitive Neuroscience, 12(3), 495-504. doi:10.1162/089892900562165

George, N., Evans, J., Fiori, N., Davidoff, J., & Renault, B. (1996). Brain events related to normal and moderately scrambled faces. Cognitive Brain Research, 4(2), 65-76. doi:10.1016/0926-6410(95)00045-3

Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. Neuropsychologia, 45(1), 32-41. doi:10.1016/j.neuropsychologia.2006.04.015

Goren, C. C., Sarty, M., & Wu, P. Y. (1975). Visual following and pattern discrimination of face-like

   stimuli by newborn infants. Pediatrics, 56(4), 544-549. Retrieved from

   http://pediatrics.aappublications.org/

Gosling, A., & Eimer, M. (2011). An event-related brain potential study of explicit face recognition.

   Neuropsychologia, 49, 2736-2745. doi:10.1016/j.neuropsychologia.2011.05.025

Grueter, M., Grueter, T., Bell, V., Horst, J., Laskowski, W., Sperling, K., … Kennerknecht, I. (2007).

   Hereditary prosopagnosia: The first case series. Cortex, 43(6), 734-749. doi:10.1016/S0010-

   9452(08)70502-1

Grunwald, T., & Jurthen, M. (2006). Novelty detection and encoding for declarative memory within

   the human hippocampus. Clinical EEG and Neuroscience, 37(4), 309-314.

   doi:10.1177/155005940603700408

Grüsser, O. J., Pause, M., & Schreiter, U. (1990). Localization and responses of neurones in the

   parieto-insular vestibular cortex of awake monkeys (Macaca fascicularis). The Journal of

   physiology, 430(1), 537. doi:10.1113/jphysiol.1990.sp018306

Grüter, T., Grüter, M., & Carbon, C. C. (2008). Neural and genetic foundations of face recognition

   and prosopagnosia. Journal of Neuropsychology, 2(1), 79-97.

   doi:10.1348/174866407X231001

Guillaume, C., Guillery-Girard, B., Chaby, L., Lebreton, K., Hugueville, L., Eustache, F., & Fiori, N.

   (2009). The time course of repetition effects for familiar faces and objects: an ERP study.

   Brain Research, 1248, 149-161. doi:10.1016/j.brainres.2008.10.069

Hadjikhani, N., & De Gelder, B. (2002). Neural basis of prosopagnosia: An fMRI study. Human

   brain mapping, 16(3), 176-182. doi:10.1002/hbm.10043

Halgren, E., Dale, A. M., Sereno, M. I., Tootell, R. B., Marinkovic, K., & Rosen, B. R. (1999).

   Location of human face-selective cortex with respect to retinotopic areas. Human Brain

   Mapping, 7(1), 29-37. doi:10.1002/(SICI)1097-0193(1999)7:1<29::AID-HBM3>3.0.CO;2-R

Halgren, E., Raij, T., Marinkovic, K., Jousmäki, V., & Hari, R. (2000). Cognitive response profile of the human fusiform face area as determined by MEG. Cerebral cortex, 10(1), 69-81. doi:10.1093/cercor/10.1.69

Harris, A. M., Duchaine, B. C., & Nakayama, K. (2005). Normal and abnormal face selectivity of the M170 response in developmental prosopagnosics. Neuropsychologia, 43(14), 2125-2136. doi:10.1016/j.neuropsychologia.2005.03.017

Harris, A., & Nakayama, K. (2008). Rapid adaptation of the M170 response: importance of face parts. Cerebral Cortex, 18(2), 467-476. doi:10.1093/cercor/bhm078

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. Trends in cognitive sciences, 4(6), 223-233. doi:10.1016/S1364-6613(00)01482-0

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

Herzmann, G., Danthiir, V., Schacht, A., Sommer, W., & Wilhelm, O. (2008). Toward a comprehensive test battery for face cognition: Assessment of the tasks. Behavior Research Methods, 40(3), 840-857. doi:10.3758/BRM.40.3.840

Hirstein, W., & Ramachandran, V. S. (1997). Capgras syndrome: a novel probe for understanding the neural representation of the identity and familiarity of persons. Proceedings of the Royal Society of London B: Biological Sciences, 264(1380), 437-444. doi:10.1098/rspb.1997.0062

Holcomb, P.J. (1993) Semantic priming and stimulus de-gradation: Implications for the role of the N400 in language processing. Psychophysiology, 30(1), 47-61. doi:10.1111/j.1469-8986.1993.tb03204.x

Hole, G. J. (1994). Configurational factors in the perception of unfamiliar faces. Perception, 23(1), 65-74. doi:10.1068/p230065

Humphreys, K., Avidan, G., & Behrmann, M. (2007). A detailed investigation of facial expression processing in congenital prosopagnosia as compared to acquired prosopagnosia. Experimental Brain Research, 176(2), 356-373. doi:10.1007/s00221-006-0621-5

Ishai, A., Haxby, J. V., & Ungerleider, L. G. (2002). Visual Imagery of Famous Faces: Effects of Memory and Attention Revealed by fMRI. NeuroImage, 17, 1729-1741. doi:10.1006/nimg.2002.1330

Itier, R. J., & Taylor, M. J. (2002). Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: a repetition study using ERPs. Neuroimage, 15(2), 353-372. doi:10.1006/nimg.2001.0982

Itier, R. J., & Taylor, M. J. (2004a). Effects of repetition and configural changes on the development of face recognition processes. Developmental Science, 7(4), 469-487. doi:10.1111/j.1467-7687.2004.00367.x

