FACE SPECIFIC VS. EXPERTISE HYPOTHESES: INSIGHTS INTO THE UNDERLYING MECHANISMS OF FACE PROCESSING IN PROSOPAGNOSIA

BY

SUMAYA LAMB

A thesis
Submitted to the Victoria University of Wellington
in fulfilment of the requirements for the degree of
Master of Science in
Cognitive and Behavioural Neuroscience

Victoria University of Wellington
2018

Abstract

A prominent debate in visual perception centers on the nature of mechanisms underlying face processing. One side of this debate argues that faces are processed by specialised mechanisms that are not involved in any form of object processing. By contrast, the other side argues that faces are processed by generic mechanisms common to all objects for which we are experts. To distinguish between these two hypotheses, I investigated whether participants with impaired face processing (developmental prosopagnosia) can acquire expertise with novel objects called greebles. To do so, I recruited 10 developmental prosopagnosics and 10 neurotypical control participants. All participants completed a standard training program for developing expertise with greebles, as well as two similar training programs with upright faces and inverted faces. Prosopagnosics were able to acquire expertise with greebles to the same extent as controls but were impaired when learning upright faces. These results demonstrate that deficits for face processing in individuals with prosopagnosia are dissociated from their ability to gain expertise with objects. Overall, the results support the hypothesis that face processing relies on specialised mechanisms, rather than generic expertise mechanisms. Despite their deficits, though, prosopagnosics still showed some evidence of learning with upright faces and showed better learning with upright faces than inverted faces. These findings suggest that prosopagnosics have face-specific mechanisms that are somewhat functional, and that training could be a useful rehabilitation tool in developmental prosopagnosia. Finally, I found substantial heterogeneity among the patterns of performance of the prosopagnosics, suggesting that further investigations into the subtypes of prosopagnosia are warranted.

Acknowledgements

To my supervisor, Dr. Tirta Susilo, thank you so much for your endless patience and support these last two years. It’s been an absolute privilege to be your student, and you have gone above and beyond as a supervisor and mentor. Under your guidance I feel as though I have developed dramatically as a researcher in every aspect. I will be forever grateful for the skills, inspiration and opportunities you have given me in what I hope will be a long and successful research career.

I’d also like to thank all the members of the Susilo Lab. I have been incredibly lucky to be part of a lab filled with caring, critically minded, and motivated individuals. Special thanks to Angus, Cassandra, Daniel, Hazel, and Lauren. The unwavering support and guidance that you all have given me and continue to give has helped me grow academically and personally. Thank you for always listening, keeping me grounded when my ideas and worries become just a bit too ridiculous, and reminding me that sleeping and taking breaks are things that normal people do. Also, an extra special thank you to Cassandra and Hazel for proof reading this thesis.

Thank you to my parents. Thank you for always encouraging me to pursue the things in life that make me happy and thank you for always supporting the many (ever changing) endeavours that followed from that advice. It’s clear that without your support I would not be in the position I am in today.

I’d also like to thank my partner, Ryan. You have been absolutely incredible and patient throughout my entire post-graduate career. An ever-calming presence, thank you for everything, from making me laugh to listening to every single idea I’ve had for this thesis.

Finally, I’d like to acknowledge the Marsden fund awarded to Tirta by the Royal Society of New Zealand and the Virginia Rose Burns award given to me by Victoria University of Wellington. My research would not have been possible without their financial support.

Table of Contents

List of Tables

List of Figures

FACE SPECIFIC VS. EXPERTISE HYPOTHESES: INSIGHTS INTO THE UNDERLYING MECHANISMS OF FACE PROCESSING IN PROSOPAGNOSIA

Every day, people encounter dozens of faces that they need to recognise and respond to. We use faces to figure out the identity of others, to recognise their emotions, and to infer their intentions. It is perhaps not surprising, then, that people tend to be very good at quickly recognising faces, compared to other objects. For example, people only need a tenth of a second to detect faces—twice as fast as other objects (Crouzet, Kircher, & Thorpe, 2010). As well as being able to identify faces quickly, people report that human faces tend to ‘pop out’ amongst other objects (Hershler & Hochstien, 2005). In fact, people’s tendency to notice faces is so strong that we often see face images in objects that are not faces, such as clouds, cars, paint splatters, and even on pieces of toast, suggesting that we are attuned to perceive faces over other patterns (Liu et al., 2014). Research has also shown that this bias to focus on faces over other objects is present from birth. For example, newborns track images of faces longer than images of other stimuli (Farroni et al., 2005). But how is it that people can process faces so quickly and efficiently?

At the moment, there are two leading hypotheses for how people are able to process faces so efficiently. One is the *face-specific hypothesis*, which suggests that face processing is special because it is carried out by dedicated mechanisms that play no role in the processing of other objects (Kanwisher, 2000; McKone & Robbins, 2011; Yin, 1969). The other is the *expertise hypothesis*, which suggests that face processing is not special, and only appears unique because people tend to have vast amounts of experience and therefore expertise with faces, compared to other objects (Diamond & Carey, 1986; Goldstein, 1975; Tarr & Gauthier, 2000). Over the last 30 years, these hypotheses have been tested using a wide range of methods, including behavioural experiments, neuroimaging, electroencephalography, and patient studies. One method that stands out in its ability to distinguish between the hypotheses is a behavioural training paradigm that uses computer-generated novel objects called greebles. In this paradigm, participants are extensively trained to recognise greebles over the course of several days so that they become experts at recognising them. Greeble studies are valuable because they allow us to examine the acquisition of expertise with novel objects in a controlled environment—which allows us to search for associations and dissociations between how people process faces and how they process other objects-of-expertise. As a result, these studies can be used to tease apart the face-specific and expertise hypotheses.

Greeble studies have been used with several cases of face-blindness, formally known as *prosopagnosia* (Behrmann, Marotta, Gauthier, Tarr, & McKeef, 2005; Bukach et al., 2012; Duchaine, Dingle, Butterworth, & Nakayama, 2004; Rezlescu, Barton, Pitcher, & Duchaine, 2014). Prosopagnosia is the inability to recognise faces despite otherwise normal vision, which can be lifelong (developmental prosopagnosia; McConachie, 1976) or acquired following brain damage (acquired prosopagnosia; Bodamer, 1947). Greeble studies with prosopagnosics are a particularly powerful tool in the debate between the face-specific and expertise hypotheses because the two hypotheses make competing predictions. The expertise hypothesis predicts that prosopagnosics should be impaired at acquiring expertise with both greebles and faces to the same extent, because the two stimulus classes are processed by the same expertise mechanisms. By contrast, the face-specific hypothesis predicts that, prosopagnosics could become greeble experts, so long as their impairment is restricted to face-specific mechanisms.

In this thesis, I aim to distinguish between the face-specific and expertise hypotheses by running a greeble training study with 10 developmental prosopagnosics. This study is valuable for several reasons. First, it provides a robust test of the face-specific and expertise hypotheses with the largest sample of prosopagnosics to date – past studies have tested only 1 or 2 participants (Behrmann et al., 2005; Bukach et al 2012; Duchaine et al., 2004; Rezlescu et al., 2014). Second, this study allows us to look for subtypes or different forms of prosopagnosia, should different results be obtained with different prosopagnosics. Third, because past greeble studies have mostly tested acquired cases of prosopagnosia, my data will provide useful insights into the nature of deficits in developmental prosopagnosia andthe development of normal face recognition mechanisms. Finally, my thesis is relevant for understanding whether developmental prosopagnosics can learn to recognise faces, because individuals with prosopagnosia will complete training with faces in addition to their training with greebles.

1.1. Face-specific vs expertise debate

As I outlined earlier, there are two opposing hypotheses for how specialised processing of faces emerges. The face-specific hypothesis suggests that face processing relies on specialised mechanisms that are dedicated to processing only faces, and which cannot be applied to the processing of other visual stimuli (Kanwisher, 2000; McKone & Robbins, 2011). These mechanisms could be innate (Morton & Johnson, 1991) or develop in early childhood (Le Grand, Mondloch, Maurer, & Brent, 2001; 2003). By contrast, the expertise hypothesis suggests that face processing only appears special because we have much more exposure to faces than to other objects (Diamond & Carey, 1986). According to the expertise hypothesis, rather than being its own specialised system, face processing is actually part of a broader system for expert object recognition, and our extensive experience with faces is the cause of what appears to be specialised mechanisms for face processing. A fundamental prediction the expertise hypothesis makes is that if people had extensive learning and exposure to other types of objects, then those objects would show the same efficient processing that we typically see in faces (Diamond & Carey, 1986). To test this prediction, researchers have searched for unique effects for faces, and then tested whether those effects are truly “face-specific,” or if they also exist for objects that subjects have a great amount of expertise with—known as objects-of-expertise. For example, birds would be considered objects-of expertise for avid birdwatchers’, or cars for dedicated car enthusiasts. Below I summarise the key evidence for the two hypotheses in studies using behavioural methods, electroencephalography, neuroimaging, and neuropsychology cases.

1.1.1 Behavioural studies. A unique property of face processing is that people predominantly integrate facial features together as a whole when processing a face, rather than processing the individual parts separately (Tanaka & Gordon, 2011). This process of integration is known as holistic processing. Because of the consistent layout of key facial features (e.g., eyes, nose, mouth), holistic integration of a face takes place along an upright T-shaped template. Numerous studies have demonstrated the importance of holistic processing for faces by disrupting this upright T-shaped template through inverting or separating faces into parts. As a result of these disruptions, holistic processing is also disrupted and so is our ability to process faces quickly and efficiently. A classic example of this effect is the Thatcher illusion (Thompson, 1980). A Thatcherised face is a face image in which the eye and mouth features are flipped upside down. When a Thatcherised face is presented upside-down (inverted) people are less able to notice the grotesque nature of the image and to identify what is wrong about the face (see Figure 1). This effect occurs because the upright T-shape template, and therefore holistic processing, is disrupted because the entire face is inverted. Only when the face is seen upright, and holistic processing is not disrupted, does its grotesque nature become obvious.

This content is unavailable.

Please consult the print version for access.

Figure 1. Example of the Thatcher face and illusion (adapted from Thompson, 1980). It is more difficult to tell that Margaret Thatcher's eyes and mouth are inverted in the first image, whereas this alteration is clearer when the face is upright. Participants tend to provide lower grotesque ratings for the first image than for the second image.

Some researchers have proposed that all objects fit somewhere along a holistic-to-feature based processing continuum (Farah, 1991). While most other objects reside somewhere in the middle of the spectrum, faces are unusual in that they rest much farther up the holistic end of this continuum (Tanaka and Gordon, 2011). The difference in the level of holistic processing required for faces compared to objects can be seen with another common holistic effect, known as the inversion effect (see Figure 2a).

The inversion effect refers to the finding that people found it much more difficult to recognise inverted faces compared to upright faces (Yin, 1969). The inversion effect for faces has been demonstrated consistently with various tasks, including recognition, detection, naming, and matching, and across various types of face representations, such as computer-generated faces, drawings, caricatures, and photographs (Valentine, 1988). While objects do show an inversion effect, the inversion effect for faces is consistently found to be disproportionately larger for faces than for other objects (Carey & Diamond, 1977; Farah, Tanaka, & Drain, 1995(a); Scapinello &Yarmey, 1970; Schwaninger, Carbon & Leder, 2003).

Several other tasks also demonstrate face-specific effects—namely the part-whole task and the composite face task. In the part-whole task (Figure 2b, Tanaka & Farah, 1993), recognition of face parts is substantially better in the context of a whole face than in isolation. For example, recognising Jim is easier when presented with two whole faces where only one face part is different, than trying to recognise Jim's eyes, out of two pairs of eyes (see figure 2b). No part-whole effect has been demonstrated with non-face stimuli such as houses or chairs, or even with distorted face stimuli such as scrambled or inverted faces (Pellicano, 2006; Tanaka & Farah, 1993; Tanaka & Sengco, 1997). In the composite face task, the top and bottom halves of two different faces are merged together to create a new composite face (Figure 2c, Young, Hellawell, & Hay, 1987). Participants then need to identify the person in the top half of the composite face, presented either aligned or unaligned. When the two halves are aligned, the face is processed as a whole and holistic processing is uninterrupted. As a result, the two halves of the face merge together, perceptually creating a new identity, and making it harder to isolate the identity in the top half. By contrast, when the faces are unaligned, this connection is broken, holistic processing is disrupted, and it is easier to identify the top half of the face (McKone, 2008; Robbins & McKone, 2003; Young, et al., 1987). The composite effect is much larger for faces compared to other objects and inverted faces (Macchi Cassia, Picozzi, Kuefner, Bricolo, & Turati, 2009; Robbins & Mckone, 2007). Taken together, the Thatcher, inversion, part-whole, and composite effects suggest that face recognition relies on holistic processing—unlike processing of other objects. This difference in reliance on holistic processing is consistent with the idea that face processing involves a separate mechanism to object processing, supporting the face-specific hypothesis.

This content is unavailable.

Please consult the print version for access.

Figure 2. Examples of trials from the Face Inversion task (A), Part-whole task (B) and composite face task (C). Adapted from McKone and Robbins (2007).

The evidence I have outlined so far suggests that faces tend to rely on holistic processing more than other objects. However, it is still possible that this holistic processing is not specific to faces per se, but instead emerges from people's extensive expertise with faces. For example, one study found an inversion effect for dog recognition among dog show judges with 10 years of experience (Diamond & Carey, 1986). Inversion effects have also been found in several studies that assess car recognition among car experts (Curby, Glazek, & Gauthier, 2009; Gauthier et al., 2002; Gauthier, Curran, Curby, & Collins, 2003). In addition, the part-whole effect was found in one study investigating biology experts' ability to identify cells (Tanaka & Gauthier, 1997). These studies are consistent with the idea that "specialised" holistic face processing could be accounted for by more generic expertise mechanisms.

However, one issue with this interpretation is that holistic effects for objects-of-expertise are typically much smaller than those found for faces. For example, the size of the inversion effect for faces ranges from 15 to 30 percent difference when comparing performance on upright and inverted faces. By contrast, the inversion effects for cars or other objects are often not significant (Busy & Vanderkolk, 2005; Robbins & McKone, 2007), or are substantially smaller than the inversion effects for faces—approximately an eight percent difference between upright and inverted conditions (Xu, Liu, & Kanwisher, 2005). One exception is the inversion effect found by Diamond and Carey (1986). In that study, Diamond and Carey found that dog experts' inversion effect for dogs was comparable to their inversion effect for faces. However, a follow-up study suggested that the large inversion effect in Diamond and Carey's study with dog experts could be attributed to the fact that the dog images were not novel to the participants (Robbins & McKone, 2007). The dog images were drawn from the archives of the American Kennel Club and the dog judges themselves were American Kennel Club judges, making it likely that they had had exposure to the exact images in the experiment prior to testing. When the dog experts were tested again using novel dog images, they showed a smaller inversion effect for dogs that was similar in size to the inversion effects seen for other objects-of-expertise, such as cars (Curby et al., 2009).

Advocates of the expertise hypothesis argue that focussing on the magnitude of these effects is irrelevant because greater effects with faces could simply be a product of greater exposure (Gauthier & Bukach, 2007). However, if the magnitude of holistic-processing effects is related to the amount of experience with the object type, experts should show greater holistic effects than novices. In fact, there is evidence against this idea—although a part-whole effect was found in one study investigating recognition of biological cells, experts actually demonstrated a smaller part-whole effect than novices (Tanaka & Gauthier, 1997). This pattern is the exact opposite of the pattern we would expect to see if expertise produces holistic processing and suggests that holistic processing cannot be entirely the result of experience.

Overall, the behavioural evidence shows that many characteristics of face processing appear to be unique to faces and are not seen to the same extent in objects-of-expertise. In turn, this evidence provides more support for the face-specific hypothesis rather than the expertise hypothesis.

1.1.2 Electroencephalography (EEG). Researchers have also found face-specific effects when looking at event related potentials (ERPs), specifically in the N170—a well-documented electrophysiological response that occurs 170 ms after the presentation of any visual stimulus (Bentin, Allison, Puce, Perez, & McCarthy, 1996). The N170 component is consistently larger for faces compared to other non-face objects. This finding has led researchers to use the N170 as a marker of face-specific processing. Consistent with the behavioural evidence, when an inverted face is shown, the N170 component is larger and delayed (Rossion, et al., 2000). This delay supports the conclusion that inverted faces are processed less efficiently than upright faces due to the disruption of holistic processing, while the relative size difference is reflective of the idea that processing of inverted faces incorporates object and face mechanisms (Rossion & Jacques, 2011). Also consistent with the behavioural evidence, inversion effects in the N170 for objects (e.g., cars, chairs, shoes, houses, or novel objects) are negligible.

Although the N170 response is not delayed when people are processing inverted everyday objects, it is still possible that face-specific N170 effects can be explained by people's expertise with faces. For example, a larger N170 when viewing images of finger prints has been found for expert finger print examiners compared to novices (Busey & Vanderkolk, 2005). Additionally, experts demonstrated a delay in their N170 for fingerprints when these images were inverted, perhaps reflecting a disruption in holistic processing. However, as with the behavioural evidence, a similar argument about the magnitude of effects can be made when looking at N170 evidence. Although some expertise N170 effects mimic the pattern of face effects, the magnitude of these expertise related N170s are substantially smaller than those found for faces. Again, the importance of the magnitude of these effects is a topic of debate. However, if greater experience does lead to processing that is similar to faces, then we should see an increase in the magnitude of effects for objects-of-expertise when compared to normal objects. For example, although a car expert may show a smaller N170 for cars when compared to a typical N170 for upright faces, their N170 for cars should still be larger than an N170 response for a normal object, such as a table. Consistent with this idea, bird and car experts show a larger N170 than novices when looking at birds and cars respectively (Tanaka & Curran, 2001). This finding suggests that face-specific N170 effects could be accounted for by expertise mechanisms.

On the other hand, caution needs to be taken when using ERP amplitude alone to make inferences about the neural mechanisms involved in face processing. As a method, EEG is more reliable determining *when* certain brain processes occur than it is at pinpointing *where* they occur in the brain. When brain activation occurs, often several different regions are activated, meaning that most ERPs have multiple brain areas contributing to their activation (Rossion & Jacques, 2011). In fact, in their source localisation review, Rossion and Jacques found that the N170 is a complex ERP response that reflects contributions not only from activation in face-selective areas, but also activation in general object areas. Taking multiple possible areas of activation into account, Tanaka and Curran (2001) found that although the N170 did increase with expertise, this increase was recorded in a set of electrodes that only overlapped with one electrode out of four that are typically associated with the N170 for faces. These results suggest that although expertise does have an effect on people's responses to objects, the N170 for objects-of-expertise is not reflecting the same processes as the N170 response for faces. This evidence suggests that the N170 response for faces cannot be accounted for by expertise mechanisms.

1.1.3 Neuroimaging. Another domain in which face-specific effects have been found is neuroimaging. Multiple regions have been found in the brain that respond more strongly to faces than other objects (known as *face-selective* areas), mostly in the occipito-temporal cortex (Duchaine & Yovel, 2015). Three regions in particular seem to play an important role: the superior temporal sulcus (STS), the occipital face area (OFA), and the fusiform face area (FFA). The presence of these face-selective areas supports the hypothesis that face processing relies on a separate mechanism to object processing. Furthermore, a number of studies in monkeys have also found face-selective cells that respond selectively to certain facial features (Tsao, Freiwald, Knutsen, Madeville, & Tootell, 2003), and whole brain regions that contain only face-selective cells (Tsao, Freiwald, Tootell, & Livingstone, 2006). Taken together, this evidence also supports the face-specific hypothesis.

Of the face-selective areas, the FFA tends to respond most selectively to faces debate (Kanwisher, McDermott, & Chun, 1997; Yovel & Kanwisher, 2004), therefore researchers have investigated how the FFA responds to objects-of-expertise as another way of attempting to resolve the face-specific versus expertise. The evidence from these studies is somewhat mixed. Some studies have found no increase in FFA activation for objects when comparing experts and novices, which suggests that the face-selectivity of this area cannot be explained by expertise (de Beeck, Baker, DiCarlo, & Kanwisher, 2006; Grill-Spector, Knouf, &

Kanwisher, 2004). However, car experts and bird experts have shown increased activation in the FFA when observing their respective objects-of-expertise, compared to novices (Gauthier et al, 2002; Rhodes, Byatt, Michie, & Puce, 2004; Xu, 2005). Similarly, Harley and colleagues (2009) found a positive correlation between FFA activation and level of expertise in radiology students and practicing radiologists looking at chest radiographs. It is important to note, however, that the increased activation seen in experts in these studies is small—processing of face stimuli still produces activation twice as high as any object of expertise (Kanwisher & Yovel, 2006).

Overall, it appears there is neuroimaging evidence to support both the face-specific and expertise hypotheses, although the evidence supporting the expertise hypothesis is not as consistent as the evidence supporting the face-specific hypothesis.

1.1.4 Neuropsychology cases. Although neuroimaging studies can show us what areas are activated by faces and objects-of-expertise, these studies cannot draw conclusions about what mechanisms are crucial and necessary for processing them. To address this issue, researchers have run studies involving patients with impairments in face or object recognition to test for associations and dissociations between object and face processes. Since Bodamer (1947) coined the term prosopagnosia, many cases of patients suffering from face-specific deficits have been reported. Of course, it could be that prosopagnosia is not face selective, but is instead characterised by general difficulty discriminating objects-of-expertise. Two case studies provide important evidence against this idea. One of these—prosopagnosic patient WJ—acquired prosopagnosia after a series of strokes, and soon after he took up a career as a sheep farmer (McNeil & Warrington, 1993). Despite exhibiting severe prosopagnosia with human faces, WJ showed extraordinary skill at recognising highly similar sheep faces. In fact, WJ performed better than controls who were matched on experience with sheep. WJ continued to perform at this high level even when his task was to learn and recognise unfamiliar sheep. All of WJ's extensive experience with sheep occurred after his stroke, suggesting he was able to acquire expertise in sheep recognition even though he was impaired with face recognition. This dissociation between WJs ability to gain expertise and his ability to process faces suggests that face processing and expertise processing rely on separate mechanisms, supporting the face-specific hypothesis.

Expanding on the results showing that the ability to gain expertise does not rely on the mechanisms used for face-processing, the second key case-study indicates that face processing does not rely on the mechanisms used for general object processing. Patient CK suffers from reading difficulties and an inability to recognise objects such as airplanes and toy soldiers—and these toy soldiers were arguably objects-of-expertise for CK, who was adept at recognising the thousands of toy soldiers in his collection before his impairment (Moscovitch, Winocur, & Behrmann, 1997). But despite CK's current deficiencies in object processing, he has no difficulties with face recognition (Kanwisher, 2000). Taken together, the evidence from the cases of WJ, and CK suggests there is a double dissociation between face and object recognition, providing strong evidence that face processing relies on separate neural mechanisms to the processing of normal objects (and arguably a separate mechanism to objects-of-expertise as well). However, reports of expertise in these cases, especially in the case of CK, are anecdotal. Without a way of systematically assessing prior or current expertise, the cases of CK and WJ provide only preliminary evidence that face, and expertise mechanisms may be separate.

