Cardiff University, School of Psychology,

Cardiff, Wales

Supporting face familiarization using perceptual and engineering frameworks

---

Scott Phillip Jones

A Thesis

submitted to the School of Psychology

for the degree

Doctor of Philosophy

2013

CONTENTS

Acknowledgements

Dr Dominic Dwyer and Dr Michael Lewis have been wonderful in their support throughout my time at Cardiff University. I am grateful that they were always very accommodating when I required just “5 minutes” of their time. I would also like to thank them and Rob Honey for giving me the opportunity to undertake a PhD.

I am also very grateful for some brilliant friends, many of whom I have made throughout the course of my PhD. Their support and enthusiasm for my research area has helped my own understanding, and all the hilarious moments we have shared has left me with some great memories.

Finally I would like to thank Katy Burgess for her on-going support: reading and correcting me on my often dreadful punctuation and grammar, and for listening to any talks that I have prepared several times over. I realise that this must have been very tedious at times. I would also like to thank my family for their continued backing and guidance it means more to me than I could ever accurately describe.

List of Publications

Jones, S. P., & Dwyer, D. M. (in press). Perceptual learning with complex visual stimuli is based on location, rather than content, of discriminating features. *Journal of Experimental Psychology: Animal Behaviour Processes.* doi: 10.1037/a0031509

Jones, S. P., Dwyer, D. M., & Lewis, M. B. (submitted). Exposure to similar comparators does not facilitate matching with unfamiliar faces. *Applied Cognitive Psychology*

Note: Chapter 2 closely reflects that of Jones, Dwyer, & Lewis (submitted), while Chapter 3 is similar to that of Jones & Dwyer, (in press).

List of Presentations

Scott P. Jones, Dominic M. Dwyer, Michael B. Lewis (2012). *Can exposure to similar comparison stimuli facilitate improved recognition of unfamiliar faces?* Experimental Psychology Society Meeting, Hull, UK.

Scott P. Jones, Dominic M. Dwyer (2012). *Perceptual learning using complex visual stimuli is based on location, rather than content, of discriminating features*. Associative Learning Symposium, Gregynog, Newtown, Powys, Wales, UK.

Scott P. Jones, Dominic M. Dwyer (2012). *Can exposure to similar comparison stimuli facilitate improved recognition of unfamiliar faces?* Cognitive Group Seminar Series, School of Psychology Cardiff University.

Scott P. Jones, Dominic M. Dwyer, Michael B. Lewis (2012). *Exposing multiple viewpoints in face recognition: A comparison of photos vs.computer generated stimuli.* Human Face Perception and Animation Meeting, School of Psychology, Cardiff University.

Scott P. Jones, Dominic M. Dwyer, Michael B. Lewis (2011). *Comparison in face processing: Can pre-exposure facilitate improved recognition of unfamiliar faces?* Society for Applied Research into Memory and Cognition, 9th bi-annual meeting, New York City, New York, USA.

Scott P. Jones, Dominic M. Dwyer, Michael B. Lewis (2010). *Comparison in face processing: Can pre-exposure facilitate improved recognition of unfamiliar faces?* Cardiff University Postgraduate Conference, Gregynog, Newtown, Powys, Wales, UK.

Summary of Thesis

The identification of unfamiliar faces is known to be inferior to the recognition of faces with which we are familiar. This can lead to undesirable consequences such as misidentification. However, there is some evidence to suggest that a brief period of familiarisation can dramatically improve our ability to recognise an unfamiliar individual. Chapter 1 outlines the previous research that has aimed to understand the mechanisms of face processing, and to improve the recognition of unfamiliar faces. Three areas that require further investigation are identified and the experimental work reported in the three empirical chapters addresses these issues.

Chapter 2 reports five experiments, using photographs of faces as stimuli, which examined whether a short training exposure promoting stimulus comparison can facilitate recognition of unfamiliar faces (c.f. Dwyer & Vladeanu, 2009). The results revealed that, contrary to expectation, any beneficial effects of comparison do not extend to improving discrimination between targets and nonexposed stimuli.

The results of Chapter 2 required a return to the mechanisms of perceptual learning thought to underpin the comparison effect. Numerous attempts to unpack this process have relied on experiments that have examined the content, but not the location, of the unique features of a stimulus (e.g., Hall, 2003; Mitchell, Nash, & Hall, 2008; Mundy, Honey, & Dwyer, 2007). Chapter 3 used checkerboards as stimuli, manipulating the placement of the unique feature, as a way of breaking the perfect correlation between content and location and assess their relative contributions to perceptual learning. The findings indicated that discrimination between similar stimuli on the basis of exposure can be explained entirely by learning where to look, with no independent effect of learning about particular stimulus features.

Chapter 4 returned to the issue of potential methods to improve recognition, and examined the possibility that training using synthesised faces created from a single view and presented at multiple yaw rotations can aid face recognition (Liu, Chai, Shan, Honma, & Osada, 2009). The findings of three experiments; strengthen the claim that identifying an individual can be improved using multiple synthesised views generated from a single front view of a face, and suggest that this improvement may be affected by the quality of synthesised material.

In summary, while the results reported within this thesis indicate that comparison between similar faces does not produce an effective way of supporting the recognition of unfamiliar faces, they do indicate that experience with a face and/or artificial faces may be a practical means of facilitating identification.

List of Tables

Chapter 2

Chapter 3

Chapter 4

List of Figures

Chapter 2

Chapter 3

Chapter 4

Chapter 1: General Introduction

1.0 Introduction

It is commonplace to carry some form of photographic identification in modern society. This form of identification is required in a variety of environments, from workplace entry to immigration. Moreover, photographs are frequently used as a means of identifying wanted persons – something exemplified in the way that, during the 2003 invasion of Iraq, soldiers were issued a set of playing cards depicting the faces of the most wanted members of Saddam Hussein’s government. However, despite the reliance on photographic identification, the ability of people to accurately match between a photograph and the individual in question is typically poor (e.g., Bindemann & Sandford, 2011). Equally, we are unlikely to be able to recognise an individual who we have had little or no experience with (see Hole & Bourne, 2010). This fallibility of human recognition highlights the limited value of current identification measures, despite their common usage.

More generally, there is a wealth of experimental evidence from laboratory tests indicating that the processing of the images of unfamiliar faces is substantially inferior to that of familiar individuals, especially when changes in image quality are involved. One of the most well-known field demonstrations was conducted by Kemp and colleagues (Kemp, Towell, & Pike, 1997) who reported that supermarket staff accepted 50% of photo-ID cards which did not depict the individual in front of them. Findings from laboratory settings yield similar effects. For example, even under optimal viewing conditions, participants can still succumb to errors in identification 10-20% of the time (Bindemann, Avetisyan, & Blackwell, 2010; Megreya, Bindemann, & Havard, 2011). This poor identification can also transcend different pictures of the same individual as demonstrated, using several forms of ID cards, by Bindemann and Sandford (2011). It is then surprising that society puts so much trust in photographs to prove identity.

One goal of psychological research in face processing is to improve the identification and processing of previously unfamiliar faces. However, any forensically useful improvement would need to prove generalizable in the presence of novel pictures and across viewpoints, because memory for unfamiliar faces has been demonstrated to be view-dependent (Longmore, Liu, & Young, 2008). That is, any improvement that was simply a case of improved picture matching rather than improved recognition of an individual *per se* would be of no practical benefit. Compounding this finding, is evidence that seeing a face in one view often fails to transfer successfully to another view (Bruce, 1982). However, there has been some success using images spanning multiple views (Liu & Ward, 2006; Pike, Kemp, Towell, & Phillips, 1997), and extended training periods to encourage familiarisation (O'Donnell & Bruce, 2001; Stevenage, 1998). Briefly put, in the studies that spanned multiple views, faces were presented in a range of angles during training that gave more structural information (see section 1.2.3 for a more detailed outline), while the studies that used extended training periods gave repeated exposure to a face that induced a change in the image-based information processed from a target (see section 1.2.2 for a more detailed outline). Both these approaches utilised methods of exposure that were thought to encourage deeper processing more like that seen with familiar faces. It has been suggested that the shift from unfamiliar to more familiar like processing is a consequence of some kind of perceptual learning mechanism (Bruce & Burton, 2002; Valentine & Bruce, 1986).

Perceptual learning has been defined as a “consistent change in the perception of a stimulus array following practice or experience with this array” (Gibson, 1963, p. 29). The key demonstration of this effect is that simple pre-exposure to two stimuli, requiring no external reinforcement, can increase their discriminability (Gibson & Walk, 1956). The lack of formally supervised training or reinforcement is particularly interesting given the application to situations which are often self-regulated, such as face processing. Albeit that the question of whether the absence of formal supervision entirely removes strategic contributions to perceptual learning based on “mere” exposure has yet to be conclusively answered (e.g., Mackintosh, 2009). Regardless, the schedule of unsupervised stimulus exposure strongly influences the quality of perceptual learning and accuracy of subsequent discrimination (e.g., Dwyer, Hodder, & Honey, 2004; Honey, Bateson, & Horn, 1994). Most relevant here, recent research has found improvements in discrimination and matching using faces as stimuli and relatively short exposures. That is, certain exposure schedules can aid discrimination between individuals, and also matching between faces, by providing an opportunity for comparison between individuals (e.g., Dwyer, Mundy, Vladeanu, & Honey, 2009; Dwyer & Vladeanu, 2009; Mundy *et al.*, 2007).

1.1 Rationale

This thesis is separated into three experimental chapters which follow three related themes. The first theme will explore exposure schedules derived from the perceptual learning literature as a potential mechanism for improving recognition of unfamiliar faces. As reviewed below, perceptual learning has been implicated in various contexts within the face-processing literature (see, O'Toole, Deffenbacher, Valentin, & Abdi, 1994; O'Toole, Vetter, & Blanz, 1999; Valentine, 1991). Many studies on the effects of perceptual learning have investigated the best conditions with which to induce learning. For example, it has been shown that intermixed exposure (e.g., AX, BX, AX, BX, …), whereby A and B represent a unique element and X a common feature, produces better discrimination performance compared to blocked exposure (AX, AX, …BX, BX, …). The advantage of intermixed scheduling has been demonstrated across a range of species including chicks (Honey *et al.*, 1994), rats (Symonds & Hall, 1995), and humans (Mundy, Dwyer, & Honey, 2006). Moreover, recent evidence has suggested that this form of unsupervised intermixed exposure can facilitate better discrimination of previously unfamiliar faces (Mundy *et al.*, 2007).

However, there is relatively little conclusive evidence of whether the effect can generalise to situations when a target face is to be discriminated from novel (nonexposed) stimuli. If this effect does generalise then it will have potential forensic applications.

The second theme will be an analysis of location-based versus content-based learning using another form of complex visual stimuli (i.e., checkerboards). The possibility that perceptual learning based on content and location might dissociate is particularly interesting given that the processing of familiar faces relies heavily on internal features. Various training-based improvements in face perception have been attributed to directing attention towards these features as the key to enhanced recognition. However, it is not possible to separate the location and content of the internal features of the face. Although few studies have investigated learning based on attention to diagnostic locations, checkerboard stimuli have been used to probe human perceptual learning (e.g., Lavis & Mitchell, 2006; Mitchell, Kadib, Nash, Lavis, & Hall, 2008; Wang & Mitchell, 2011). These checkerboards provide one potential stimulus for assessing the contribution of location separate from that of content, albeit that existing studies, using checkerboards, leave open the question of whether learning is based on the content or location of the unique features because they have typically been perfectly correlated. While the basic schedule effects underpinning many analyses of perceptual learning are present in stimuli that are not open to strategic spatial attention, for example: other probabilistically defined checkerboard stimuli (McLaren, 1997), flavours (Dwyer *et al.*, 2004), and morphed faces (Dwyer *et al.*, 2009); It remains important to ascertain whether location- and content-based perceptual learning can be separated.

The final theme of this thesis returns to the forensic concerns of poor performance in unfamiliar face recognition to investigate an alternative means of improving recognition using an engineering framework. One of the difficulties associated with unfamiliar face processing is that it is largely image-dependent, thus, being able to recognise a face in one view under certain set of conditions does not guarantee recognition under other conditions. That said, there is evidence that recognition performance is enhanced after exposure to multiple static images, taken from different viewpoints, of a single individual (Pike *et al.*, 1997). A variety of computer-based face modelling approaches that support the generation of multiple views from even a single 2D photographic input and thus might offer a mechanical means to supply multiple viewpoint information when only a single veridical image is available (see, Blanz & Vetter, 1999). However, the question remains as to whether this can genuinely support human face recognition to overcome the fact that unfamiliar face recognition is highly view-dependant (Longmore *et al.*, 2008). Promisingly, there is already preliminary evidence that this may be the case (Liu *et al.*, 2009; Liu & Ward, 2006), albeit that the generality of such effects has not been explored.

The following sections of the introduction will outline the literature relevant to these three themes. These sections concentrate on how faces are represented, theoretical analyses of how faces move from unfamiliar to familiar, and attempts to influence and improve face recognition. Moreover, it will briefly outline why face recognition is fallible, and attempts to improve this using engineering frameworks. In addition, I will explore the links between perceptual learning and face recognition, in the context of some of the main theories of perceptual learning and the mechanisms which underpin these established effects.

1.2 Face Perception

Our ability to process faces is an important aspect of our lives. Vital information for our social and emotional responses can be detected from the face. However, identifying an individual can be a challenging task, given that all faces have the same general configuration of internal features (i.e., a pair of eyes above a nose and mouth) and thus this basic configuration cannot serve as any basis for recognition. In addition, faces can go through various changes which can impair this process further such as the growth of facial and head hair, or age related changes in skin tone and texture. Moreover, recognition of a face often has to cope with changes in angle and illumination. The central questions at hand here are how experience allows us to build representations of familiar faces that transcend these differences, and whether we can exploit these experience induced changes in situations where accurate identification of otherwise unfamiliar faces is crucial.

Despite the challenges to our ability to identify an individual it has been suggested that we are "experts" at identification (Bahrick, Bahrick, & Wittlinger, 1975; c.f. Mondloch, Le Grand, & Maurer, 2002). While this is not the case for all faces, especially ones we are unfamiliar with, we can be considered experts at recognising our family and friends: It is often the case that we can meet and recognise people from our childhood (e.g., friends from our school years). Indeed, high school graduates have demonstrated 90% correct identification or matching rate for names and faces at least 15 years from their last encounter (Bahrick *et al.*, 1975). This demonstrates a finely tuned ability to recognise familiar faces. Laboratory research confirms that it is easier to detect differences between two pictures when familiar faces are used compared to when the faces are novel (Buttle & Raymond, 2003). In addition, near perfect recognition can be observed despite poor resolution, lighting, and viewing angle (Burton, Bruce, & Hancock, 1999), and with changes to expression (Bruce, 1982).

In contrast, our ability with unfamiliar faces is very poor. This has been demonstrated in a variety of field and laboratory experiments. In the field, Logie, Baddeley, and Woodhead (1987) found substantial error rates when asking people in the street to identify an unfamiliar target or decide whether a person matches a credit card photograph. Moreover, this is supported by laboratory demonstrations with unfamiliar faces. For example, from several 1-in-10 line-ups, participants were asked: firstly, if the target person was present in the array, and finally, if they could identify the matching person in the array. It was found that performance on this task was surprisingly error prone: that is, despite the high quality nature of the photographs, there was a 30% rate of incorrect responses on both target present and target absent arrays (Bruce *et al.*, 1999; Bruce, Henderson, Newman, & Burton, 2001).

The superiority of recognising a familiar individual suggests that face processing for familiar and unfamiliar faces may rely on different processes (see section 1.2.4). Indeed, models of face processing typically differentiate between familiar and unfamiliar recognition processes (Burton *et al.*, 1999; Hancock, Bruce, & Burton, 2000). Moreover, there is evidence of different processing strategies for familiar face compared to unfamiliar faces. For example, the internal features of a face (e.g., the eyes, nose and mouth) have more influence than the external features (e.g., hair or face outline) in the recognition of familiar faces than in unfamiliar faces (Campbell *et al.*, 1999; Ellis, Shepherd, & Davies, 1979; Young, Hay, McWeeny, Flude, & Ellis, 1985).

Further to this, unfamiliar face representations are more image-specific and recognition accuracy is susceptible to various image changes. That is, even when little or no memory load is involved, two different images of the same individual can often be misinterpreted as the images of two similar looking people, an error that seldom happens with familiar individuals (Young, Hay, & Ellis, 1985). Indeed, substantial errors are made when participants are asked match photo array images of an unfamiliar individual shown previously on high quality video. Conversely, they are found to be highly accurate for familiar faces even with poor quality videos (Bruce *et al.*, 1999; Bruce *et al.*, 2001). In addition, the idea that unfamiliar faces are image-specific is supported by the disruption induced by changing image conditions. Hill and Bruce (1996) demonstrated this in a same/different task whereby performance was reduced, in same trials, when two images of the same individual were lit from different directions compared to lighting from the same direction. Bruce *et al.* (1999) also found that matching video images to full face photo arrays was more difficult when the video depicted the face at a 30° angle. Given that the face matching tasks require no memory component, the image-specificity of unfamiliar faces appears to be a product of the way in which faces are processed.