Itier, R. J., & Taylor, M. J. (2004b). Effects of repetition learning on upright, inverted and contrast-reversed face processing using ERPs. Neuroimage, 21(4), 1518-1532. doi:10.1016/j.neuroimage.2003.12.016

Itier, R. J., & Taylor, M. J. (2004c). Face recognition memory and configural processing: a developmental ERP study using upright, inverted, and contrast-reversed faces. Journal of cognitive neuroscience, 16(3), 487-502. doi:10.1162/089892904322926818

Itier, R. J., & Taylor, M. J. (2004d). N170 or N1? Spatiotemporal differences between object and face processing using ERPs. Cerebral Cortex, 14(2), 132-142. doi:10.1093/cercor/bhg111

Itier, R. J., Alain, C., Sedore, K., & McIntosh, A. R. (2007). Early face processing specificity: It's in the eyes! Journal of Cognitive Neuroscience, 19(11), 1815-1826. doi:10.1162/jocn.2007.19.11.1815

Itier, R. J., Latinus, M., & Taylor, M. J. (2006). Face, eye and object early processing: what is the face specificity? Neuroimage, 29(2), 667-676. doi:10.1016/j.neuroimage.2005.07.041

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. Journal of Memory & Language, 30(5), 513-541. doi:10.1016/0749-596X(91)90025-F

Jaeschke, R., Singer, J., & Guyatt, G. H. (1989). Measurement of health status: Ascertaining the minimal clinically important difference. Controlled Clinical Trials, 10(4), 407-415. doi:10.1016/0197-2456(89)90005-6

Jasper, H. H. (1958). The ten-twenty electrode system of the international federation. Electroencephalography and Clinical Neurophysiology, 10, 371-375. Retrieved from http://ci.nii.ac.jp/

Jemel, B., Schuller, A. M., Cheref-Khan, Y., Goffaux, V., Crommelinck, M., & Bruyer, R. (2003). Stepwise emergence of the face-sensitive N170 event-related potential component. Neuroreport, 14(16), 2035-2039. doi:10.1097/01.wnr.0000092465.31470.2f

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. Cognition, 40(1), 1-19. doi:10.1016/0010-0277(91)90045-6

Jones, R. D., & Tranel, D. (2001). Severe developmental prosopagnosia in a child with superior intellect. Journal of Clinical and Experimental Neuropsychology, 23(3), 265-273. doi:10.1076/jcen.23.3.265.1183

Joordens, S., & Hockley, W. E. (2000). Recollection and familiarity through the looking glass: When old does not mirror new. Journal of Experimental Psychology: Learning, Memory, & Cognition, 26(6), 1534-1555. doi:10.1037/0278-7393.26.6.1534

Joyce, C., & Rossion, B. (2005). The face-sensitive N170 and VPP components manifest the same brain processes: the effect of reference electrode site. Clinical Neurophysiology, 116(11), 2613-2631. doi:10.1016/j.clinph.2005.07.005

Kahane, P., Hoffmann, D., Minotti, L., & Berthoz, A. (2003). Reappraisal of the human vestibular cortex by cortical electrical stimulation study. Annals of neurology, 54(5), 615-624. doi:10.1002/ana.10726

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. The Journal of Neuroscience, 17(11), 4302-4311. doi:10.3410/f.717989828.793472998

Kaufmann, J. M., Schulz, C., & Schweinberger, S. R. (2013). High and low performers differ in the use of shape information for face recognition. Neuropsychologia, 51(7), 1310-1319. doi:10.1016/j.neuropsychologia.2013.03.015

Kennerknecht, I., Grueter, T., Welling, B., Wentzek, S., Horst, J., Edwards, S., & Grueter, M. (2006). First Report of Prevalence of Non-Syndromic Hereditary Prosopagnosia (HPA). American Journal of Medical Genetics Part A, 140A(15), 1617-1622. doi:10.1002/ajmg.a.31343

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. doi:10.1002/ajmg.a.32552

Khan, S., & Chang, R. (2013). Anatomy of the vestibular system: a review. NeuroRehabilitation, 32(3), 437-443. doi:10.3233/NRE-130866

Kracke, I. (1994). Developmental prosopagnosia in Asperger syndrome: Presentation and discussion of an individual case. Developmental Medicine and Child Neurology, 36, 873-886. doi: 10.1111/j.1469-8749.1994.tb11778.x

Kress, T., & Daum, I. (2003). Event-related potentials reflect impaired face recognition in patients with congenital prosopagnosia. Neuroscience Letters, 352(2), 133-136. doi:10.1016/j.neulet.2003.08.047

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language

comprehension. Trends in Cognitive Sciences, 4(12), 463-470. doi:10.1016/S1364-

6613(00)01560-6

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400

component of the event related brain potential (ERP). Annual review of psychology, 62, 621-

647. doi:10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic

incongruity. Science, 207(4427), 203-205. doi:10.1126/science.7350657

Lai, J., Pancaroglu, R., Oruc, I., Barton, J. J. S., & Davies-Thompson, J. (2014). Neuroanatomic

correlates of the feature-salience hierarchy in face processing: An fMRI-adaptation study.

Neuropsychologia, 53, 274-283. doi:10.1016/j.neuropsychologia.2013.10.016

Le Grand, R., Cooper, P. A., Mondloch, C. J., Lewis, T. L., Sagiv, N., de Gelder, B., & Maurer, D.