1.2. Studying expertise with novel objects.

One problem with studies that use real world experts is that the test stimuli used in those studies tend to be common every-day objects, such as cars. As a result, it is likely that even "novices" have substantial experience at processing and recognising those objects—which could explain why the expertise effects found in these studies tend to be smaller than the equivalent face effects. As well as being unable to control the experience of novices, researchers also do not have control over the experience that the "experts" in these studies have with their objects-of-expertise. For example, two expert dog show judges recruited with a 10-years of experience criteria could have wildly different levels of exposure to dogs over that 10-year period. This lack of control makes it difficult to interpret the size of the expertise effects seen in these studies. Furthermore, expertise effects depend on how similar the stimuli in the study are to the exact objects the experts are highly experienced with. For example, Tanaka, Curran and Sheinberg (2005) demonstrated that expert owl recognisers were comparable to novices when it came to identifying wading birds, and Diamond and Carey (1986) found that expertise effects were larger when dog show judges were identifying breeds of dogs they had the most experience with. Because the exact domain of expertise will vary from expert to expert, stimuli would need to be tailored for each participant to properly assess their performance with objects-of-expertise. For all these reasons, the use of real world experts makes the design and interpretation of expertise studies difficult, which in turn makes it difficult to distinguish between the face-specific and expertise hypotheses.

One solution to these problems is to train participants to become experts with completely novel objects A number of novel objects (sheinbugs from Richler, Wilmer, & Gauthier, 2017; ziggerins from Wong, Palmeri, & Gauthier, 2009, and greebles from Gauthier & Tarr, 1997; see Figure 3) have been carefully designed to share similar properties with faces—variation in features across different individuals, but a consistent configuration of those features across individuals. Participants are then extensively trained so they acquire lab-based expertise with these novel objects. By training expertise this way, researchers can control the amount of exposure participants have to these objects. As an added advantage, this design allows for within-subjects comparisons (i.e., the same participants before and after training), which allows for more power than studies using separate novice and expert groups.

For my thesis, I focus on using the novel object greebles, because they are used in the original training paradigms (Gauthier & Tarr, 1997, Gauthier, Williams, Tarr, & Tanaka, 1998) and have been used extensively in the training literature (see Appendix E).

This content is unavailable.

Please consult the print version for access.

Figure 3. Examples of different novel objects created for training studies. From left to right: Sheinbugs, Greebles, and Ziggerins.

1.3. Greeble studies

When creating the greeble stimuli, Gauthier and Tarr (1997) argued that there were three important characteristics of faces that may be important for the way they are processed. The first characteristic is that faces all have similar features to one another, and—more importantly—that the configuration of these features is the same. For example, every face has eyes, a nose, and a mouth, and the relative position of those features is similar from one face to the next. To match this characteristic, every greeble has similar features—bogues, a quiff and a dunth (see Figure 4). People use this individual information because broadly labelling an individual as a "person" usually does not provide enough information to be useful in our everyday lives. By contrast, when people interact with objects, they tend to label them with a broad category such as "chair" or "car" because more specific or individual information is often not required or useful for carrying out everyday activities (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). Therefore, the greeble paradigm was designed to mirror this tendency for identifying people at an individual-level. For example, every greeble has an individual name, as well as a family name that is shared with other similarly shaped greebles. The third proposed key characteristic of faces is that people are experts at recognising them. To address this characteristic, researchers included a criterion to serve as a marker for expertise. As mentioned above, we have a bias for naming objects by their broad category label, like chair or car. However, when someone becomes an expert with a type of object, they become just as fast at recognising objects of that type by their individual or specific labels as by their broad category label (Tanaka & Taylor, 1991). Therefore, the criterion proposed by the researchers was that when individual and family naming are equally fast, it means that expertise has been achieved with greebles (Gauthier & Tarr, 1997). Because the greeble paradigm mimics these three main characteristics of faces, the expertise hypothesis predicts that expertise with greebles should elicit ‘face-specific’ effects.

This content is unavailable.

Please consult the print version for access.

Figure 4: Image of a greeble with part labels. Image adapted from Gauthier and Tarr (1997).

Using this paradigm, researchers have provided evidence that the expertise hypothesis could account for many of the supposedly face-specific effects. For example, people show a larger inversion effect for greebles after training, compared to before training, which suggests holistic processing of greebles increased with experience (Gauthier et al., 1998; Rossion, Gauthier, Goffaux, Tarr, & Crommelinck, 2002). The composite effect and whole-part advantages typically seen with faces have also been found in greebles after training (Gauthier et al., 1998; Gauthier & Tarr, 2002). Likewise, researchers have found that the N170 for greebles was larger after training than before training (Rossion et al., 2002), and was of similar magnitude to the N170 response to faces. The time delay for the peak amplitude of the N170 commonly associated with inverted faces has also been found for inverted greebles after training (Rossion, et al 2002). Furthermore, processing of greebles after training has been shown to tax the same neural mechanisms as face processing (Rossion, Kung, & Tarr, 2004). More specifically, showing faces immediately after greebles results a significantly reduced N170 for the faces compared to when faces are presented in the absence of greebles—perhaps because neural mechanisms involved in processing both faces and greebles do not have sufficient time to reset and activate again. Finally, participants show increased FFA activation to greebles after training (Gauthier & Tarr, 1997; Gauthier, Tarr, Skudlarski, & Gore, 1999). Crucially, the same 2 or 3 voxels that are most active for face processing also show the largest increase in activity over the course of expertise training on greebles, suggesting that the FFA is recruited as participants learn to discriminate similar novel objects, and is used for processing objects where a person has expertise (Tarr & Gauthier, 2000).

Taken together, these results suggest that processing of objects-of-expertise involves similar mechanisms to those used when processing faces, which provides support for the expertise hypothesis. However, although these expertise effects mimic those found for faces, it is once again the case that these effects tend to be smaller than those seen for faces. To get past this stalling point in the debate, some researchers are taking a new approach—investigating the acquisition of expertise in individuals with impaired face processing due to prosopagnosia.

1.4. Greeble studies in prosopagnosia

Although the above greeble studies demonstrate that people show similar patterns of responding for faces and objects after training, these studies don't necessarily suggest the processes behind these patterns are one and the same. One solution to this problem is to look for dissociations between face and expertise processing in patient samples, which provides a stronger test of whether face and expertise processes operate using the same or separate mechanisms. Prosopagnosics are impaired with face recognition despite otherwise normal vision and broader cognitive functioning. The expertise hypothesis states that the face processing is part of larger mechanisms for wider expert recognition, therefore, this hypothesis predicts that those with prosopagnosia should also be unable to acquire expertise with greebles. By contrast, the face-specific hypothesis—which suggests that expertise and face processing mechanisms are separate—predicts that impairments in face processing should not impair people's ability to acquire expertise with novel objects. Because the two hypotheses make competing predictions about prosopagnosics' ability to acquire expertise, the greeble paradigm offers a compelling way to tease apart the face-specific and expertise hypotheses.

Four studies have looked at whether prosopagnosics can acquire expertise with greebles. In two of these studies, acquired prosopagnosics SM and LR were unable to learn to recognise greebles as accurately as controls —although with additional training LR was eventually able to (Behrmann, et al., 2005; Bukach et al., 2012). SM and LR's deficits for faces and their failure to gain expertise with greebles suggest that both types of stimuli could be handled by the same impaired mechanism, consistent with the expertise hypothesis. However, SM and LR performed worse than controls even as novices (before learning took place). This prior deficit in performance raises the possibility that their failure to obtain expert recognition with greebles may have nothing to do with expertise acquisition but is rather caused by broader deficits with object recognition (Gauthier, et al., 1999) or learning names (Bukach et al 2012). In other words, it might not be that faces and objects-of-expertise are processed using the same mechanism, but rather that faces and objects-of-expertise are processed by separate mechanisms that both happened to be impaired.

A third study tested developmental prosopagnosic Edward (Duchaine, et al., 2004). Edward appears to have “pure” prosopagnosia since he performed well above the normal range on a wide variety of non-face visual recognition tasks. Unlike SM and LR, Edward was able to become a greeble expert at the same rate of learning as controls. Edward’s ability to acquire expertise in identifying greebles despite his face recognition impairments suggests that the mechanisms for face and expert object processing are dissociated, which is consistent with the face-specific hypothesis. However, this study suffers from “circular analysis”, in that the same data (signalling Edward’s prosopagnosia) were used for selecting Edward into the study and for making the inference that Edward would have been impaired with faces had he been asked to learn them—Edward did not ever complete a training procedure with faces. As a result, this study leaves an important question unanswered—if Edward had been presented with a similar face training paradigm, would he have shown impairments in learning faces, or would he have learnt equally as quickly and accurately as he did with greebles? If Edward were able to learn faces as quickly as greebles, then his ability to learn greebles could not be taken as evidence that face and expert object processing are dissociated.

A fourth study remedied this problem by running the greeble training paradigm alongside a matched face training paradigm. This study tested two acquired prosopagnosics, Herschel and Florence (Rezlescu et al., 2014). Both Herschel and Florence were impaired when learning faces, yet they matched controls in their ability to expertly learn to identify greebles. These results suggest that specialised mechanisms for faces and expert objects are not the same and can be selectively impaired, which the researchers suggest provides compelling support for the face-specific hypothesis of face processing. One criticism of this interpretation, though, is that Herschel and Florence might have succeeded at learning greebles using atypical strategies such as distinguishing between individual greebles based on a particular feature rather than holistically (Gauthier, 2014). These strategies could not be applied to faces because faces are highly similar to one another.

1.5. Overview of thesis

In this thesis, I aim to address whether face processing relies on face-specific or expertise mechanisms by conducting an extended replication of Rezlescu and colleagues (2014) study using a larger sample of 10 developmental prosopagnosics (from now on prosopagnosics). If face processing relies on face-specific mechanisms, prosopagnosics would be able to acquire expertise with greebles similar to controls, and that prosopagnosics would perform better with greebles than with faces. However, should face processing rely on expertise mechanisms, then prosopagnosics should be impaired in their ability to acquire expertise with greebles, relative to controls, and that performance with greebles should be the same as performance with faces.

Previous training literature provides some evidence of face learning in individuals with prosopagnosia (Rezlescu et al., 2014). Therefore, my second aim is to investigate whether prosopagnosics, despite their lifelong deficits at recognising faces, are able to learn to identify faces in this training paradigm. If there is any evidence of learning, it would indicate that training programs could be an effective way of rehabilitating the deficits experienced by prosopagnosics. In addition, I want to investigate *how* prosopagnosics are able to learn faces. Specifically, I want to investigate whether they are using an alternate strategy to controls to learn faces. Therefore, I will also include an inverted faces training paradigm and compare prosopagnosics' performance with upright faces to their performance with inverted faces. Inverted faces are more difficult to learn and recognise for controls, which is reflected in a large inversion effect for faces. If prosopagnosics do not show this typical inversion effect, those results would suggest that prosopagnosics are using an alternative strategy to learn faces. However, should prosopagnosics show a similar inversion effect to controls such that their performance drops when learning inverted faces, those results would suggest that prosopagnosics are using similar (but compromised) mechanisms to process faces.

My third aim is to identify potential subtypes of prosopagnosia. There is some evidence that prosopagnosia is a heterogeneous condition. For example, some prosopagnosics only have face recognition impairments, while others have trouble recognising objects as well (Duchaine & Nakayama, 2005). However, research into potential subtypes of prosopagnosia is scarce, which means our understanding of the different causes and consequences of the disorder is limited. I will attempt to address this gap, by investigating individual differences in the patterns of performance among prosopagnosics. By doing so, I will be able to further investigate possible variations and subtypes of developmental prosopagnosia.

2. Method

2.1. Participants

Prosopagnosics. Ten prosopagnosics (seven women, two men, one other) were sourced from the prosopagnosia database www.faceblind.org. Their mean age was 25.20 (*SD* = 3.52). All prosopagnosics completed a standard diagnostic battery which includes the self-report Prosopagnosia Index 20 (PI20, Shah, Gaule, Sowden, Bird, & Cook, 2015), two objective tests of face recognition namely the Cambridge Face Memory Test (CFMT; Duchaine & Nakayama, 2006) an in-house famous face test, and the low/mid-level visual battery the Leuven Perceptual Organisation Screening Test (L-POST; Torfs, Vancleef, Lafosse, Wagemans, de-Wit, 2014). All met my diagnostic criteria demonstrating impairment (1.75 standard deviations below control mean) on PI-20, CFMT and famous face test, but normal performance on L-POST (no more than 4 sub-tests failed). Additionally, participants needed to report no prior history of brain injury. Prosopagnosics were selected from the prosopagnosia database based on the above exclusion and inclusion criteria (see Table 1). All prosopagnosics were currently living in the US, UK, or Canada.

The prosopagnosics also completed the Cambridge Car Memory Test (CCMT; Dennett et al., 2012) to ensure their ability to recognise objects is normal. This is critical because if participants are impaired in their ability to recognise objects, then failure to acquire expertise with objects could be due to impairments in basic object recognition, rather than to deficits to expertise mechanism. The CCMT has been widely used as a proxy for object recognition (Palermo et al., 2016; Kumfor et al., 2015; Rezlescu, Pitcher, & Duchaine, 2012; Richler, Wilmer, & Gauthier, 2017). All prosopagnosics performed within the normal range on the CCMT (no lower than 1 standard deviation below the general population mean; Dennett et al., 2011), suggesting they had no impairments in their general object recognition abilities.

Controls. Ten neurotypical participants (eight women, two men) were sourced from the Victoria University of Wellington community in Wellington, New Zealand. The mean age of participants was 25.60 (*SD* = 3.44). All participants completed the CFMT and the CCMT to ensure normal face and object recognition, and all performed within the normal range on both tests (z-scores between 1 and -1; see Table 2).

Controls were given a total of 150 NZD worth of grocery vouchers for their participation, while prosopagnosics received 150 USD Amazon gift vouchers. Average and individual performance on all screening measures is displayed in Tables 1 and 2.

Table 1**.** *Average and individual demographic information and screening task scores for prosopagnosics. CFMT, CCMT, FFT and PI20 performance are presented as z-scores. L-POST scores indicate how many out of 15 sub-sets participants failed. Standard deviations are presented in parentheses.*

Participant	Age	Gender	CFMT	CCMT	FFT	PI20	L-POST
Prosopagnosic Average	25.20 (3.52)	-	-2.70 (0.62)	0.14 (0.60)	-4.55 (1.37)	4.06 (0.27)	-
DP1	25	Female	-1.83	-0.82	-5.72	3.78	0/15
DP2	24	Other	-3.07	-0.04	-6.19	4.70	0/15
DP3	30	Female	-1.97	0.71	-6.03	3.87	0/15
DP4	22	Female	-2.52	0.52	-4.50	3.78	1/15
DP5	21	Female	-3.48	-0.37	-3.50	4.05	1/15
DP6	29	Female	-2.10	0.64	-1.78	4.05	1/15
DP7	21	Female	-3.51	0.75	-3.33	4.05	0/15
DP8	23	Male	-2.79	0.61	-5.43	3.96	0/15
DP9	27	Female	-3.48	0.72	-5.42	4.42	0/15
DP10	31	Male	-2.24	-0.88	-3.63	4.24	0/15

Table 2. *Average and individual demographic and screening task scores for controls. CFMT and CCMT performance presented as z-scores.*

Participant	Age	Gender	CFMT	CCMT
Control Average	25.60 (3.44)	-	0.69 (1.00)	0.05 (0.37)
C1	30	Female	1.82	-0.60
C2	21	Female	-0.38	0.19
C3	26	Female	1.13	-0.04
C4	22	Female	-0.65	0.08
C5	29	Male	1.40	0.21
C6	22	Male	-0.92	0.41
C7	23	Female	0.31	-0.15
C8	30	Female	0.31	-0.25
C9	26	Female	1.68	0.72
C10	27	Female	1.82	-0.09

2.2. Design

The current experiment used a 2 (Group: control, prosopagnosics) x 3 (Training Condition: greebles, upright faces, inverted faces) mixed-model experimental design, where group is a between-subjects factor, and the training condition is a within-subjects factor.

2.3. Stimuli

Images. The image stimuli comprised one set of greeble images and two sets of face images (Figure 5). Each set contained 20 target images and 60 distractor images. The greeble set and one of the face sets were identical to those used in a previous study (Rezlescu et al., 2014), and the greeble set was a subset of a larger set used in the original greeble study (Gauthier, et al., 1997). A second face set was created using FaceGen Modeller (Singular Inversions, 2017) to allow different sets of faces to be used for the upright and inverted face conditions. To ensure that the two face sets were comparable, I created a large set of face images that were as similar to one another as the face images in the first set. Similarity was quantified on a pixel-by-pixel basis where each pixel was assigned a number from 0 to 255 based on the gradient of the pixel from white to black respectively. For a thorough description of these face image comparisons and the results of the comparisons, see Appendix A. The orientation (upright vs inverted) of the two face sets was counterbalanced across participants.

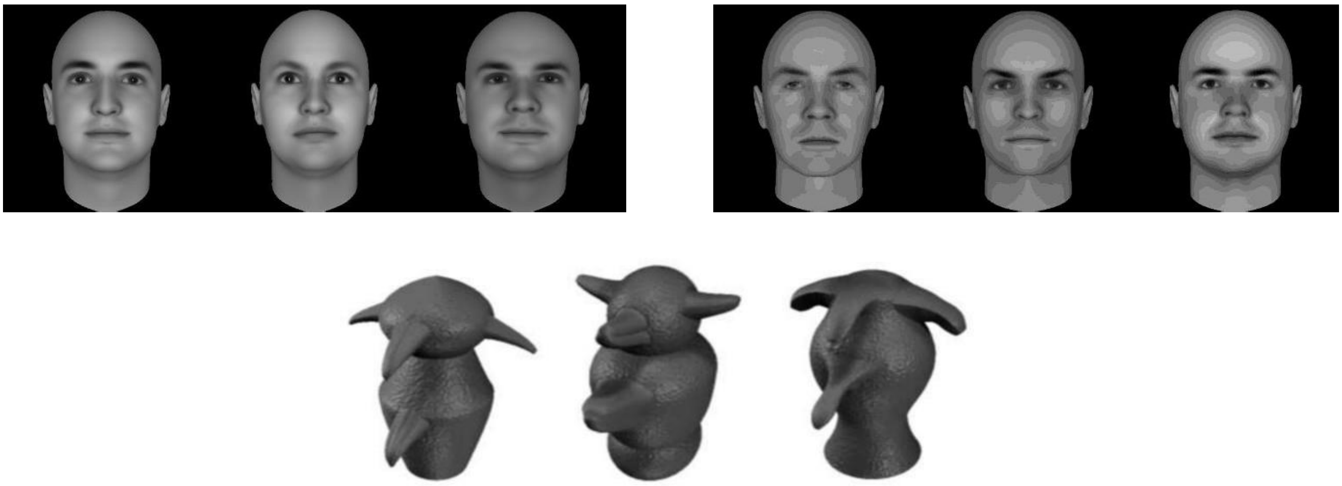


Figure 5. Exemplar images of individual from Face Set 1 (A), Face Set 2 (B), and Greeble Set (C). See Appendix B for entire sets.

Names. Three separate individual name sets, and one family name set were used in this study (see Table 3). Each individual name set was comprised of 20 names. Two of the individual name sets were identical to those used previously (Rezlescu et al., 2014), while the third was created for this study. For each participant, each condition (upright faces, inverted faces, greebles) was assigned one of the three individual name sets at random. All names were randomly generated four letter nonsense names. Each name within a set started with a different letter, and the first letters of each individual name were consistent between individual name sets. Individual name sets were only used once per participant, so that each participant saw all three name sets across the entire experiment.

The family name set was used only in the greeble condition. There were five family names in total in this set, each beginning with a different letter, with each family name corresponding to a specific body shape (see Figure 6). Greebles within the same family all share a similar body shape and are different from one another in a combination of parts and body shape. Note that because family conditions were unable to be created for human face conditions, all comparisons between training conditions were done for individual names only.

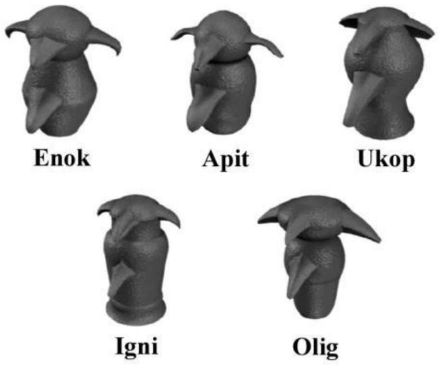


Figure 6: Image of the five different types of greeble families. Each family has its own distinct body shape.

Table 3. *Name sets used in this study. Each participant saw each name set once throughout the entire experiment.*

Name Set 1	Name Set 2	Name Set 3	Greeble Family Name Set
Bugy	Bogo	Buru	Apit
Crin	Corg	Curn	Enok
Dron	Dorl	Drib	Igni
Fisu	Feka	Fati	Olig
Gris	Grup	Gorl	Ukop
Haki	Halk	Hato	
Ibby	Illo	Inli	
Jora	Julu	Jiti	
Kuru	Kani	Koki	
Lorc	Lodi	Logi	
Maur	Mora	Mara	
Puri	Pria	Pebu	
Quom	Quop	Quol	
Remi	Riap	Rive	
Sarl	Sina	Suru	
Troc	Triz	Trul	
Vank	Veni	Vanu	
Wali	Wolo	Waru	
Yulu	Yaju	Yani	
Zeli	Zolt	Zunk	

2.4. Tasks

These tasks are similar in methodology to the ones use in the original greeble training paradigm (Gauthier and Tarr, 1997). The goal of each training condition was for participants to learn the names of 20 target individuals. Training with faces involved four separate tasks, outlined below. Training with greebles involved the same four tasks, plus three additional family training tasks.

Individual Inspect. Participants were presented with the image of an individual along with the name of the individual, which was presented below the image. Participants could take as long as they needed to memorise the image.

Naming with Feedback. Participants were shown only the image of an individual. Participants pressed the key on the keyboard that corresponded to the first letter of the individual's name (e.g., if an image of Dron was shown, participants would press the letter 'd'). The individual remained on screen until the participant responded. If the response was incorrect, the participant was provided feedback that their response was incorrect. Irrespective of whether the response was correct or incorrect, the participant was then shown an image of the same individual a second time, this time with the individual's name below the image. Participants were given as much time as they needed to review the image.