Despite these differences in recognition performance there are similarities in the processing of familiar and unfamiliar faces. For example, both unfamiliar and familiar faces show categorical perception effects (Beale & Keil, 1995; Levin & Beale, 2000), and are disturbed by inverting the face (Rhodes, Brake, & Atkinson, 1993; Young, Hellawell, & Hay, 1987). These similarities suggest that there is at least some commonality between the processing of familiar and unfamiliar faces. However, the differences highlighted earlier remain, and relatively little is known about how faces move from unfamiliar to familiar.

1.2.1 A special type of stimuli?

The question of whether faces are a special class of stimuli relates to the idea that there may be specific visual mechanisms for processing facial identity independent of that of other classes of objects. Several strands of research have provided evidence for the special nature of faces. First is the suggestion that some face processing abilities are innate (Morton & Johnson, 1991). Indeed, Charles Darwin questioned whether our adult skill of face perception is due primarily to species adaptation or by developmental experience (as cited in Johnson, 2011). Second is the notion of a unique processing style attributed to faces (Robbins & McKone, 2003) reflected by various effects that disrupt and impede recognition of faces more than that of other objects (e.g.,Yin, 1969). Finally, there is an abundance of evidence of face selective regions within the brain and that certain neurons show face-specific responses (Kanwisher, McDermott, & Chun, 1997).

Developmental research has suggested that infants are born with an innate disposition to look at faces, and such attention to faces above other objects is evidence of the special nature of this class of stimuli (Morton & Johnson, 1991). Indeed, infants that are just 10 minutes old display a preference to stimuli which most resembled a human face (Goren, Sarty, & Wu, 1975). Moreover, both new born babies and young monkeys raised without any face exposure display an ability to differentiate between two similar faces, even with a viewpoint change between exposure and test (Sugita, 2008; Turati, Bulf, & Simion, 2008; Turati, Cassia, Simion, & Leo, 2006). These results suggest that there are some innate features of face processing. However, it has been shown that encoding unfamiliar faces is less efficient than that of adulthood (Chung & Thomson, 1995), and that adult-level recognition and discrimination abilities take time to develop (Carey, 1992; Mondloch *et al.*, 2002). This is supported by evidence of a perceptual narrowing period of infants during development which effectively disrupts the initial ability to characterise faces of other-races or species (Kelly *et al.*, 2007; Kelly *et al.*, 2005; Sugita, 2008). Thus, although there is evidence for an innate preference for detecting faces, this system can be modified and refined to alter the way faces are recognised and discriminated. For this reason it has been suggested that the innate system is "experience-expectant" (McKone & Robbins, 2011 p. 150) although this does not diminish the distinction between the innate dispositions towards faces and that towards other visual objects.

The special status afforded to faces has also been supported from behavioural experiments examining adult face perception. The inversion effect, first described by Yin (1969), refers to the fact that presenting faces upside-down severely disrupts recognition. In light of the fact that our expertise with faces is thought to rely on processing in a holistic or configural manner, the disruptive effect of inversion on faces compared to other objects has been taken as evidence that faces are special (Farah, Wilson, Drain, & Tanaka, 1998; Robbins & McKone, 2007; Tanaka & Gauthier, 1997). In particular, the disruptive effects of inversion have been attributed to the disruption of global relationships (Bartlett & Searcy, 1993; Robbins & McKone, 2003). However, while a disproportionately large effect of inversion has been demonstrated for faces over other objects (McKone & Robbins, 2011; Robbins & McKone, 2007), it should be noted that there is evidence that inversion affects the processing of other stimuli with which people demonstrate expertise. Indeed, these large inversion effects have been observed with dog (Diamond & Carey, 1986), bird (Gauthier, Skudlarski, Gore, & Anderson, 2000) and car (Curby, Glazek, & Gauthier, 2009) experts, suggesting that our ability to process faces is a result of expertise with a stimulus class rather than expertise with faces *per se*. Moreover, people with similar levels of expertise with certain objects show similar processing traits to those of faces. That is, the ability to distinguish between homogenous classes of stimuli consisting of the same parts is thought to require a more configural style of processing that can account for the "second order relational features" (Diamond & Carey, 1986; Farah *et al.*, 1998; Rhodes *et al.*, 1993). This expertise has also been demonstrated using laboratory training with other types of homogenous stimuli. For example, Gauthier and Tarr (1997) trained participants to become experts with "Greebles" and found that this class of stimuli were treated in a similar way to faces. Trained greeble experts displayed faster identification of greeble configurations compared to novices. This indicates that training increased the configural processing of non-face objects. In all these cases it appears that configural processing of objects can be acquired through experience, which suggests that the processing of faces cannot be special merely because it relies on configural mechanisms. That said, recent reviews have argued that only faces display holistic processing on "gold standard measures" i.e., composite, part-whole, and part-in-spacing-changed-whole task (for a detailed discussion see, Bukach, Gauthier, & Tarr, 2006; Gauthier & Bukach, 2007; McKone & Robbins, 2011; Robbins & McKone, 2007). Thus, the exact relationship between inversion effects and configural processing remains to be established.

Neuroimaging studies have suggested that there are at least three principal cortical regions where activity is greatest for faces relative to other non-face objects. These findings are based on neural activation of participants passively viewing faces verses objects. The regions that are consistently found to be activated by presentation of faces are: in the fusiform gyrus (e.g., the fusiform face area or FFA); lateral inferior and middle occipital gyrus (e.g., occipital face area or OFA); and in the superior temporal sulcus (e.g., STS, the face-selective region: for reviews see, Kanwisher & Yovel, 2006; Natu & O'Toole, 2011). Moreover, evidence for the "specialness" of face processing within the brain comes from studies of individuals with prosopagnosia (face blindness). The patient known as P.S. demonstrates the critical role of the right OFA in judging identity. P.S. shows typical level performance for discriminating faces from objects and discriminating between objects, but is impaired on serveral tests of face recognition ability (see, Ramon, Busigny, & Rossion, 2010). The deficit in identification for P.S. is associated with lesions encompassing left mid-ventral and right inferior occipital cortex (Ramon, Busigny, & Rossion, 2010). The rarity of this case is highlighted by the observation that P.S. has problems in recognising and naming faces but not objects, therefore indicating a selective brain region whereby faces are processed and the special nature of these stimuli in comparison with other objects.

While the evidence summarised above does suggest faces are in some ways "special" as stimuli, it is important to note that the processing of faces is not fixed or immutable, and that McKone and Robbins (2011) suggest that certain basic properties can improve perceptual efficiency. For example, recently repeated stimuli will be identified faster and using fewer resources. Indeed, familiarization with face stimuli improved the accuracy with which they were identified in noise (Gold, Bennett, & Sekuler, 1999a, 1999b; Gold, Sekuler, & Bennett, 2004). Moreover, it was demonstrated that this perceptual learning was due to an increase in the internal signal strength of the familiar faces. Interestingly, these studies also demonstrated that the same was true of non-face stimuli, implying continuity between the mechanisms underpinning perceptual learning with face and non-face stimuli (Gold *et al.*, 1999a, 1999b; 2004).

The empirical work of Chapters 2 and 4 will use pictures of human faces as stimuli. This work will examine whether it is possible to build upon the “expertise” of face processing by giving exposure of different types to encourage more in-depth processing. That is, (in Chapter 2) whether training based on perceptual learning principles can be extended to familiarisation of faces, to ameliorate certain deficits in our ability to recognise faces, and (in Chapter 4) whether artificially generated faces can support learning.

1.2.2 Improving face recognition - exposure helps

Numerous studies have focused on improving recognition of unfamiliar faces; however, according to some, deliberate efforts to improve the face recognition skills which most of us develop, are “doomed to failure” (Ellis, Jeeves, Newcombe, & Young, 1986, p. 9). This statement stems from a review of identification training programmes by Malpass (1981a). The review suggested that while identification training can facilitate short-term improvements for other-race faces, there is no evidence for training-based improvements for own-race faces. Moreover, there is evidence that suggests that training programmes[1] can adversely affect own-race recognition (Malpass, 1981b). For example, when participants were required to classify specific features, post-training performance decreased (Woodhead, Baddeley, & Simmonds, 1979). That said, it had been suggested that that these techniques focused too heavily on featural information of a face, and so the configural information that

---

[1] Training programmes can encompasses a variety of methods such as explicitly noting similarities and differences between faces (Malpass, Lavigueur, & Weldon 1973), exposing photofits of a target (Woodhead *et al.*, 1979) or concept learning (Lavrakas, Buri, & Mayzner, 1976). However, while the content of the training varies between experiments, broadly speaking, most training programmes involve exposure with a set of faces that participants will be later be asked to identify.

was available was not capitalised upon (Ellis *et al.*, 1979). However, while, attributional judgements on qualities such as honesty or likeableness have, in some cases, improved recognition (e.g., Bower & Karlin, 1974), no lasting enhancement to performance was observed when incorporating a global strategy within a training programme (Malpass, 1981a).

Despite the idea that training cannot improve recognition, it has been suggested that people who are accurate at matching unfamiliar faces may be able to direct their attention towards the more diagnostic information contained in the internal features of the eyes, nose and mouth (Fletcher, Butavicius, & Lee, 2008). This directed attention is closer to the processing strategy used for recognising familiar faces (see, Ellis *et al.*, 1979). Indeed, a more recent strategy to improve recognition of novel faces has been to use an extended period of familiarisation training. For example, using a training period consisting of successive exposures across nine days, participants displayed a switch from external feature dominance to a more internal based supremacy (Angeli, Bruce, & Ellis, 1999). Furthermore, Bonner, Burton, and Bruce (2003) found a shift in matching performance for faces learned from video-based images. In this experiment, participants were familiarised during a three-day period. After which, familiarised participants showed equivalent face-matching performance for internal features compared to external features, such that internal feature matching showed a significant improvement while external feature matching remained consistent. Moreover, O’Donnell and Bruce (2001) demonstrated that long exposure periods induced a shift in attention similar to that of familiar face recognition. They asked participants to learn a number of individuals during an exposure period. Faces with associated names were presented repeatedly, in video format, until all individuals could be named correctly. Participants were then given a matching task. It was found that for the control group of unfamiliar faces only changes in the hair were readily detected, whereas faces that had been familiarised were more sensitive to changes in internal features, particularly the eyes. These results suggest that the shift from external features towards internal features is apparent in the way face representations change during familiarisation. These findings also indicate that improvements in recognition can be gained under laboratory conditions, at least when using extended exposure periods.

The fact that experience appears to be critical in the transition from unfamiliar to familiar styles of face processing has often been described as a form of perceptual learning (Bruce & Burton, 2002; O'Toole *et al.*, 1994; Valentine, 1991). However, there have been few attempts to utilise this process for any applied gain. That is, in an applied sense it would be inappropriate (and unworkable) to advocate long exposure periods to improve recognition. The problem is the time constraints placed on most individuals involved in identification procedures. A relatively brief period of familiarisation, however, would be more practical should there be a means for making it effective and in some cases there have been changes based on comparatively small periods of familiarisation. For example, Clutterbuck and Johnston (2005) demonstrated that 10 two-second exposures to a face produced better performance on a face-matching task than was seen with completely novel images. This improvement was selective to the internal features of the face. In addition, studies in perceptual learning have found that, relatively brief exposure (as short as 5 × 2 s) can facilitate discrimination between pairs of faces made similar by morphing them with each other (e.g., Dwyer *et al.,* 2009; Mundy *et al.,* 2007). Brief familiarisation will be considered later following an outline of perceptual learning and exposure schedules. The possible applied benefit of this brief familiarisation training is considered in Chapter 2.

1.2.3 Technological developments versus human performance

As noted, the nature of facial memory is far from perfect and compromised by the passage of time (see section 1.2), highlighting the need for an effective and accurate method of improvement. Indeed, law enforcement agents have been using photographs of suspected criminals in the hope that they can aid the process of recognising suspects since the 19th century and it remains an important process in the fight against crime and terror (Cole, 2001).

In applied situations, face composites are often used. A composite is an attempt to recreate the appearance of a criminal created with the assistance of a witness, used in police investigations, to help identify and/or find the criminal. Advances in this area of facial reconstruction have progressed from artists' impressions, through mechanical systems (e.g., Identikit or Photokit) to software systems (e.g., E-FIT), and more recently genetic algorithms (e.g., Principle Component Analysis (PCA)). However, studies investigating the effectiveness of these methods on recognition have produced divergent results. These systems are limited by the number of available features within a set, the memorial ability of the witness (Lindsay, Mansour, Bertrand, Kalmet, & Melsom, 2011), and the utility of this approach is low (Ellis, Davies, & Shepherd, 1978; Frowd *et al.*, 2005). Moreover, systems like E-fit, that use a feature selection approach, only give correct identification rates at a mean level of 20 per cent (Davies, Van Der Willik, & Morrison, 2000; Frowd *et al.*, 2005). According to Lindsay *et al.* (2011), a majority of these systems fail because they are feature based, unlike facial memory which relies on more holistic processes (Robbins & McKone, 2003). One system which uses a more holistic approach is genetic algorithms. This technique uses PCA to produce a multidimensional similarity space (Face space; see also section 1.2.5.)[2] from a set of faces from which the algorithm can build a likeness (see, Lindsey *et al.*, 2011). Essentially the algorithms generate faces from different sets of input

---

[2] It is important to note that a face space can be defined as a either psychological or physical space. Generated faces are located on all dimensions within a physical face space that are produced by applying models to “empirical image data” (see, Vetter & Walker, 2011 p. 390). The reference to face space within this section (1.2.3) will refer to the physical space whereas section 1.2.5 will describe the psychological space.

faces aiming to get progressively closer to a target in terms of likeness. However, it has been argued that all the methods mentioned above need further refinement given that they still do not produce identification at the same level of accuracy as artists' impressions (Davies & Valentine, 2007). Even so, there can be great variations in quality of these artists' sketches depending on the skill of the artist (Davies & Valentine, 2007). Therefore, one advantage of genetic methods is that an effective algorithm could remove this variation.

Returning to the genetic algorithms process for creating a face, one application of this method has been the generation of better face stimulus materials for laboratory experiments. For example, the morphable face model (Blanz & Vetter, 1999, 2003)[3] was constructed on the basis of laser scans from 100 males and 100 females, with each scan representing two different kinds of information (see, O'Toole, Vetter, Troje, & Bülthoff, 1997). The two kinds of information represent the three-dimensional head surface data and texture average (sometimes referred to as a two-dimensional reflectance map). An average of each dimension was then computed and every face coded, upon a continuous scale, which represents deviation from this given 3D and texture average (for a more in-depth discussion on how the two kinds of information correspond, see Blanz & Vetter, 1999; Vetter & Poggio, 1997). PCA (Principle component analysis) is then conducted to find the eigenvectors allowing a new range of faces to be synthesised. What this method of construction (and others like it) allows is each face to be rendered under clearly defined lighting conditions or views. That is, computer-based face modelling approaches allow the generation of multiple views from even a single 2D photographic input.

---

[3] For an updated version of the model using a similar method of construction see The Basel Face Model (Paysan, Knothe, Amberg, Romdhani, & Vetter, 2009). The update includes better original scans using a more advanced scanning device and an improved correspondence algorithm. However the method of construction remains almost identical to the previous versions.

The failure of many previous approaches to using photographic images to support the recognition of otherwise unfamiliar individuals is perhaps due to the view-dependent nature of faces (e.g., Bruce, 1982; Longmore *et al.*, 2008). A problem exacerbated in applied settings by the fact that it is common to have only a very limited array of images of a target individual (e.g., the police with a single mugshot of a suspect). Indeed, an unfamiliar individual is seldom caught displaying the best pose for future recognition or indeed the same pose as they might be displaying when seen at a later time.

Perhaps unsurprisingly, recognition performance is better when multiple views of a face are made available. For example, it has been shown that multiple static presentations of a face produce better recognition than single static photos (see Experiment 2 of Pike *et al.*, 1997), a similar yet non-significant trend was reported in an experiment examining recognition of a photo after exposure to a single image, multiple images or a video sequence by Bruce and Valentine (1988)[4]. While these results suggest multiple views of a face assist in identifying a face this opportunity is seldom afforded in more applied settings. Critically, this method of generating faces (see description above) allows a face to be generated onto an average from even a single 2D photographic input. The face can then be manipulated within the computer programme to present multiple views of a face. Some studies have compared the recognition of original photographs to how accurately faces are recognised from reconstructed busts and found that, although not as high as the original photograph, performance was very accurate (Bailenson, Beall, & Blascovich, 2003; Bailenson, Beall, Blascovich, & Rex, 2004). More recently, it has been demonstrated that computer modelled faces, created from a single 2D photograph input, can produce more accurate recognition than controls in an old/new task (Liu *et al.*, 2009). People clearly have some ability to generalise

---

[4] It is worth noting that in some demonstrations motion has been shown to facilitate improved recognition compared to multiple static views (e.g., Pike *et al.*, 1997; Roark, O'Toole, Abdi, & Barrett, 2006).

from a single photographic image to a different view of the target face, so the fact that generated views support recognition implies that these models must (in some circumstances at least) perform this transformation more effectively than an unaided person. The possibility that recognition can be aided by this mechanical improvement will be explored in more detail in Chapter 4.