(2006). What aspects of face processing are impaired in developmental prosopagnosia? Brain

and Cognition, 61(2), 139-158. doi:10.1016/j.bandc.2005.11.005

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Neuroperception: Early visual

experience and face processing. Nature, 410, 890. doi:10.1038/35073749

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2004). Impairment in holistic face

processing following early visual deprivation. Psychological Science, 15(11), 762-768.

doi:10.1111/j.0956-7976.2004.00753.x

Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural

information in face recognition. The Quarterly Journal of Experimental Psychology Section

A: Human Experimental Psychology, 53(2), 513-536. doi:10.1080/713755889

Lee, Y., Duchaine, B., Wilson, H. R., & Nakayama, K. (2009). Three cases of developmental

prosopagnosia from one family: Detailed neuropsychological and psychophysical

investigation of face processing. Cortex, 46(8), 949-964. doi:10.1016/j.cortex.2009.07.012

Lefly, D. L., & Pennington, B. F. (2000). Reliability and Validity of the Adult Reading History Questionnaire. Journal of Learning Disabilities, 33(3), 286-296. doi:10.1177/002221940003300306

Lissauer, H. (1890). Ein Fall von Seelenblindheit nebst einem Beitrage zur Theorie derselben. Archiv für Psychiatrie und Nervenkrankheiten, 21(2), 222-270. doi:10.1007/BF02226765

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: an MEG study. Nature neuroscience, 5(9), 910-916. doi:10.1038/nn909

Lobel, E., Kleine, J F., Bihan, D. L., Leroy-Willig, A., & Berthoz, A. (1998). Functional MRI of Galvanic Vestibular Stimulation. Journal of Neurophysiology, 80(5), 2699-2709. Retrieved from http://jn.physiology.org/

Logan, A. J., Wilkinson, F., Wilson, H. R., Gordon, G. E., & Loffler, G. (2016). The Caledonian face test: A new test of face discrimination. Vision research, 119, 29-41. doi:10.1016/j.visres.2015.11.003

Lopez, C., & Blanke, O. (2014). Nobel Prize centenary: Robert Bárány and the vestibular system. Current Biology, 24(21), R1026-R1028. doi:10.1016/j.cub.2014.09.067

Lopez, C., Blanke, O., & Mast, F. W. (2012). The human vestibular cortex revealed by coordinate-based activation likelihood estimation meta-analysis. Neuroscience, 212, 159-179. doi:10.1016/j.neuroscience.2012.03.028

Luck, S. J. (2005). An Introduction to the Event-Related Potential Technique. UK: MIT Press.

MacKenzie, G., and Donaldson, D. I. (2007). Dissociating recollection from familiarity: electrophysiological evidence that familiarity for faces is associated with a posterior old/new effect. Neuroimage, 36(2), 454–463. doi:10.1016/j.neuroimage.2006.12.005

Macrae, C. N., & Lewis, L. L. (2002). Do I know you? Processing orientation and face recognition. Psychological Science, 13(2), 194-196. doi:10.1111/1467-9280.00436

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. Psychological Review, 87(3), 252-271. doi:10.1037/0033-295X.87.3.252

Mast, F. W., Merfeld, D. M., & Kosslyn, S. M. (2006). Visual mental imagery during caloric vestibular stimulation. Neuropsychologia, 44(1), 101-109. doi:10.1016/j.neuropsychologia.2005.04.005

Matson, J. L., Gonzalez, M. L., & Rivet, T. T. (2008). Reliability of the Autism Spectrum Disorder-Behavior Problems for Children (ASD-BPC). Research in Autism Spectrum Disorders, 2(4), 696-706. doi:10.1016/j.rasd.2008.02.003

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. Trends in Cognitive Sciences, 6(6), 255-260. doi:10.1016/S1364 6613(02)01903-4

Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011). Impaired acuity of the approximate number system underlies mathematical learning disability (dyscalculia). Child Development, 82(4), 1224-1237. doi:10.1111/j.1467-8624.2011.01608.x

McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. Journal of Cognitive Neuroscience, 9(5), 605-610. doi:10.1162/jocn.1997.9.5.605

McClelland, J. L., & Chappell, M. (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effects of experience in recognition memory. Psychological Review, 105(4), 724-760. doi:10.1037/0033-295X.105.4.734-760

McClelland, J. L., & Rumelhart, D. E. (1981). An Interactive Activation Model of the Effect of Context in Perception: Part 1. An account of basic findings. Psychological Review, 88(5), 375-407. Doi:10.1037/0033-295X.88.5.375

McConachie, H. R. (1976). Developmental Prosopagnosia. A Single Case Report. Cortex, 12(1), 76-82. doi:10.1016/S0010-9452(76)80033-0

McGeoch, P. D., & Ramachandran, V. S. (2008). Vestibular stimulation can relieve central pain of spinal origin. Spinal Cord, 46(11), 756-757. doi:10.1038/sc.2008.47

McKone, E., Hall, A., Pidcock, M., Palermo, R., Wilkinson, R. B., Rivolta, D., ... & 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. doi:10.1080/02643294.2011.616880

McNeill, A., & Burton, A. M. (2002). The locus of semantic priming effects in person recognition. The Quarterly Journal of Experimental Psychology, 55A(4), 1141-1156. doi:10.1080/02724980244000189

Miller, S. M. & Ngo, T. T. (2007). Studies of caloric vestibular stimulation: Implications for the cognitive neurosciences, the clinical neurosciences and neurophilosophy. ActaNeuropsychiatrica, 19(3), 183-203. doi:10.1111/j.1601-5215.2007.00208.x