Individual Naming. Participants were shown an image of an individual and were tasked with pressing on the keyboard the first letter of the name that corresponded to that image. Again, the individual remained on screen until the participants responded, and participants were provided feedback if they responded incorrectly. However, no review image was shown after their response. Participants were informed that on some trials, the individual would be an individual that had not been learned (distractors), and that the correct response to a distractor was to respond "No Name" by pressing 'n' on the keyboard.

Individual Verification. Participants were presented with a name for 1000ms, followed by an image of an individual which stayed on the screen until participants responded. Participants indicated whether the name and image matched or mismatched by pressing 1 (match) or 0 (mismatch) on the keyboard.

Family Inspect (greeble training only). Similar to the individual inspect trials, four individuals from each family (20 in total) were shown for as long as participants needed, so that they could memorise the shared characteristics for each greeble family.

Family Naming (greeble training only). This task followed the same procedure as the individual naming task, except that participants were tasked with pressing the first letter of the appropriate family name for the greeble displayed on the screen.

Family Verification (greeble training only). This task followed the same procedure as the individual verification task, except that participants were presented with a family name followed by an image of a greeble. These family verification trials were intermixed with individual verification trials in the greeble condition.

2.5. Sessions

Each training condition consisted of eight sessions spread across eight consecutive days (one session per day). Sessions 1 to 4 made up the *learning phase*. Sessions 5 to 8 made up the *testing phase*.

Sessions 1 to 4 consisted of varying numbers of tasks and trials (see Table 4 for details about the number of total trials presented in each session). The full 20 target individuals were gradually introduced such that five individuals were shown in Session 1, 10 in Session 2, 15 in Session 3, and all 20 by Session 4. Sessions 1 to 4 took approximately one hour each to complete. Sessions 5 to 8 consisted of 60 naming and 60 verification trials in the face conditions, and 60 naming and 120 verification trials in the greeble condition (60 individual verification trials and 60 family verification trials). Sessions 5 to 8 each took approximately 15 minutes to complete. In total, participants were training for approximately 5 hours spread over the course of eight days for each training condition.

Table 4. *Total number of trials for each task in each session across all conditions*[1]*. Trial numbers are the same as those used in Rezlescu et al. (2014). Tasks with asterisks are included in the greeble condition only. See Appendix C for each sessions structure.*

Task	Session 1	Session 2	Session 3	Session 4	Sessions 5 - 8
Individual Inspect	20	30	30	40	-
Naming with Feedback	15	30	45	60	-
Individual Naming	120	120	180	180	60
Individual Verification	130	210	195	180	60
Family Inspect*	20	-	-	-	-
Family Naming*	60	-	-	-	-
Family Verification*	120	180	180	180	60

2.6 Procedure

Training was completed online using Testable (Testable S.R.L, 2017). The three training conditions were counterbalanced across participants. Within one condition (upright

[1] Due to technical difficulties, 60 individual naming trials in session 3 of the greeble condition were missing across all participants, meaning that participants had less time training with naming individual greebles than in the other two conditions. Although problematic, the missing trials cannot explain the patterns of results found in this study. Prosopagnosics performed better with greebles than they did with faces—if anything, these missing trials would be working against the dissociation I observed between prosopagnosics performance on greebles and their performance on upright faces.

faces, inverted faces or greebles) one session was sent to participants each day for eight consecutive days. For control participants, each daily session was sent to participants at 9am NZT. Reminders were also sent out at 6pm each day. For US, Canadian and UK prosopagnosics, sessions were sent out at 8pm NZT (4am EDT, 1am PDT and 9am GMT). Session reminders for US and Canadian participants were sent at 6pm EDT and 3pm PDT, while session reminders for UK participants were sent at 8pm GMT. Participants could complete their session at any point during the day but were asked to complete each session in a space that was comfortable for them and that was free of distractions. Participants were not sent the next session until they had completed the previous session.

Between each training condition, participants were given a minimum of one-weeks break. Due to scheduling constraints, some participants had a longer break between conditions—the longest of which was three weeks.

3. Results

Similar to previous literature (Duchaine et al., 2012; Gauthier et al., 1998; Gauthier et al., 1999; Rezlescu et al., 2014), I focused my analysis on the testing phase (sessions 5 - 8), because it is typically assumed that expertise has been reached by the fifth session. Another advantage of focusing on the testing phase is that the number of trials across sessions in the testing phase sessions is consistent—unlike the learning phase. Average accuracy and response time were calculated for each participant for each session across all conditions. These calculations were done separately for naming and verification trials. For the greeble condition, there were separate calculations for naming, overall verification, verification family, and verification individual trials. For each participant, in each session, response times that were two standard deviations above the calculated mean for that participant in verification and naming tasks were excluded. All statistical analyses were conducted in Jamovi (Jamovi Project, 2017). Mean scores for each training condition for each participant were calculated by averaging the scores from the four testing phase sessions. Due to technical errors, participant Control7's (C7) upright face session two, and inverted face session one data were not recorded. For graphing purposes only, a stand in score was created by averaging performance from the two sessions either side of the missing data (upright faces session two) or using the same score as the score in next session (inverted faces session one; Rezlescu et al., 2014).

3.1. Group-level analysis.

3.1.1 Did prosopagnosics show a dissociation between upright faces and greebles? The face-specific hypothesis allows for a dissociation between performance at recognising upright faces and greebles, whereas the expertise hypothesis predicts an association between performance for the two stimuli types. I therefore begin by testing whether prosopagnosics show comparable learning of upright faces and greebles, across both naming and verification trials, relative to controls (Figure 6).

For naming trials, I conducted a 2(Group: control, prosopagnosic) x 2 (Condition: greeble, upright faces) ANOVA. This ANOVA revealed a significant Group x Condition interaction ($F(1,18) = 17.4$, $p < .001$, partial $\eta^2 = .492$). Follow up paired samples t-tests show that in the greeble condition controls performed similarly to prosopagnosics ($t\ (29.9) = .701$, $p = .896$) (See Table 5). In contrast, in the upright face condition, controls performed better than prosopagnosics ($t(29.9) = 5.071$, $p <.001$; see Table 5). A similar result was obtained with the verification trials. The 2 x 2 ANOVA produced a significant Group x Condition interaction ($F(1,18) = 5.62$, $p < .001$, partial $\eta^2 = .143$). This significant interaction was driven by two trends in the data. More specifically, control and prosopagnosics appear to perform more similarly in the greeble condition (*MDiff* = 1.10%, $t(29.9) = .394$, $p = .979$), than in the upright face condition (*MDiff* = 7.4%, $t(29.9) = 2.651$, $p = .060$; see Table 5), although neither of these post hoc analyses are significant.

Additionally, follow up paired-samples *t*-tests shows that there was no difference in controls' performance in the greeble and upright face conditions for either naming ($t(18) = 0.284$, $p = .992$) or verification ($t(18) = 1.489$, $p = .464$). This result demonstrates that the greeble and upright face conditions were matched for difficulty.

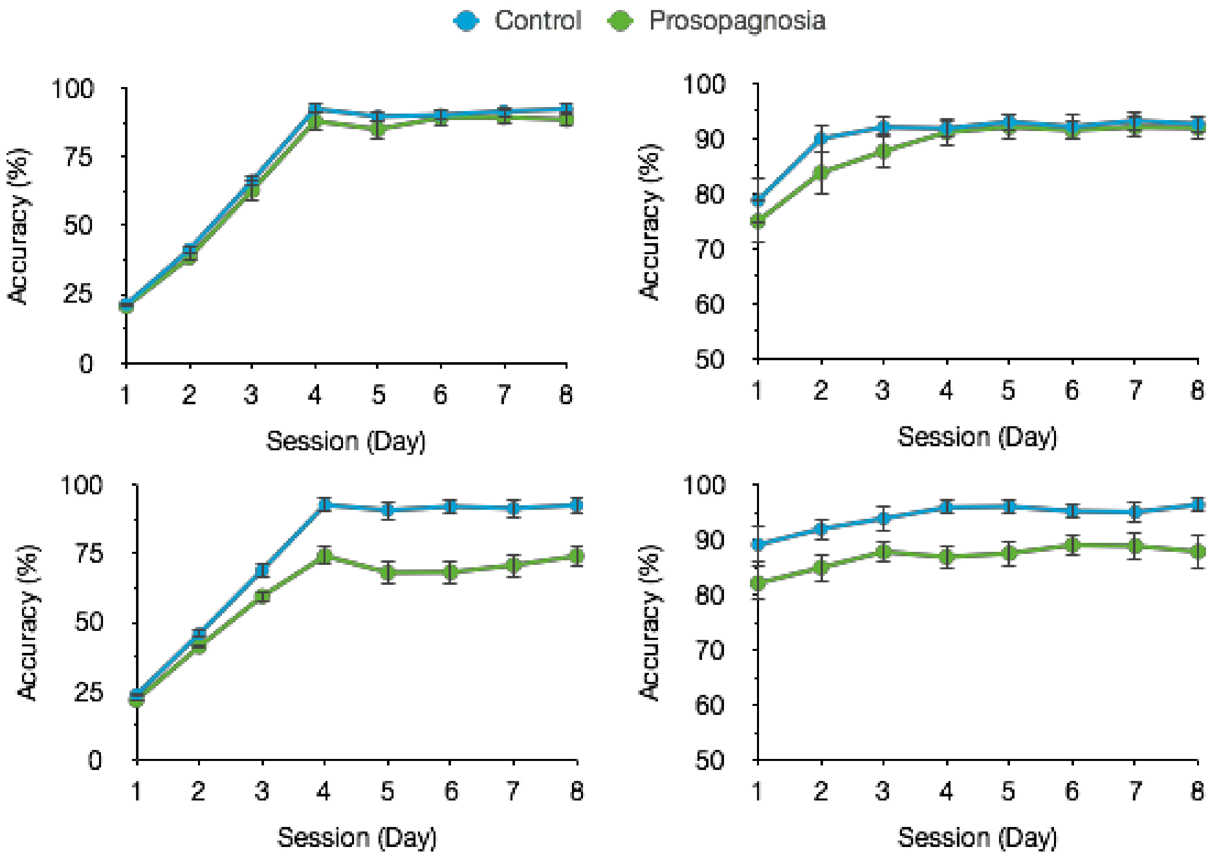


Figure 6. Top Panel: Accuracy performance of controls (blue) and prosopagnosics (green) in the naming (left) and verification (right) tasks in the greeble condition. Bottom panel: Accuracy performance of controls (blue) and prosopagnosics (green) in the naming (left) and verification (right) tasks for upright faces conditions. Naming scores have been scaled corresponding to the number of individuals learned in that session to reflect the varying difficulty in each session (five individuals in session 1, ten in session 2, fifteen in session 3, and twenty from session 4 onwards). For example, in session 1, because only 5 individuals (25% of the total 20) were presented, the maximum possible naming score is 25%. Chance performance is at 50% for verification trials.

Next, I analysed reaction times to check whether speed-accuracy trade-offs could explain the dissociation between performance for upright faces and greebles observed in the prosopagnosic group. Specifically, I examined whether prosopagnosics were slower than controls when naming and verifying greebles, and faster when naming and verifying upright faces. Figure 7 displays the RT of controls and prosopagnosics in the naming and verification tasks for greeble and upright face conditions. For naming trials, there was no significant main effect of Group (F(1,18) = 0.079, p = .783, η_p^2 = .004) or Condition (F(1,18) = 0.333, p =

.571, $\eta_p^2 = .018$). No significant interaction was found between Group and Condition either ($F(1,18) = 0.279$, $p = .604$, $\eta_p^2 = .015$). For verification trials, there was no significant main effect of Group ($F(1,18) = 0.197$, $p = .663$, $\eta_p^2 = .011$), but a significant main effect of condition ($F(1,18) = 8.641$, $p = .009$, $\eta_p^2 = .323$) was found such that the verification of upright faces was slower than the verification of greebles (see Table 5). No significant interaction was found between Group and Condition ($F(1,18) = 0.133$, $p = .720$, $\eta_p^2 = .005$). These results suggest there was no speed-accuracy-trade-offs among prosopagnosics for naming or verification.

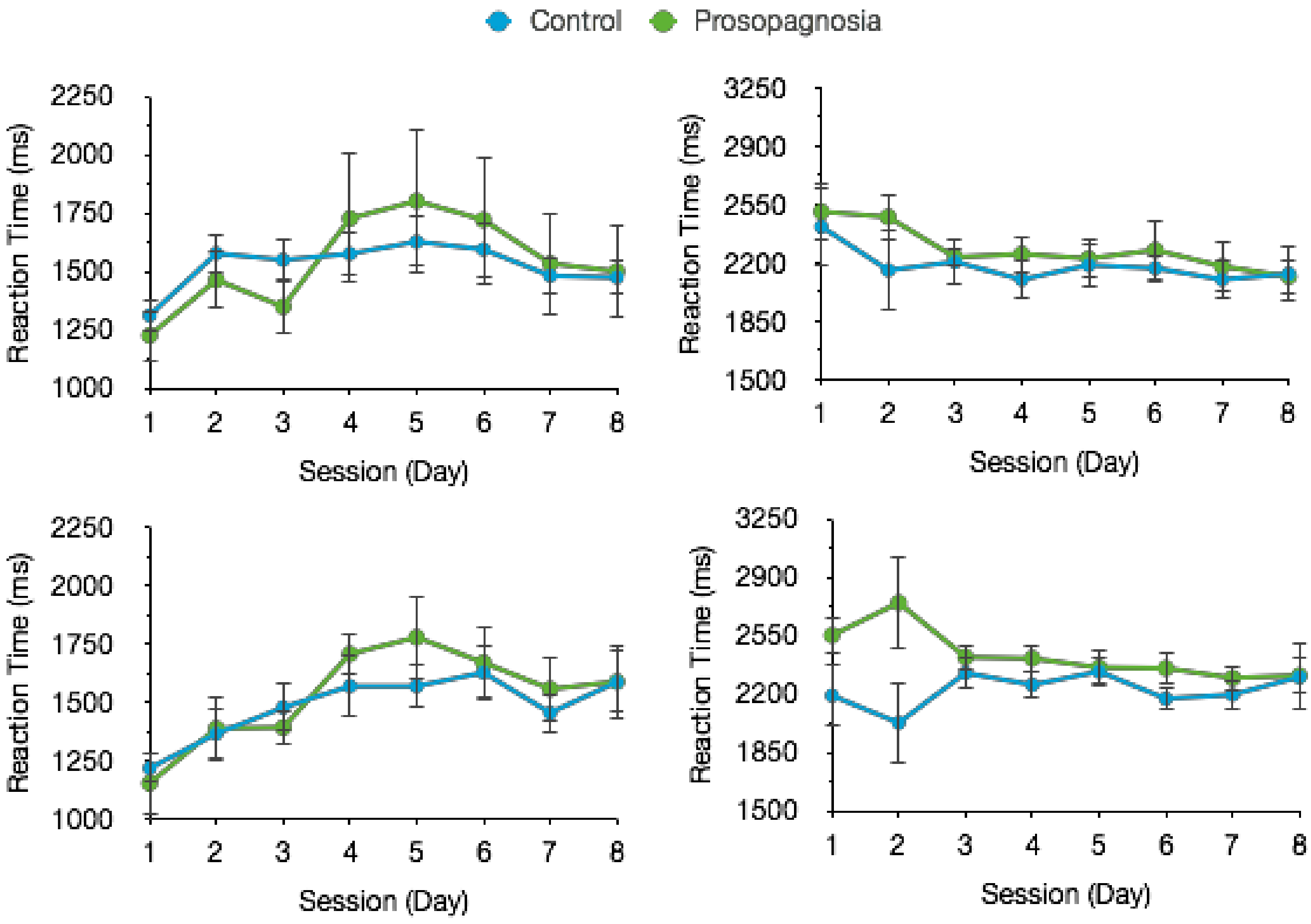


Figure 7. Top panel: Average RT of controls (blue) and prosopagnosics (green) in the naming tasks for greeble condition in each session. Bottom panel displays average RT of controls (blue) and prosopagnosics (green) in the verification tasks for upright face conditions in each session.

In sum, prosopagnosics showed a dissociation between upright faces and greebles across both naming and verification trials, such that they always performed better with recognising greebles than recognising faces. In contrast, controls learned both types of stimuli equally well. Response time analyses show that the prosopagnosics' dissociation cannot be accounted by speed-accuracy trade-offs. Taken together, these results provide compelling evidence for a dissociation between upright faces and greebles in prosopagnosia, which supports the face-specific hypothesis and is inconsistent with the expertise hypothesis.

Table 5. *Average accuracy rate (%) and reaction time (ms) in naming and verification tasks for prosopagnosics and controls in the upright face and greeble conditions. Standard deviations are presented in parentheses.*

	Upright Faces				Greebles			
	Accuracy - %		Reaction Time - ms		Accuracy - %		Reaction Time - ms	
	Proso	Control	Proso	Control	Proso	Control	Proso	Control
Naming	70.21	91.90	1643	1617	88.00	90.92	1640	1546
	(12.37)	(9.39)	(489)	(331)	(9.52)	(5.87)	(781)	(243)
Verification	88.30	95.70	2330	2248	91.80	92.90	2200	2147
	(7.94)	(4.52)	(300)	(306)	(6.56)	(5.40)	(413)	(364)

3.1.2. Did prosopagnosics learn greebles as well as did controls? Despite the dissociation found in prosopagnosics between performance for upright faces and performance for greebles, it may be that prosopagnosics were still impaired in their ability to learn greebles. To address this issue, I compared performance on the greeble condition between prosopagnosic and control groups on both family and individual verification. I also investigated whether performance of prosopagnosics met the expertise criterion proposed by the creators of the greeble paradigm (Gauthier et al., 1998).

To test whether prosopagnosics learned greebles to the same extent as controls, I conducted independent samples *t*-tests to assess group differences in accuracy and RT. Analyses were run for naming, overall verification (including both family and individual trials), verification family, and verification individual trials separately, and only on sessions trials during the testing phase (sessions 5-8). For each participant, a mean accuracy score was computed for each trial type across the four sessions of the testing phase. A mean score was then computed for both the control and prosopagnosic groups for each trial type.

Table 6 displays the means, standard deviations, and *p*-values for accuracy and reaction time for both controls and prosopagnosics for greeble naming, overall verification, family verification, and individual verification. Performance was comparable between controls and prosopagnosics across all four tasks (*p* = .407, .687, .895, and .443 respectively).

Next, I analysed response times to check whether speed-accuracy trade-offs could explain the similar performance found between controls and prosopagnosics. As shown in Table 6, no significant differences in RT were found between control and prosopagnosic performance in any of the four greeble tasks. These results suggest that speed-accuracy trade-offs cannot explain the similar performance of controls and prosopagnosics.

Table 6. *Average accuracy (%) and reaction time (ms) for all naming and verification tasks for prosopagnosics and controls in the greeble condition. p-values for comparisons between prosopagnosics and control for each task. Standard deviations are presented in parentheses.*

	Accuracy (%)			Reaction Time (ms)		
	M (SD)		*p*	*M (SD)*		*p*
	Prosopagnosia	Control		Prosopagnosia	Control	
Naming	88.00	90.92	.407	1640	1546	.719
	(9.52)	(5.87)		(781)	(243)	
Overall	91.80	92.90	.687	2200	2147	.766
Verification	(6.56)	(5.40)		(413)	(364)	
Verification	92.50	92.00	.895	2068	2186	.408
Family	(8.91)	(7.77)		(237)	(370)	
Verification	91.30	93.50	.443	2342	2238	.650
Individual	(9.33)	(4.63)		(649)	(301)	

3.1.3 Expertise Criteria. To be considered an expert with greebles, according to criteria proposed by Gauthier et al. (1998), participants' RT for family and individual verification need to be the same during the testing phase. To assess whether participants met this expertise criteria, the average RT for each participant was taken for each testing phase session. RTs two standard deviations above the mean were excluded from analysis. The average for each session was then calculated for both the control and prosopagnosic groups. Independent samples *t*-tests were conducted comparing the RT for family verification and individual verification in each session separately for controls and prosopagnosics.

Figure 8 displays the mean RT across each testing phase session for control and prosopagnosic participants. Table 7 displays the mean, standard deviations, and *p*-values for the greeble verification family and individual tasks. Both my control and prosopagnosic groups met this criteria by session 5 (Control: $t(9) = 1.01$, $p = .339$, prosopagnosics: $t(9) = 1.71$, $p = .122$) and maintained no difference in reaction time for the remainder of the testing phase.

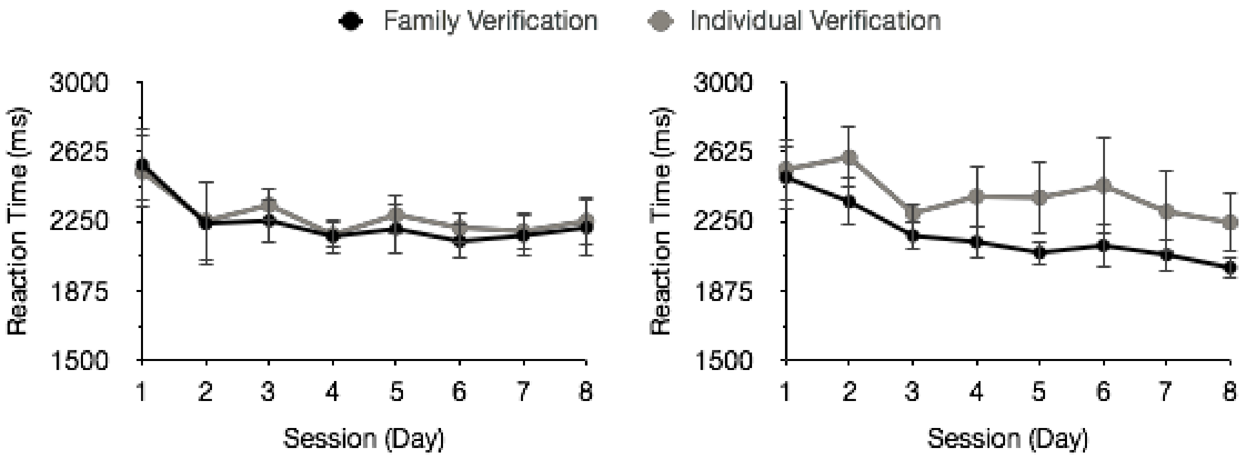


Figure 8: Average RT for greeble family and individual verification tasks for control (left) and prosopagnosic (right) groups. Error bars represent the standard error.