1.2.4 Models of face learning

Bruce and Young (1986) suggested that face processing involves generating and assessing certain codes, depending on the nature of recognition. These codes convey different types of information depending on the stage of processing. For example, pictorial codes contain information of the image which is being presented. Structural codes are more abstract, in that they contain information needed to recognise a face despite image changes. It is possible that the period of familiarisation is effective due to the different codes which are activated for familiar and unfamiliar faces. Bruce (1982) ran two experiments which have provided evidence for the existence of these codes and supported the idea that familiar and unfamiliar processing relies on different codes. In the first, participants were given a set of unfamiliar faces and asked to recognise them. Performance was best for images which remained the same between exposure and test. Moreover, responses and accuracy were slower and more error prone when a change in pose or expression was applied to the images between exposure and test. In the second, using familiar faces accuracy was unaffected by the change in pose or expression. It is thought that because familiar faces are represented by stronger, better developed, structural codes then the change in image format was easier to compensate. On the contrary, participants will use pictorial codes of the image as well as some structural information of the face in unfamiliar face recognition (Bruce, 1982).

The Bruce and Young model of face processing (1986) follows a sequential pattern through the process of recognition. First the model is concerned with the process of encoding the face and producing a representation. This is the structural encoding phase which involves producing a number of different representations. If a face is unfamiliar then the representation will rely on the more view-specific information from which it has been formed. However, if the face is familiar, then the initial representation will activate a Face Recognition Unit (FRU) for that face. These units are individual for each face and are activated by any view of that face. Units that are activated then induce the feeling of familiarity. Following activation of an FRU, a Person Identity Node (PIN) acts as the next stage in the model, these nodes are stimulated by FRUs, which in turn provide access to Semantic Information Units (SIU) and subsequently name information (Name units). The semantic units contain information regarding the owner of a face e.g., their age, occupation, and your relationship towards that person.

While the Bruce and Young model (1986) differentiates between familiar and unfamiliar representations of a face, it does not attempt to explain how a face moves from unfamiliar to familiar. Specifically it fails to accurately account for the effect that repeated exposure will have on the stored representations (see, Hole & Bourne, 2010, for a detailed discussion). However, in more recent hybrid models such as the Interactive Activation and Competition model (IAC), proposed by Burton, Bruce, and Johnston (1990), it is thought that repeated exposure strengthens the connections between the FRU and corresponding PIN. Thus if a face has been seen recently the link is strengthened such that on subsequent viewings, of the same face, processing will proceed faster. Burton (1994) proposed an update of this model in order to capture how faces are learned through adulthood. The Interactive Activation and Competition with Learning model (IACL), provides a mechanism with which faces are learned. Within the model, Burton emphasises three key components to account for the fact that learning occurs from simple repeated exposure to a face. It should be “automatic, gradual and cumulative” (Bruce & Burton, 2002, p.333). Being “automatic” means that it should be an unsupervised process whereby the action of recognising a familiar face should also be one that allows identification and learning of an unfamiliar face (the idea of unsupervised learning will be briefly discussed later: in section 1.3). The model should also be “gradual” in that, as discussed above, familiarisation occurs through a process of experience. According to Bruce and Burton (2002), there are stages of familiarity that should be included into a learning model to reflect the fact that some faces are better known than others. Finally is the notion that the model should be “cumulative”, that is, any new representation should not impede any previous representation given that we seem to be able to learn new faces without detriment to recognition of known faces. To achieve this, the model proposes an increased pool of FRUs and PINs that allows for the formation of new patterns which will become selected as the new most excitatory unit. Using a Hebbian update process the model is able to link the new pattern of activation with the specialised input. This new unit will now become active and strengthened for repeated presentations and thus arguably learn to recognise a new face.

Overcoming the fact that unfamiliar faces rely on pictorial codes rather than more structural based representations will be an important feature of any familiarisation training programme. This view specificity will be examined more thoroughly within the empirical work of Chapter 4. Moreover, the suggestion from the model of adult face learning is that representations of a face are built through a gradual and largely unsupervised experience with an individual. This is broadly consistent with aspects of perceptual learning theory (see section 1.3). The ability of any improvement based on perceptual learning to generalise to novel stimuli (e.g., that do not rely on pictorial cues) is addressed within Chapter 2.

1.2.5 Face space and the own race bias

The concept of a multidimensional “face space”, that was introduced by Valentine (1991), serves as a metaphor for how faces are thought to be represented. A face space is an internal construct which generates representations for previously stored or perceived faces. The idea of an abstract face space was initially introduced to account for some well-established factors in face perception, specifically the effect of distinctiveness on recognition (Light, Kayra-Stuart, & Hollander, 1979) and classification (Valentine & Bruce, 1986). There are several different interpretations of face space, however, the frameworks of these models comprise of a few central concepts which are constant despite other variations. Firstly, faces are defined as points within a multidimensional space. Secondly, the axes of the multidimensional space define a feature set in which individual faces can be encoded. That is, each face has a value on each feature axis; the combination of these feature values defines the position of a face within the space. Finally, the similarity of two faces in the space can be measured as the Euclidean distance between them (for a more detailed discussion, see O'Toole, 2011). Valentine suggests that an individual's life time experience will contribute to the distribution of faces within the space, and that the dimensions of the space are thought to be scaled through a process of perceptual learning based on the population of faces which have been experienced (Valentine, Chiroro, & Dixon, 1995).

Within these broad parameters, two families of face-space frameworks (exemplar and norm-based models) have been proposed as alternative accounts (Valentine & Endo, 1992). Exemplar based models of face space suggest that faces are encoded within a Euclidean space. The identity of a face is then calculated as a vector distance from that of others in the space. Typical faces are clustered relatively close together near the origin of the space producing a high exemplar density while distinctive faces are encoded further from the central tendency. It is worth noting that exemplar models are silent on how familiarity is gained. One way that this has been interpreted is that faces within an exemplar model have equal strength representations and exemplars only exist for faces that are familiar (e.g., Lewis 2004). As such, the move from unrecognisable (i.e., unfamiliar) to recognisable (i.e., familiar) must be thought of as an all-or-nothing process (c.f., Lewis, 2004). In contrast to exemplar models, norm-based models of face space (e.g., Valentine, 1991) encode face representations by a deviation on a number of dimensions from that of an average face. That is, each face within the model is encoded along a vector in relation to a norm or average face. This average face is based on an individual's prior experience with faces. While a number a recent studies have found evidence to consistent with a norm-based approach (Leopold, O'Toole, Vetter, & Blanz, 2001; Ross, Hancock, & Lewis, 2010), the original norm-based model still lacks any mechanism whereby faces move from typical (i.e., harder to discriminate) towards distinctive (i.e., easier to discriminate). That is, there is no mention of how a face moves within this space from unfamiliar to familiar or what information is extracted from a face in order to create these representations. Lewis' (2004) Face-space R model, adds a familiarity dimension to the representation of each individual face, which means that the strength of the representation increases in steps presumably due to exposure to that individual, while the location of an individual within the face-space remains the same.

Even though Lewis (2004) recognises the need to differentiate between familiar and unfamiliar faces, the means of doing so within his model says nothing about the underlying mechanism. This is unsurprising given how little is known about how face representations gain familiarity. While some authors have suggested that this occurs through perceptual learning (e.g., Valentine *et al.*, 1995), there is little elaboration of the processes that may be involved as opposed to a specification of the output of these presumed mechanisms. For example, perceptual learning has been cited in explanations of the own-race bias, but this is a reference to the effects of learning rather than the mechanism, as it is merely assumed that because experience shapes an individual's face-space, then the relative expertise of processing our own-race compared to other-race faces is a product of experiencing faces of our own-race more often (Bruce & Burton, 2002). This is clearly consistent with the idea that perceptual learning influences the discriminability of a face (Buttle & Raymond, 2003; Mundy *et al.*, 2007; Valentine *et al.*, 1995), but does not expand upon how perceptual learning produces such an effect. In a similar vein, the way that experience shapes the tendency for internal and external feature processing has also been attributed to perceptual learning. For example, Egyptian children show the same level of performance as the British adults on the internal feature matching task but Egyptian adults perform a face matching task using internal features more accurately. Egyptian adults are thought to have developed this proficiency from an extended cultural experience of women who frequently adopt headscarves thus only displaying internal face features (Megreya & Bindemann, 2009). This finding appears to point toward a general effect of experience without any implications regarding the mechanisms involved.

Despite the absence of a detailed discussion of perceptual learning mechanisms within the face processing literature, the general idea of a link between perceptual learning and the processing of faces is entirely reasonable due to the many studies that have demonstrated that experience with a face can lead to better discrimination (e.g., Stevenage, 1998). There are also a number of studies that have demonstrated that the way a face is processed is altered by experience and that as a face becomes more familiar, a shift in attention towards the internal features become more apparent (Bonner *et al.*, 2003; O'Donnell & Bruce, 2001; Osborne & Stevenage, 2008). The idea that exposure or experience can produce changes in an ability to perceive certain features of a stimuli is a basic premise of perceptual learning. However, while studies have reported changes in discrimination ability (Mundy *et al.*, 2007), and utilised them as the most effective means with which to improve upon an existing percept following a course of perceptual learning (Dwyer & Vladeanu, 2009), few have investigated how general this change in representation is. That is, the limits of how effective these methods are in modifying face processing have not been explored. The aim of Chapter 2 will be to examine if the perceptual learning literature can inform a training programme that successfully aids recognition of previously unfamiliar faces and whether this improvement can effectively generalise to nonexposed (novel) stimuli.

1.3 Perceptual Learning

Perceptual learning can be defined as a relatively long lasting change to an organism's perceptual system, which improves the ability to respond to their environment (Goldstone, 1998). There have been numerous demonstrations, across a variety of stimuli and species, that exposure both with and without explicit training will enhance discrimination between otherwise confusable stimuli (for reviews see, Goldstone, 1998; Hall, 1991). A familiar example comes from William James (1890) who discussed how learning might affect the ability to discriminate between wines. James suggested that learning to distinguish claret from burgundy could be aided, in part, by attaching different names to the wine (i.e., by providing feedback), as the different name associations would stretch the difference between the flavours and makes them more discriminable. While this illustrates a form of supervised perceptual learning that has been explained in terms of acquired distinctiveness of cues because of the feedback provided (e.g., Goldstone, 1998; Hall, 1991; Hall, Mitchell, Graham, & Lavis, 2003; Mackintosh, 1975; see section 1.3.3 for a more detailed explanation), it has also been shown that mere (unsupervised) exposure to two stimuli is sufficient to aid discriminability between them. For example, Gibson and Walk (1956) demonstrated perceptual learning using two groups of rats, the first group had metal shapes (e.g., circles and triangles) hung from home cages, the control group had ordinary home cages. It was found that when trained on a discrimination between the shapes, the group that had lived with the shapes in their home cages learned faster than controls. That is, the opportunity to simply experience the stimuli resulted in better discrimination at test. Moreover, this unsupervised learning has also been demonstrated using human subjects (e.g., Dwyer, Mundy, & Honey, 2011; Mundy *et al.*, 2006; Mundy *et al.*, 2007; McLaren, 1997; Wills, Suret, & McLaren, 2004), which is important given that face processing has also been thought rely on unsupervised learning (e.g., Burton, 1994). Indeed, if faces are learnt through individual experience, then that experience is likely to involve no external reinforcement.

The specific mechanisms involved in any observed improvement through mere exposure have been debated. Some associative models of learning have been used to explain the effects of unsupervised perceptual learning e.g., the elemental model proposed by McLaren and Mackintosh, (2000, 2002). Separate to this, some authors have also questioned the validity of this unsupervised learning (e.g., Mackintosh, 2009), which has led to the assertion that some instances of perceptual learning can be explained in terms of associative theories of reinforcement such as acquired equivalence or distinctiveness (see, Honey & Hall, 1989). Regardless of the type of mechanism attributed to this change in learning it should be noted that the schedule of exposure has been shown to affect the quality of learning (e.g., Dwyer *et al.*, 2004; Honey *et al.*, 1994).

1.3.1 The effect of schedule on exposure

The schedule by which stimuli are exposed also influences the development of perceptual learning over and above the amount of exposure. Intermixed exposure (i.e., AX, BX, AX, BX….) will result in superior subsequent discrimination between AX and BX than will the same amount of blocked exposure (i.e., AX, AX, …. BX, BX…), and in turn both forms of exposure will support superior discrimination than no pre-test exposure at all. This experimental finding was first demonstrated by Honey *et al.* (1994) in an experiment with domestic chicks, but has since been replicated in rats (e.g., Symonds & Hall, 1995), and numerous human experiments (e.g., Mitchell, Kadib *et al.*, 2008; Mundy *et al.*, 2006).

There have been various explanations for these schedule effects on the magnitude perceptual learning effect. One perspective suggests that during the intermixed exposure phase various associations will be established. In particular, excitatory associations develop between A and X and between B and X, such that, on AX trials X will retrieve a representation of B and, conversely, on BX trials X will retrieve a representation of A. As such, on AX trials A will signal the absence of B, and, equally, on BX trials B signals the absence of A. These conditions will therefore establish mutually inhibitory associations between A and B and thus reduce generalisation between AX and BX. However, if all AX presentations precede BX trials (e.g., a blocked exposure schedule) then B will only become a weak inhibitor of A (e.g., McLaren & Mackintosh, 2000). Alternatively, a more recent explanation by Hall (2003) suggests that the advantage of intermixed exposure is due to a process of habituation which leads to reduced salience of the elements which constitute these stimuli, due to the repeated presentation of AX and BX. In addition, an extra dimension of reverse dishabituation produces at least a partial restoration of the lost salience when the representation of the stimulus is retrieved but the stimulus itself is absent. Briefly put, the excitatory associations formed between A and X, and B and X, means that, during intermixed exposure the representation of B is retrieved on each AX trial and vice-versa for A on the BX trial, which partially restore the salience of A and B, but not of X because it is present on every trial.

A non-associative account proposed by Gibson (1969) suggested that the opportunity to compare two similar stimuli will enhance one's ability to discriminate between them. This is because it is a particularly effective means of drawing attention towards their unique differentiating features - a process of stimulus differentiation. Such a process is thought to be aided by the intermixed schedule, although Gibson remained silent on any possible mechanism which could underpin this process. However, while Gibson's own presentation of this idea left the mechanisms underpinning stimulus differentiation relatively undefined, there have been numerous subsequent attempts to unpack this process. For example the advantage of intermixed exposure (e.g., AX, BX, AX, BX) has been explained in terms of the effects of habituation on representation formation (e.g., Dwyer *et al.*, 2011; Honey *et al.*, 1994; Mundy *et al.*, 2007). The intermixed exposure schedule encourages greater habituation of X than A or B on each trial. That is, while the number of stimulus presentations within intermixed and blocked exposure is equated, the intervals between the various stimuli are not: in the intermixed condition the time elapsed between presentations of A (and B) is longer than those of X. Assuming habituation produces a loss of salience this means that intermixed exposure to AX and BX will result in X losing salience faster than A or B. In turn, if X is less salient than A or B, then attention will be drawn towards A and B, with the consequence that these differentiating features will dominate the stored representations of AX and BX.

1.3.2 Gibson's theory of perceptual learning

Gibson suggested that perceptual learning is the product of learning to extract readily available information from a given structure (Gibson, 1969). That is, perceptual learning is a process whereby an individual's experience and practice with an environment increases its ability to extract information from it. Gibson saw this process of learning as an active process of search and exploration without the need of external reinforcement. Key to this process is the discovery of distinctive features within a given structure which change the response to a set of stimuli. Importantly, this change in response is an increase in specificity (i.e., the detection of properties, patterns or distinctive features) and not an addition or subtraction of any previous information, which allow them to be discriminated.

Gibson (1969) suggests that perceptual learning is due to a process of comparison-driven stimulus differentiation whereby exposure to the stimuli enhanced the effectiveness of the unique features (which distinguish similar stimuli) relative to the common features (which do not). Gibson's suggestion implies that perceptual learning will result in the salience of, and/or the attention towards, the unique features A and B being greater than for the common features X. These elements are not associated with any outcome given that the organism is simply exploring the stimulus via a process of active search. Moreover, there is no active incentive to search for the distinctive features, and no external rewards. Therefore, this form of unsupervised perceptual learning is not dependent on any external correction of responses, but of a self-regulated adaption through search and observation. That results in a better understanding or representation of the unique features which are diagnostic of a stimulus. The active process of searching and exploring a stimulus without the need for external reinforcement that Gibson advances seems to link well to the notion that faces are generally learned without any obvious supervised learning process. Moreover the suggestion that the internal features become more important as a face becomes more familiar (e.g., Ellis *et al.*, 1979) might reflect the enhancement of unique features as advocated by Gibson. Conversely, if perceptual learning does influence face processing then it is not unreasonable to assume that, in some cases, attention to location is underpinning some discrimination improvements in perceptual learning. That is, if attention to the internal and external features of faces changes with a function of their familiarity, then perhaps this indicates a process whereby exposure produces an increase in the amount of attention to the location of the critical stimulus features.