Minnebusch, D. A., Suchan, B., Köster, O., & Daum, I. (2009). A bilateral occipitotemporal network mediates face perception. Behavioural brain research, 198(1), 179-185. doi:10.1016/j.bbr.2008.10.041

Minnebusch, D. A., Suchan, B., Ramon, M., & Daum, I. (2007). Event-related potentials reflect heterogeneity of developmental prosopagnosia. European Journal of Neuroscience, 25(7), 2234-2247. doi:10.1111/j.1460-9568.2007.05451.x

Mogg, K., & Bradley P, B. (2000). Selective Attention and Anxiety: A Cognitive-Motivational Perspective. In T. Dalgleish & M. J. Power (Eds.), Handbook of Cognition and Emotion (pp. 145–170). UK: John Wiley & Sons.

Mooney, C. M. (1957). Age in the development of closure ability in children. Canadian Journal of Psychology, 11(4), 219-226. doi:10.1037/h0083717

Mooney, C. M. (1960). Recognition of ambiguous and unambiguous visual configurations with short and longer exposures. British Journal of Psychology, 51(2), 119-125. doi:10.1111/j.2044-8295.1960.tb00732.x

Morgan, H. M., Klein, C., Boehm, S. G., Shapiro, K. L., & Linden, D. E. (2008). Working memory load for faces modulates P300, N170, and N250r. Journal of cognitive neuroscience, 20(6), 989-1002. doi:10.1162/jocn.2008.20072

Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. Journal of cognitive neuroscience, 9(5), 555-604. doi:10.1162/jocn.1997.9.5.555

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. Cognitive Psychology, 9(3), 353-383. doi:10.1016/0010-0285(77)90012-3

Németh, K., Zimmer, M., Schweinberger, S. R., Vakli, P., & Kovács, G. (2014). The background of reduced face specificity of N170 in congenital prosopagnosia. PloS one, 9(7), e101393. doi:10.1371/journal.pone.0101393

Neumann, M. F., & Schweinberger, S. R. (2009). N250r ERP repetition effects from distractor faces when attending to another face under load: Evidence for a face attention resource. Brain research, 1270, 64-77. doi:10.1016/j.brainres.2009.03.018

Neumann, M. F., Mohamed, T. N., & Schweinberger, S. R. (2011). Face and object encoding under perceptual load: ERP evidence. NeuroImage, 54(4), 3021-3027. doi:10.1016/j.neuroimage.2010.10.075

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

Öhman, A. (1999). Distinguishing unconscious from conscious emotional processes: Methodological considerations and theoretical implications. In T. Dalgleish & M. J. Power (Eds.), Handbook of Cognition and Emotion (pp. 321-352). UK: John Wiley & Sons.

Olichney, J. M., Taylor, J. R., Gatherwright, J., Salmon, D. P., Bressler, A. J., Kutas, M., & Iragui-Madoz, V. J. (2008). Patients with MCI and N400 or P600 abnormalities are at very high risk for conversion to dementia. Neurology, 70(19), 1763-1770. doi:10.1212/01.wnl.0000281689. 28759.ab

Olivares, E. I., Iglesias, J., Saavedra, C., Trujillo-Barreto, N. J., & Valdés-Sosa, M. (2015). Brain Signals of Face Processing as Revealed by Event-Related Potentials. Behavioural Neurology, 2015(514361), 1-16. doi:10.1155/2015/514361

Olivares, E. I., Saavedra, C., Trujillo-Barreto, N. J. and Iglesias, J. (2013) Long-term information and distributed neural activation are relevant for the "internal features advantage" in face processing: Electrophysiological and source reconstruction evidence. Cortex, 49(10), 2735-2747. doi:10.1016/j.cortex.2013.08.001

Palermo, R., Willis, M. L., Rivolta, D., McKone, E., Wilson, C. E., & Calder, A. J. (2011). Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. Neuropsychologia, 49(5), 1226-1235. doi:10.1016/j.neuropsychologia.2011.02.021

Paller, K. A., Gonsalves, B., Grabowecky, M., Bozic, V. S., & Yamada, S. (2000). Electrophysiological Correlates of Recollecting Faces of Known and Unknown Individuals. NeuroImage, 11, 98-110. doi:10.1006/nimg.1999.0521

Palmer, S. E. (1977). Hierarchical structure in perceptual representation. Cognitive psychology, 9(4), 441-474. doi:10.1016/0010-0285(77)90016-0

Parketny, J., Towler, J., & Eimer, M. (2015). The activation of visual face memory and explicit face recognition are delayed in developmental prosopagnosia. Neuropsychologia, 75, 538-547. doi:10.1016/j.neuropsychologia.2015.07.009

Perfect, T. J., Weston, N. J., Dennis, I., Amelia, S. (2008). The effects of precedence on Navon-induced processing bias in face recognition. The Quarterly Journal of Experimental Psychology, 61(10), 1479-1486. doi:10.1080/17470210802034678

Peterson, B. W., Fukushima, K., Hirai, N., Schor, R. H., & Wilson, V. J. (1980). Responses of vestibulospinal and reticulospinal neurons to sinusoidal vestibular stimulation. Journal of neurophysiology, 43(5), 1236-1250. Retrieved from http://jn.physiology.org/

Pfütze, E. M., Sommer, W., & Schweinberger, S. R. (2002). Age-related slowing in face and name recognition: evidence from event-related brain potentials. Psychology and aging, 17(1), 140-160. doi:10.1037/0882-7974.17.1.140