Recently, expertise researchers have proposed an additional expertise criterion—that the difference between reaction times for individual and family naming must be no more than 95 ms (Bukach et al, 2012). Applying this restriction to my analysis, the control group still achieves expertise by session 5 and maintain this for the rest of the testing period. However, the prosopagnosic group does not meet this restriction at any point of the testing phase (see Table 7). While this result suggests that prosopagnosics may not be entirely normal at learning greebles, it should be noted that prosopagnosics have comparable performance and reaction times for all tasks in the greeble paradigm (see Table 6) and fulfil the original criteria of expertise (Gauthier & Tarr, 1997). Therefore, prosopagnosics demonstrate control-like learning across many other criteria, suggesting normal or almost normal ability to acquire expertise with greebles, in contrast to their impaired learning of upright faces.

Table 7. *Comparisons of average reaction time (ms) between verification family (Fam) and verification individual (Ind) tasks for prosopagnosics and controls in the greeble condition. p-values for comparisons between family and individual verification for controls and prosopagnosics. Standard deviations are presented in parentheses*

	Session 5			Session 6			Session 7			Session 8		
	M (SD)		*p*	*M (SD)*		*p*	*M (SD)*		*p*	*M (SD)*		*p*
	Fam	Ind		Fam	Ind		Fam	Ind		Fam	Ind	
Control	2210 (429)	2286 (354)	.339	2143 (298)	2217 (275)	.137	2173 (345)	2197 (316)	.580	2216 (509)	2252 (401)	.419
DP	2081 (205)	2379 (641)	.112	2120 (388)	2444 (832)	.107	2070 (287)	2302 (703)	.149	2001 (199)	2244 (525)	.086

In sum, the prosopagnosic group demonstrated that they had learned greebles to the same extent as controls across naming, verification, family verification, and individual verification task. Additionally, there was no evidence of any speed-accuracy trade-offs in performance. The controls and prosopagnosics also met the original criteria for greeble expertise proposed by Gauthier and Tarr (1998). However, prosopagnosics did not meet the additional 95-ms criteria, whereas the control group met this criteria by session 5. One interpretation of this result is that prosopagnosics in this study never achieved expertise with greebles. However, one recent study found that even controls often fail to meet these revised criteria for expertise (Rezlescu et al., 2014). Because of this finding and other problems with the original criteria, I have reservations about the validity of these expertise criterion as an indicator of expertise (see Specificity of face recognition mechanisms in Discussion complete explanation). Overall, I believe that the similarities in performance and RT between control and prosopagnosic groups provide sufficient evidence to conclude that prosopagnosics learned greebles to the same extent as controls.

3.1.4. How well did prosopagnosics learn upright faces, and how did they do it? My second aim was to examine whether prosopagnosics demonstrate any ability to learn to recognise upright faces. Unfortunately, the format of the training sessions makes it difficult to assess evidence of learning across all eight sessions because during the learning sessions (sessions 1-4) participants received feedback after each trial, and the tasks were made progressively more difficult each session. However, I found that performance of prosopagnosics improved over the course of the testing sessions (sessions 5-8; $F(3,27) = 3.58$, $p = .027$). More specifically, prosopagnosics were better able to name upright faces at session 8 than at session 5 (*MDiff* = 6%, $p < .05$). Furthermore, by the end of the training programme prosopagnosics were able to name upright faces at a level well above chance (74%). Taken together, these results suggest that my training program helped prosopagnosics learn to recognise the upright faces in my study, although less accurately than controls.

I next investigated whether prosopagnosics were learning faces in a similar way to controls. Specifically, are prosopagnosics able to use residual holistic face processing mechanisms, or do they rely on alternate mechanisms, such as featural processing? If prosopagnosics are using normal holistic processing, I would expect to see similar sized inversion effects between the control and prosopagnosic groups. I conducted four 2(Group: prosopagnosics, control) x2 (Condition: upright faces, inverted faces) mixed-measures ANOVAs to assess accuracy and reaction time performance for each of the naming and verification tasks separately (Figure 9).

For the naming task, there was evidence of a similar inversion effect for prosopagnosics and controls. Consistent with the results discussed earlier, the ANOVA revealed a significant main effect of condition such that participants performed better in the upright faces condition than in the inverted ($F(1,18) = 9.34$, $p = .007$, $\eta_p^2 = .342$; see Table 8). Likewise, a main effect of group was found ($F(1,18) = 9.73$, $p = .006$, $\eta_p^2 = .351$) such that controls performed better than prosopagnosics (see Table 8). Crucially, there was no significant interaction was found between Group and Condition for accuracy ($F(1,18) = 0.603$, $p = .447$, $\eta_p^2 = .032$). Therefore, there is no evidence that the inversion effect is smaller for prosopagnosics than for controls (see Figure 9).

Similarly, for the verification tasks a significant main effect of condition was found such that participants performed better in the upright faces condition than in the inverted face condition ($F(1,18) = 10.49$, $p = .005$, $\eta_p^2 = .368$; see Table 8). However, no main effect of group was found (F $(1,18) = 2.41$, $p = .138$, $\eta_p^2 = .118$). Again, no significant interaction was found between Group and Condition for RT ($F(1,18) = 1.61$, $p = .221$, $\eta_p^2 = .082$), indicating that the inversion effect was not smaller for prosopagnosics compared to control.

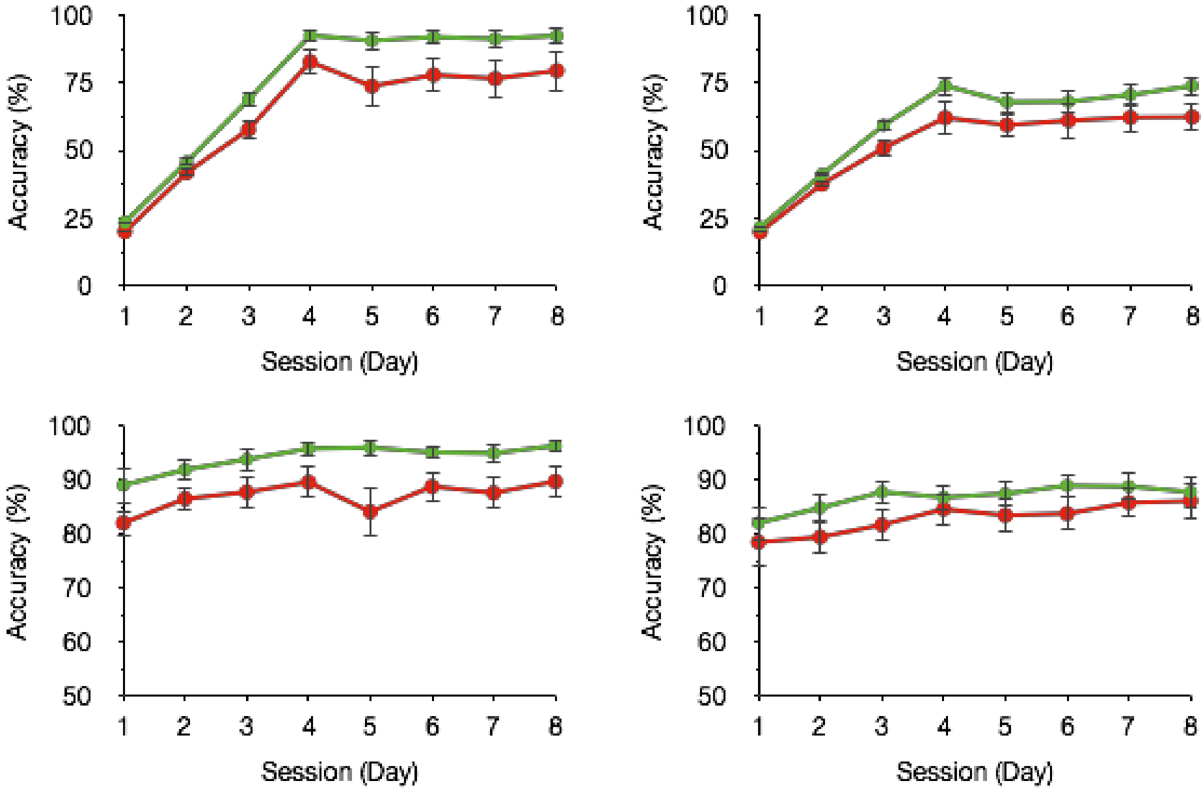


Figure 9. Top panel displays accuracy of controls (left) and prosopagnosics (right) in the naming tasks for upright (green) and inverted (red) face conditions. Naming scores have been scaled corresponding to the number of individuals learned in that session to reflect the varying difficulty in each session (five individuals in session 1, ten in session 2, fifteen in session 3, and twenty from session 4 onwards). Bottom panel displays accuracy of controls (left) and prosopagnosics (right) in the verification tasks for upright (green) and inverted (red) face conditions. Chance performance is at 50%.

Next, I analysed RT in both conditions for prosopagnosics and controls to check for speed-accuracy trade-offs (Figure 10). Specifically, I wanted to ensure that both groups had similar RTs across both upright and inverted testing phases to rule out the possibility that the similar inversion effects could be explained by speed-accuracy trade-offs. For the naming tasks, no significant main effect of Group (F(1,18) = 0.386, p = .542, η_p^2 = .021) was found. However, there was a significant main effect of Condition (F(1,18) = 5.734, p = .028, η_p^2 = .238), such that naming in the inverted face condition was slower than naming in the upright face condition overall. Importantly, no significant interaction was found between Group or Condition either (F(1,18) = 0.364, p = .554, η_p^2 = .015), suggesting that speed-accuracy trade-offs cannot account for the inversion effect results.

Likewise, for the verification tasks, there was no significant main effect of group (F(1,18) = 0.343, p = .565, η_p^2 = .417). However, again a significant main effect of condition (F(1,18) = 12.890, p = .002, η_p^2 = .417) was found, such that the verification of upright faces was faster than for inverted faces. Once again, no significant interaction was found between Group and Condition for RT (F(1,18) = 0.016, p = .902, η_p^2 = .001).

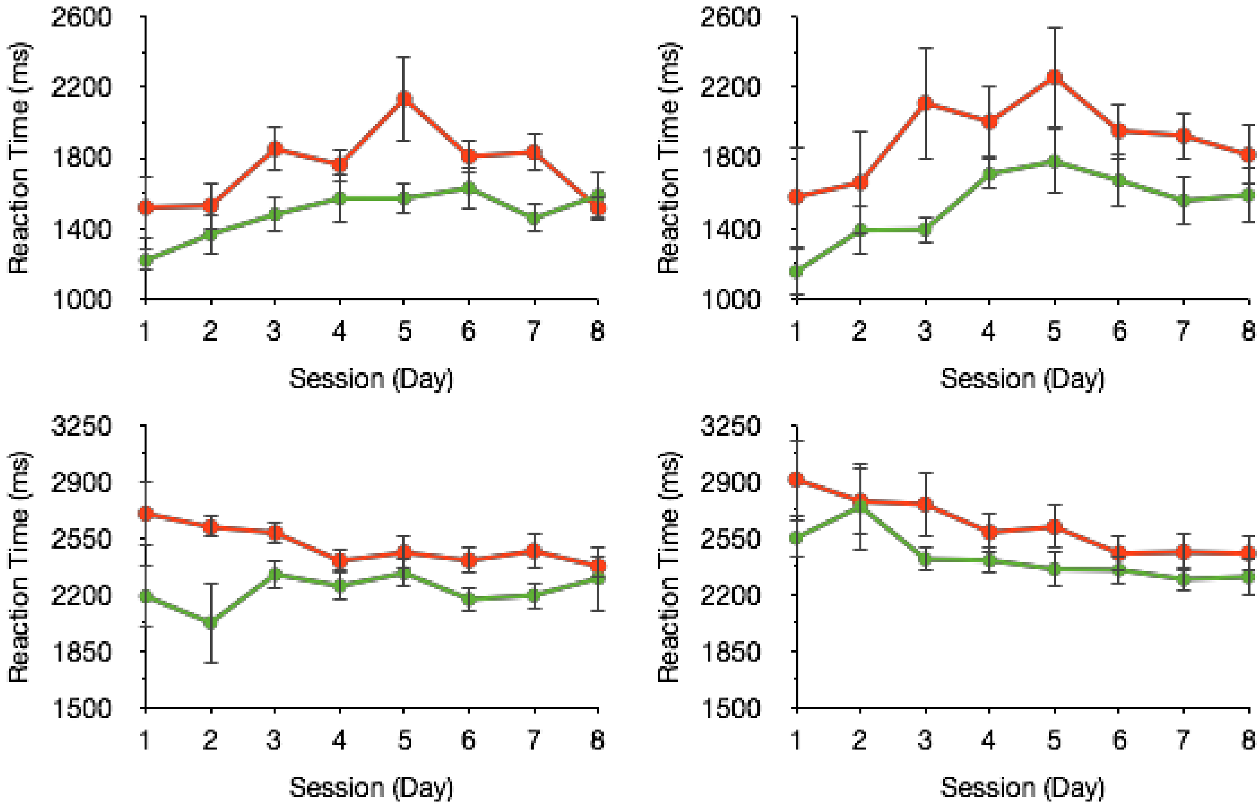


Figure 10.. Top panel displays average RT of controls (left) and prosopagnosics (right) in the naming tasks for upright (green) and inverted (red) face conditions in each session. Bottom panel displays average RT of controls (left) and prosopagnosics (right) in the verification tasks for upright (green) and inverted (red) face conditions in each session.

Table 8. *Average accuracy rate (%) and reaction time (ms) in naming and verification tasks for prosopagnosics (proso) and controls in the upright and inverted face conditions. Standard deviations are presented in parentheses*

	Upright Faces				Inverted Faces			
	Accuracy - %		Reaction Time - ms		Accuracy - %		Reaction Time - ms	
	Proso	Control	Proso	Control	Proso	Control	Proso	Control
Naming	70.20	91.90	1643	1617	61.40	77.10	1988	1823
	(12.37)	(9.39)	(489)	(331)	(16.95)	(22.11)	(550)	(300)
Verification	88.30	95.70	2330	2248	84.80	87.70	2499	2430
	(7.94)	(4.52)	(300)	(306)	(9.64)	(10.32)	(371)	(235)

In sum, participants performed better in the upright face condition than in the inverted face condition overall. However, although prosopagnosics performed worse than controls overall for naming of upright and inverted faces, they were not worse overall for verification of upright nor inverted faces (see Table 8). This discrepancy could be explained by a ceiling effect for verification—controls were accurate at upright face verification 96% of the time. The naming task is also more applicable to real life situations in that people are probably far more likely to see a face and need to recall the name than they are to be presented with a name and a face and need to verify that they match. Therefore, I propose that the naming task is a more useful indicator of whether prosopagnosics are impaired in their ability to recognise faces.

3.2. Individual-level Analysis.

My third aim is to examine individual differences in prosopagnosics' abilities to learn greebles and faces, I ran individual level analyses to compare each prosopagnosic's performance to the mean performance of controls using Crawford's revised standardised difference tests (RSDT) and single-score comparison tests (Crawford, Garthwaite, & Porter, 2010). The RSDT allows for a comparison between the difference in performance on two tasks in a single-case with the distribution of this difference in controls. The single-score comparison test tests for a deficit and allows for an assessment of whether a single-case performs significantly below the control distribution on a given task (see Appendix D for more information and formulas for these two tests).

3.2.1 Did individual prosopagnosics show a dissociation between upright faces and greebles? To examine whether individual prosopagnosics demonstrated a dissociation between greeble and upright face performance I conducted Crawford's RSDT analysing accuracy and RT in naming and verification conditions separately. In the naming task, 5/10 prosopagnosics demonstrated a significant dissociation between greeble and upright face performance, while only 1/10 prosopagnosics demonstrated a significant dissociation in the verification tasks (see Tables 9 and 10). These results show there is substantial heterogeneity in the pattern of deficits seen in individuals with developmental prosopagnosia.

Next, I analysed RT performance to see if each individual prosopagnosics' dissociation or lack thereof could be due to speed-accuracy trade-offs. Again, I conducted Crawford's RSDT to compare the difference between RT scores from the greeble and upright face condition for each prosopagnosic to the difference between average control RT scores in these two conditions. This analysis was conducted for naming and verification separately. In the naming task, only one prosopagnosic showed a pattern of responding consistent with a speed accuracy trade-off—DP6 responded much slower in the greeble condition (M = 3816 ms) compared to the upright face condition (M = 2935 ms). In the verification task, there were no significant difference between individual prosopagnosics' RT scores and that of controls (see Tables 9 and 11). Overall, these results suggest speed-accuracy trade-offs do not explain most of the dissociations (or lack thereof) found in the prosopagnosics in this study.

Table 9. *Summary of whether individual prosopagnosics demonstrated a dissociation between greeble and upright face performance for both naming and verification trials, and whether a speed-accuracy trade off can account for these scores. See Tables 10 and 11 for z-scores and statistics.*

Participant	Dissociation Naming	Speed-accuracy trade off present	Dissociation Verification	Speed-accuracy trade off present
DP1	No	No	No	No
DP2	No	No	No	No
DP3	Yes	No	No	No
DP4	Yes	No	No	No
DP5	No	No	No	No
DP6	No	Yes	No	No
DP7	No	No	No	No
DP8	Yes	No	No	No
DP9	Yes	No	Yes	No
DP10	Yes	No	No	No

Table 10. *Comparison of individual prosopagnosic (DP) accuracy and RT between the greeble and upright faces naming trials. Table shows respective z-scores to control mean, t-scores, p-values, and effect sizes (*Z_{DCC}*) from Crawford's RSDT. See Table 17 for descriptive statistics. Significant scores in bold (p < .05)*

	Accuracy				RT			
Participant	Greeble *Z*-Score	Upright Face *Z*-Score	*t* (*p*)	Z_{DCC}	Greeble *Z*-Score	Upright Face *Z*-Score	*t* (*p*)	Z_{DCC}
DP1	0.89	-.202	1.189 (.265)	1.382	**-1.967**	**-0.024**	**3.32 (.009)**	**-4.062**
DP2	-2.255	-2.465	0.227 (.794)	0.263	-0.502	0.665	2.05 (.071)	-2.439
DP3	**-0.028**	**-2.951**	**3.093 (.013)**	**3.701**	-0.403	-0.517	0.20 (.844)	0.237
DP4	**1.121**	**-1.045**	**2.304 (.047)**	**2.717**	-0.284	-0.079	0.37 (.722)	-0.429
DP5	-3.907	-3.305	0.657 (.528)	-0.762	-1.588	-1.465	0.22 (.831)	-0.258
DP6	-1.178	-1.798	0.677 (.502)	0.785	**9.342**	**3.982**	**7.75 (<.001)**	**11.205**
DP7	-1.107	-1.842	0.803 (.442)	0.931	-0.444	-0.779	0.60 (.565)	0.700
DP8	**1.048**	**-1.355**	**2.568 (.031)**	**3.042**	-0.342	-0.692	0.63 (.548)	0.732
DP9	**-0.31**	**-4.546**	**4.347 (.002)**	**5.362**	-0.156	-0.474	0.57 (.584)	0.665
DP10	**0.537**	**-3.571**	**4.23 (.002)**	**5.201**	0.214	0.384	0.30 (.769)	-0.355

Table 11: *Comparison of individual prosopagnosic (DP) accuracy and RT between the greeble and upright faces verification trials. Table shows respective z-scores to control mean, t-scores, p-values, and effect sizes (*Z_{DCC} *) from Crawford's RSDT. See Table 17 for descriptive statistics. Significant scores in bold (p < .05)*

	Accuracy				RT			
Participant	Greeble Z-Score	Upright Face Z-Score	*t* (*p*)	Z_{DCC}	Greeble Z-Score	Upright Face Z-Score	*t* (*p*)	Z_{DCC}
DP1	0.659	0.305	0.32 (.756)	0.367	-0.83	0.261	2.031 (.073)	-2.42
DP2	-2.698	-2.184	-0.534 (.653)	-0.534	0.522	-0.092	1.156 (.278)	1.361
DP3	0.622	1.631	2.011 (.075)	2.336	-0.654	-0.222	0.816 (.436)	-0.957
DP4	0.891	0.215	0.61 (.557)	0.701	0.201	0.317	0.221 (.830)	-0.258
DP5	-1.541	-3.013	1.324 (.218)	1.527	-0.236	-0.431	0.37 (.720)	0.433
DP6	0.467	-0.615	0.975 (.355)	1.122	3.132	2.935	0.374 (.717)	0.437
DP7	-0.267	-1.077	0.732 (.483)	0.841	-0.077	0.196	0.517 (.618)	-0.606
DP8	1.083	-1.077	1.931 (.086)	2.241	-0.604	-0.023	1.096 (.302)	-1.29
DP9	**-0.922**	**-5.593**	**4.027 (.003)**	**4.843**	-0.255	-0.497	0.457 (.659)	0.535
DP10	-0.152	-1.721	1.41 (.192)	5.201	0.269	0.225	0.083 (.936)	0.097

3.2.2. Did each prosopagnosic learn greebles as well as did controls? In order to investigate this question, I compared scores across all different greeble trials between each individual DP and the control mean using Crawford's single-case comparisons test. I found that for each of the greeble naming, overall verification, and family verification tasks, 9 out of the 10 prosopagnosics demonstrated similar performance to controls (see Tables 12 and 13). DP5 performed worse than controls on the greeble naming task, while DP2 performed worse than controls on the Greeble verification overall, and on Greeble family verification. For the greeble individual verification task, 8 out of 10 prosopagnosics demonstrated comparable performance to controls (see Tables 12 and 13)—DP's 5 and 9 performed worse than controls.

Next, I compared individual prosopagnosic and control RTs to assess whether any speed-accuracy trade-offs could account for prosopagnosics' individual scores on each of the greeble tasks. Specifically, if prosopagnosics were similarly accurate to controls, but were responding more slowly, those results could suggest that prosopagnosics were impaired or using a different strategy to controls when recognising greebles. Conversely, if prosopagnosics were less accurate than controls, I wanted to ensure that this was not because they were responding faster than controls. Compared to controls, only DP6 demonstrated a significant difference between their RT on the greeble naming (M =3816), overall verification (M = 3287), and individual verification (M = 4140) tasks. In DP6's case, their reaction time is consistent with a speed-accuracy trade off because their performance was comparable to controls on all greeble tasks, yet their RT was much slower than the average control RT (greeble naming: M = 1546, SD = 243; greeble overall verification: M = 2147, SD = 364; greeble individual verification: M = 2238, SD = 301).