1.3.3 Goldstone's review of perceptual learning

Goldstone's (1998) review of perceptual learning identifies a number of different mechanisms which contribute to perceptual learning. Through this approach, Goldstone advocates three mechanisms that mediate perceptual learning between the external world and cognition, via an extensive review across a broad literature spanning neuroscience, cross cultural studies, and importantly face processing. These mechanisms include attentional weighing, differentiation and unitisation, demonstrating the broad application of perceptual learning. Goldstone also acknowledges the distinction between training mechanisms which require feedback of information to the organism (supervised training); and those which require no feedback (unsupervised training), operating instead on the "statistical structure inherent in the environmentally supplied stimuli" (Goldstone, 1998, p. 588). Moreover, perceptual learning mechanisms can also be labelled quick or more time-consuming, and more generally distinguished between strategic or peripheral adaptations.

Goldstone (1998) suggests that attentional weighing is a mechanism for perceptual learning, and that by increasing attention to the important features of a stimulus, and reducing attention to those features or dimensions which are less informative, allows for precepts to adapt to tasks or environments. Crucially, during different stages in information processing, attention can be selectively directed towards important stimulus features. Evidence suggests that dimensions that are particularly useful for tasks are often the focus of a shift in attention (Lawrence, 1949). These situations have been described as acquiring distinctiveness if they have been indicative of different outcomes. This has also been described as psychologically stretching dimensions useful in categorisation (Nosofsky, 1986). Alternatively, the irrelevant dimensions acquire equivalence, that is, they become less distinguishable by virtue of signalling the same outcome (e.g., Honey & Hall, 1989). These shifts have been described as more of a strategic choice rather than perceptual in nature (Goldstone, 1998; but see, Honey & Hall, 1989, 1991 for an alternative view). However, it is worth noting that these shifts are not always completely voluntary such that it has been found that attentional highlighting can be detrimental to the observer. For example, it has been shown that when letter consistency was switched from target to a distractor attention is still captured (Shiffrin & Schneider, 1977).

Perceptual adaptations through attentional weighing are exhibited within categorical perception. That is, it is easier to discriminate between two stimuli that straddle a category boundary than between two stimuli belonging to the same perceptual category. For instance, the finding that familiar faces exhibit these effects between two faces which straddle between two identities than two belonging to the same identity (Beale & Keil, 1995). It has also been suggested that individuals are better able to distinguish between physically different stimuli, when the stimuli come from different categories, rather than the same category of origin (Calder, Young, Perrett, Etcoff, & Rowland, 1996). Goldstone (1994; 1995) has also indicated that category differences can be enhanced through training, and that there are three influences on category learning. Firstly, those dimensions relevant to particular category are sensitized. Secondly, variations which are not relevant are deemphasised. And finally, as outlined, dimensions that are relevant are selected and sensitised at the category boundary. Although not as detailed in her description, Gibson proposed a similar process in the form of "selective attention" (see, Gibson, 1969 p. 115). This relates to the exploratory activity, which direct attention towards a stimulus. For example, the exploratory activity of the visual system includes fixating the eyes, scanning and head turning. Clearly this action is needed for selective perception, e.g., direction of gaze towards a unique stimulus element. Critically this selective attention is thought to expose stimuli receptors "to chosen aspects of potential stimulation" (Gibson 1969 p.115). A process which Gibson describes as vital for selecting and rejecting stimulus input. This is an active process, adaptively selecting or rejecting areas of interest, what is salient from the mass of information available. This mechanism of perceptual learning may indicate, at least in visual tasks, a process whereby areas of interest (locations) are utilised – a possibility that is explored further in Chapter 3. There are certainly many studies within the face processing literature that suggest attending to the most critical locations (e.g., the internal features) is a result of some familiarity with a face and therefore that familiarity is underpinned by some attentional improvement through learning. That is, when a face is familiar then the areas of interest which are the most critical in identification are the internal features (O'Donnell & Bruce, 2001) and that a possible explanation for our poor performance with unfamiliar faces is that this diagnostic information is not being attended. The idea that attention to location underpins some discrimination improvements within perceptual learning will be examined further in this Chapter and in Chapter 3 (see sections 1.3.4 and 1.3.5 below).

Another mechanism of perceptual learning is differentiation. According to Goldstone (1998) this occurs when percept's become increasingly differentiated from each other. Thus, stimuli that were once psychologically combined become separated to enable originally indistinguishable discriminations to be made between each percept. Accordingly, it has been suggested that simple pre-exposure to a stimulus facilitates differentiation. This process is achieved by promoting separate processing of individual features within a stimulus.

It is worth noting that while differentiation requires that wholes are separated into different parts, another mechanism, outlined by Goldstone, is that of unitization, which operates in direct opposition. Via unitization, stimuli that initially require several elements to be detected can be accomplished by detecting a single entity. Therefore, if differentiation leads to separate percepts, unitization is a complete integrated view of the singular parts. The idea that experience can chunk separate elements into a single unit (unitization) and conversely separate single units into separate elements (differentiation) would expectedly produce results which directly oppose each other. Importantly, Goldstone (1998) addresses the idea that differentiation and unitization could be considered mutually inconsistent by proposing that each mechanism is recruited depending on the requirements of either the stimulus or task. Thus, differentiation occurs if the parts represent variations relevant to the discrimination, while unitisation will take place when separate parts co-occur and often require the same response.

Differentiation has been studied across many types of stimuli. Indeed, the expertise in face processing has been attributed to some form of differentiation. For example, the own-race bias whereby identifying faces belonging to the same race as the participant is easier than with faces from other races (Shapiro & Penrod, 1986). The difficulty in recognising faces of other-races has been attributed to the features of commonly experienced faces not corresponding to the distinctive aspects of other race faces (O'Toole, Peterson, & Deffenbacher, 1996). This effect can be considered as a process whereby familiar objects undergo differentiation. Moreover, it has been suggested that entire categories can become differentiated. According to Goldstone (1996), differentiation relies on the developing qualities that uniquely identify novel objects from familiar objects. For example, experts can categorise domain specific objects at basic and subordinate levels faster than novices (e.g., McLaren, 1997; Tanaka & Taylor, 1991). Indeed, O'Toole *et al.* (1996) has shown that Japanese and white participants are quicker at making gender discriminations for individuals from their own race.

1.3.4 Vernier discrimination and transfer

Training can also induce differentiation in psychophysical tasks. For example, vernier discrimination tasks which require participants to respond to one line being displaced below or above a second line can display large improvements from the baseline level, to the level finer than the spacing between photoreceptors following training (e.g., Poggio, Fahle, & Edelman, 1992). A trademark of these acuity tasks and visual perceptual learning has established that, at least, some of what is learned is stimulus specific or to retinal location (e.g., Fahle, Edelman, & Poggio, 1995; Poggio *et al.*, 1992; Shiu & Pashler, 1992). However recent evidence has suggested that the degree of specificity can be affected by the amount of training (Jeter, Dosher, Liu, & Lu, 2010), task difficulty (Ahissar & Hochstein, 1997), and type of training schedule (Xiao *et al.*, 2008). For example, Xiao *et al.* (2008) examined whether learning to discriminate a specific stimulus features (feature learning) is separate from learning of location factors (location learning), and demonstrated that additional location training enabled feature learning to transfer to a new location (Xiao *et al.*, 2008).

While these findings are somewhat restricted by the fact that these visual hyperacuity experiments appear to rely on mechanisms largely centred on the sensory cortex that are separate from the general mechanisms involved in associative learning, they still highlight the potential importance of location in a given stimulus. In contrast, more general processes of perceptual learning are typically investigated using complex stimuli, and traditional associative learning mechanisms focus on changes in content rather than location (see below). Critically Xiao *et al*. (2008) highlights the idea of attention to location could underpin some discrimination improvements in perceptual learning. Certainly, the experiments of Xiao *et al*. (2008), and parts of the accounts suggested by Gibson (1969) (e.g., selective attention) and Goldstone (1994) (e.g., selective sensitisation in category learning), allude to the importance of location. Chapter 3 will explore, in detail, the role of location within certain stimuli used to explore the mechanisms of human perceptual learning.

1.3.5 Associative explanations of perceptual learning

The above sections suggest that perceptual learning is a result of discovering unique features of a stimulus which alter the perception of a given stimuli. Thus, perceptual learning is considered as a process of discovery of the unique features that leads to an adjustment in the representation of the given stimulus. The mechanisms of this change have been attributed to early perceptual processes such as comparison driven differentiation (Gibson, 1969), or perceptual differentiation such as that described by Goldstone (1998). Both of these mechanisms suggest that stimuli can be broken down in order to identify and separate the diagnostic and common features between two stimuli. However, once a stimulus can be broken down into constituent elements, then there is the possibility of learning about the relationships between these elements through associative processes. At its most general, associative learning suggests that when two stimuli co-occur they form an association (a directional connection) such that, the presentation of one stimulus can excite the representation of the other. Such principles have been used to examine how the existence of an association can allow the presence or absence of one element to modify the state of another in a manner that might produce perceptual learning effects (see, Hall, 1991).

One strategy used in the investigation of perceptual learning, and of the mechanisms which underlie it, has been to use stimuli that afford the direct manipulation of their constituent elements. For example, studies of perceptual learning in rats have often used compound flavours where a common element (e.g., sucrose, X) is combined with one of two unique elements (e.g., salt/lemon, A/B) (e.g., Mackintosh, Kaye, & Bennett, 1991). This allows testing of the elements alone as a means of assessing the effects of perceptual learning on those elements themselves, and has provided evidence for both changes in stimulus salience (e.g., Blair & Hall, 2003; Blair, Wilkinson, & Hall, 2004), and for the development of mutual inhibition between the unique elements of stimuli presented in alternation (e.g., Dwyer, Bennett, Mackintosh, 2001; Dwyer & Mackintosh, 2002). The same strategy has also been used in the study of perceptual learning with complex visual stimuli in humans. That is, using checkerboards consisting of a common background X and unique features (e.g., A/B/C etc…). These visual stimuli also allow for the separate analysis of the unique stimuli, and such analyses have suggested, amongst other things, that intermixed exposure to such checkerboards results in better memory for the unique features than does blocked exposure (Lavis, Kadib, Mitchell, & Hall, 2011) (see Chapter 3).

Associative explanations of these effects suggest that the results of these studies can be explained in terms of associations forming between the unique and common elements. For example, the elemental model of McLaren and Mackintosh (2000) was designed specifically as an account of perceptual learning. Put briefly, the model instantiates three particular associative mechanisms: Latent inhibition of common features, mutual inhibition between unique features, and unitization. Consider two similar stimuli, AX and BX (where A and B refer to their unique elements and X to the elements they have in common). Latent inhibition is a phenomenon in which conditioning procedures produce a detriment to the rate of learning (Lubow, 1973). If participants are given exposure to AX and BX, the X elements will be present on twice the amount of trials leading to a greater opportunity for X to be subject to latent inhibition (see also Hall, 2008). Further, McLaren and Mackintosh (2002) describe the notion that exposure to compounds AX and BX will result in elements comprising both A, and B predicting the elements containing X, but only the elements comprising A will be reliably predicted by other elements from A (and only the elements B will be reliably predicted by other elements from B). Thus, to the extent that latent inhibition depends on the degree to which a stimulus is predicted, X elements will be affected more than elements in either A or B. As described in section 1.3.1, this account also instantiates the idea that intermixed exposure to AX and BX should result in mutual inhibition between the unique features A and B (e.g., McLaren, Kaye, & Mackintosh, 1989; McLaren & Mackintosh, 2000). Finally unitization, whereby individual elements become associated and thus the activation of a subset of elements will allow the whole representation to be activated (McLaren & Mackintosh, 2000).

As noted previously, other associative accounts that have been described in relation to the intermixed-blocked effect (see section 1.3.1) include Hall's (2003) suggestion that in the absence of any significant consequences, direct activation of a memory of a stimulus leads to a reduction in its salience. However, in the absence of the stimulus, the representation will recover previously lost salience through associative activation. Additionally, Mundy *et al.* (2007) suggest that comparison driven differentiation is the result of short term adaptation or habituation and Mitchell *et al.* (2008) assert that higher level memory mechanisms are involved in the representation of a stimulus.

Although there are important differences between these accounts, they all consider only the changes to the content of the unique elements and neglect the possible importance of location. That is, all of the above accounts consider changes through activation of the unique features relative to their common features without considering the role of attentional processes towards the location of the unique feature. However, as has been noted above (section 1.2.2), location is important for the perception of faces and in at least some instances of perceptual learning. Moreover, many studies of the associative analyses of perceptual learning use stimuli whereby the unique features A/B always appear in the same place on a common background X, and so any exposure-dependent influence on the discriminability of AX and BX might reflect learning about the content of those unique features (e.g., a learnt change in their salience) or about their location (e.g., learning where to look for discriminating features) (e.g., Lavis & Mitchell, 2006; Mitchell, Kadib, *et al.,* 2008; Wang & Mitchell, 2011). It is important to note that other studies that have investigated the general mechanisms implicated by these theories may not all be susceptible to idea of location based learning. For example, studies that have used stimuli such as flavours that obviously cannot be discriminated on location (e.g., Blair & Hall, 2003; Dwyer & Honey, 2007; Symonds & Hall, 1995). Despite this, it remains a possiblity that location is central to at least some demonstrations of visual perceptual learning. If this is the case, then this directly challenges the associative accounts which focus solely on content of the unique element, and also the interpretation of studies examining the transfer of learning from one situation to another. For this reason, Chapter 3 will examine this issue using checkerboards. This choice of stimulus is a function of how this type of stimulus is constructed in that unlike faces they allow for direct manipulation of critical features. That is, faces have a distinct configuration, thus none of the critical features can be moved without disrupting this basic pattern (e.g., the eyes must sit above the nose etc).

1.4 Perceptual Learning and Faces

As noted earlier, perceptual learning processes have been identified as contributing to face processing, this possibility is reinforced by the fact that many studies which have examined the intermixed blocked effect have utilised morphed face stimuli. While examining the mechanisms behind the intermixed blocked effect Mundy *et al.* (2007) found that, alternating or simultaneous exposure of different faces was more effective at facilitating discrimination compared to presentations of each face in separate blocks. This pattern of results has been supported by other investigations into perceptual learning that have demonstrated that unsupervised, brief, intermixed exposure can produce key characteristics of familiar face processing during subsequent tests (see Dwyer *et al.*, 2009; Dwyer & Vladeanu, 2009). Moreover, these improvements, taken alongside other findings of perceptual learning, suggest that general mechanisms of perceptual learning are applicable to face processing (Gold *et al.*, 1999a, 1999b; Gold *et al.*, 2004).

In all of the perceptual learning studies cited in the previous paragraph, the key measure of performance was the ability to distinguish between particular stimuli that were presented during both exposure and test. Critically, in none of these cases was the ability to distinguish between an exposed stimulus and a novel foil compared to the ability to distinguish between two novel stimuli. Indeed, there is some evidence that implies comparison does not facilitate discrimination between exposed and nonexposed flavour stimuli in rats (e.g., Blair & Hall, 2003). In the context of face processing, this means that the intermixed exposure schedule may only help to distinguish between two people from that exposure schedule and not between the learned face and a member of the general population. If true, this would limit the forensic application of perceptual learning to learning faces.

Another issue with applying perceptual learning to face recognition is that unfamiliar face recognition is image-based (e.g., Bruce, 1982; Longmore *et al.*, 2008). As such, it is sensitive to changes in lighting (Hill & Bruce, 1996; Kemp, 1996), expression (Bruce, 1982), and viewpoint (Hill, Schyns, & Akamatsu, 1997; for a review see Johnston & Edmonds, 2009). It is difficult to gauge, therefore, whether any effect of comparison can reflect image-based learning or a more general improvement in recognition. Thus, before any potential forensic application of stimulus differentiation can be assessed, it is important to ascertain whether comparison improves recognition of a stimulus *per se*, or whether it only improves performance with the particular images that were presented during the comparison process.

There is, however, some recent evidence that suggests that it is possible to generalise comparison-based learning to a novel stimulus. Dwyer and Vladeanu (2009) reported that matching performance, using artificially generated faces, can be facilitated by alternating a target face with similar comparators during exposure. The improvement in matching of faces was attributed to a process of stimulus differentiation, as discussed in the literature on perceptual learning (section 1.3.2). That is, Dwyer and Vladeanu (2009) suggested that the similar comparators and the target face shared a set of common features, and that adaptation of those common characteristics by exposure to the similar comparators would have enhanced the unique features of the target face, thus facilitating subsequent matching performance.