Pickering, E. C., & Schweinberger, S. R. (2003). N200, N250r, and N400 Event-Related Brain Potentials Reveal Three Loci of Repetition Priming for Familiar Names. Journal of Experimental Psychology: Learning, Memory, and Cognition, 29(6), 1298-1311. doi:10.1037/0278-7393.29.6.1298

Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. The Journal of Neuroscience, 16(16), 5205-5215. Retrieved from http://www.jneurosci.org/

Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember–know judgments in a continuous recognition paradigm. Journal of Experimental Psychology: Learning, Memory, & Cognition, 26(2), 294-320. doi:10.1037/0278-7393.26.2.294

Rhodes, G. (1985). Lateralized processes in face recognition. British journal of Psychology, 76(2), 249-271. doi:10.1111/j.2044-8295.1985.tb01949.x

Righart, R.,& DeGelder,B.(2007). Impaired face and body perception in developmental prosopagnosia. Proceedings of the National Academy of Sciences, 104(43), 17234-17238. doi:10.1073/pnas.0707753104

Rivolta, D., Palermo, R., Schmalzl, L., & Coltheart, M. (2012). Covert face recognition in congenital prosopagnosia: A group study. Cortex, 48(3), 344-352. doi:10.1016/j.cortex.2011.01.005

Rivolta, D., Schmalzl, L., Coltheart, M., & Palermo, R. (2010). Semantic information can facilitate covert face recognition in congenital prosopagnosia. Journal of Clinical and Experimental Neuropsychology, 32(9), 1002-1016. doi:10.1080/13803391003662710

Ro, T., Russell, C., & Lavie, N. (2001). Changing Faces: A Detection Advantage in the Flicker Paradigm. Psychological Science, 12(1), 94–99. doi:10.1111/1467-9280.00317

Robbins, R., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioural tasks. Cognition, 103(1), 34-79. doi:10.1016/j.cognition.2006.02.008

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. Acta Psychologica, 128(2). 274-289. doi:10.1016/j.actpsy.2008.02.003

Rossion, B., & Caharel, S. (2011). ERP evidence for the speed of face categorization in the human brain: Disentangling the contribution of low-level visual cues from face perception. Vision Research, 51(12), 1297-1311. doi:10.1016/j.visres.2011.04.003

Rossion, B., & Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. Neuroimage, 39(4), 1959-1979. doi:10.1016/j.neuroimage.2007.10.011

Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., & Crommelinck, M. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: an electrophysiological account of face-specific processes in the human brain. Neuroreport, 11(1), 69-72. doi:10.1097/00001756-200001170-00014

Rotshtein, P., Geng, J. J., Driver, J., & Dolan, R. J. (2007). Role of features and second-order spatial relations in face discrimination, face recognition, and individual face skills: Behavioral and functional magnetic resonance imaging data. Journal of Cognitive Neuroscience, 19(9), 1435-1452. doi:10.1162/jocn.2007.19.9.1435

Rugg, M. D. (1987). Dissociation of semantic priming, word and non-word repetition effects by event-related potentials. The quarterly journal of Experimental Psychology, 39(1), 123-148. doi:10.1080/02724988743000060

Rutter, D. R. (1989). The role of cuelessness in social interaction: An examination of teaching by telephone. In D. Roger & P. Bull (Eds.), Conversation: An Interdisciplinary Perspective (pp. 294-312). Avon, UK: Multilingual Matters.

Sagiv, N., & Bentin, S. (2001). Structural encoding of human and schematic faces: holistic and part-based processes. Journal of Cognitive Neuroscience, 13(7), 937-951. doi:10.1162/089892901753165854

Saladin, K. S. (2012). Anatomy and Physiology: The Unity of Form and Function. New York: McGraw Hill.

Sang, F. Y. P., Jauregui-Renaud, K., Green, D. A., Bronstein, A. M., & Gresty, M. A. (2006). Depersonalisation/derealisation symptoms in vestibular disease. Journal of Neurology, Neurosurgery & Psychiatry, 77(6), 760-766. doi:10.1136/jnnp.2005.075473

Schmalzl, L., Palermo, R., & Coltheart, M. (2008). Cognitive heterogeneity in genetically based prosopagnosia: A family study. Journal of Neuropsychology, 2(1), 99-117. doi:10.1348/174866407X256554

Schmalzl, L., Palermo, R., Green, M., Brunsdon, R., & Coltheart, M. (2008). Training of familiar face recognition and visual scan paths for faces in a child with congenital prosopagnosia. Cognitive Neuropsychology, 25(5), 704-729. doi:10.1080/02643290802299350

Schweinberger, S. R. (1996). How Gorbachev primed Yeltsin: Analyses of associative priming in person recognition by means of reaction times and event-related brain potentials. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22(6), 1383-1407. doi:10.1037/0278-7393.22.6.1383

Schweinberger, S. R., & Burton, A. M. (2003). Covert recognition and the neural system for face processing. Cortex, 39(1), 9-30. doi:10.1016/S0010-9452(08)70071-6

Schweinberger, S. R., & Neumann, M. F. (2016). Repetition effects in human ERPs to faces. Cortex, 80, 141-153. doi:10.1016/j.cortex.2015.11.001

Schweinberger, S. R., Huddy, V. & Burton, A. M. (2004). N250r: a face-selective brain response to stimulus repetitions. Neuroreport, 15(9), 1501-1505. doi:10.1097/01.wnr.0000131675.00319.42