Table 12. *Summary of whether individual prosopagnosics were comparable to control on all greeble tasks, and if speed-accuracy trade-off accounts for their scores.*

Participant	Greeble Naming	Speed Acc Trade off	Greeble Verification	Speed Acc Trade off	Greeble Family Verification	Speed Acc Trade off	Greeble Individual Verification	Speed Acc Trade off
DP1	Yes	No	Yes	No	Yes	No	Yes	No
DP2	Yes	No	No	No	No	No	Yes	No
DP3	Yes	No	Yes	No	Yes	No	Yes	No
DP4	Yes	No	Yes	No	Yes	No	Yes	No
DP5	No	No	Yes	No	Yes	No	No	No
DP6	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
DP7	Yes	No	Yes	No	Yes	No	Yes	No
DP8	Yes	No	Yes	No	Yes	No	Yes	No
DP9	Yes	No	Yes	No	Yes	No	No	No
DP10	Yes	No	Yes	No	Yes	No	Yes	No

Table 13. *Assessment of the difference of individual prosopagnosic (DP) scores to a control mean across all greeble tasks. The relevant single-score comparisons test t-scores, p-values (in parentheses), and effect sizes are presented below. See Table 17 for descriptive statistics. Significant differences in bold (p < .05).*

	Greeble Naming				Greeble Verification				Greeble Family Verification				Greeble Individual Verification			
	Accuracy		RT		Accuracy		RT		Accuracy		RT		Accuracy		RT	
Participant	*t* (*p*)	Z_{DCC}	*t* (*p*)	Z_{DCC}	*t* (*p*)	Z_{DCC}	*t* (*p*)	Z_{DCC}	*t* (*p*)	Z_{DCC}	*t* (*p*)	Z_{DCC}	*t* (*p*)	Z_{DCC}	*t* (*p*)	Z_{DCC}
DP1	0.853 (.416)	0.894	-1.876 (.093)	1.967	0.629 (.545)	0.659	-0.791 (.449)	0.830	0.42 (.685)	0.44	-1.178 (.269)	1.235	0.824 (.431)	0.864	-0.887 (.398)	0.930
DP2	-2.125 (.063)	2.228	-0.479 (.644)	0.502	**-2.573 (.030)**	**2.698**	0.498 (.631)	0.522	**-2.598 (.029)**	**2.725**	0.268 (.795)	0.281	-1.579 (.149)	1.657	0.418 (.686)	0.439
DP3	-0.028 (0.979)	0.029	-0.385 (.710)	0.403	0.593 (.568)	0.622	-0.623 (.548)	0.654	0.777 (.457)	0.815	-0.915 (.384)	0.959	0.138 (.893)	0.145	-0.767 (.463)	0.804
DP4	1.056 (.319)	1.107	-0.271 (.793)	0.284	0.849 (.418)	0.891	0.191 (.853)	0.201	0.777 (.457)	0.815	-0.237 (.818)	0.249	0.737 (.480)	0.773	0.371 (.720)	0.389
DP5	-3.682 (.003)	3.862	-1.515 (.164)	1.588	-1.469 (.176)	1.541	-0.225 (.827)	0.236	-0.297 (.733)	0.311	-0.526 (.612)	0.511	**-2.867 (.019)**	**3.006**	-0.288 (.780)	0.302
DP6	-1.109 (.296)	1.164	**8.907 (<.001)**	**9.342**	0.445 (.667)	0.467	**2.986 (.015)**	**3.132**	0.777 (.457)	0.815	0.814 (.436)	0.854	-0.206 (.841)	0.216	**6.025 (<.001)**	**6.319**
DP7	-1.043 (.324)	1.094	-0.424 (.682)	0.444	-0.254 (.805)	0.267	-0.073 (.943)	0.077	0.675 (.517)	0.708	-0.402 (.697)	0.422	-1.664 (.130)	1.745	-0.051 (.961)	0.053
DP8	0.988 (.349)	1.036	-0.326 (.752)	0.342	1.033 (.323)	1.083	-0.576 (.579)	0.604	0.726 (.486)	0.762	-0.812 (.438)	0.851	1.252 (.242)	1.313	-0.811 (.438)	0.850
DP9	-0.297 (.773)	0.312	-0.149 (.885)	0.156	-0.879 (.402)	0.922	-0.240 (.813)	0.255	0.266 (.796)	0.279	-0.309 (.764)	0.324	**-2.436 (.038)**	**2.550**	-0.634 (.542)	0.664
DP10	0.515 (.629)	0.54	0.204 (.843)	0.214	-0.145 (.888)	0.152	0.257 (.803)	0.269	-0.961 (.362)	1.008	0.253 (.806)	0.265	1.339 (.214)	1.404	-0.079 (.939)	0.083

3.2.3. How well did individual prosopagnosics learn upright faces, and how did they do it? To examine whether individual prosopagnosics demonstrated comparable inversion effects to controls, I conducted Crawford RSDTs testing for accuracy and RT differences between upright and inverted face conditions. These analyses were done for naming and verification tasks separately. In the naming task 8 out of 10 prosopagnosics demonstrated a comparable inversion effect to controls (see Table 14 and 15). In addition, two prosopagnosics (DPs 2 and 6) demonstrated an inverted inversion effect, in that their upright face performance was less accurate than their performance in the inverted face condition (see Table 17), that was comparable in size to control inversion effects. Consequently, this demonstrates that 6 out of 10 prosopagnosics show a typical inversion effect (upright better than inverted), while 2 out of 10 demonstrate an inverted inversion effect. Similarly, in the verification task, 9 out of 10 prosopagnosics demonstrated a comparable inversion effect to controls (see Table 14 and 16). Again, DPs 2 and 6 demonstrated an inverted inversion effect in the verification task.

Next, I compared individual prosopagnosic and control RTs to assess whether any speed-accuracy trade-offs could account for prosopagnosics' individual inversion effects. Compared to controls, only one prosopagnosic demonstrated a pattern of responding consistent with speed-accuracy trade-offs between the upright and inverted naming tasks (see DP9 in Table 17). There was no evidence of a speed-accuracy trade-off for any of the prosopagnosics in the verification tasks.

Table 14. *Summary of whether individual prosopagnosics' inversion effects were comparable to control inversion effects, and if speed-accuracy trade-off account for their scores.*

Participant	Comparable Inversion Effect Naming	Speed-accuracy trade off present	Comparable Inversion Effect Verification	Speed-accuracy trade off present
DP1	Yes	No	Yes	No
DP2	Yes – inverted	No	Yes – inverted	No
DP3	Yes	No	Yes	No
DP4	Yes	No	Yes	No
DP5	Yes	No	Yes	No
DP6	Yes - inverted	No	Yes - inverted	No
DP7	Yes	No	Yes	No
DP8	Yes	No	Yes	No
DP9	No	Yes	No	No
DP10	No	No	Yes	No

Table 15. *Comparison of individual prosopagnosic (DP) accuracy and RT between the upright and inverted face naming trials. Table shows respective z-scores to control mean, t-scores, p-values, and effect sizes* (Z_{DCC}) *of Crawford's RSDT. See Table 17 for descriptive statistics. Significant scores in bold (p < .05)*

	Accuracy				RT			
Participant	Upright Face Z-Score	Inverted Face Z-Score	*t* (*p*)	Z_{DCC}	Upright Face Z-Score	Inverted Face Z-Score	*t* (*p*)	Z_{DCC}
DP1	-0.202	0.118	0.326 (.752)	0.372	-0.024	1.736	2.252 (.051)	2.672
DP2	-2.465	0.037	2.065 (.069)	2.386	**0.665**	**-3.034**	**4.482 (.002)**	**5.615**
DP3	-2.951	-1	1.618 (.140)	1.861	**-0.517**	**-3.038**	**3.167 (.011)**	**3.828**
DP4	-1.045	-1.264	0.183 (.859)	0.021	-0.079	1.374	1.870 (.094)	2.206
DP5	-3.305	-1.924	1.15 (.280)	1.318	-1.465	-3.111	2.110 (.064)	2.498
DP6	-1.798	0.339	1.77 (.110)	2.039	**3.982**	**0.477**	**4.274 (.002)**	**5.321**
DP7	-1.842	-0.0642	1 (.343)	1.146	**-0.779**	**-4.034**	**4.000 (.003)**	**4.941**
DP8	-1.355	-0.623	0.611 (.556)	0.698	-0.692	-1.106	0.540 (.602)	0.629
DP9	**-4.546**	**-1.698**	**2.345 (.044)**	**2.718**	**-0.474**	**-4.885**	**5.218 (.001)**	**6.696**
DP10	**-3.571**	**-0.529**	**2.50 (.034)**	**2.904**	**0.384**	**-3.183**	**4.340 (.002)**	**5.414**

Table 16. *Comparison of individual prosopagnosic (DP) accuracy and RT between the upright and inverted face verification trials. Table shows respective z-scores to control mean, t-scores, p-values, and effect sizes* (Z_{DCC}) *of Crawford's RSDT. See Table 17 for descriptive statistics. Significant scores in bold (p < .05)*

	Accuracy				RT			
Participant	Upright Face Z-Score	Inverted Face Z-Score	*t* (*p*)	Z_{DCC}	Upright Face Z-Score	Inverted Face Z-Score	*t* (*p*)	Z_{DCC}
DP1	0.305	0.709	0.315 (.760)	0.329	0.261	3.04	1.722 (.119)	1.919
DP2	-2.184	0.628	2.173 (.058)	2.499	-0.092	1.877	1.222 (.253)	1.359
DP3	-1.631	-0.182	1.129 (.288)	1.288	-0.222	1.077	0.807 (.440)	0.897
DP4	0.215	-0.060	0.215 (.835)	0.244	0.317	3.237	1.809 (.104)	2.016
DP5	-3.013	-1.678	1.042 (.325)	1.187	-0.431	1.95	1.477 (.174)	1.644
DP6	-0.615	0.750	1.064 (.315)	1.213	2.935	4.597	1.032 (.329)	1.148
DP7	-1.077	-0.343	0.574 (.580)	0.653	0.196	1.547	0.839 (.423)	0.933
DP8	-1.077	-0.303	0.605 (.560)	0.688	-0.023	3.297	2.054 (.070)	2.292
DP9	**-5.593**	**-2.042**	**2.727 (.023)**	**3.156**	-0.497	0.743	0.771 (.461)	0.856
DP10	-1.721	-0.262	1.137 (.285)	1.297	0.225	1.167	0.585 (.573)	0.65

Table 17. *Accuracy and RT scores for control mean and individual prosopagnosics (DP) across all tested conditions and trials. Standard deviations presented in parentheses*

	Greeble Naming		Greeble Verification		Greeble Family Verification		Greeble Individual Verification		Upright Naming		Upright Verification		Inverted Naming		Inverted Verification	
	Acc	RT	Acc	RT	Acc	RT	Acc	RT	Acc	RT	Acc	RT	Acc	RT	Acc	RT
Control Mean	90.92 (5.87)	1546 (243)	92.90 (5.40)	2147 (364)	92.00 (7.77)	2186 (370)	93.50 (4.63)	2238 (301)	91.90 (9.39)	1643 (489)	95.70 (4.52)	2330 (300)	77.10 (22.11)	1988 (550)	87.70 (10.32)	2499 (371)
DP1	96.25	1068	96.46	1845.09	95.42	1729	97.50	1958	90	1609	97.08	2328	81.25	2838.15	95.00	2734.66
DP2	77.92	1424	78.33	2336.57	70.83	2290	85.83	2370	68.75	1837	85.83	2220	77.92	1717.38	94.17	2385.81
DP3	90.83	1448	96.25	1908.51	98.33	1831	94.17	1996	64.17	1446	88.33	2180	55.00	1715.93	85.83	2146.39
DP4	97.50	1477	97.71	2220.24	98.33	2094	97.08	2355	82.08	1591	96.67	2345	49.17	2753.31	87.08	2794.09
DP5	68.33	1160	84.58	2060.54	89.58	1982	79.58	2147	60.83	1132	82.08	2116	34.58	1699.35	70.42	2408.01
DP6	84.17	3816	95.42	3286.72	98.33	2502	92.50	4140	75	2935	92.92	3146	84.58	2542.04	95.42	3202.28
DP7	84.58	1438	91.46	2118.55	97.50	2030	85.42	2222	74.58	1359	90.83	2308	62.92	1482.46	84.17	2286.74
DP8	97.08	1463	98.75	1926.69	97.92	1871	99.58	1982	79.17	1388	90.83	2241	63.33	2170.09	84.58	2811.63
DP9	89.17	1508	87.92	2053.93	94.17	2066	81.67	2038	49.17	1460	70.42	2096	39.58	1281.80	66.67	2045.82
DP10	94.17	1598	92.08	2244.91	84.17	2284	100.00	2213	58.33	1744	87.92	2317	65.42	1682.14	85.00	2173.42

4. Discussion

The current study had three main aims. First, I sought to determine the extent to which prosopagnosics are able to learn faces and greebles. To address this question, I conducted training with upright faces and greebles in a group of prosopagnosics. The results showed a dissociation between upright face and greeble performance in prosopagnosia. Second, I assessed whether prosopagnosics were able to learn faces at all, and if so, whether they were able to learn by engaging normal face processes. To do so I analysed prosopagnosic performance in the upright face condition across the testing phase, finding that prosopagnosics did have some capacity to learn faces. Then, I compared the inversion effects of the prosopagnosic group to the inversion effects of the control group to assess whether there were any differences between the groups. The results showed that there was no difference in inversion effect between these two groups. Finally, I explored the heterogeneity of prosopagnosia by comparing individual prosopagnosic performance against control means. Specifically, I investigated whether each prosopagnosic demonstrated a dissociation between upright face and greeble performance, and whether each prosopagnosic was engaging normal face processing mechanisms when recognising faces.

Below I discuss the implications of my results for a broad range of issues concerning face recognition and prosopagnosia. These include the specificity of face recognition mechanisms, the development of face recognition mechanisms, limitations of the greeble paradigm, face learning and mechanisms of face learning in prosopagnosia, and finally, varieties of prosopagnosia.

4.1 Specificity of face recognition mechanisms

My results extend our understanding of the mechanisms behind face recognition in several ways. In this study, prosopagnosics completed the same training paradigm with upright faces and greebles, similar to a recent study run with acquired prosopagnosics (Rezlescu, et al., 2014). This study helps distinguish between the expertise and face-specific hypothesis because the two hypotheses make competing predictions about the results this study should produce. The expertise hypothesis suggests face and expertise processes are associated, so that disruptions to face processing will also disrupt the processing of objects-of-expertise. On the other hand, the face-specific hypothesis suggests that face and expertise process are dissociated, so that impairments restricted to face processing will have no effect on the processing of objects-of-expertise. The prosopagnosics in this study were impaired in their ability to learn to recognise faces, relative to controls, but were able to learn to recognise greebles normally. In other words, there was a dissociation between upright face and greeble performance in individuals with prosopagnosia. These results demonstrate that the impaired ability of prosopagnosics to recognise faces does not carry over to their ability acquire expertise with objects. Therefore, these results suggest that the mechanisms used for processing faces are separate to the mechanisms required for acquiring generic expertise.

One caveat to this conclusion is that although the prosopagnosics in this study did meet the original criterion for greeble expertise (Gauthier & Tarr, 1997), they did not meet one proposed criterion for greeble expertise—that the difference between family and individual verification needs to be less than 95 ms (Bukach et al., 2012). However, this 95ms criteria has little theoretical basis—it appears to be an arbitrary criterion based on the 95% confidence interval of RT in one sample of nine people (Bukach et al., 2012). Furthermore, having a concrete RT criterion is problematic because the difference in reaction times between individual and family verification depend not only on how expert a participant is, but on the relative difficulty of the individual and family verification tasks used in the study. For example, in a study where the greeble families are easily distinguishable, but greebles within a family are highly similar, even "experts" might not meet this criterion because of the difficulty they would have for individual verification. Other researchers have criticised this criterion in the past, noting that even control participants failed to meet this 95ms criterion, which suggests suggesting this criterion may not be a useful marker of expertise. Taking all measures into consideration, I still conclude that there is strong evidence of a dissociation between face and expertise performance in prosopagnosia.

4.2 Development of face recognition mechanisms

In addition to demonstrating that face recognition relies on face-specific mechanisms, my results also speak to how these mechanisms might develop in early life. Previous dissociations between recognition of faces and greebles in acquired prosopagnosics (e.g., Rezlescu et al., 2014) provide evidence for the specificity of face recognition mechanisms in the adult brain, but it is possible that these mechanisms originate from domain-general mechanisms with broader functionality that only later become specialised for faces (Johnson, 2011). However, the face-greeble dissociation I observed with developmental prosopagnosics in this study suggests not only that face recognition depends on face-specific mechanisms, but also that these mechanisms *develop independently* from other mechanisms required for obtaining expertise with non-face objects, and that they might be face-specific from birth.

The idea that face recognition depends on face-specific mechanisms that develop independently and mature early in life is consistent with data from preverbal infants and non-human animals. For example, new-born infants prefer face-like patterns (i.e., two “eyes” above one “mouth”) over a variety of control patterns (Goren, Sarty & Wu, 1975; Johnson, Dziurawiec, Ellis, & Morton, 1991; Farroni, et al., 2005), suggesting that infants have an innate preference for detecting and attending to faces at a time where neither expertise for faces nor objects could have been achieved. A recent study using laser projection and 4D ultrasound with pregnant mothers even reports that foetuses in the womb are more likely to track face-like stimuli over an inverted control, suggesting that the bias for face detection is present prenatally (Reid et al., 2017). Additionally, monkeys demonstrate a similar preference for face images over object images and are able to make subtle discrimination between pairs of faces, even after being deprived of face input (but not visual input) for the first two years of life (Sugita, 2007). Finally, several studies with congenital cataract patients show that early visual input is necessary for normal development of face recognition. Specifically, individuals deprived of visual input to the right hemisphere (due to left-eye cataracts) are less sensitive to configural information in faces (like spacing between the eyes; Le Grand, et al., 2003). These patients are also less able to integrate information across the whole face (Le Grand, Mondloch, Maurer, & Brent, et al., 2004). However, these patients are deprived not only of face input but of *all* visual inputs. Therefore, these studies are consistent with the idea that when the visual system receives normal input early on in life, face recognition mechanisms require very little if any exposure to face stimuli. In contrast, expert object recognition requires substantial experience to develop. In addition to my study, these results support the idea that not only do face-specific mechanisms underlie face processing, but that expertise and face mechanisms develop separately.

4.3 Limitations of the greeble training paradigm

My thesis also offers important insights into the greeble training paradigm itself. Several aspects of the greeble training paradigm have been previously questioned. Specifically, issues around within condition similarity, greebles’ potential similarity to faces, and concerns around the criteria used for expertise have been cited as potential counter explanations for this paradigm’s ability to induce expertise.

First, a key assumption of the greeble training paradigm is that training leads people to become experts with greebles. However, there is reason to think that the original criterion I, and others, have used to determine if participants have acquired expertise is not a valid indicator of expertise. Going by my criteria, both controls and prosopagnosics acquired "expertise" with greebles in the first session (after 120 trials). In the wider expertise literature, researchers typically use much more stringent criteria for who is considered an "expert". Typically, in studies involving real-world domains of expertise, researchers require their "experts" to have at least ten years' experience. For example, in Diamond and Carey's (1986) study, researchers required participants to have ten years' experience as a judge in a prominent American kennel club dogs show to be defined as a dog expert; and in Tanaka and Taylor's (1991) study, participants were bird experts if they had ten years' experience as avid bird watchers, were active in their clubs, and recommended by members of their club. Given this ten-year criterion, it is unlikely, that people could become experts with greebles after only one training session. If the participants are not "experts" with greebles after training, perhaps I am not dissociating between face-processing and expertise-processing, but between face-processing and normal object processing. If so, my data cannot distinguish between the face-specific and expertise hypotheses.

In future studies, then, it is vital that the expertise criteria are more theoretically grounded. The current criteria were based on Tanaka and Taylor's (1991) observation that experts are equally fast at pairing an object with its broad category (e.g. "car") and its more specific sub-category (e.g. "Toyota"). In the greeble paradigm, someone is considered an expert when they are as fast at verifying family names of greebles as they are at verifying the individual names. However, both family and individual names could arguably be considered specific sub-categories. If we were to more appropriately map on the criteria used by Tanaka and Taylor to the greeble paradigm, the broad category would be the label "greeble" rather than a family name. Perhaps, then, the criterion set out by Gauthier and Tarr (1997) is not sensitive enough to detect true expertise with greebles. Therefore, a criterion that more appropriately follows Tanaka and Taylor's (1991) observations of experts would be to consider people experts with greebles only when they are equally fast at pairing the label "greeble" with a greeble and pairing the family or individual name with the corresponding greeble. A valuable step forward would be to implement and assess this adapted criterion in future training paradigms.

Second, there is the assumption that greebles are novel objects. However, some researchers have suggested that greebles look somewhat like faces, which raises the possibility that face processing mechanisms are activated when processing greebles prior to expertise training (Kanwisher & Yovel, 2006). This similarity in turn would undermine their use as a novel, non-face object to test between face-specific and expertise hypotheses. For example, greebles produced an inversion effect in the fusiform face area similar to that observed with faces, even prior to training (Brants, Wagemans, & Op de Beeck, 2011)—a pattern that is not usually seen with novel objects (Haxby et al., 1999). However, this notion predicts that individuals with prosopagnosia should have difficulty with greebles because they are processed by the same face recognition mechanisms that are impaired in prosopagnosia. However, prosopagnosics in this study were able to recognise greebles but not faces, suggesting that—at least for these prosopagnosics—greebles do not look like faces. It is still possible, though, that these prosopagnosics perceive greebles to be face-like but can recognise greebles using object processing mechanisms—whereas faces need to be recognised using face-specific mechanisms. Consistent with this possibility, a patient with object agnosia, but normal face recognition was impaired at learning greebles (Gauthier, Behrmann, & Tarr, 2007). Taken together, the evidence suggests that greebles may *activate* face areas, but that only general object processing mechanisms are *crucial* for processing greebles.

Third, even if these assumptions are valid, there remains a possible counter-explanation for why prosopagnosics were impaired in their ability to learn faces, but not their ability to learn greebles—that prosopagnosics use alternative strategies to controls when learning faces and greebles, and that this strategy happens to be more effective for greebles than for faces. For example, perhaps prosopagnosics rely on more feature-based strategies than controls. If so, are there reasons to think that alternative strategies might be more effective for greebles than for faces? One possibility is that greeble stimuli are less similar to one another than the face stimuli used in this study. As a result, perhaps greebles—unlike faces—have distinctive enough features that people could identify each greeble by its features in isolation. In that case, prosopagnosics who are relying on feature-based strategies might be able to effectively recognise greebles but would still have difficulty recognising faces. If this counter-explanation were true, the dissociation observed between the ability to learn greebles and the ability to learn faces could be a product of the strategies prosopagnosics are using rather than separate mechanisms for recognising faces and objects-of-expertise. In turn, it's possible that prosopagnosics would be impaired in their ability to gain expertise with objects if those objects were as similar to one another as faces are.

One way to address this counter-explanation may be to make individual greebles more similar to one another, ideally as similar to one another as the individual faces are to one another. That is, one could match faces and greebles in terms of the *perceptual similarity* between exemplars within each stimulus category. However, because controls have prior experience with faces but not greebles, this approach would have resulted in controls performing better with faces than greebles, which would create the issue of differential task difficulty between conditions.