Although the findings of Dwyer and Vladeanu (2009) certainly seem to support the idea that the enhancement produced by comparison might extend to novel test stimuli, there are some caveats which limit the interpretation and potential generality of their findings. The first relates to the stimulus set used in the experiments. Artificially generated faces were used, and therefore any improvements may be limited to such stimuli. Indeed, if we briefly consider the face-space metaphor whereby face representations are locations within a multidimensional psychological space (Valentine, 1991), then the artificial nature of stimuli used in Dwyer and Vladeanu (2009) means that they may not have occupied any of this higher level space at all. That is, they may not have been truly processed as faces (although there is no particular reason to think that this would be the case). Alternatively, even if the stimulus set were processed as faces, then it may be that the distinctive nature of the stimulus set meant that they occupied a proportion of face space which made them more recognisable. Certainly some models of face-space would predict this outcome, given that the distinctiveness of a face is thought to refer to the distance between the representation and the average (Burton, 1994; see Lewis, 2004 for a discussion). That is, typical faces suffer from misidentification because these representations occupy a dense proportion of face space, whereas distinctive faces are in a less populated area and are easier to recognise, because there is less competition from other surrounding face representations. Furthermore, empirical evidence supports this notion, suggesting that the more distinctive a face then the more easily recognised it is (Light *et al.*, 1979). It has been argued that this is because they provide unusual cues that may encourage a more in-depth processing strategy (Fleishman, Buckley, & Klosinsky, 1976; Shapiro & Penrod, 1986). Secondly, Dwyer and Valdeanu (2009) generated the non-target comparators used in the exposure phase, and the novel test foils used in the test phase, in exactly the same fashion – both similar comparators and test foils were created by morphing away from the target face. As such, all non-targets will have the same common features, and so the ability for comparison-facilitated exposure to generalise to novel test stimuli might be restricted to situations in which the same common features are present in all stimuli – an assumption that may not be true for real faces.

To summarise, it is unclear whether the beneficial effects of comparison will be seen using real photographic stimuli, and if this effect can genuinely transfer to novel test stimuli in a manner that makes it a useful applied practice.

1.5 Unanswered Questions

Questions remain within the field of face perception and perceptual learning that are of both theoretical and forensic interest. I have identified three main areas that require further empirical work. First, whether perceptual-learning-type training can contribute to improving recognition of an unfamiliar target that can generalise to novel nonexposed stimuli (Chapter 2). The second aim to explore the role of location and content based learning within complex visual checkerboard stimuli. That is, the perfect correlation between the content of a unique feature (e.g., its colour or shape) and its location (i.e., where it appears on the background X) will be broken, in order to ascertain the relative contributions of these aspects of the unique stimuli to the learning effects observed (Chapter 3). The third and final goal of the thesis is to examine the benefits of using synthetic views to compensate for the view-dependent nature of unfamiliar recognition. That is, whether technology could be used to support human face recognition within a tractable training schedule, and if this assistance can manage this transition between viewpoints above that of our own general level of performance (Chapter 4).

Chapter 2: Exposure training with faces

2.0 Abstract

Chapter 2 reports four experiments examining whether comparison between one set of stimuli can facilitate the discrimination of these exposed target faces from other nonexposed faces. Experiment 1 found that selection of a target from an array of novel foils was not facilitated by intermixed exposure to the target face and similar comparators (defined as images of the same gender). Experiment 2 also found no advantage for similar comparators (morphed towards the target) over dissimilar comparators, or repeated target exposure alone. But all repeated exposure conditions produced better performance than a brief presentation of the target. Experiment 3 again demonstrated that repeated exposure produced equivalent learning in Similar and Dissimilar conditions, and also showed that varying the number of similar or dissimilar comparators failed to improve recognition. Experiment 4 demonstrated no difference between internal feature presentation and whole face presentations within a Similar or Dissimilar exposure schedule. In all four experiments exposure to a target alongside similar comparators failed to support selection of the target from novel test stimuli to a greater degree than exposure alongside dissimilar comparators or repeated target exposure alone. The current results suggest that the facilitatory effects of comparison during exposure are limited to improving discrimination between the exposed stimuli and thus that comparison is not an effective means for improving the general recognition of faces.

2.1 Introduction

One of the themes explored in Chapter 1 was that experimental exposure to a face can improve recognition of unfamiliar individuals and that some unspecified process of perceptual learning may underpin how we learn new faces. Gibson (1969) suggests that perceptual learning operates through a process of comparison, and that the opportunity to compare two stimuli (AX and BX) allows for a process of stimulus differentiation by which attention is increased for the unique features of a given stimulus (A and B) compared to the common features (X). This explanation of perceptual learning is consistent with findings demonstrating that an intermixed schedule of exposure (e.g., AX, BX, AX, BX) leads to better discrimination than blocked exposure (e.g., AX, AX, BX, BX) (e.g., Honey *et al*., 1994; Mundy *et al*., 2006; Symonds & Hall, 1995). This effect has also been demonstrated using faces as stimuli (e.g., Dwyer *et al*., 2009; Mundy *et al*., 2007), leading some authors to suggest that utilising such schedules may be an effective way to improve recognition (Dwyer & Valdeanu, 2009).

However, previous experiments have failed to observe any benefit of comparison between exposed and nonexposed stimuli in rats (e.g., Blair & Hall, 2003). Moreover, the single demonstration of a generalizable effect to nonexposed faces by Dwyer and Vladeanu (2009) has been questioned (see section 1.4). Thus, it is unclear whether the beneficial effects of comparison will be seen using real photographic stimuli and if this effect can genuinely transfer to novel test stimuli in a manner that makes it a useful applied practice.

The main aim of the current experiments was to explore the conditions by which the effect of comparison can be applied to learning a previously unfamiliar individual and to explore the ability of any such learning to generalise to novel test stimuli. All experiments involved an exposure period whereby a target individual was familiarised before a recognition task. The task involved participants identifying the target from a line-up of faces (henceforth known as arrays). Arrays always displayed the target and previously unseen foils. To ensure any transfer was not merely based on the recognition of particular images, the faces displayed in the test arrays were always subject to expression and/or contrast changes.

2.2 Experiment 1

The main purpose of Experiment 1 (see Table 1 for the full design) was to examine whether alternating a target face with four other similar comparison faces (i.e., same gender) during exposure affected participant's ability to select that target from a test array of novel foil faces (This corresponds to condition Similar shown in Table 1). Control conditions comprised of the following: the target faces alternated with dissimilar comparison faces (different gender – Dissimilar condition); the target faces presented in alternation with a fixation cross (No-Comparator condition), and target faces presented for one extended period equalling the total exposure time for the other conditions (Single exposure condition).

In addition, the extent to which the similarity between the target and the test foils influenced performance was also examined by manipulating whether the test foils were created by morphing from the target face to novel faces (see Panel A of Figure 1), or by using novel faces as foils (see Panel B of Figure 1). If comparison to similar faces does improve processing of a target face, as suggested by Dwyer and Vladeanu (2009), then it is expected that targets alternating with similar comparators will be better recognised than targets in other conditions.

2.2.1 Method

Participants

Thirty-two participants between the ages of 18-34, were recruited from the School of Psychology at Cardiff University. Participants received course credit in return for their participation. All had normal or corrected-to-normal vision.

Apparatus and Stimuli

Eight white faces, four male and four female, were randomly selected from the Centre for Vital Longevity Face Database (Minear & Park, 2004), to be targets. Faces were between the ages of 19 and 45 years old and were selected to avoid the presence of non-face cues (e.g., glasses and facial hair). Each target was grouped with twelve faces of the same gender. Four became similar comparators, and eight were used as test array foils. Dissimilar comparators were the opposite gender to the target face. External features of all faces were removed by applying a mask function in Adobe Photoshop™ image editing software. Images were displayed centrally on a white background at 360 × 504 pixels subtending to an approximate visual angle of 12.5° × 17.2°.

All test arrays were constructed from faces of the same gender as the target. Two types of test array were used, both displayed the internal features of five individual faces. The first type was a Morph array, in which the stimuli consisted of the target plus four morphed faces (Panel A of Figure 1). The second type was a Non-Morph array, in which the target face and four other non-morphed faces all displaying happy expressions were presented (Panel B of Figure 1). None of the foils in either array had been seen in any exposure condition. Faces in the Morph array were created using a software package called Morpheus v3.10 professional™, by morphing four previously unseen individuals away from the target face. Morphs were 50% blends of the target face and a non-target face.

Table 1. Design of Experiment 1

Condition	Training sequence	Test Arrays (Morph)	Test Array (Non-Morph)
Single exposure	A, fp	Select A from an array of A, A5, A6, A7, A8	Select AH from an array of AH, AH9, AH10, AH11, AH12
No-Comparator	B, fp, B, fp, B, fp, B, fp. (× 4)	Select B from an array of B, B5, B6, B7, B8	Select BH from an array of BH, BH9, BH10, BH11, BH12
Similar	C, C1, C, C2, C, C3, C, C4. (× 4)	Select C from an array of C, C5, C6, C7, C8	Select CH from an array of CH, CH9, CH10, CH11, CH12
Dissimilar	D, X1, D, X2, D, X3, D, X4. (× 4)	Select D from an array of D, D5, D6, D7, D8	Select DH from an array of DH, DH9, DH10, DH11, DH12

Note: A-D indicates target faces (these were counterbalanced across conditions as described in section 2.1.1). 1–4 refer to comparator faces (e.g., C1-C4 in indicate similar comparator faces to target C, while X1 to X4 illustrate dissimilar comparator faces to target D). Exposure was repeated four times as indicated by × 4. fp denotes fixation points which were represented by a cross on screen. 5-8 in the Morph array indicate previously unseen faces used as foils (e.g., A5 to A8 are morphed foils for target A). AH-DH represents target individuals with an expression change in the Non-Morph array. 9-12 represent unseen foils present in the Non-Morph array (e.g., BH9 to BH12 are non-morphed foils for target B).

Both types of array were displayed on-screen at 1029 × 367 pixels subtending an approximate visual angle of 32.4° × 12.8°. The stimuli were presented centrally on a 17 inch monitor. A custom-written programme using DirectRT software (Version 2008.1.13; Empirisoft, New York, NY) was used to control the presentation of the stimuli on a PC. Responses were registered using a computer keyboard with QWERTY layout. For the experiment, the letters A, S, D, F and G were covered with coloured labels A, B, C, D and E to match the letter depicting each face in the test array.

Design and Procedure

Participants completed all four conditions (Single, No-comparator, Similar, and Dissimilar: see Table 1) in a within-subject design. Single exposure gave one presentation of a target face which remained on screen for thirty-two seconds (this was time-matched to the total amount of exposure time in the repeated exposure conditions). The repeated conditions (e.g., No-comparator, Similar and Dissimilar conditions) presented each target a total of sixteen times for two seconds each. For the Similar and Dissimilar conditions, each target had four comparator faces; each comparator was presented four times per condition, equating to a total of sixteen comparator presentations. Following exposure, participants completed a recognition task with both the Morphed and Non-Morphed arrays. This task was accompanied by a measure asking participants how confident they felt about their selection. There were two repetitions of each condition (with different target faces) before moving on to the next condition. Participants were sat approximately 70 cm from the computer screen. A brief practice trial then gave an opportunity for participants to be familiarised with the general procedure of exposure and test phases.

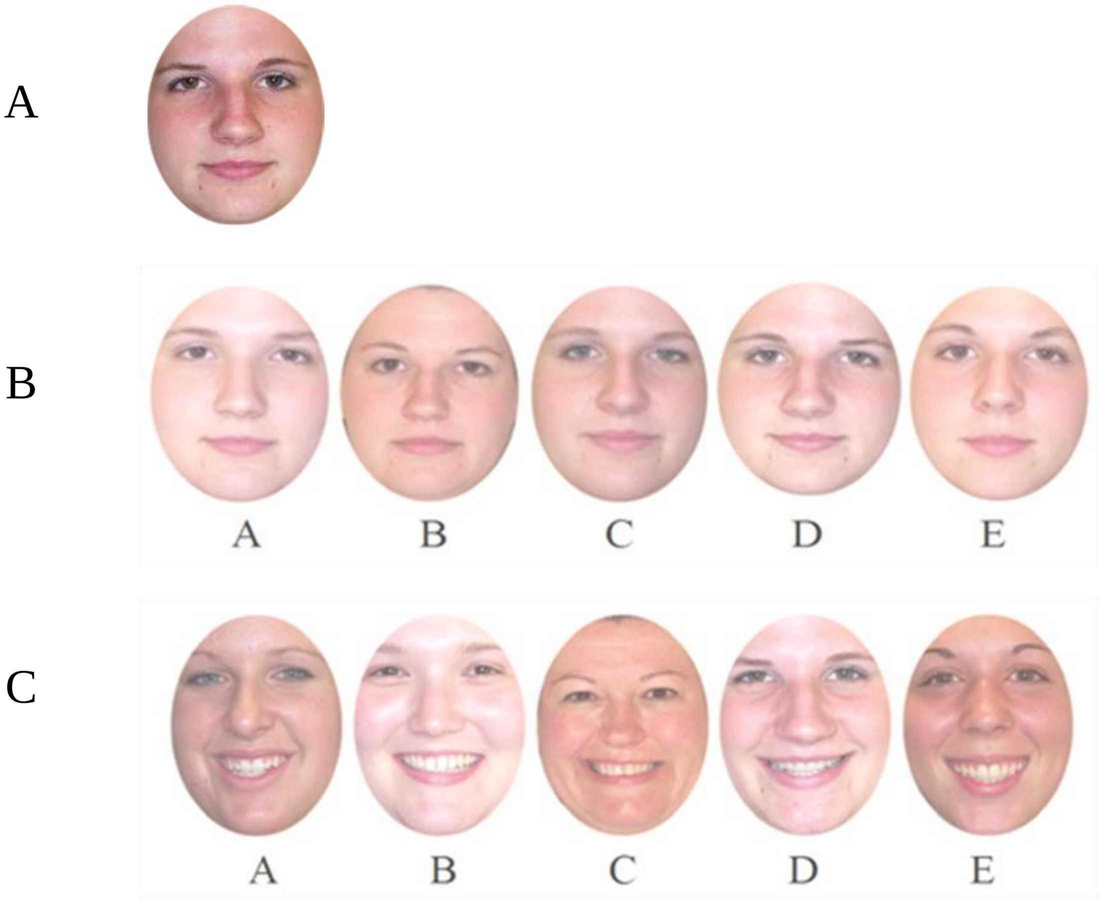

Figure 1: Examples of stimuli used in Experiments 1, 2, and 4. Panel A displays an example of the target face presented during training exposure. Panel B represents the Morph arrays associated with the example target face. Panel C represents the Non-Morph arrays associated with the example target. Participants in Experiment 1 and 2 were given a 3s presentation of an array and then asked to choose the letter that corresponded to the target seen during exposure.

During the exposure phase, a randomly assigned name such as “Matthew” was presented under each target face, while comparators for this target were labelled “Not Matthew”. Names were used to facilitate instructions to participants who were asked to identify named individuals during the recognition task (e.g., “*Please select the target face Matthew from the following array*”). Participants were then given an array of faces which were displayed for 3s. Following this array exposure participants were asked to make their identification. Confidence was then measured using a button response to a 7-point Likert scale (1: “Not at all confident”, 7: “Extremely confident”). The assignment of faces to condition was counterbalanced so that all faces were presented an equal number of times within each condition. This was implemented using a Latin square manipulation. Furthermore, the presentation of arrays were counterbalanced such that half the participants received test trials with the Morph array first. The other half of the participants saw the Non-Morph array first. Moreover, the order of exposure conditions was also counterbalanced such that each condition was presented equally often first, second, third or fourth across participants.

Data Analysis

The primary measure used in all the experiments in this chapter was the accuracy of recognition. In addition, a Confidence-Accuracy (CA) measure was calculated by multiplying accuracy (negatively scored for incorrect answers so 1 = Correct and -1 = Incorrect) by the confidence score (less 0.5). This gives a score between -6.5 and +6.5 in 13 equal steps. This CA score highlights the fact that a highly confident incorrect answer demonstrates worse performance than low confidence incorrect answers while highly confident correct answers represent the best performance.