Schweinberger, S. R., Kaufmann, J. M., Moratti, S., Keil, A., & Burton, A. M. (2007). Brain responses to repetitions of human and animal faces, inverted faces, and objects—an MEG study. Brain research, 1184, 226-233. doi:10.1016/j.brainres.2007.09.079

Schweinberger, S. R., Pfütze, E. M., & Sommer, W. (1995). Repetition priming and associative priming of face recognition: Evidence from event-related potentials. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21(3), 722. doi:10.1037/0278-7393.21.3.722

Schweinberger, S. R., Pickering, E. C., Jentzsch, I., Burton, A. M., & Kaufmann, J. M. (2002). Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. Cognitive Brain Research, 14(3), 398-409. doi:10.1016/S0926-6410(02)00142-8

Shallice, T., & Jackson, M. (1988). Lissauer on Agnosia. Cognitive Neuropsychology, 5(2), 153-192. doi:10.1080/02643298808252931

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM-retrieving effectively from memory. Psychonomic Bulletin & Review, 4, 145-166. doi:10.3758/BF03209391

Silverman, W. K., Saavedra, L. M., & Pina, A. A. (2001). Test-Retest Reliability of Anxiety Symptoms and Diagnoses With the Anxiety Disorders Interview Schedule for DSM-IV:

Child and Parent Versions. Journal of the American Academy of Child and Adolescent Psychiatry, 40(8), 937-944. doi:10.1097/00004583-200108000-00016

Smith, P. F., Geddes, L. H., Baek, J., Darlington, C. L., & Zheng, Y. (2010). Modulation of Memory by Vestibular Lesions and Galvanic Vestibular Stimulation. Frontiers in Neurology, 1(141). doi:10.3389/fneur.2010.00141

Soria Bauser, D. A., Schriewer, E., & Suchan, B. (2015). Dissociation between the behavioural and electrophysiological effects of the face and body composite illusions. British Journal of Psychology, 106(3), 414-432. doi:10.1111/bjop.12101

Stollhoff, R., Jost, J., Elze, T., & Kennerknecht, I. (2011). Deficits in long-term recognition memory reveal dissociated subtypes in congenital prosopagnosia. PLoS ONE, 6(1), e15702. doi:10.1371/journal.pone.0015702

Susilo, T., McKone, E., Dennett, H., Darke, H., Palermo, R., Hall, A., ... & Rhodes, G. (2010). Face recognition impairments despite normal holistic processing and face space coding: Evidence from a case of developmental prosopagnosia. Cognitive neuropsychology, 27(8), 636-664. doi:10.1080/02643294.2011.613372

Suzuki, M., Kitano, H., Ito, R., Kitanishi, T., Yazawa, Y., Ogawa, T., Shiino, A., & Kitajima, K. (2001). Cortical and subcortical vestibular response to caloric stimulation detected by functional magnetic resonance imaging. Cognitive Brain Research, 12(3), 441-449. doi:10.1016/S0926-6410(01)00080-5

Takeda, N., Hashikawa, K., Moriwaki, H., Oku, N., Koizuka, I., Kitahara, T., … Nishimura, T. (1996). Effects of caloric vestibular stimulation on parietal and temporal blood flow in human brain: a consecutive technetium-99m-HMPAO spect study. Journal of Vestibular Research, 6(2), 127-134. doi:10.3233/VES-1996-6207

Tanaka, J. W., & Farah, M. J. (1993). Parts and Wholes in Face Recognition. The Quarterly Journal of Experimental Psychology, 46A(2), 225-245. doi:10.1080/14640749308401045

Tanaka, J. W., Curran, T., Porterfield, A. L., & Collins, D. (2006). Activation of Pre-existing and Acquired Face Representations: The N250 Event-related Potential as an Index of Face Familiarity. Journal of Cognitive Neuroscience, 18(9), 1488-1497. doi:10.1162/jocn.2006.18.9.1488

Tascioglu, A. B. (2005). Brief review of vestibular system anatomy and its higher order projections. Neuroanatomy, 4, 24-27. Retrieved from http://www.neuroanatomy.org/

Temple, C. (1992). Developmental memory impairment: Faces and patterns. In R. Campbell (Ed.), Mental Lives: Case Studies in Cognition (pp.199-215). London: Blackwell.

Theeuwes, J., & Van der Stigchel, S. (2006). Faces capture attention: Evidence from inhibition of return. Visual Cognition, 13(6), 657–665. doi:10.1080/13506280500410949

Thierry, G., Martin, C. D., Downing, P., & Pegna, A. J. (2007). Controlling for interstimulus perceptual variance abolishes N170 face selectivity. Nature neuroscience, 10(4), 505-511. doi:10.1038/nn1864

Thomas, A. L., Lawler, K., Olson, I. R., & Aguirre, G. K. (2008). The Philadelphia face perception battery. Archives of Clinical Neuropsychology, 23(2), 175-187. doi:10.1016/j.acn.2007.10.003

Tippett, L. J., Miller, L. A., & Farah, M. J. (2000). Prosopamnesia: A selective impairment in face learning. Cognitive Neuropsychology, 17(1/2/3), 241-255. doi:10.1080/026432900380599

Towler, J., & Eimer, M. (2012). Electrophysiological studies of face processing in developmental prosopagnosia: neuropsychological and neurodevelopmental perspectives. Cognitive Neuropsychology, 29(5-6), 503-529. doi:10.1080/02643294.2012.716757