A further reason not to match on similarity is that matching for difficulty across greeble and upright face conditions provides a cleaner and straightforward means to interpreting results. Specifically, in the current study, greeble and upright performance is the same in the control group. Therefore, a dissociation between greeble and upright performance in prosopagnosics indicates that face and expertise processes rely on separate mechanisms (supporting the face-specific hypothesis), while similar performance in the greeble and upright face condition indicates that face and expertise processing relies on one system (supporting the expertise hypothesis). By contrast, in order to match on similarity, greebles would need to be substantially more similar to each other than those used in my study. A consequence of using more similar greebles would be that the difficulty of distinguishing those greebles would also increase—meaning that performance with greebles would be lower than with faces, even for controls. This baseline difference makes interpreting the results more complicated—it is no longer enough to look solely for a within-subjects dissociation between greeble and face learning. Instead, the size of the difference between greeble and face performance in the prosopagnosic group would need to be compared to the size of the difference in controls— which also reduces the power of the experiment. Furthermore, when matching on similarity, several patterns of performance would provide ambiguous evidence—if performance of prosopagnosics dropped for both greebles and faces, but less so for greebles than faces, that pattern could be interpreted as evidence for either hypothesis. Therefore, matching based on difficulty provides a more straightforward and clean way of testing whether face-specific or expertise mechanisms underlie face processing.

Rather than attempting to increase the similarity of greebles, a better way to match on both difficulty and similarity might be to reduce the perceptual similarity of the face stimuli, while at the same time using faces that people find more difficult distinguish. More specifically, future studies could use faces that participants have less experience with, namely “other-race” faces. Human adults are significantly worse at recognising other-race compared to own-race faces (Feingold, 1914; Malpass & Kravitz, 1969; Wan, Crookes, Reynolds, Irons & McKone, 2015) Therefore, using other-race faces instead of Caucasian faces could provide a useful way of matching face and greeble stimuli in terms of perceptual similarity *and* task difficulty. In addition, because people have less exposure to other race faces, both prosopagnosics and controls would be more akin to "novices" with those faces prior to training. Therefore, other race faces could provide a useful way of addressing the counter-explanation that people could be using alternative strategies to recognise.

4.4 Face learning in developmental prosopagnosia

A second aim of this study was to examine the extent to which prosopagnosics can learn to recognise upright faces. Although prosopagnosics showed poorer overall performance with upright faces than controls, they did demonstrate some learning effects. By the end of training (sessions 5-8), prosopagnosics performed well above chance for both naming trials (74%, chance = 5%) and verification trials (87%, chance = 50%), suggesting that they retain some ability to learn and recognise faces. Notably, these learning effects are much higher than those achieved by two acquired prosopagnosics who trained with the same paradigm: Herschel (naming: 29.6%, verification: 68.3%) and Florence (naming: 41.3%, verification: 58.3%; Rezlescu et al., 2014). These results are in line with the suggestion that developmental prosopagnosics may benefit more from rehabilitation and training strategies than acquired prosopagnosics (DeGutis, Chiu, Grosso, & Cohan, 2014).

Consistent with my findings, several studies have already shown that rehabilitation efforts with developmental prosopagnosics can have positive outcomes. For example, developmental prosopagnosic MZ completed a training program where she was required to allocate faces to one of two categories based on the relative positioning of facial features such as eyes and mouth (DeGutis, Bentin, Robertson, D'Esposito, 2007). MZ's performance on a wide variety of face recognition tasks increased after training. MZ also showed increased functional connectivity between face-selective regions in the brain, as well as a face selective N170 response that was not evident before training. This training was later scaled up to a group of 24 developmental prosopagnosics who also demonstrated moderate improvements in face recognition (DeGutis, Cohan, Nakayama, 2014). In another perceptual training study, several prosopagnosics demonstrated improved recognition in a task where they had to identify which face (out of two choices) most resembled the target face (Davies-Thompson, et al., 2015). Like MZ, these prosopagnosics increased performance on face recognition tasks after training, but they also demonstrated generalisation of face learning to untrained faces.

The above training studies are all similar in that they focus on *perceptual* aspects of face processing (i.e., matching or discriminating between face images or between different arrangements of features in a face). In contrast, the training paradigm I used in this study focuses on the *conceptual* aspects of face processing, where participants have to pair face images with names. The addition of name pairings arguably creates a scenario for face recognition that is more parallel to everyday learning of new faces. There are reasons to think that training programmes with a conceptual component (in the form of pairing faces with names) could be even more effective—evidence suggests that people are better able to recognise faces that had earlier been paired with names, compared to faces they learned without names (Schwartz & Yovel, 2016). Moreover, the kind of names matters: faces paired with real names, like Mike or John, are better recognised than faces paired with symbols (like %%%%% or *****) or object names (like Table). Consistent with the idea that adding a conceptual component to training could be beneficial to everyday face recognition, two prosopagnosic children, AL and KD, demonstrated impeccable face recognition ability with familiar faces months after completing training programmes that involved pairing faces with names (94% and 100% respectively). Their improvement during the training programmes also translated to their ability to assess unfamiliar faces as novel (Brunsdon, Coltheart, Nickels, & Joy, 2006; Schmalzl, Palermo, Green, Brunsdon, & Coltheart, 2008).

Future research should also endeavour to find out which aspects of training are important for improving the ability of prosopagnosics to recognise faces. For example, there are reasons to think that mere exposure to faces is not enough to help to recognise faces (Yovel, et al., 2012)—many prosopagnosics have difficulty recognising their own spouses, parents, or even themselves, despite having repeated exposure to those faces and having made efforts to remember them (Yardley, McDermott, Pisarski, Duchaine, & Nakayama, 2008). One possible explanation is that in daily life, prosopagnosics don't actually practice explicitly matching faces with the identities of those faces. Instead—perhaps because of the negative consequences of misidentifications—prosopagnosics rely on non-face cues that are more likely to result in successful identifications, such as clothing, hairstyles, gait, and voice (Cook & Biotti, 2016). By contrast, in controlled training paradigms, stimuli are designed so that participants are unable to use external cues and therefore must focus on the face itself to identify the individual (Brunsdon et al., 2006; DeGutis et al., 2007; Schmalzl et al., 2008). Furthermore, because these studies often use faces that are highly similar to one another, participants must focus not only on individual features but also on their relation to one another (configural processing) to accurately discriminate between individuals. Therefore, it could be that these training studies encourage participants to engage configural processing, which in turn might improve this important element of face processing (DeGutis, et al., 2007).

Regardless of the mechanism behind its effectiveness, the evidence that training appears to improve face-recognition in prosopagnosics is consistent with neuropsychological evidence about the neuroanatomy of prosopagnosia. More specifically, there is evidence that impairments seen in prosopagnosics could be related to compromised connections between areas important for face processing (Thomas et al., 2009). Of course, it is currently unknown whether deficits in these connections are a cause or effect of face recognition deficits (Behrmann, Avidan, Thomas, & Nishimura, 2011). However, issues with connectivity are implicated in other developmental disorders such as autism, attention-deficit hyperactivity disorder, and dyslexia (Fields, 2008), and evidence from these areas demonstrates that training can strengthen the connections between key areas (Mackey, Whitaker & Bunge, 2012; Meinzer et al., 2010). In prosopagnosia, training has been shown to ameliorate other neuronal deficits, such as a reduced N170 for faces (DeGutis, et al., 2007), which is perhaps a result of strengthened connections in the face processing system. Overall, this evidence suggests training studies could improve the connections between key face processing areas of the brain.

However, there is possible limitation to my conclusion that the training in this study improves face recognition— perhaps my study is not showing improvements in face recognition, but rather in image or photo recognition (Burton, 2013). My training paradigm uses the same images of individuals throughout learning and testing phases. Perhaps prosopagnosics learned only to recognise the specific images shown to them, rather than the faces behind those images. Future research should examine whether the effects of training programmes generalise to face recognition across different images. Additionally, some researchers have already begun to test general population face recognition capabilities after learning faces in different contexts, under different lighting and angles, and with different expressions, hairstyles, and clothing (Ritchie & Burton, 2017). These studies provide tentative evidence that high variable exposure produces better face recognition performance than low variable exposure. Training programs incorporating these variations—which individuals are likely to encounter in their everyday life—might prove even more effective in improving the face recognition abilities of people with prosopagnosia.

4.5. Mechanisms of face learning

Given that my results suggest prosopagnosics have some ability to engage in face processing, the next question to examine was whether prosopagnosics learned upright faces using typical face recognition mechanisms (i.e. those used by controls) but were impaired at using them, or whether they instead used alternative strategies. A classic hallmark of typical face recognition mechanisms is the face inversion effect, which is that people are substantially worse at recognising inverted faces than upright faces (Yin, 1969; Rossion, 2008). These inversion effects suggest that people process faces in a way that is orientation specific. If prosopagnosics process faces using these same orientation-specific mechanisms, we should see similar inversion effects in prosopagnosics to the inversion effects we see in controls, despite worse performance overall.

To investigate the extent to which prosopagnosics used typical face recognition mechanisms during training, I compared controls and prosopagnosics on the difference in between their performance on the upright and inverted face conditions. As expected, controls showed a sizeable inversion effect when learning upright and inverted faces. Crucially, prosopagnosics also showed an inversion effect, and there was no significant difference in the size of the inversion effects between the two groups. This finding suggests that prosopagnosics were processing faces using the same orientation-specific mechanisms as healthy controls. More broadly, these results suggest that these face mechanisms are not completely non-functional in prosopagnosia, only somewhat impaired. Furthermore, if prosopagnosics are using residual face processing mechanisms it could explain why patients with developmental prosopagnosia still show activation of important face areas in the brain (including the FFA and OFA; Avidan, Thomas & Behrmann, 2008).

One limitation of looking at inversion effects as an index of normal face processing is that the presence of inversion effects alone is not a robust indicator of typical face recognition mechanisms. A growing body of research demonstrates that inversion, part-whole, and composite face effects do not correlate with one another when looking within the normal variation in the general population (Rezlescu, Susilo, Wilmer, & Caramazza, 2017), suggesting that these effects are not measuring the same underlying face recognition mechanisms. In turn, these results suggest that multiple mechanisms are required for recognising faces successfully. Future studies should test whether prosopagnosics demonstrate similar, albeit impaired, patterns of responding to controls on other tests that tap into typical face recognition, such as the part-whole and composite face tasks

4.6. Varieties of developmental prosopagnosia

The final aim of this thesis was to take a preliminary look at the heterogeneity of prosopagnosia and identify potential prosopagnosia subtypes. To do this I ran individual-level analyses to capture different patterns of results among the ten prosopagnosics. My results showed that although there was a dissociation for five out of ten prosopagnosics between their performance on upright faces and their performance on greebles, the other five prosopagnosics performed similarly on those two conditions, suggesting potential subtypes of prosopagnosia. One subtype may be characterised by face-specific impairments, while the other may be characterised by both face-specific impairments and expertise-specific impairments that cannot be accounted by broader problems with object and visual recognition (because all prosopagnosics in this study performed normally on object recognition tests as part of the screening protocol).

Further evidence for the heterogeneity of prosopagnosia comes from variation in the inversion effects of prosopagnosics in this study. Six prosopagnosics demonstrated face inversion effects that were comparable to control inversion effects—that is that their inverted face performance was lower than their upright face performance. However, two prosopagnosics showed an “inverted” inversion effect, namely, better performance with inverted faces than with upright faces. These differences in the pattern of performance with upright and inverted faces suggest a few possible different types of developmental impairments that could occur. However, in order to fully interpret these results, we first need to understand how inverted faces are treated by the face processing system.

Initial studies with inverted faces suggest that they are mostly processed by object recognition mechanisms rather than face recognition mechanisms. In fact, in one study, prosopagnosics performed worse than controls on upright face tasks only, whereas inverted face performance remained mostly normal (Farah, Wilson, Drain, & Tanaka, 1995(b)). This result suggests that perhaps the mechanisms utilised for upright and inverted faces are separate, with only the upright mechanism being impaired in prosopagnosia. Additionally, patient CK, who suffers from object agnosia but has intact face processing, demonstrates comparable performance on upright face tasks, yet shows worse performance than controls when attempting to recognise inverted faces (Movscovitch, et al., 1997). More recent studies however suggest that inverted faces are likely processed by *both* object and face recognition mechanisms. A TMS study found that inverted face recognition decreased not only when stimulation was delivered to the object-selective area Lateral Occipital Cortex (LOC), but also when it was delivered to the face-selective area OFA (Pitcher, Duchaine, Walsh, Yovel,

& Kanwisher, 2011). Conversely, only TMS applied to the OFA reduced performance with upright faces. Additionally, inverted faces were shown to substantially activate the FFA, suggesting that inverted faces can engage ‘face-specific’ mechanisms (Kanwisher, Tong, & Nakayama 1998).

The aforementioned results leave us with a few possibilities for what kind of impairments could be occurring in prosopagnosia. One explanation for a reduction in both upright and inverted face performance is that there could be face recognition mechanisms that are not orientation-specific, such as those that analyse the local shape of individual face features like eyes (Pitcher, et al., 2011). Impairments to this mechanism would therefore cause reductions in performance for both upright and inverted faces. This trend is one that can be seen in six out of the ten prosopagnosics in my sample. An explanation for inverted inversion effect could be to do with damage to a differing part of the face processing system. While inverted faces have shown to be processed to some extent by certain face areas, this does not rule out the possibility that there are face areas that are not engaged when inverted faces are viewed. Therefore, it is possible that impairments to one of these areas causes impairments with upright faces, leaving inverted face performance untouched. A similar idea has been proposed by Farah and colleagues (1995b), where we have a mandatory mechanism that is activated for upright faces. In some cases of prosopagnosia, it is this mechanism that is impaired. However, crucially, this mechanism only engages for upright faces, leaving inverted faces to be processed by other mechanisms. This idea explains trends of where prosopagnosics perform similar to controls with inverted faces, but their upright face performance is reduced, which is the case with prosopagnosics showing inverted inversion effect in my sample.

4.7. Conclusions

In this thesis, I aimed to investigate the extent to which developmental prosopagnosia reflect impairments of a face-specific mechanism. I recruited individuals with developmental prosopagnosia, as well as matched controls, and investigated their ability to learn faces and novel objects. Overall, the results showed that prosopagnosics are impaired in their ability to learn faces but are able to gain expertise with novel objects. These results support the hypothesis that face processing relies on a face-specific mechanism, rather than a general expertise mechanism. Despite being impaired, the results also show that prosopagnosics have some ability to learn faces, which raises the possibility that training programs could help prosopagnosics learn to recognise at least some faces. I also found that most prosopagnosics show normal face-inversion effects, supporting the idea that prosopagnosics still use normal face-processing strategies/mechanisms to some extent. Finally, I found substantial heterogeneity within the prosopagnosics, suggesting further investigations into subtypes of prosopagnosia are warranted. These results have important implications for our understanding of developmental prosopagnosia as a disorder, as well as the mechanisms used in general face-processing.

References

Avidan, G., Thomas, C., & Behrmann, M. (2008). An integrative approach towards understanding the psychological and neural basis of congenital prosopagnosia in *Cortical mechanisms of vision* (ed.) Jenkin, M. & Harris, L. R (2008). 241-270.

Behrmann, M., Avidan, G., Thomas, C., & Nishimura, M. (2011). Impairments in face perception.

Behrmann, M., Marotta, J., Gauthier, I., Tarr, M. J., & McKeeff, T. J. (2005). Behavioral change and its neural correlates in visual agnosia after expertise training. *Journal of cognitive neuroscience*, *17*(4), 554-568.

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.

Bodamer, J. Die Prosop-Agnosie. Arch. Psychiat. Neroenkrankheit. 179,653, 1947.

Brants, M., Wagemans, J., & de Beeck, H. P. O. (2011). Activation of fusiform face area by Greebles is related to face similarity but not expertise. *Journal of Cognitive Neuroscience*, *23*(12), 3949-3958.

Brunsdon, R., Coltheart, M., Nickels, L., & Joy, P. (2006). Developmental prosopagnosia: A case analysis and treatment study. *Cognitive Neuropsychology*, *23*(6), 822-840.

Bukach, C. M., Gauthier, I., Tarr, M. J., Kadlec, H., Barth, S., Ryan, E., Turpin, J., & Bub, D. N. (2012). Does acquisition of Greeble expertise in prosopagnosia rule out a domain-general deficit? *Neuropsychologia*, *50*(2), 289-304.

Burton, M. (2013). Why has research in face recognition progressed so slowly? The importance of variability. *Quarterly Journal of Experimental Psychology*, *66*(8), 1467-1485.

Busey, T. A., & Vanderkolk, J. R. (2005). Behavioral and electrophysiological evidence for configural processing in fingerprint experts. *Vision research*, *45*(4), 431-448.

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science*, *195*(4275), 312-314.

Cassia, V. M., Picozzi, M., Kuefner, D., Bricolo, E., & Turati, C. (2009). Holistic processing for faces and cars in preschool-aged children and adults: Evidence from the composite effect. *Developmental science*, *12*(2), 236-248.

Cook, R., & Biotti, F. (2016). Developmental prosopagnosia. *Current Biology*, *26*(8), R312-R313.

Crawford, J. R., Garthwaite, P. H., & Porter, S. (2010). Point and interval estimates of effect sizes for the case-controls design in neuropsychology: rationale, methods, implementations, and proposed reporting standards. *Cognitive neuropsychology*, *27*(3), 245-260.

Crouzet, S. M., Kircher, H., & Thorpe, S. J. (2010). Fast Saccades Towards Faces: Face Detetion in just 100ms. Journal of Vision, 10(4), 16-16.

Curby, K. M., Glazek, K., & Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 94.

Davies-Thompson, J., Fletcher, K., Hills, C., Corrow, S., Pancaroglu, R., & Barton, J. J. (2015, August). Perceptual training of faces in rehabilitation of acquired prosopagnosia. In *38th European Conference on Visual Perception*.

Damasio, A. R., Damasio, H. and Van Hoesen, G. W. (1982) Prosopagnosia: Anatomic basis and behavioral mechanisms. Neurology 32, 331-341.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: an effect of expertise. *Journal of Experimental Psychology: General*, *115*(2), 107.

de Beeck, H. P. O., Baker, C. I., DiCarlo, J. J., & Kanwisher, N. G. (2006). Discrimination training alters object representations in human extrastriate cortex. *Journal of Neuroscience*, *26*(50), 13025-13036.

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.

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.

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

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.

Duchaine, B. C., Dingle, K., Butterworth, E., & Nakayama, K. (2004). Normal greeble learning in a severe case of developmental prosopagnosia. *Neuron*, *43*(4), 469-473.

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.

Duchaine, B., & Yovel, G. (2015). A revised neural framework for face processing. *Annual Review of Vision Science, 1*, 393-416.

Farah, M. J. (1991). Cognitive neuropsychology: Patterns of co-occurrence among the associative agnosias: Implications for visual object representation. *Cognitive Neuropsychology*, *8*(1), 1-19.

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995a). What causes the face inversion effect? *Journal of Experimental Psychology: Human perception and performance*, *21*(3), 628.

Farah, M. J., Wilson, K. D., Drain, H. M., & Tanaka, J. R. (1995b). The inverted face inversion effect in prosopagnosia: Evidence for mandatory, face-specific perceptual mechanisms. *Vision research*, *35*(14), 2089-2093.

Farroni, T., Johnson, M. H., Menon, E., Zulian, L., Faraguna, D., & Csibra, G. (2005). Newborns' preference for face-relevant stimuli: Effects of contrast polarity. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(47), 17245-17250.

Fields, R. D. (2008). White matter in learning, cognition and psychiatric disorders. *Trends in neurosciences*, *31*(7), 361-370.

Feingold, G. A. (1914). Influence of environment on identification of persons and things. *J. Am. Inst. Crim. L. & Criminology*, *5*, 39.

Gauthier, I. (2014, March). A commentary on Rezlescu, Barton, Pitcher, and Duchaine (2014) Normal acquisition of expertise with greebles in two cases of acquire prosopagnosia PNAS [blogpost]. Retrieved from http://gauthier.psy.vanderbilt.edu/a-commentary-on-rezlescu-barton-pitcher-duchaine-2014-normal-acquisition-of-expertise-with-greebles-in-two-cases-of-acquired-prosopagnosia-pnas/

Gauthier, I., Behrmann, M., & Tarr, M. J. (2004). Are Greebles like faces? Using the neuropsychological exception to test the rule. *Neuropsychologia*, *42*(14), 1961-1970.

Gauthier, I., & Bukach, C. (2007). Should we reject the expertise hypothesis? *Cognition*, *103*(2), 322-330.

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature neuroscience*, *6*(4), 428.

Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: Exploring mechanisms for face recognition. *Vision research*, *37*(12), 1673-1682.

Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(2), 431.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nature neuroscience*, *2*(6), 568-573.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2002). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature neuroscience, 3*(2), 191-197.

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training 'greeble' experts: a framework for studying expert object recognition processes. *Vision research, 38*(15), 2401-2428.

Goldstein, A. G. (1975). Recognition of inverted photographs of faces by children and adults. *The Journal of genetic psychology*, *127*(1), 109-123.Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature neuroscience*, *7*(5), 555-562.

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.

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature neuroscience*, *7*(5), 555.

Harley, E. M., Pope, W. B., Villablanca, J. P., Mumford, J., Suh, R., Mazziotta, J. C., Enzmann, D., & Engel, S. A. (2009). Engagement of fusiform cortex and disengagement of lateral occipital cortex in the acquisition of radiological expertise. *Cerebral Cortex*, *19*(11), 2746-2754.

Haxby, J. V., Ungerleider, L. G., Clark, V. P., Schouten, J. L., Hoffman, E. A., & Martin, A. (1999). The effect of face inversion on activity in human neural systems for face and object perception. *Neuron*, *22*(1), 189-199.

Hershler, O., & Hochstein, S. (2005). At first sight: A high-level pop out effect for faces. *Vision research*, *45*(13), 1707-1724.

jamovi project. (2017). jamovi (version 0.8) [computer software]. Retrieved from https://www.jamovi.org

Johnson, M. H. (2011). Interactive specialization: a domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience*, *1*(1), 7-21.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1-2), 1-19.

Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience*, *3*(8), 759.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*(11), 4302-4311.

Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition*, *68*(1), B1-B11.

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1476), 2109-2128.

Klein, S. B., Gabriel, R. H., Gangi, C. E., & Robertson, T. E. (2008). Reflections on the self: a case study of a prosopagnosic patient. *Social Cognition*, *26*(6), 766.