Standard null-hypothesis significance testing only assesses how unlikely the observed data are given the assumption of the null hypothesis. As such, it does not provide a direct assessment of whether the absence of a significant difference can be taken as positive evidence for there being no true difference between conditions. In contrast, Bayesian tests are based on calculating the relative probability of the null and alternative hypotheses, and thus afford the assessment of whether the evidence is in favour of either of these hypotheses. As will be seen below, a number of inferences from my experimental data rely on the absence of a difference between conditions. Therefore, Bayesian analyses were also conducted as a means of assessing the strength of empirical support *for* the hypothesis that two conditions do not differ (Rouder, Speckman, Sun, Morey and Iverson, 2009). The Bayes factor (denoted as B_{01}) relates to the probability that the null is true to the probability that the alternative is true given the data observed. The calculation of the Bayes factor requires the specification of an effect size for the alternate hypothesis (although the exact value has relatively little influence on the output of the calculations (see Rouder *et al*., 2009)). In the absence of any directly comparable and reliable demonstrations of comparison that could be used to set an effect size the analyses reported here were based on the specified effect size of one standard deviation between the treatment and control means (as suggested by Rouder *et al*. (2009)). While it could be argued that the effect of repeated exposure compared to brief exposure is well established in the experiments that are reported here (the most comprehensive analysis is demonstrated by Experiment 3), using this effect size would give a less conservative estimate that the suggested default. Therefore, utilising the suggested default across all experiments reported in this chapter gave an estimate of whether the absence of a difference between two conditions genuinely supported the conclusion that there was indeed no effect (Rouder *et al*., 2009). The analysis was performed using the web-based calculator (http://pcl.missouri.edu/) and utilised the Jeffreys-Zellner-Siow (JZS) prior suggested by Rouder *et al*. (2009). As of yet, there are no published algorithms for factorial ANOVA procedures, thus key comparisons reported below have been reduced, where appropriate, to paired or unpaired Bayes t-tests, equivalent to the comparisons between two groups or conditions. The resulting

Bayes factors can then be interpreted as either supporting the null or alternative (or as inconclusive evidence for either) according to the convention suggested by Jeffreys (1961) and recommended by Rouder *et al*. (2009): a Bayes factor over 3 suggests there is some evidence to support the null hypothesis, while a factor of 10 indicates strong evidence for the null. On the other hand, a factor less than 1/3 suggests some evidence for the alternative and less than 1/10 indicates strong evidence favouring the alternative. Any value between 1/3 and 3 constitutes inconclusive evidence in support of either the null or alternative.

2.2.2 Results and Discussion

Figure 2 shows the test data as percentage correct as a function of exposure condition (Single exposure, No-Comparator, Similar exposure and Dissimilar exposure), and test type (Morph and Non-Morph). It is noticeable that there was little, if any, improvement for identification when target stimuli were exposed with similar comparators. Thus, learning appeared to be unaffected by the opportunity for comparison between targets and similar faces.

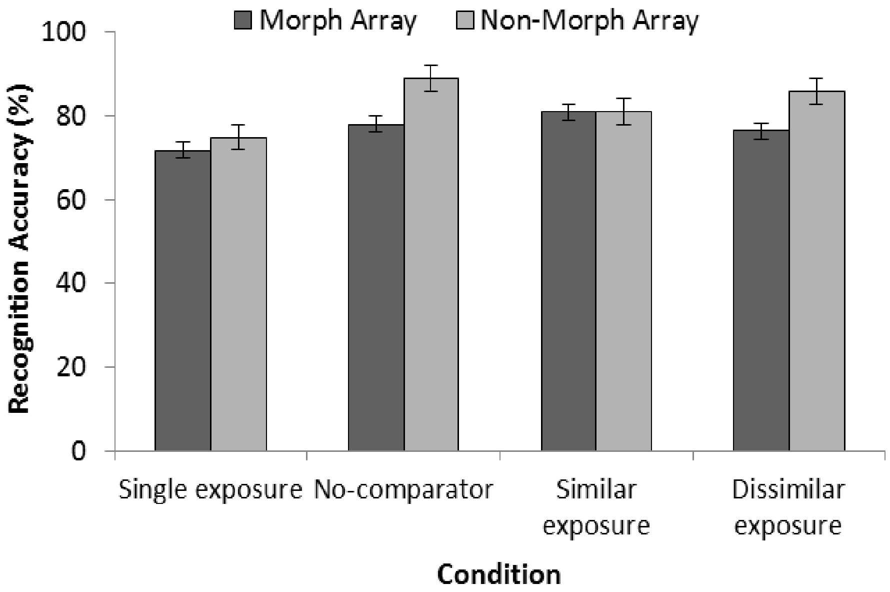

Figure 2: Recognition accuracy as percentage correct, with Standard Error Mean (SEM) from Experiment 1. Data are organised by exposure condition (Single, No-comparator, Similar, and Dissimilar), and are presented as a function of array type (Morph or Non-Morph). Performance at chance level is 20%.

A within-subject ANOVA[5] confirmed this observation indicating that there was no significant main effect on accuracy depending on exposure condition, $F(3, 93) = 1.53$, $p = .212$, MSE = 0.082. Despite there being no difference in the main ANOVA, pairwise comparisons were computed in order to explore the performance of the individual exposure conditions and to facilitate Bayesian analysis. Turning first to the Similar comparator condition, which might have been expected to produce the best performance, there were no significant differences compared to: Single, $F(1, 31) = 1.77$, $p = .194$, MSE = 0.221, $B_{01} = 3.114$, No-Comparator, $F < 1$, $B_{01} = 6.365$, or Dissimilar comparators, $F < 1$, $B_{01} = 7.190$. There was also no difference between the Single exposure condition and the other conditions (No-Comparator, $F(1, 31) = 3.43$, $p = .074$, MSE = 0.192, $B_{01} = 1.476$, and Dissimilar, $F(1, 31) = 2.259$, $p = .143$, MSE = 0.173, $B_{01} = 2.486$). Nor was there a difference between No-comparator and Dissimilar comparators ($F < 1$, $B_{01} = 6.264$). There was a significant main effect of test $F(1, 31) = 4.54$, $p = .041$, MSE = 0.048, with the Morph array producing less accurate responses compared to the Non-Morph array. No significant interaction was observed between type of exposure condition and type of test, $F < 1$. In all of the

---

[5] Although ANOVA methods are robust with respect to violations of its underlying assumptions (see, Howell, 2006; Field, 2009), it should be noted that in some of the experimental conditions in this experiment the data was negatively skewed. In order to assess whether this factor had a meaningful effect on the ANOVA reported above, the data was re-analysed using Friedman's non-parametric methods for assessing differences between multiple within-subject conditions, and follow-up analyses were performed using Wilcoxon matched-pairs tests. There were no significant differences in recognition accuracy depending on exposure condition, $\chi^2(32) = 3.19$, $p = .362$. Post-hoc analysis revealed there were no significant differences in recognition between exposure conditions (Largest $Z = -1.67$, $p = .095$ between No-Comparator and Single exposure). As these non-parametric analyses matched that of the ANOVA it would appear that the presence of some skew in the dataset has not unduly influenced the analysis in a way that affects the theoretical interpretation of the results. This issue will not be rehearsed at length in subsequent experiments where skewed data was present, but the non-parametric analysis will be reported in a footnote to confirm that the general argument made here is applicable.

experiments reported in this chapter, the pattern of results observed with the CA scores was the same as with accuracy alone. This measure will not be reported further in this chapter.

In short, the current experiment failed to observe a facilitation of recognition for target faces exposed alongside similar comparators compared to any of the control conditions. Critically, Bayes analysis suggested genuine evidence in favour of the null in all of these cases. This contrasts with previous studies using artificially generated faces (Dwyer & Vladeanu, 2009) or with studies of perceptual learning with morphed faces (Mundy *et al.*, 2007).

2.3 Experiment 2

While Experiment 1 did not reflect previous evidence suggesting that comparison aids recognition, numerous differences between the details of the experiments could have contributed to the dissociation in results. Perhaps the most theoretically interesting difference between the experiments was in the images used as comparators. The comparators used by Dwyer and Vladeanu (2009) in their Similar condition were created by morphing away from the target face. Thus, there was a strong degree of similarity between the target and similar comparators. Moreover, as the test foils were created in the same fashion, then both similar comparators and test foils were constrained to have some of the same features in common (and in common with the target face). This level of similarity may have enhanced any process of differentiation between the target and foils. In Experiment 1, Similar and Dissimilar were based on gender, and so even the supposedly similar comparator faces may have been too dissimilar to the target face to support a useful comparison process. Indeed, the stimuli used by Mundy *et al.* (2007) were also morphed faces that possessed a high degree of similarity. In addition, the original demonstration by Dwyer and Vladeanu (2009) included a control whereby the target was presented for a brief period to provide a baseline for the accuracy of recognition. The point of a Brief exposure (or some other non-learning control) is that without such measure it is possible that the performance seen at test in Experiment 1 is simply the baseline level, that would be seen without any exposure, and that the entire exposure phase was irrelevant.

The design of Experiment 2 addressed these issues in a within-subjects design while otherwise retaining the same general procedural details as the previous experiment. Brief exposure gave a single (2s) presentation of a target while the No-comparator condition remained unchanged from Experiment 1. Similar comparators were created by morphing away from the target (as with the test foils in the Morph array from Experiment 1, see Figure 1, Panel A) while dissimilar comparators were the same gender as the target but with no other treatment (as with the foils in the Non-Morph array from Experiment 1, see Figure 1, Panel B). Following exposure, participants were required to identify the target from two separate test arrays (one with morphed foils, the other with unmorphed foils), and give confidence ratings for their identity choices. If the results of Experiment 1 are reliable then there should be no difference between repeated exposure conditions on recognition performance. Moreover, following the results of Dwyer & Vladeanu (2009) all repeated exposure conditions should produce better recognition performance than Brief exposure.

2.3.1 Method

Participants

Thirty-two participants, between the ages of 18-32, were recruited from the School of Psychology at Cardiff University. No participant had taken part in the previous experiment. All had normal or corrected-to-normal vision.

Stimuli

The stimuli consisted of the same base faces as used in Experiment 1. From this set targets were then separately morphed towards each comparator face. Similar comparators were 50% morphs of the target and a distractor face. Dissimilar comparators were unmorphed distractor faces. Thus, unlike Experiment 1, similar comparators were now defined as a 50% morph between a target and comparator, and dissimilar comparators were defined as the same gender as the target. As in Experiment 1, morphed faces had the external features removed by applying a mask function in Adobe Photoshop™ image editing software. Each photograph was displayed centrally on a white background at 360 × 504 pixels subtending to an approximate visual angle of 12.5° × 17.2°. All other details, such as the creation of arrays, were identical to Experiment 1.

Design and Procedure

Participants completed all four conditions in a within-subject design: Brief exposure, No-comparator, Similar, and Dissimilar exposure. Brief exposure consisted of a single 2s exposure to a target followed by the recognition task. The design of the repeated conditions (e.g., No-Comparator, Similar and Dissimilar conditions) remained the same as Experiment 1 other than the differences in stimulus definition (e.g., each target was presented a total of sixteen times and both Similar and Dissimilar conditions had four comparator faces; each comparator was presented four times per condition, equating to a total of sixteen comparator presentations). All other procedural details remained consistent with Experiment 1. The assignment of faces to condition, the order in which the exposure conditions were presented, and the order of testing in the Morph and Non-Morph conditions, were counterbalanced as in Experiment 1.

2.3.2 Results

Figure 3 displays percentage of correct responses as a factor of exposure condition (Brief, No-Comparator, Similar and Dissimilar) and test type (Morph and Non-Morph). Performance was generally better following repeated exposure compared to Brief exposure on both test arrays, but there is little or no difference between the repeated conditions on either array type.

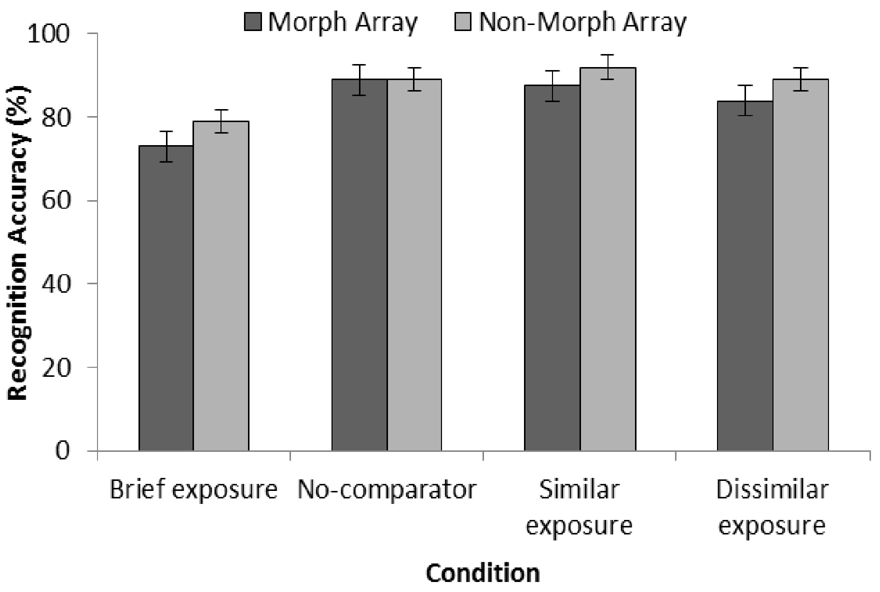

Figure 3: Recognition accuracy as percentage correct (with SEM) from Experiment 2. Data are organised by exposure condition (Brief, No-comparator, Similar, and Dissimilar), and are presented as a function of array type (Morph or Non-Morph). Performance at chance level is 20%.

A within-subject ANOVA[6] with factors of exposure condition (Brief exposure, No-Comparator, Similar exposure and Dissimilar exposure) and test type (Morph and Non-Morph) indicated a significant main effect on accuracy depending on exposure condition $F(3, 93) = 3.60$, $p = .016$, MSE = 0.067, but no other main effects or interactions (largest $F(1, 31) = 1.90$, $p = .177$, MSE = 0.051, for the main effect of test type). Pairwise analyses suggested that Brief exposure produced lower recognition scores than all other conditions: Similar exposure, $F(1, 31) = 6.67$, $p = .015$, MSE = 0.169, $B_{01} = .382$, Dissimilar exposure, $F(1, 31) = 5.16$, $p = .030$, MSE = 0.128, $B_{01} = .705$ and the No-comparator condition, $F(1, 31) = 10.33$, $p = .003$, MSE = 0.097, $B_{01} = .095$. No differences were observed between other exposure conditions. That is, comparing the Similar to No-comparator, $F < 1$, $B_{01} = 7.103$, and Dissimilar conditions, $F < 1$, $B_{01} = 5.080$, found no advantage for the Similar condition. Furthermore, there were no differences between the No-Comparator and Dissimilar comparator conditions, $F < 1$, $B_{01} = 6.479$.

2.3.3 Discussion

As in Experiment 1, exposing the target along with similar comparators failed to produce superior recognition of a target face over exposure without similar comparators. Brief exposure to a target produced lower levels of accurate matching compared to other conditions. Thus, the absence of facilitation in the Similar comparator condition cannot be attributed to a failure of learning *per se*. Similar patterns of recognition have been observed

---

[6] As in Experiment 1 a negative skew was present in the data of Experiment 2. Non-parametric analysis these results confirmed the conclusions of ANOVA analysis. That is, there was a significant difference in accuracy between exposure conditions $\chi^2(32) = 11.43$, $p = .010$. Post-hoc analysis revealed that the Brief exposure condition was recognised less accurately than the No-comparator $Z = -2.82$, $p = .005$, Similar $Z = -2.38$, $p = .017$, and Dissimilar exposure conditions $Z = -2.29$, $p = .022$. There were no differences between the other exposures (Largest $Z = -.85$, $p = .396$ between Similar and Dissimilar exposures). This analysis again closely reflected that of the parametric analysis reported in the main text.

in other learning tasks utilising a single vs. multiple exposure design (e.g., Experiment 1 of Longmore *et al*., 2008). The absence of a comparator-similarity effect was observed despite the fact that a morphing procedure was used to ensure that the comparators were genuinely similar to the targets. These conclusions were supported by the Bayesian analysis which again provided evidence for the absence of an effect. Moreover, both the comparator faces and test foils were based on the target, ensuring that the general level of similarity between the target and both foils and comparators was the same. These methods closely replicate those used by Dwyer and Vladeanu (2009), and thus the current results imply that insufficient comparator similarity alone cannot explain the absence of facilitation by similar comparators.

That said, the artificially generated faces used by Dwyer and Vladeanu may well come from a more restricted set of dimensions than the real face images used in the current experiments. Indeed, as face space has been estimated to contain between 15 and 22 dimensions (Lewis, 2004), then simply training a target against four comparators may not have spanned enough of this space to ensure that any particular dimensions which differentiated the target and comparators were the same dimensions which differentiated the target and novel test foils. Therefore, Experiment 3 examined whether the number, rather than simply the type, of potential comparators influenced learning of a novel face.

2.4 Experiment 3A

In Experiment 3A, all participants were exposed to Similar (same gender) and Dissimilar (different gender) comparator conditions. Within each of these, target stimuli were shown in alternation with 0, 1, 2, 4, or 16 different comparators. As noted above, exposure to multiple different comparators should maximise the possibility that the dimensions (or features) on which the target differs from the comparators overlap with the dimensions on which the target differs from the test foils (Table 2 summarises the design of Experiment 3A).