Towler, J., Gosling, A., Duchaine, B., & Eimer, M. (2012). The face-sensitive N170 component in developmental prosopagnosia. Neuropsychologia, 50(14), 3588-3599. doi:10.1016/j.neuropsychologia.2012.10.017

Tranel, D., & Damasio, A. R. (1985). Knowledge without Awareness: An Autonomic Index of Facial
    Recognition by Prosopagnosics. Science, 228(4706), 1453-1454.
    doi:10.1126/science.4012303

Tulving, E. (1985). Memory and consciousness. Canadian Psychology/Psychologie canadienne,
    26(1), 1-12. doi:10.1037/h0080017

Utz, K. S., Dimova, V., Oppenländer, K., & Kerkhoff, G. (2010). Electrified minds: Transcranial
    direct current stimulation (tDCS) and Galvanic Vestibular Stimulation (GVS) as methods of
    non-invasive brain stimulation in neuropsychology: A review of current data and future
    implications. Neuropsycholgia, 48(10), 2789-2810.
    doi:10.1016/j.neuropsychologia.2010.06.002

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition.
    British Journal of Psychology, 79(4), 471-491. doi:10.1111/j.2044-8295.1988.tb02747.x

Valenza, E., Simion, F., Cassia, V. M., & Umiltà, C. (1996). Face preference at birth. Journal of
    experimental psychology: Human Perception and Performance, 22(4), 892.
    doi:10.1037/0096-1523.22.4.892

Vilberg, K. L., & Rugg, M. D. (2007). Dissociation of the neural correlates of recognition memory
    according to familiarity, recollection, and amount of recollected information.
    Neuropsycholgia, 45(10), 2216-2225. doi:10.1016/j.neuropsychologia.2007.02.027

Vitte, E., Derosier, C., Caritu, Y., Berthoz, A., Hasboun, D., & Soulié, D. (1996). Activation of the
    hippocampal formation by vestibular stimulation: a functional magnetic resonance imaging
    study. Experimental Brain Research, 112, 523-526. doi:10.1007/BF00227958

Von Kriegstein, K., Kleinschmidt, A., & Giraud, A. L. (2006). Voice recognition and cross-modal
    responses to familiar speakers' voices in prosopagnosia. Cerebral Cortex, 16(9), 1314-1322.
    doi:10.1093/cercor/bhj073

Vuilleumier, P. (2002). Facial expression and selective attention. Current Opinion in Psychiatry, 15(3), 291–300. doi:10.1097/00001504-200205000-00011

Weston, N. J., & Perfect, T. J. (2005). Effects of processing bias on the recognition of composite face halves. Psychonomic Bulletin & Review, 12(6), 1038-1042. doi:10.3758/BF03206440

Wiese, H. (2011). The structure of semantic person memory: Evidence from semantic priming in person recognition. British Journal of Psychology, 102(4), 899-914. doi:10.1111/j.2044-8295.2011.02042.x

Wigan, A. L. (1844). The duality of the mind. London, England: Longman.

Wilkinson, D. T., Ferguson, H. J., & Worley, A. (2012). Galvanic vestibular stimulation modulates the electrophysiological response during face processing. Visual Neuroscience, 29(4-5), 255-262. doi:10.1017/S0952523812000235

Wilkinson, D., & Halligan, P. (2004). The relevance of behavioural measures for functional-imaging studies of cognition. Nature Reviews Neuroscience, 5(1), 67-73. doi:10.1038/nrn1302

Wilkinson, D., Ferguson, H. J., & Worley, A. (2012). Galvanic vestibular stimulation modulates the electrophysiological response during face processing. Visual Neuroscience, 29(4-5), 255-262. doi:10.1017/S0952523812000235

Wilkinson, D., Ko, P., Kilduff, P., McGlinchey, R., & Milberg, W. (2005). Improvement of a face perception deficit via subsensory galvanic vestibular stimulation. Journal of the International Neuropsychological Society, 11(7), 925-929. doi:10.1017/S1355617705051076

Wilkinson, D., Ko, P., Wiriadjaja, A., Kilduff, P., McGlinchey, R., & Milberg, W. (2009). Unilateral damage to the right cerebral hemisphere disrupts the apprehension of whole faces and their component parts. Neuropsychologia, 47(7), 1701-1711. doi:10.1016/j.neuropsychologia.2009.02.008

Wilkinson, D., Morris, R., Milberg, W., & Sakel, M. (2013). Caloric vestibular stimulation in

    aphasic syndrome. Frontiers in integrative neuroscience, 7(99), 1-9.

    doi:10.3389/fnint.2013.00099

Wilkinson, D., Nicholls, S., Pattenden, C., Kilduff, P., & Milberg, W. (2008). Galvanic vestibular

    stimulation speeds visual memory recall. Experimental Brain Research, 189(2), 243-248.

    doi:10.1007/s00221-008-1463-0

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., ... & Duchaine,

    B. (2010). Human face recognition ability is specific and highly heritable. Proceedings of the

    National Academy of sciences, 107(11), 5238-5241. doi:10.1073/pnas.0913053107

Wilson, B. (1987). Rehabilitation of memory. London: Guilford Press.

Wilson, V. J., Peterson, B. W., Fukushima, K., Hirai, N., & Uchino, Y. (1979). Analysis of

    vestibulocollic reflexes by sinusoidal polarization of vestibular afferent fibers. Journal of

    neurophysiology, 42(2), 331-346. Retrieved from http://jn.physiology.org/

Yin, R. K. (1969). Looking at upside-down faces. Journal of experimental psychology, 81(1), 141-

    145. doi:10.1037/h0027474

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a

    dual-process model. Journal of Experimental Psychology: Learning, Memory, & Cognition,

    20(6), 1341-1354. doi:10.1037/0278-7393.20.6.1341

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research.