Kumfor, F., Hutchings, R., Irish, M., Hodges, J. R., Rhodes, G., Palermo, R., & Piguet, O. (2015). Do I know you? Examining face and object memory in frontotemporal dementia. *Neuropsychologia*, *71*, 101-111.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Neuroperception: Early visual experience and face processing. *Nature*, *410*(6831), 890-890.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2003). Expert face processing requires visual input to the right hemisphere during infancy. *Nature Neuroscience*, *6*(10), 1108-1112.

Le Grand, R. L., Mondloch, C. J., Maurer, D., & Brent, H. P. (2004). Impairment in holistic face processing following early visual deprivation. *Psychological Science*, *15*(11), 762-768.

Liu, J., Li, J., Feng, L., Li, L., Tian, J., & Lee, K. (2014). Seeing Jesus in toast: neural and behavioral correlates of face pareidolia. *Cortex*, *53*, 60-77.

Mackey, A. P., Whitaker, K. J., & Bunge, S. A. (2012). Experience-dependent plasticity in white matter microstructure: reasoning training alters structural connectivity. *Frontiers in neuroanatomy*, *6*, 32.

Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology*, *13*(4), 330.

McConachie, H. R. (1976). Developmental prosopagnosia. A single case report. *Cortex*, *12*(1), 76-82.

McKone, E. (2008). Configural processing and face viewpoint. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(2), 310.

McKone, E., & Robbins, R. (2011). Are faces special. *The Oxford Handbook of Face Perception*, 149-176.

McNeil JE and Warrington EK (1993) Prosopagnosia: A face-specific disorder. Quarterly Journal of Experimental Psychology A, Human Experimental Psychology 46A: 1–10.

Meinzer, M., Mohammadi, S., Kugel, H., Schiffbauer, H., Flöel, A., Albers, J., Kramer, K., Ricarda, M., Baumgartner, A., Knecht, S., Breitenstein, C., & Deppe, M. (2010). Integrity of the hippocampus and surrounding white matter is correlated with language training success in aphasia. *Neuroimage*, *53*(1), 283-290.

Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: a two-process theory of infant face recognition. *Psychological review*, *98*(2), 164.

Palermo, R., Rossion, B., Rhodes, G., Laguesse, R., Tez, T., Hall, B., ... & Al-Janabi, S. (2016). Do people have insight into their face recognition abilities?. *The Quarterly Journal of Experimental Psychology*, *70*(2), 218-233.

Pellicano, E., Rhodes, G., & Peters, M. (2006). Are preschoolers sensitive to configural information in faces?. *Developmental science*, *9*(3), 270-277.

Pitcher, D., Duchaine, B., Walsh, V., Yovel, G., & Kanwisher, N. (2011). The role of lateral occipital face and object areas in the face inversion effect. *Neuropsychologia*, *49*(12), 3448-3453.

Reid, V. M., Dunn, K., Young, R. J., Amu, J., Donovan, T., & Reissland, N. (2017). The human fetus preferentially engages with face-like visual stimuli. *Current Biology*, *27*(12), 1825-1828.

Rezlescu, C., Barton, J. J., Pitcher, D., & Duchaine, B. (2014). Normal acquisition of expertise with greebles in two cases of acquired prosopagnosia. *Proceedings of the National Academy of Sciences*, *111*(14), 5123-5128.

Rezlescu, C., Pitcher, D., & Duchaine, B. (2012). Acquired prosopagnosia with spared within-class object recognition but impaired recognition of degraded basic-level objects. *Cognitive Neuropsychology*, *29*(4), 325-347.

Rezlescu, C., Susilo, T., Wilmer, J. B., & Caramazza, A. (2017). The inversion, part-whole, and composite effects reflect distinct perceptual mechanisms with varied relationships to face recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(12), 1961.

Rhodes, G., Byatt, G., Michie, P. T., & Puce, A. (2004). Is the fusiform face area specialized for faces, individuation, or expert individuation?. *Journal of Cognitive Neuroscience*, *16*(2), 189-203.

Richler, J. J., Wilmer, J. B., & Gauthier, I. (2017). General object recognition is specific: Evidence from novel and familiar objects. *Cognition*, *166*, 42-55.

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

Robbins, R., & McKone, E. (2003). Can holistic processing be learned for inverted faces?. *Cognition*, *88*(1), 79-107.

Robbins, R., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition*, *103*(1), 34-79.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive psychology*, *8*(3), 382-439.

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta psychologica*, *128*(2), 274-289.

Rossion, B., Gauthier, I., Goffaux, V., Tarr, M. J., & Crommelinck, M. (2002). Expertise training with novel objects leads to left-lateralized facelike electrophysiological responses. *Psychological Science*, *13*(3), 250-257.

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.

Rossion, B., & Jacques, C. (2011). The N170: understanding the time-course of face perception in the human brain. *The Oxford handbook of ERP components*, 115-142.

Rossion, B., Kung, C. C., & Tarr, M. J. (2004). Visual expertise with nonface objects leads to competition with the early perceptual processing of faces in the human occipitotemporal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(40), 14521-14526.

Scapinello, K. F., & Yarmey, A. D. (1970). The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli. *Psychonomic Science*, *21*(6), 329-330.

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.

Schwaninger, A., Carbon, C. C., & Leder, H. (2003). Expert face processing: Specialization and constraints. *Development of Face Processing*, 81-97.

Schwartz, L., & Yovel, G. (2016). The roles of perceptual and conceptual information in face recognition. *Journal of Experimental Psychology: General*, *145*(11), 1493.

Shah, P., Gaule, A., Sowden, S., Bird, G., & Cook, R. (2015). The 20-item prosopagnosia index (PI20): A self-report instrument for identifying developmental prosopagnosia. *Royal Society Open Science*, 6: 140343.

Sugita, Y. (2008). Face perception in monkeys reared with no exposure to faces. *Proceedings of the National Academy of Sciences*, *105*(1), 394-398.

Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychological science*, *12*(1), 43-47.

Tanaka, J. W., Curran, T., & Sheinberg, D. L. (2005). The training and transfer of real-world perceptual expertise. *Psychological Science*, *16*(2), 145-151.

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

Tanaka, J. W., & Gauthier, I. (1997). Expertise in object and face recognition. *Psychology of Learning and Motivation*, *36*, 83-125.

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition*, *25*(5), 583-592.

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder?. *Cognitive Psychology*, *23*(3), 457-482.

Tarr, M. J., & Gauthier, I. (2000). FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature neuroscience*, *3*(8), 764.

Thomas, C., Avidan, G., Humphreys, K., Jung, K. J., Gao, F., & Behrmann, M. (2009). Reduced structural connectivity in ventral visual cortex in congenital prosopagnosia. *Nature neuroscience*, *12*(1), 29.

Thompson, P. (1980). Margaret Thatcher: a new illusion. *Perception*.

Torfs, K., Vancleef, K., Lafosse, C., Wagemans, J., & de-Wit, L. (2014). The Leuven Perceptual Organization Screening Test (L-POST), an online test to assess mid-level visual perception. *Behavior Research Methods*, *46*(2), 472-487.

Tsao, D. Y., Freiwald, W. A., Knutsen, T. A., Mandeville, J. B., & Tootell, R. B. (2003). Faces and objects in macaque cerebral cortex. *Nature neuroscience*, *6*(9), 989-995.

Tsao, D. Y., Freiwald, W. A., Tootell, R. B., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science*, *311*(5761), 670-674.

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British journal of psychology*, *79*(4), 471-491.

Wan, L., Crookes, K., Reynolds, K. J., Irons, J. L., & McKone, E. (2015). A cultural setting where the other-race effect on face recognition has no social–motivational component and derives entirely from lifetime perceptual experience. *Cognition*, *144*, 91-115.

Wong, A. C. N., Palmeri, T. J., & Gauthier, I. (2009). Conditions for facelike expertise with objects: Becoming a Ziggerin expert—but which type?. *Psychological Science*, *20*(9), 1108-1117.

Xu, Y. (2005). Revisiting the role of the fusiform face area in visual expertise. *Cerebral Cortex*, *15*(8), 1234-1242.

Xu, Y., Liu, J., & Kanwisher, N. (2005). The M170 is selective for faces, not for expertise. *Neuropsychologia*, *43*(4), 588-597.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of experimental psychology*, *81*(1), 141.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, *42*(11), 1166-1178.

Yovel, G., Halsband, K., Pelleg, M., Farkash, N., Gal, B., & Goshen-Gottstein, Y. (2012). Can massive but passive exposure to faces contribute to face recognition abilities?. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(2), 285.

Yardley, L., McDermott, L., Pisarski, S., Duchaine, B., & Nakayama, K. (2008). Psychosocial consequences of developmental prosopagnosia: A problem of recognition. *Journal of psychosomatic research*, *65*(5), 445-451.

Yovel, G., & Kanwisher, N. (2004). Face perception: domain specific, not process specific. *Neuron*, *44*(5), 889-898.

Appendix

Appendix A – Similarity Correlation Matrix

For this thesis a program was created to calculate the similarity between two, same sized, greyscale images. This program was created in Java and is available on request. Similarity is assessed based on pixel colour value. In both images each pixel is assigned a number from 0 to 255, corresponding to the lightness of that pixel on a spectrum from black to white, respectively. Each individual pixel value in one image is then listed from the top right pixel to bottom left. The same is then calculated for the second image so that there is one list of pixel values for each image. A correlation is then run between these two lists to assess the similarity between the two images. For the face sets in the current study, correlations ranged from .99 to .90.

To assess how well my face set (Face Set 2) matched the original face set (Face Set 1; Rezlescu et al., 2014), I ran a Pearson's Chi Square test comparing the distribution of correlations between Face Set 1 and Face Set 2. No significant difference was found between the distributions of correlations between images in Face Set 1 and and the distribution of correlations between images in Face Set 2 (X^2 (10, N = 6320) = 9.854, p = .453; see Figure 11). Additionally, no significant difference was found between the distribution of correlations between these two faces sets when looking only at the target images (X^2 (5, 380) = 7.161, p = .209; see Figure 12)

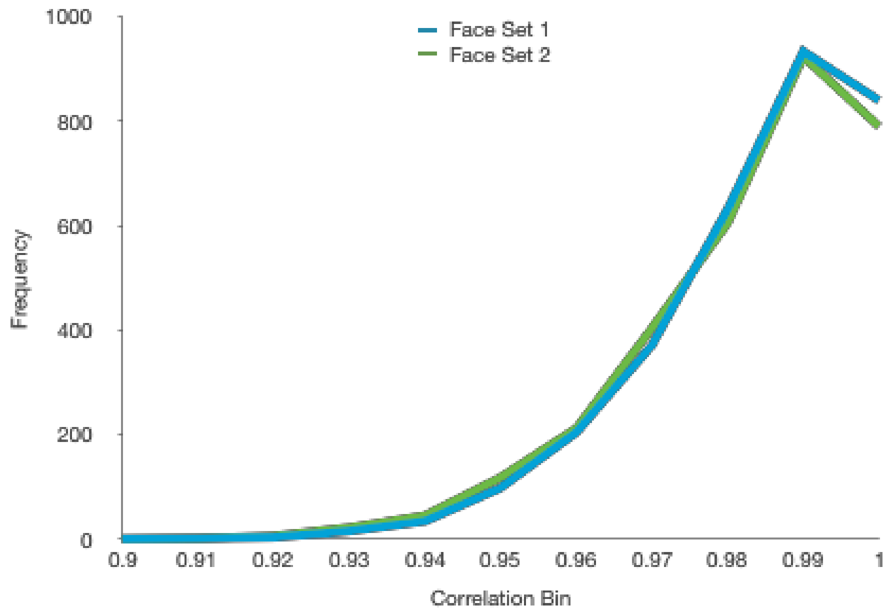


Figure 11. The distribution of all images in Face Set 1 and 2 based on the frequency of correlations in the respective correlation bins.

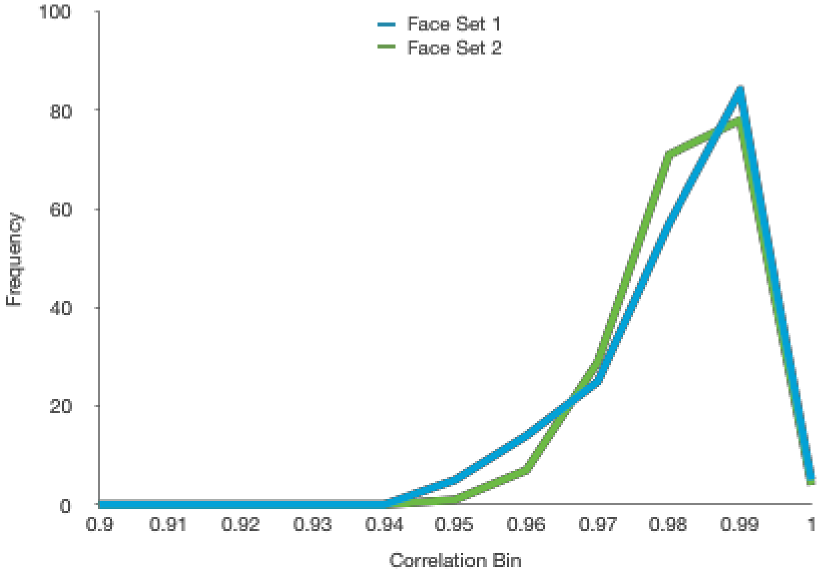
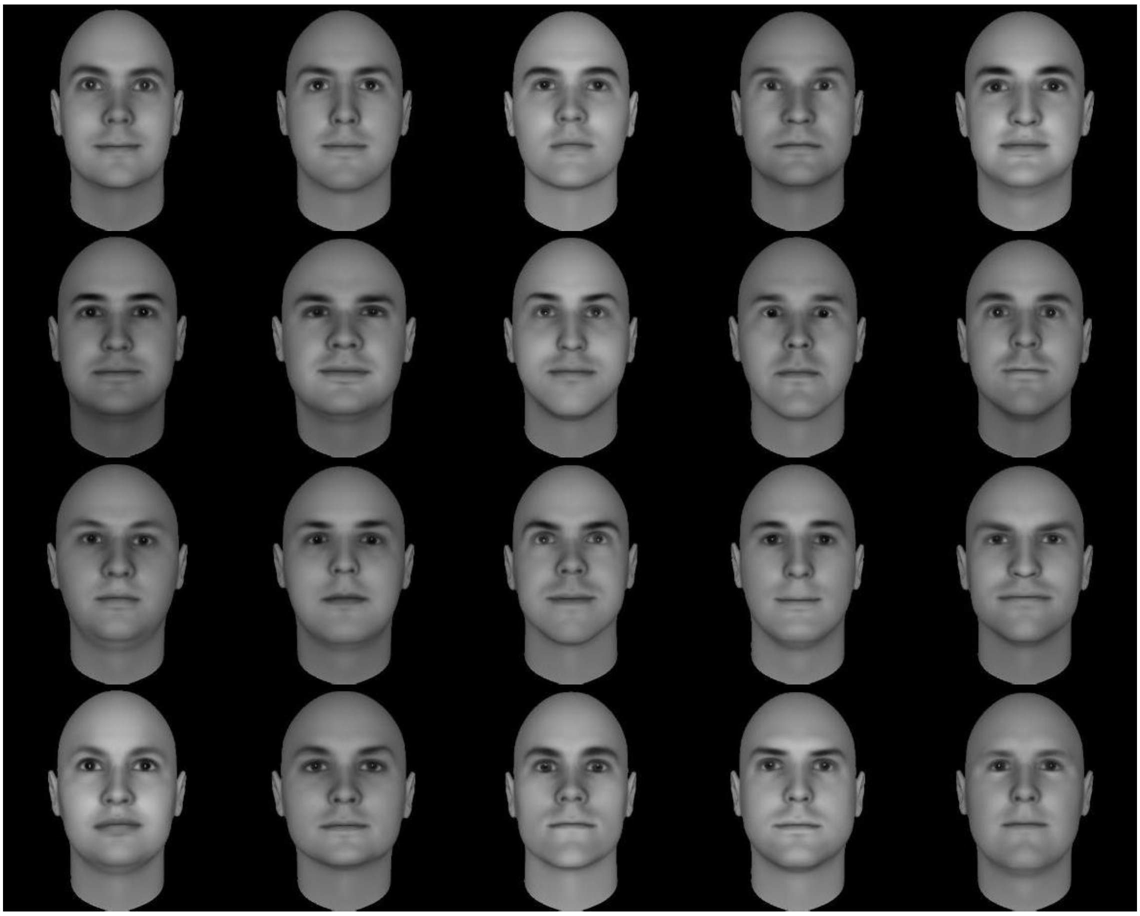


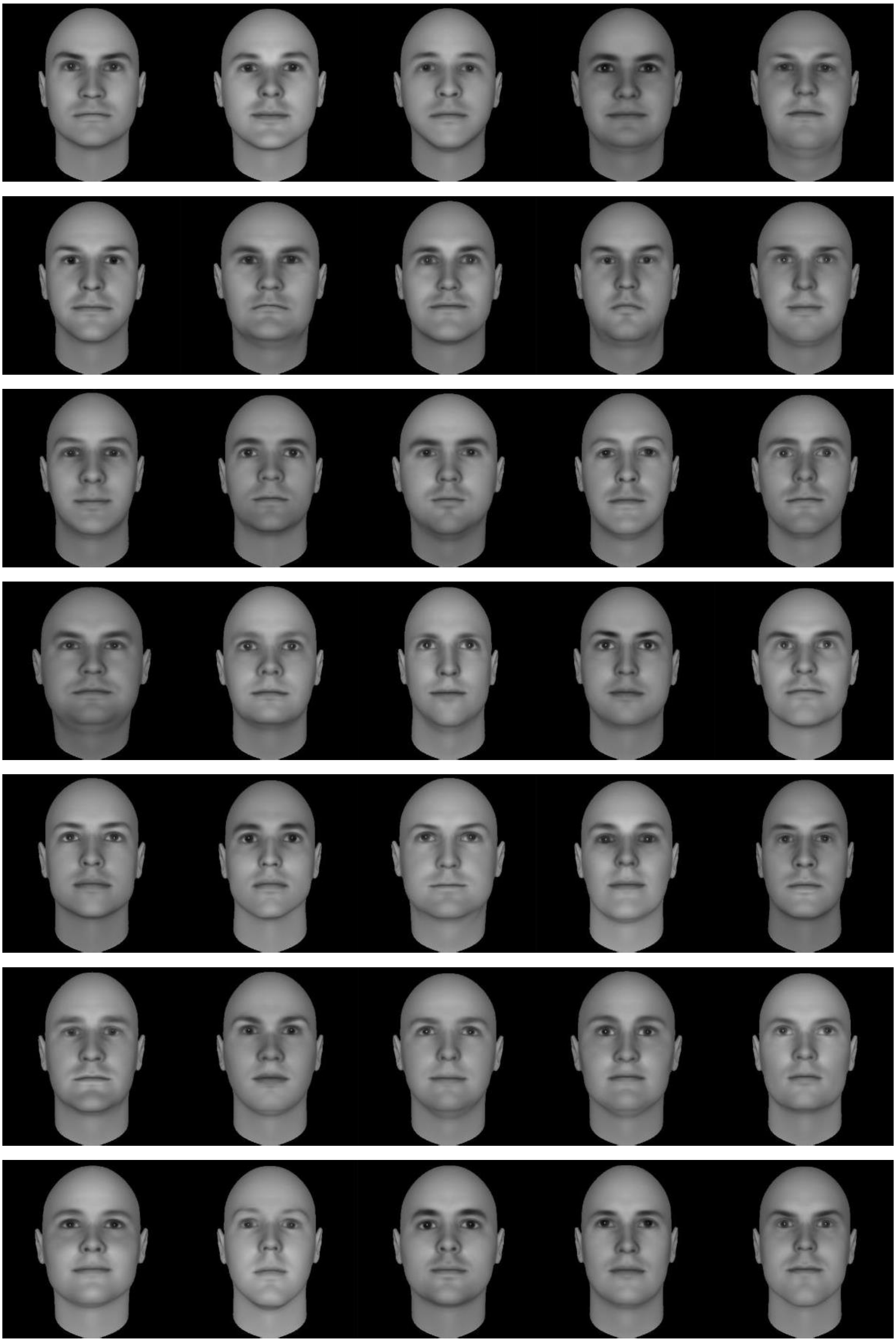
Figure 12. The distribution of target images in Face Set 1 and 2 based on the frequency of correlations in the respective correlation bins.