Table 2

Design of Experiment 3

Condition	Target	Number of Comparators	Test Array
Similar	A	0	Select AH from a range of AH, AH1, AH2, AH3
	B	1	Select BH from a range of BH, BH1, BH2, BH3
	C	2	Select CH from a range of CH, CH1, CH2, CH3
	D	4	Select DH from a range of DH, DH1, DH2, DH3
	E	16	Select EH from a range of EH, EH1, EH2, EH3
Dissimilar	F	0	Select FH from a range of FH, FH1, FH2, FH3
	G	1	Select GH from a range of GH, GH1, GH2, GH3
	H	2	Select HH from a range of HH, HH1, HH2, HH3
	I	4	Select IH from a range of IH, IH1, IH2, IH3
	J	16	Select JH from a range of JH, JH1, JH2, JH3

Note: 0-16 represents the number of comparator faces displayed in alteration with the Target faces A-J (these faces were counterbalanced across conditions as described in section 2.4.1). AH-JH represents target individuals with an expression change. 1-3 in the test arrays represents the different faces used as present in each array (e.g., AH1 to AH3 are the test foils for target A). Note: this design was performed twice, once each with male and female target faces.

2.4.1 Method

Participants

Forty participants, 33 females and 7 males, between the ages of 18 and 25, were recruited from the School of Psychology at Cardiff University, none of whom had previously completed the first two experiments. All had normal or corrected-to-normal vision.

Stimuli

A total of twenty front-view photographs, 10 male and 10 female, were randomly selected to become target faces. The images were selected from those freely available in the public domain. For each target, a further 19 other faces were selected that were of the same gender and had similar hair colour and style. From the set, 16 became similar comparators and 3 were used as foils. All individuals were white and aged between 18-25 years old; half were male, and the other half female. Each image was cropped, resized, and converted to an 8-bit quality so that images had a standard width and height of 400 × 600 pixels, subtending to an approximate visual angle of 13.9° × 20.3°, during exposure. Each exposure stimulus displayed a neutral expression. The test arrays displayed a target alongside three foils with an expression of happiness. Each array was displayed at 764 × 282 pixels subtending to a visual angle of 25.2° × 9.8°. All arrays were subjected to a contrast change. All images retained some background information and the external features.

Design and Procedure

Participants completed all ten conditions in a within-subject design. There were two conditions of exposure: Similar exposure, in which a target alternated with similar comparators, and Dissimilar exposure, in which a target alternated with dissimilar comparators. Comparators were defined in the same way as Experiment 1 (e.g., similar comparators were the same gender and dissimilar comparators were a different gender to the target). Within each comparison condition there were five target faces which differed in the number of comparison faces that were presented with them (0, 1, 2, 4, or 16). In every case, the target was presented 16 times, and the comparison stimuli were interleaved between these presentations, with repetition of the comparators in the 1, 2, and 4 comparator conditions (see Table 2). All conditions were presented with both male and female target faces.

At the start of the experiment, participants were seated approximately 70 cm from the screen and instructed to examine the faces carefully and try to remember the target face presented. The presentation began with a set of standardised instructions shown on screen explaining the study. The presentation format and timings of stimuli were identical to those of Experiment 1 and 2. That is, exposure for each face was consistent with Experiments 1 and 2 (i.e., 2s with a 1s ISI) and each exposure condition was followed by a recognition task and finally a measure of choice confidence for each condition. For this experiment there was only one test array (unmorphed faces). Responses were made in the same way as previous experiments, but with the exception that response time was unlimited and a response could be made during the array presentation. Arrays disappeared when a response was made. The experiment was run in four blocks. Each block comprised one comparator condition (i.e., Similar or Dissimilar) and gender of the targets (i.e., Male or Female). A block consisted of exposure, followed by test, for five target faces and the appropriate comparators. Each of the five targets in a condition was exposed with a different number of comparators (0, 1, 2, 4, & 16). Between blocks, participants were able to pause before a new block was initiated by the experimenter. Within each block, the order in which the five different comparators were presented was counterbalanced according to a Latin square procedure. In addition, across all participants, half were given the Similar conditions first and the other half were given the Dissimilar condition first. Within this, male and female targets were presented first equally often.

2.4.2 Results

Figure 4 displays the percentage of correct responses as a factor of exposure condition (Similar and Dissimilar) and the number of different comparator presented (0, 1, 2, 4, & 16). Inspection of the figure suggests that performance was equivalent across conditions and number of comparators. A within-subjects ANOVA with the factors, condition type (Similar and Dissimilar), and number of comparators, confirmed that there were no significant effects on accuracy due to exposure condition, $F < 1$, $B_{01} = 6.535$; or to the number of comparators $F(4, 156) = 1.00$, $p = .409$, MSE = 0.096. No significant interaction between condition type and number of comparators was found $F < 1$. To the extent that adding comparators should increase recognition, the simple effects analysis comparing the performance of each number of comparator exposures are reported. Beginning with the control condition (0 comparators), no differences were found between: 0 and 1 ($F < 1$, $B_{01} = 7.188$), 0 and 2 ($F(1, 39) = 1.585$, $p = .215$, MSE = .099, $B_{01} = 3.748$), 0 and 4 ($F < 1$, $B_{01} = 5.443$), or 0 and 16 comparators ($F < 1$, $B_{01} = 6.786$). In addition no differences were found between the maximum number of comparators (16) and the following amount of comparators: 1 ($F(1, 39) = 1.106$, $p = .299$, MSE = 0.090, $B_{01} = 4.700$), 2 ($F(1, 39) = 4.057$, $p = .051$, MSE = 0.075, $B_{01} = 1.215$), or 4 ($F(1, 39) = 2.566$, $p = .117$, MSE = 0.074, $B_{01} = 2.377$). No other differences between any other number of comparator exposures were observed: 1 vs. 2 ($F < 1$, $B_{01} = 6.494$), 1 vs. 4 ($F < 1$, $B_{01} = 7.592$), 2 vs. 4 ($F < 1$, $B_{01} = 7.415$).

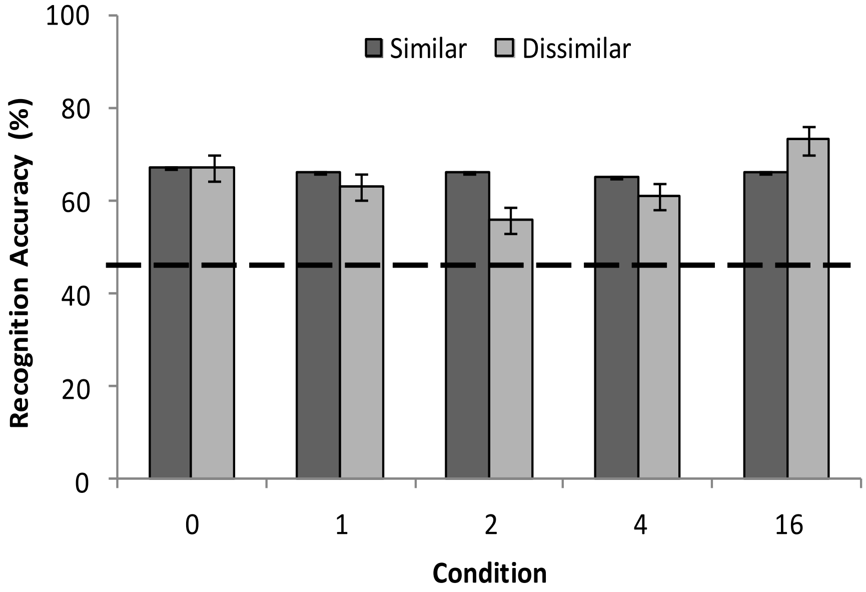

Figure 4: Recognition accuracy as percentage correct (with SEM) from Experiment 3. Data are organised by number of different comparators presented (0, 1, 2, 4, or 16), and are presented as a function of exposure condition (Similar and Dissimilar). Note: the dotted line represents the baseline Brief exposure control from Experiment 3B. Accuracy for the Brief exposure equalled 45% (SEM 3.7). Performance at chance level is 25%.

2.5 Experiment 3B

Experiment 3B used the same 20 target and associated test procedures as in Experiment 3A. However, participants only received brief (2s) exposure to the target stimuli prior to test (for half of the stimuli it was followed by a similar comparator and for the other half by a dissimilar comparator). This corresponds to the Brief exposure condition from Experiment 2, and so provides a baseline to which performance in Experiment 3A can be compared.

2.5.1 Method

Participants, Apparatus and Stimuli

Twelve participants between the ages of 18-24 were recruited from the School of Psychology at Cardiff University. All were undergraduates who participated in return for course credit. None of the 12 had participated in any of the previous experiments described here. All apparatus and stimuli were identical to those used in Experiment 3A.

Design and Procedure

There were two conditions of exposure; both displayed a target and comparator for 2s before the test arrays were presented. These two conditions were Brief Similar, in which a target was followed by similar comparator, and Brief Dissimilar, in which a target was followed by a dissimilar comparator. There were eight targets, 4 male and 4 female, with two target faces from each gender per condition. Blocks of testing were counterbalanced such that each condition was presented first and second equally often. Similarly each gender was presented first and second equally often. All other details, including definitions of similar and dissimilar, were identical to those of Experiment 3A.

2.5.2 Results

A preliminary analysis from Experiment 3B confirmed that, after Brief exposure, there were no significant differences in accuracy between the stimuli exposed with either a single similar or a dissimilar comparator, $t(11) = -1.65$, $p = .125$, $B_{01} = 1.426$.

Comparisons between Brief and Repeated exposure can been seen in Figure 4. The dotted line represents performance across conditions from Experiment 3B. A between-subjects *t*-test revealed that performance, collapsed across all conditions in Experiment 3A, was superior to performance across conditions in Experiment 3B, $t(50) = -5.22$, $p < .001$, $B_{01} < .001$. This difference was also observed in a separate analysis comparing each condition from Experiment 3A to the baseline of Brief exposure from Experiment 3B. That is, repeated exposure to similar stimuli produced better recognition than the Brief exposure, $t(50) = -4.51$, $p < .001$, $B_{01} < .001$, as did exposure to dissimilar stimuli, $t(50) = -4.62$, $p < .001$, $B_{01} < .001$. It is also worth noting that further analysis examined separate performance for each number of comparators to the baseline performance of 3B. Although not reported fully here (see appendix 1) the results demonstrate, as anticipated, that all repeated exposure produced better recognition to that of the baseline, (smallest $t(33) = -3.07$, $p = .004$, $B_{01} = .131$, unequal variances, $F = 6.06$, $p = .017$, between 2 comparators (3A) and the baseline Brief control (3B)).

2.5.3 Discussion

To summarise, the results of Experiments 3A and 3B demonstrate that repeated exposure produced an advantage over brief exposure, and that recognition performance between target exposures with similar comparators was equivalent to those with dissimilar comparators. Moreover, recognition of a target was not influenced by altering the number of comparators during exposure. In other words, recognition performance following exposure with the maximum number of comparators (e.g., sixteen) showed no advantage over targets without any comparators. Again these interpretations were strengthened by the Bayesian analysis for the critical comparisons. Therefore, taken together, the recognition task of Experiment 3A and 3B replicated the key findings of Experiment 1 and 2. Equivalent learning was observed across repeated exposures regardless of comparator type. Moreover, the comparator faces and test foils displayed the whole features of a face like that of Dwyer and Vladeanu (2009).

That said, it is possible that the effect was diminished due to the uncontrolled nature of the stimuli. The stimuli used in Experiment 3A and 3B were captured from the public domain and exhibited a high degree of variation between photos. This variation in pose and background may have directed attention away from forming accurate representations with which to compare the target. This may also explain the lower overall performance observed in this experiment compared to those reported earlier in this chapter. While the artificial stimuli used by Dwyer and Vladeanu (2009) utilised whole faces, they did so in a controlled pose. Thus, it may be that exposing the whole features of a face which promotes a process of stimulus comparison. Indeed, there is evidence to suggest that attention towards to the internal features increases with familiarity (O’Donnell & Bruce, 2001). Therefore simply removing the external features in the previous experiments (i.e., 1 and 2) may have nullified any learning through comparison. It is plausible that the whole face is required for this process given that is it rarely the case (at least in this culture) that faces are seen with just the internal features on display. Experiment 4 compared the use of exposure to the whole face versus exposure to the internal features.

2.6 Experiment 4

In Experiment 4, the same intermixed exposure presented similar (same gender) or dissimilar (different gender) comparators in alternation with target individuals. Targets and comparators were presented displaying the internal features only, or the whole face. An interval was inserted between internal and whole face exposures to alleviate any carry over effects (Table 3 summarises the design). As in the previous experiments, after each exposure stage, recognition of a target was assessed using an array of faces. If exposure to a whole face is an important mechanism by which comparison operates, then it is expected that the exposure to the Whole condition will produce the most accurate recognition. In addition, those in the Whole-Similar condition should perform better overall in accordance with the findings of Dwyer and Vladeanu (2009).

2.6.1 Method

Participants and Apparatus

Sixty-four participants between the ages of 18-47 were recruited from the School of Psychology at Cardiff University. Each participant received course credit in return for their participation. None had taken part in the previous experiments. All other apparatus details were consistent with previous experiments.

Stimuli

The same eight targets used in Experiment 1 and 2 were used for this experiment. Faces in the internal feature conditions had external features removed, using Adobe Photoshop™ image editing software. Those in the Whole face conditions were cropped, removing any background information. Faces in all conditions were displayed on a white background. During exposure, each photograph had a standard width of 360 × 504 pixels subtending a visual angle of 12.5° × 17.2°.

The test arrays displayed a target alongside four foils, at a size of 1029 × 367 pixels, subtending an approximate visual angle of 32.4° × 12.8°. Other array details such as contrast change and cropping are consistent with the Non-Morph array from Experiment 1 (Figure 1: Panel B).

Table 3

Design of Experiment 4

Condition	Training sequence	Test Array
Similar Internal	Aint, A1int, Aint, A2int, Aint, A3int, Aint, A4int. (×4)	Select AH from a range of AH, AH5, AH6, AH7, AH8
Dissimilar Internal	Bint, X1int, Bint, X2int, Bint, X3int, Bint, X4int. (×4)	Select BH from a range of BH, BH5, BH6, BH7, BH8
Reading interval (1 Min).		
Similar Whole	C, C1, C, C2, C, C3, C, C4. (×4)	Select CH from a range of CH, CH5, CH6, CH7, CH8
Dissimilar Whole	D, X1, D, X2, D, X3, D, X4. (×4)	Select DH from a range of DH, DH5, DH6, DH7, DH8

Note: A-D indicates target faces (these faces were counterbalanced across conditions as described in section 2.6.1). 1–4 refer to comparator faces (e.g., C1-C4 in indicate similar comparator faces to target C, while X1 to X4 illustrate dissimilar comparator faces to target D). int (e.g., Aint) refers to the target faces and comparators exposed with the internal features only. Exposure was repeated four times as indicated by × 4. AH-DH represents target individuals with an expression change in the Non-Morph array. 5-8 represent unseen foils present in the Non-Morph array (e.g., BH9 to BH12 are non-morphed foils for target B).

Design and Procedure

Participants completed a within-subject design comprising four exposure conditions: Similar Internal, in which all faces (e.g., the target and alternating similar comparators), displayed internal features. Dissimilar Internal, in which internal target displays alternated with internal displays of dissimilar comparators. Similar Whole, which gave alternating exposure to a target and similar comparators using whole face stimuli, and Dissimilar Whole, in which displays of whole face alternated between a target and dissimilar comparators. Comparator definitions were consistent with the previous experiments (e.g., similar comparators were the same gender and dissimilar comparators were a different gender to the target).

At the start of the experiment, participants were seated approximately 70cm from the screen and instructed to examine the faces carefully and try to remember the target face presented. The presentation began with a set of standardised instructions shown on screen explaining the study. The presentation format and timings of stimuli were identical to those of previous experiments in this chapter. That is, exposure for each face was consistent with Experiments 1 and 2 (i.e., 2s with a 1s ISI) and was followed by a recognition task and finally a measure of choice confidence. This experiment gave only one test array (non-morphed faces: see Figure 1, Panel B). Responses were made in the same way as previous experiments, during the array presentation, but with the exception that response time for selecting a face from an array was limited to 10s. Arrays disappeared when a response was made. Participants completed eight trials, a male and female version of each of the four conditions. Between Internal and Whole conditions, participants read an unrelated passage for a minute before completing the following two conditions. The order of presentation was counterbalanced such that, across participants, the Internal and Whole conditions were presented first and second equally often. Likewise, Similar and Dissimilar comparator conditions were presented first and second equally often, as was each gender.

2.6.2 Results and Discussion

Figure 5 shows percentage of correct recognition as a function of stimulus type (Internal and Whole) and exposure condition (Similar and Dissimilar). Examination of the figure suggests that there is little, if any difference between similar comparators between stimulus types. However, there is (perhaps) a suggestion of an improvement in performance for dissimilar comparators when utilising internal features rather than the whole face. A within-subject design used a two-way ANOVA[7] on the factors condition type (Similar or Dissimilar), and face exposure (Internal features and Whole face). Accuracy of recognition was not significantly affected due to exposure condition, $F < 1$, $B_{01} = 9.940$, nor due to stimulus type, $F(1,63) = 3.60$, $p = .062$, MSE = .069, $B_{01} = 1.812$. Furthermore, the interaction between exposure condition and stimulus type was not significant, $F(1,63) = 1.69$, $p = .192$, MSE = .083.