    Journal of Memory & Language, 46, 441-517. doi:10.1006/jmla.2002.2864

Young, A. W. (1999). Simulating face recognition: Implications for modelling cognition. Cognitive

    Neuropsychology, 16(1), 1-48. doi:10.1080/026432999380960

Young, A. W., & Ellis, H. D. (1989). Childhood prosopagnosia. Brain and Cognition, 9(1), 16-47.

    doi:10.1016/0278-2626(89)90042-0

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. Perception, 16, 747-759. doi:10.1068/p160747n

Yovel, G., & Duchaine, B. (2006). Specialized face perception mechanisms extract both part and spacing information: evidence from developmental prosopagnosia. Journal of Cognitive Neuroscience, 18(4), 580-593. doi:10.1162/jocn.2006.18.4.580

Zheng, X., Mondloch, C. J., & Segalowitz, S. J. (2012). The timing of individual face recognition in the brain. Neuropsychologia, 50(7), 1451-1461. doi:10.1016/j.neuropsychologia.2012.02.030

Zu Eulenburg, P., Caspers, S., Roski, C., & Eickhoff, S. B. (2012). Meta-analytical definition and functional connectivity of the human vestibular cortex. Neuroimage, 60(1), 162-169. doi:10.1016/j.neuroimage.2011.12.032

APPENDIX A
Participant Information Sheet and Consent Form

Participant Information Sheet:

You are being invited to take part in a research study. Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Please ask if anything is unclear, and do not feel rushed into making a decision.

Study Aim

This is a research study to examine how the electrical activity in your brain responds when you view faces. It will also test whether mild temperature changes in the ear canal affect your ability to process faces. The changes stimulate the organs in the ear which activate many areas in the brain. We will examine if this improves your ability to perceive and remember faces. We will also examine if there are differences between those who do and do not report difficulty recognising faces.

What will happen to me in the study?

Before you begin the experiment, a number of small electrodes will be attached to your scalp to detect any corresponding changes in underlying neural activity. It is necessary to clean the area under each electrode site to ensure good electrode-skin contact. This may cause mild discomfort (due to the abrasiveness of the cleaner), and sometimes causes a little redness. We should also add that to ensure good contact between the skin and all the electrodes, it is necessary to apply conductive, soluble gel which is best removed by washing your hair. The application of electrodes may take up to 45 minutes. We will then ask you to wear a set of headphones that features a small cone that sits in your left ear and which will sometimes change in temperature (which not everyone notices).

Over the course of two sessions, we will apply stimulation at various times while you perform decisions about visual stimuli that appear on either a computer screen or paper. We want to emphasise that the stimulation procedure should not be painful and is not typically associated with ill-effect. During stimulation, you may feel a cold sensation in your ear. This is normal and should not distract you. In other instances, the stimulation will be too light for you to notice. Some people have reported feelings of dizziness or nausea and there is always a small risk of unanticipated side-effects. However these effects are unlikely for the low levels of stimulation used in this study.

What will happen if I don't want to carry on with the study?

I understand that I may withdraw from the study at any time, and that if I refuse to take part or if I decide to withdraw, I will not suffer any penalty or loss of rights. If I do choose to withdraw or am no longer able to participate, then unless verbally directed otherwise the study investigators will keep the data collected up to that point. I also understand that my participation can be withdrawn by the study investigators.

Are there any negative side-effects?

We do not anticipate any significant side-effects although some participants may feel a little tired after the test session.

What will happen to the results of the research study?
The results of this study will be used to better understand how humans perceive and recognise facial information. The results of the present study may be published for scientific purposes, but your records or identity will not be revealed unless required by law.

What happens to the information I provide?
Participation in this study guarantees confidentiality of the information you provide. No one apart from the researcher and research supervisor will have any access to the information you provide. Your name and any other identifying information will be stored separately from your data in a securely locked filing cabinet. Questionnaires will be stored in a securely locked room for as long as is required by the Data Protection Act, and then they will be destroyed by our confidential shredding service. You may withdraw from the study at any time, and if you refuse to take part or if you decide to withdraw, you will not suffer any penalty, loss of rights, or loss of medical benefits that you have a right to receive. Your participation in the study can be withdrawn by the study investigators.

What if there is a problem?
If you have a concern about any aspect of the study then you should speak with Dr David Wilkinson, who is a supervisor of the study. He can be reached on 01227 824772. If you remain unhappy and wish to complain formally, you can do this through the School of Psychology Chair of Ethics. Further details can be obtained from the School of Psychology General Office on 01227 824775.

Who is organising and funding the research?
The research is organised by Philip Ulrich MSc and supervised by Prof Bob Johnston and Dr David Wilkinson. The study is part of a PhD research project funded by a department Graduate Teaching Assistantship and receives input from psychologist Dr Heather Ferguson.

Who has reviewed this study?
The study has been approved by the School of Psychology, University of Kent Research Ethics committee.

Participant Consent Form:

Title of project: Investigation of electrophysiological modulation during vestibular stimulation
Name of Chief Researcher: Bob Johnston

1. I Confirm that I have read and understand the information sheet for the above study and have had the opportunity to ask questions.

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.

3. I agree to take part in the above study

Name of Participant: _____
Signature: _____ Date: _____
Please retain a copy for your records