Appendix B - Image Sets

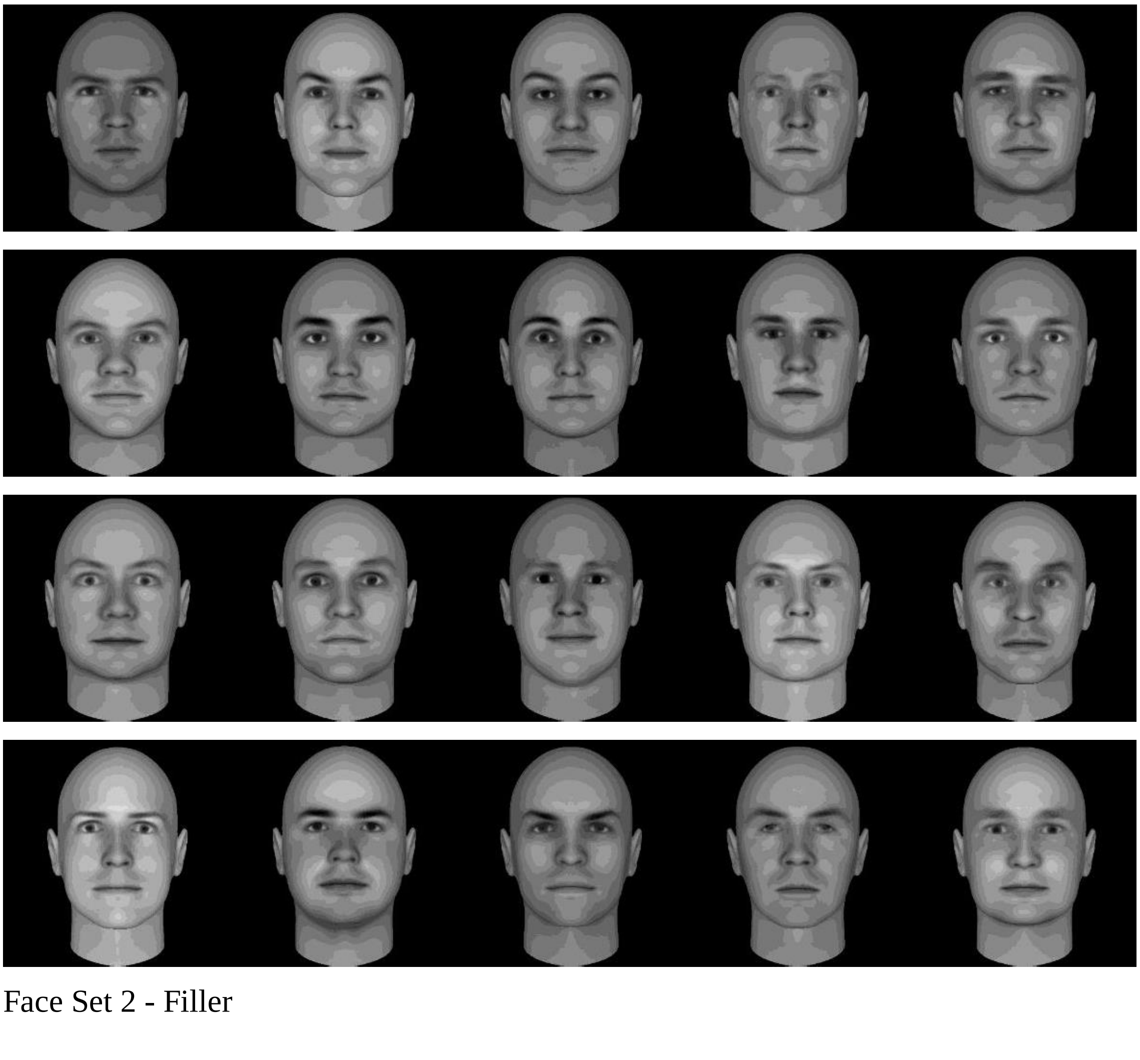
Face Set 1 - Targets



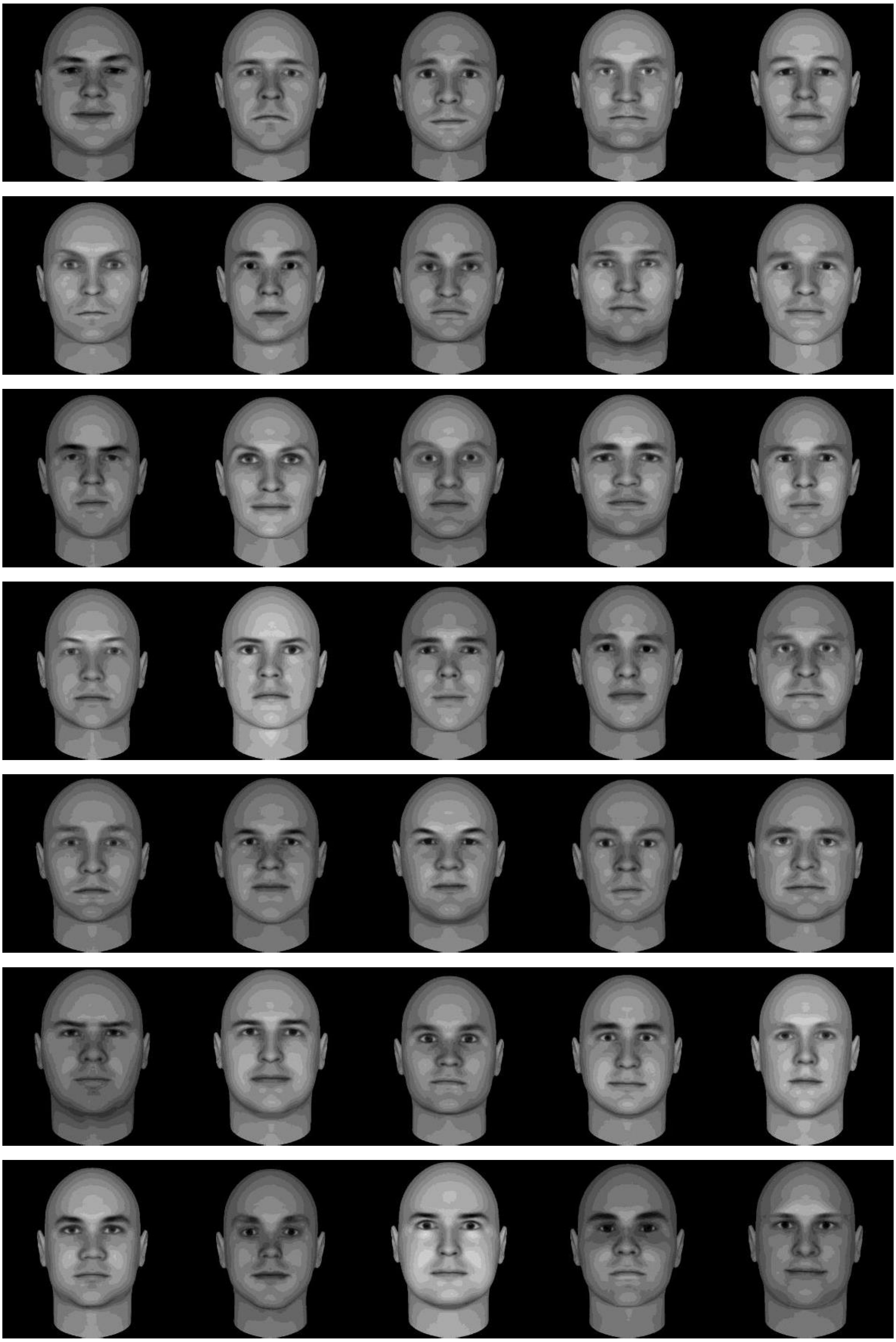
Face Set 1- Filler

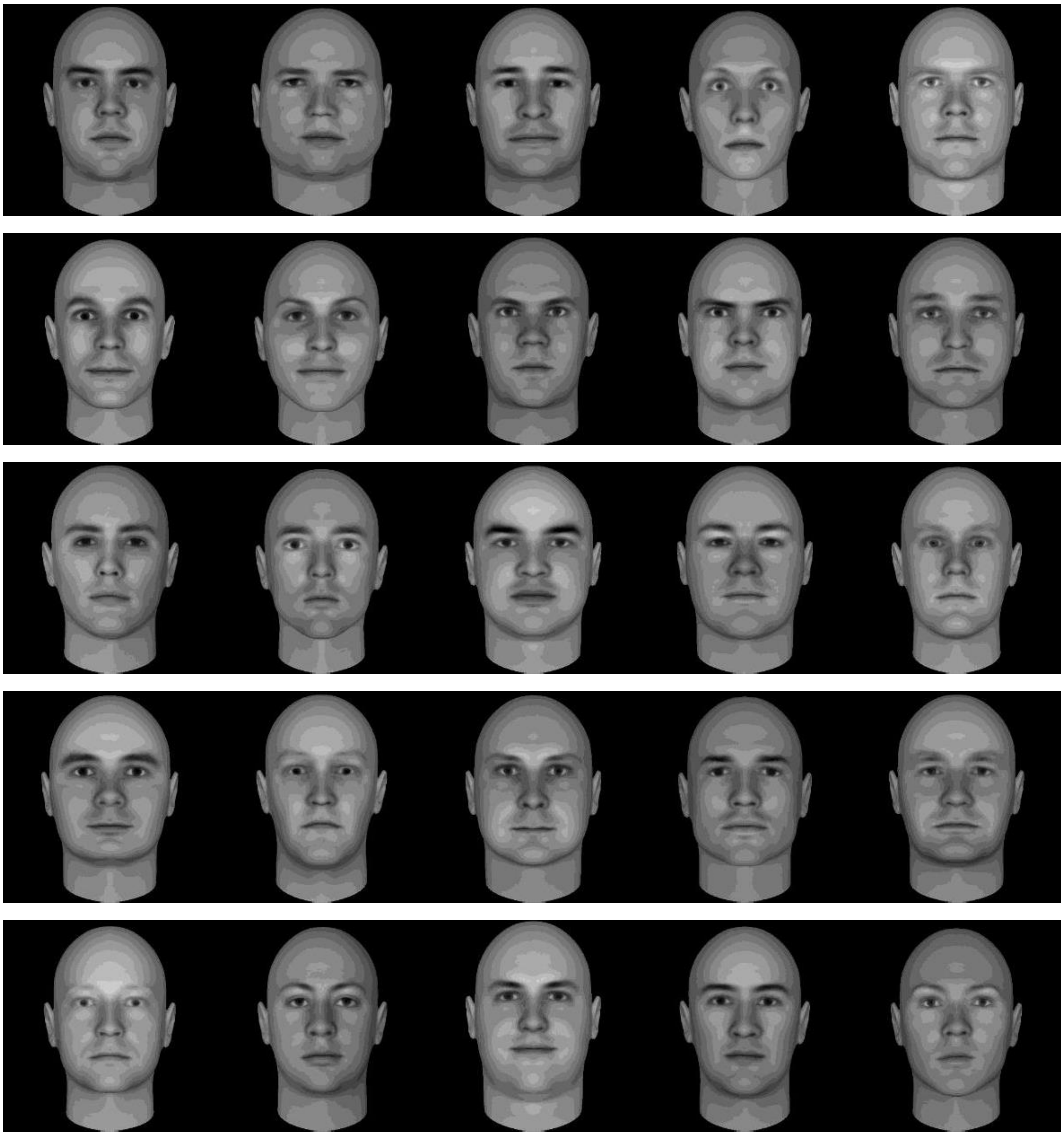


Face Set 2 - Targets

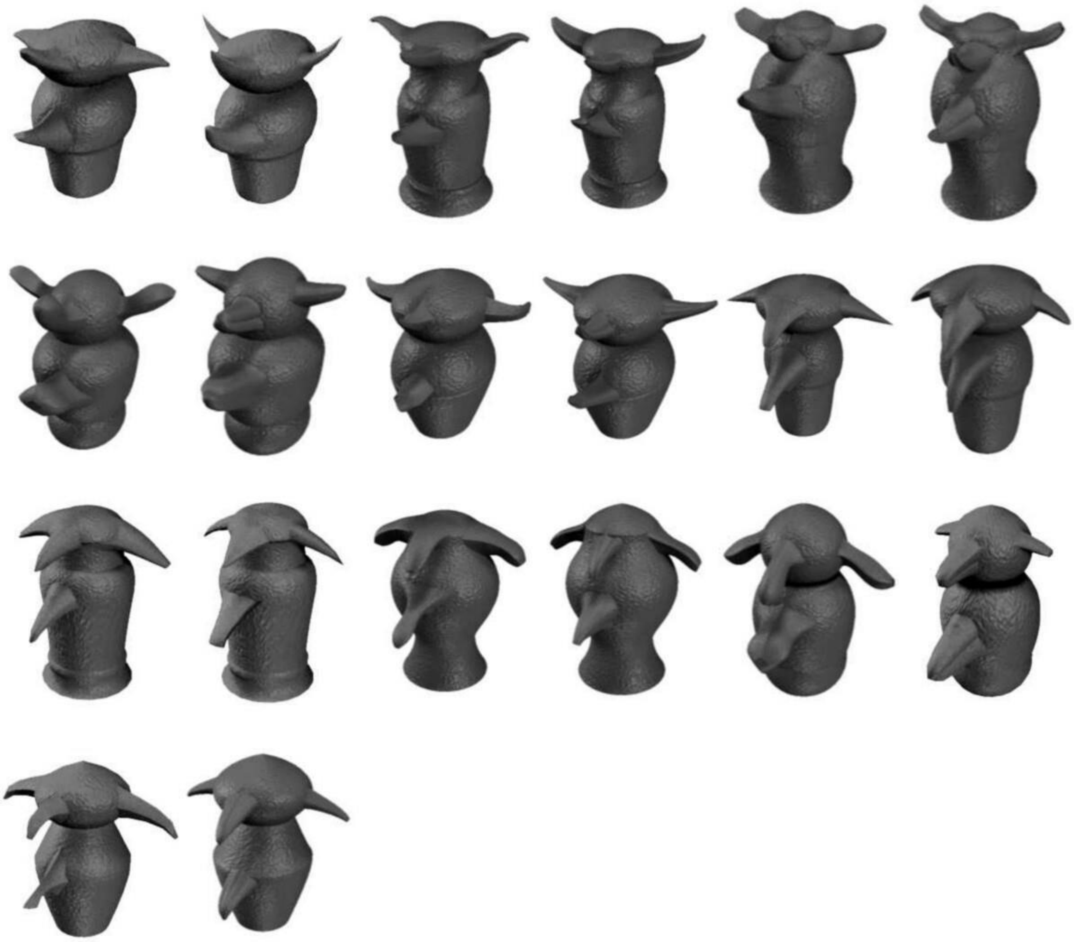


Face Set 2 - Filler

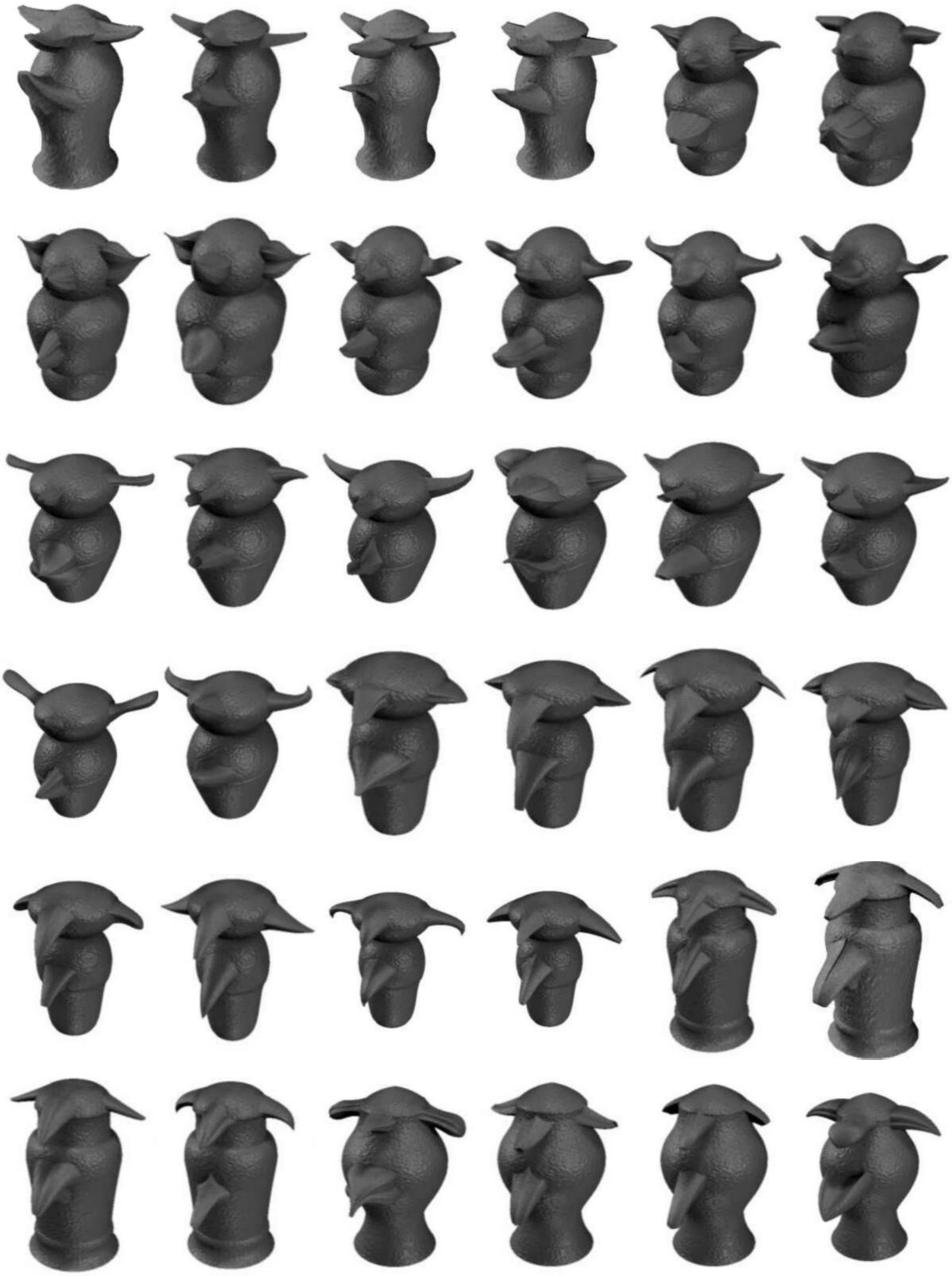


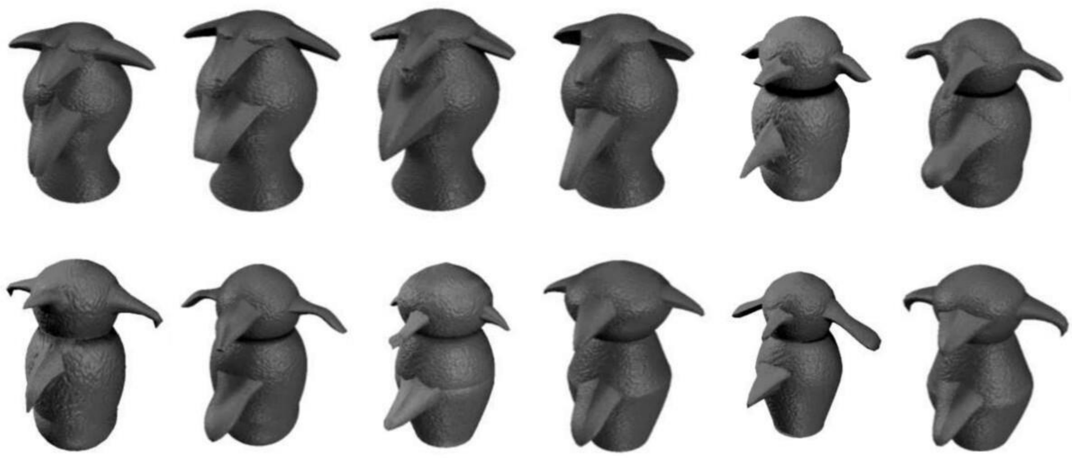


Greeble Set – Targets



Greeble Set – Filler





Appendix C – Session Structure

Structure of each session for each condition in this study. * = greeble training condition only task

Session 1	Session 2	Session 3	Session 4	Session 5 - 8
Family Inspect*	Revision	Revision	Revision	Individual Naming
Family Naming*	Individual Inspect	Individual Inspect	Individual Inspect	Individual Verification
Individual Inspect	Naming with Feedback	Naming with Feedback	Naming with Feedback	Family Verification*
Naming with Feedback	Individual Naming	Individual Naming	Individual Naming	
Family Naming*	Verification	Verification	Verification	
Individual Naming	Individual Naming	Individual Naming	Individual Naming	
Verification	Verification	Verification	Verification	
Individual Naming		Individual Naming	Individual Naming	
Verification		Verification	Verification	

Appendix D - Individual Analyses Statistics

Crawford's revised standardised difference test (RSDT) The RSDT is used when investigating whether the relative performance on two tasks is different for a single-case than for a control group. The RSDT tests whether the difference in performance on two tasks for a single-case is abnormal given the control group distribution of the differences in performance on those same two tasks. The test statistic *t* is calculated using the following formula:

$$t = \frac{Z_{X_1} - Z_{Y_1}}{\sqrt{(2-2r_{xy})\left(\frac{N_2+1}{N_2}\right)}}$$

Where Z_x and Z_y are the individual's Z-score for performance on the two tasks, based on the distribution of performance in the control group on each of those measures, r is the correlation between scores on the two tests in the control group, and N_2 is the sample size of the control group.

Single-score Comparison Test (Crawford, Garthwaite, & Porter, 2010). The Single-Score Comparison Test is used to test for a deficit in a single case. It tests whether a single-case performs significantly below the control distribution on a given task. The test statistic *t* is calculated using the following formula:

$$\frac{X_1 - \overline{X_2}}{S_2\sqrt{\frac{N_2+1}{N_2}}}$$

Where X_1 is the individual's score on the test, X_2-bar is the mean of the control sample, S_2 is the standard deviation of the control sample, and N_2 is the sample size in the control group.

Appendix F – Previous Training Studies Summary

List of all published novel object training studies to date based on the original greeble training paradigm. Number of sessions, length of sessions (hours), length of study (days), number of participants, and stimuli used are listed. Numbers listed in parenthesis for number of sessions is formatted as (learning phase, testing phase).

- = information not provided.

Researcher	Year	Training Tasks	Number of Sessions	Length of Session (hrs)	Length of Study (days)	Participants	Stimuli
Gauthier & Tarr	1997	Inspect (gender, family, and individual), naming with response, naming with feedback, Naming, Verification.	7-10	1	-	32 Controls	Greebles
Gauthier, Williams, Tarr, & Tanaka	1998	Inspect (Gender, Individual), Naming with response, Naming with Feedback, Naming, Verification	10 (4, 6)	~ 1	14	12 Controls	Greebles
Gauthier, Tarr, Anderson, Skudlarski, & Gore	1999	Inspect (Gender, Individual), Naming with response, Naming with Feedback, Naming, Verification	-	~ 7 total	Minimum 4 days (no maximum listed)	5 Controls	Greebles
Rossion, Gauthier, Gauffaux, Tarr, Crommelinck	2002	Inspect (Gender, Individual), Naming with response, Naming with Feedback, Naming, Verification	14 (4, 10)	Learning: 1 Criterion: -	9	10 Controls	Greebles
Gauthier & Tarr	2002	Inspect (Gender, Individual), Naming with response, Naming with Feedback, Naming, Verification	4-6 (4, 0-2)	1.5	14	10 Controls	Greebles

Duchaine, Dingle, Butterworth, Nakayama	2004	Inspect (family and individual), Naming with Feedback, Naming, Verification	8 (4, 4)	Learning: 1 Criterion: .25	8	1 Developmental Prosopagnosic (Edward) 6 Controls	Greebles
Behrmann, Marotta, Gauthier, Tarr, & McKeef	2005	Inspect (Gender, Family, Individual), Naming with response, Naming with Feedback, Naming, Verification	31	-	31 weeks	1 Acquired Prosopagnosic - SM	Greebles, Objects, Faces
Wong, Palmeri, & Gauthier	2009	Inspect (family and individual), Naming with Feedback, Naming, Verification, Matching	10	1	-	18 Controls	Ziggerins
Brants, Wagemans, & Op de Beeck	2011	Inspect (gender, family, and individual), naming with response, naming with feedback, Naming, Verification	10	~ 1	14	8 Controls	Greebles
Bukach et al	2012	Inspect (gender, family and individual), Naming with Response, Naming with Feedback, Naming, Verification	Controls - 9 (5, 4) AP – 21 (8, 13)	1	-	1 AP (LR) 5 Controls	Greebles
Rezlescu Barton Pitcher & Duchaine	2014	Inspect (family and individual), Naming with Feedback, Naming, Verification	8 (4, 4)	Learning: 1 Criterion: .25	8	2 APs (Herschel and Florence) 12 Controls	Greebles and Faces