---

[7] As in Experiment 1 and 2 a negative skew was present in the data of Experiment 4. Non-parametric analysis these results confirmed the conclusions of ANOVA analysis. That is, there was no significant difference in accuracy between exposure conditions $\chi^2(32) = 6.97$, $p = .073$. This analysis again closely reflected that of the parametric analysis reported in the main text.

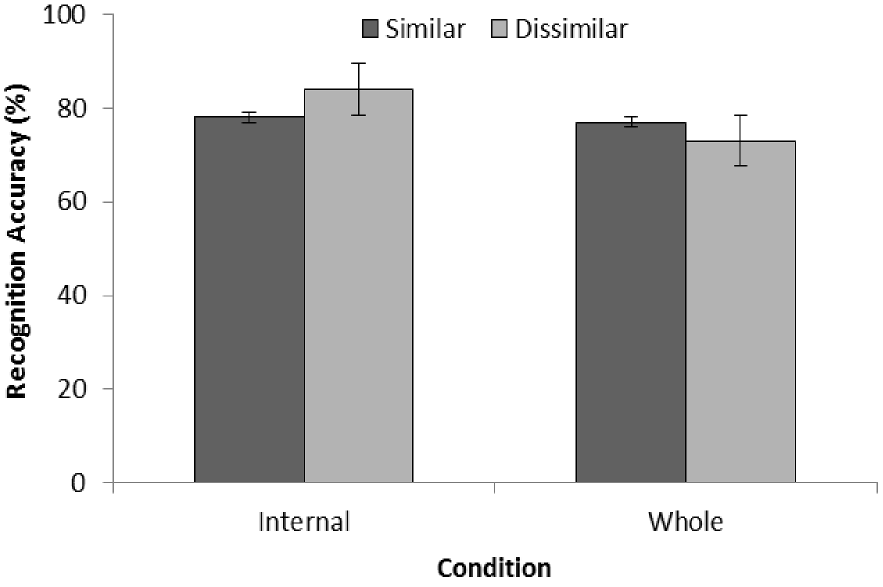

Figure 5: Recognition accuracy as percentage correct (with SEM) from Experiment 4. Data are organised by stimulus type (Internal and Whole) and are presented as a function of exposure condition (Similar and Dissimilar). Performance at chance level is 20%.

The results of this experiment indicate no benefit for exposing similar comparators. Moreover, this lack of any comparator effect was observed regardless of whether the stimuli were exposed as internal or whole faces. Again, this analysis was strengthened by the Bayes value for the critical comparison. In addition, there was no benefit of exposing internal features compared to whole faces.

Therefore, even incorporating a whole face exposure within the design failed to produce a similar comparator advantage like the one observed by Dwyer and Vladeanu (2009). Although it is plausible that participants became aware that external features were not a useful means to recognising a face – particularly as the test arrays included no external features. Moreover, the current finding is somewhat inconsistent with a proportion of the literature, which suggests that recognition of an unfamiliar faces is heavily reliant on processing of external features (Ellis *et al*, 1979; Young *et al*, 1985). In this light, some effect of manipulating whether external features were available might be expected. That said, the failure to do so may well be due to the aforementioned processing strategy adopted by participants because of the lack of external features present during test.

2.7 Chapter Discussion

The four experiments reported here examined the potential beneficial effects of similarity and comparison on the recognition of a familiarised individual. In Experiments 1 and 2, exposure to a target face along with comparators of the same gender (either morphed to be explicitly similar to the target or not) produced no facilitation in selecting the target from an array of novel foils, relative to target faces that were exposed with either different gender comparators or no comparators at all. In Experiment 3, increasing the number of comparators during the exposure phase also produced no improvement in recognition. Experiment 4 suggested that comparison was not affected by whether exposure contained the internal or external features. Importantly, the failure of comparison to improve recognition cannot be attributed to a general failure of learning as recognition performance in both similar comparator conditions, and repeatedly exposed controls, was superior to brief exposure conditions. The fact that there was no facilitation in recognition, despite the differing levels of similarity and the varying numbers of comparators used during exposure, suggests that there is no practical beneficial effect of comparison on the recognition of unfamiliar faces above that of repeating exposure to a face.

The idea that comparison-based exposure effects do not enhance matching when testing against novel foils (as demonstrated here) does not mean that this process cannot contribute to discriminating between a set of faces. The results reported by Mundy *et al.* (2007) reflected a change in the processing internal features. Moreover, similar exposure periods have produced advantages that were selective to the internal features of a face (e.g., Dwyer *et al.*, 2009). Thus, it seems that comparison may aid certain distinctions, but this does not extend to better identification against nonexposed foils. That is, the comparisons during exposure may have covered certain dimensions, but these dimensions may not be informative with respect to discriminating the target from other faces. Another way to phrase this situation is that while comparison may well help identify how Tom differs from Bob it is unlikely to be informative for how Tom differs from Fred. Therefore, the above analysis suggests that comparison to a limited number of comparators is unlikely to identify the critical aspects of a face on all relevant dimensions. Each target face was indeed presented with a name (e.g. Matt/not Matt). As such, people could both be learning about the face itself and the face-name association. However, the test task does not rely on name memory, and so any apparent confound between exposure and face-name training is not going to affect the current results. However the extent that learning the association between a face and the name could influence the processing of a face is entirely moot given that any recognition task does not require face-name memory. The implications of these issues will be further discussed in the General Discussion (Chapter 5).

Given the lack of evidence for a generalizable comparison effect within the experimental work of this chapter, and the suggestion that comparison can aid only certain distinctions, Chapter 3 will turn to a more theoretical analysis of perceptual learning. That is, comparison has facilitated discrimination between sets of faces (e.g., Mundy *et al.*, 2007; Dwyer *et al.*, 2009), conferring an advantage for the internal features for this type of stimulus. Thus, is this advantage reflective of a general mechanism of perceptual learning whereby location of the unique features are important? In Chapter 3, the meaning of what is leaned about the unique feature of a stimulus from a process of perceptual learning will be examined.

Chapter 3: Location and Content of unique features

3.0 Abstract

Exposure to complex checkerboards (comprising a common background, e.g. X, with unique features, e.g. A-D, that are placed in particular locations on the background) improves discrimination between them (perceptual learning). Such stimuli have been used previously to probe human perceptual learning but these studies leave open the question of whether learning is based on the content or location of the unique stimuli. In Experiment 5, exposure produced equivalent perceptual learning in stimuli that differed in both location and content of the unique features, in their location alone, or in their content alone. Experiment 6 suggests that perceptual learning transferred to stimuli that had new unique features (e.g., C, D) in the position that had been occupied by A and B during exposure. However, there was no transfer to stimuli that retained A and B as the unique features but moved them to a different location on the background. Experiment 7 replicated the key features of Experiment 6: no transfer of exposure learning based on content, but perfect transfer of exposure learning based on location, using a fully factorial design. In all the experiments reported here, superior discrimination between similar stimuli on the basis of exposure can be explained entirely by learning where to look, with no independent effect of learning about particular stimulus features. These results directly challenge the interpretation of practically all prior experiments using the same style of stimuli.

3.1 Introduction

The previous chapter examined the effects of perceptual learning inspired comparison-based training schedules on improving the recognition of familiarised faces. Findings indicated that repeated exposure to a stimulus consistently improved recognition, but that comparison, which should promote stimulus differentiation, is not a useful means of improving discrimination of a target from nonexposed foils. While comparison has been suggested as a mechanism that promotes stimulus differentiation, Gibson's own presentation of this idea left the mechanisms underpinning stimulus differentiation relatively undefined (see Gibson 1969). There have been numerous subsequent attempts to unpack this process (e.g., Hall, 2003; Mitchell, Nash, *et al.,* 2008; Mundy, *et al*., 2007), yet despite the differences in detail between these accounts (and there are many) they all rely on the idea that exposure can by some means produce changes to the relative salience of the unique and common features of the critical stimuli. For example, the fact that people fixate on the unique features A and B after intermixed exposure (Wang & Mitchell, 2011) may be due to these features being particularly salient and thus able to attract attention (see also Wang, Lavis, Hall & Mitchell, 2012). Moreover, there is evidence that experience or familiarisation with a face promotes better recognition through an internal feature advantage (e.g., O'Donnell & Bruce, 2001; see also Clutterbuck & Johnston 2002, 2005). It appears that attending to the location of these diagnostic features is a product of experience with a stimulus, and it may be that this is due to experience increasing the salience of the most diagnostic aspect of a stimulus. Indeed Xiao *et al.* (2008) found that training can influence the attention towards location in some instances of perceptual learning.

Gibson's (1969) suggestion implies that perceptual learning will result in the salience of, and/or the attention towards, the unique features A and B being greater than for the common features X. However, the theoretical analyses of perceptual learning noted in

Chapter 1 (see section 1.7) are all silent with respect to the location of unique features, and instead are expressed in terms of the effects of exposure on the content and relationship between these features. Any exposure-dependent influence on the discriminability of AX and BX might reflect learning about the content of those unique features (e.g., a learnt change in their salience) or about their location (e.g., learning where to look for discriminating features).

In particular, a series of studies (e.g., Lavis & Mitchell, 2006; Mitchell, Kadib, *et al.*, 2008; Wang & Mitchell, 2011) used checkerboard stimuli (see Figures 6 and 8 for examples) that were created by taking a 20 × 20 grid of multi-coloured squares (these were the common features: X) and then adding, to a particular place on the background, features made of blocks of 4-6 squares, consisting one or two colours (the unique features: A/B – although the exact details of both the unique and common features differed slightly between experiments). This method of stimulus creation means that there is a perfect correlation between the content of a unique feature (e.g., its colour or shape) and its location (i.e., where it appears on the background, X). In turn, this means that it is impossible to ascertain the relative contributions of these aspects of the unique stimuli to the learning effects observed. However, as the unique features can, in fact, be placed anywhere on the background, they allow for independent manipulation of the content and location of unique features. In contrast, changing the location of the critical features within a face would be to change the configuration and thus confound an analysis of any potential effects on learning given the relatively special nature of faces as stimuli (see section 1.2.1).

It should be noted that there are two studies that give some indication as to what may occur should this correlation be broken. Lavis and Mitchell (2006) used checkerboards in which all 400 of the squares in the 20 × 20 grid were coloured and unique features A and B were identical in content, but differed only in location. Despite there being no difference in the content of the A/B features exposure to AX and BX improved subsequent discrimination (relative to control conditions using novel stimuli) and intermixed exposure produced better discrimination than did blocked exposure. The similarity of results to experiments where A/B did differ in content suggests that differences in the location of a unique stimulus, in the absence of differences in content, are sufficient to support perceptual learning. Such a result raises the possibility that the improvements in discrimination prompted by exposure to these checkerboard stimuli depends on learnt changes in where to look for discriminating features, rather than learning about the content of those features themselves. However, a closer consideration of the stimuli used by Lavis and Mitchell (2006) suggests an alternative possibility. As all squares on the background were coloured, different patterns would be obscured or revealed as the unique feature was moved from place to place on the background. That is, although the explicitly manipulated feature had the same content at two different locations, the underlying parts of the background differed at these points so there were some content differences between AX and BX. More recently, in Experiment 3 of Wang *et al*. (2012) it was found that after exposure to AX and BX, eye gaze during test is directed to the location at which A and B appeared during training, regardless of whether the exposed features A and B, or novel features C and D, were present at these locations. Moreover, discrimination performance with both novel and exposed features was better when they appeared at the location at which A and B were presented during initial exposure than when they appeared elsewhere[8]. Taken together, these demonstrations raise the possibility that people can learn about location rather than just about content, but leave open the question about whether anything is learnt about content which influences discrimination performance at all.

---

[8] Note that the experimental work presented by Wang *et al.* (2012) was published after the experiments outlined in this chapter were performed.

Therefore, the main aim of the current studies was to break this perfect correlation between content and location and to begin to assess their relative contributions to perceptual learning. This aim was approached in two general ways: The first (Experiment 5) was to manipulate the stimuli during exposure, so that they differed in either the content or location of the unique features (or both). The second (Experiments 6 and 7) was to examine whether exposure-produced improvements in discrimination of stimuli that differed in both the content and location of the unique features would transfer to test stimuli that used either the same unique content (but at a different location) or different unique content (but at the same location). It is important to note that in some cases the absence of a difference between conditions will be of particular theoretical relevance. Thus, as in Chapter 2, Bayesian analysis will be implemented as a means as assessing the relative strength of the evidence for the accepting or rejecting the null (for a discussion, see Rouder *et al.,* 2009).

3.2 Experiment 5

In the majority of experiments using this type of checkerboard (e.g., Mitchell, Kadib, *et al*., 2008; Mitchell, Nash, *et al*., 2008) the stimuli were created such that 156 of the 400 of the squares in the 20 × 20 grid comprising X were coloured, and the unique features A and B differed both in content and location. However, Lavis and Mitchell (2006) used a slightly different approach whereby all 400 of the squares in the 20 × 20 grid were coloured, and unique features A and B were identical in content, but differed only in location. Despite the differences in the detail of the methods of stimulus construction, the same general pattern of results has emerged: exposure to AX and BX improved subsequent discrimination (relative to control conditions using novel stimuli) and intermixed exposure produced better discrimination than did blocked exposure. The similarity of results suggests that differences in the location of a unique stimulus, in the absence of differences in content, are sufficient to support perceptual learning. In turn, the fact that perceptual learning effects are seen in stimuli that do not differ in the content of their unique features raises the possibility that the improvements in discrimination prompted by exposure to these checkerboard stimuli depends on learnt changes in where to look for discriminating features, rather than learning about the content of those features themselves. Indeed, a closer examination of the stimuli used by Lavis and Mitchell (2006) coupled with the recent findings of Wang *et al*., (2012) supports the suggestion of location-based learning with this type of stimuli.

In order to directly address the issue of whether content differences are necessary, Experiment 5 sought to examine perceptual learning as a factor of whether the exposed stimuli differed in content alone, location alone, or in both content and location. Figure 6 shows examples of the stimuli used in each of the three exposure conditions. Panel I shows the Content and Location condition where one unique feature (A) has been added to the top left of the background (X) while a different unique feature (B) has been added to the top right. Panel II shows the Location only condition, where the same unique feature (A) has been added to two different places on the background (corresponding to the same places used in the Content and Location condition). Panel III shows the Content only condition, where the unique features (A and B - corresponding to the same features used in the Content and Location condition) have both been added to the same place on the background. A between-subjects design was used whereby all participants were given non-reinforced intermixed exposure to a pair of checkerboards created according to one of the three conditions described above prior to performing a same-different discrimination task with both the exposed stimuli and a novel pair of stimuli that had been created in the same fashion (Table 4 summarises the design). A perceptual learning effect will be evident to the extent that discrimination performance at test is superior with the exposed stimuli than the novel stimuli and this should be seen in the Content and Location condition. If, as suggested by the results of Lavis and Mitchell (2006), differences in the location of a unique feature is sufficient for perceptual learning, then performance in the Location only condition should also differ between exposed and novel stimuli. However, if differences in content are necessary for perceptual learning to take place, then there should be an effect of stimulus exposure in the Content only condition, but not the Location only condition.

3.2.1 Method

Participants

Thirty-six undergraduate students, aged between 18-23 years, were recruited from Cardiff University and received course credit in return for participation. Participants were randomly assigned to the three experimental conditions.

Apparatus and Stimuli

Stimuli consisted of eight different 20 x 20 colour checkerboards. All had a common element (4 used X and 4 used Y); created by colouring 156 of the 400 squares (blue, green, purple, red, or yellow on X; blue, green, pink, purple, and orange on Y). The remaining squares were grey (for X this grey was lighter than the background which filled the remainder of the screen, and for Y it was darker than the surround). Thus the common elements X and Y differed in the colour, pattern, and placement of the grey and coloured squares. Unique features (A-D) were added by changing six adjacent blocks of grey squares to two of the brighter colours. Each unique feature differed from all others in colour and shape. These unique features could be added to the backgrounds in one of four locations (roughly top left, top right, bottom left, bottom right). The application of the unique stimuli to the backgrounds

Table 4

Design of Experiment 5

Condition	Exposure	Example Feature Placement	Test
Content and Location	AX/BX	A top left/B top right	AX/BX CY/DY
Location	*AX/A*X	*A top left/A* top right	*AX/A*X CY/DY
Content	*AX/*BX	*A top left/ *B top left	*AX/*BX CY/DY

Note: A-D represent unique features while X and Y represent common background checkerboards. * represents placement of unique feature upon the common background, for example *A in the Location condition indicates that the unique feature A was positioned in the top right hand corner of the common background X (see Experiment 5 method for a detailed description and Figure 6 and 8 for examples). Stimuli AX/BX comprise different unique features in different locations (Figure 6 Panel I). Stimuli *AX/A*X comprise the same unique features in the different locations (Figure 6 Panel II). Stimuli *AX/*BX comprise different unique features but in the same location (Figure 6 Panel III).

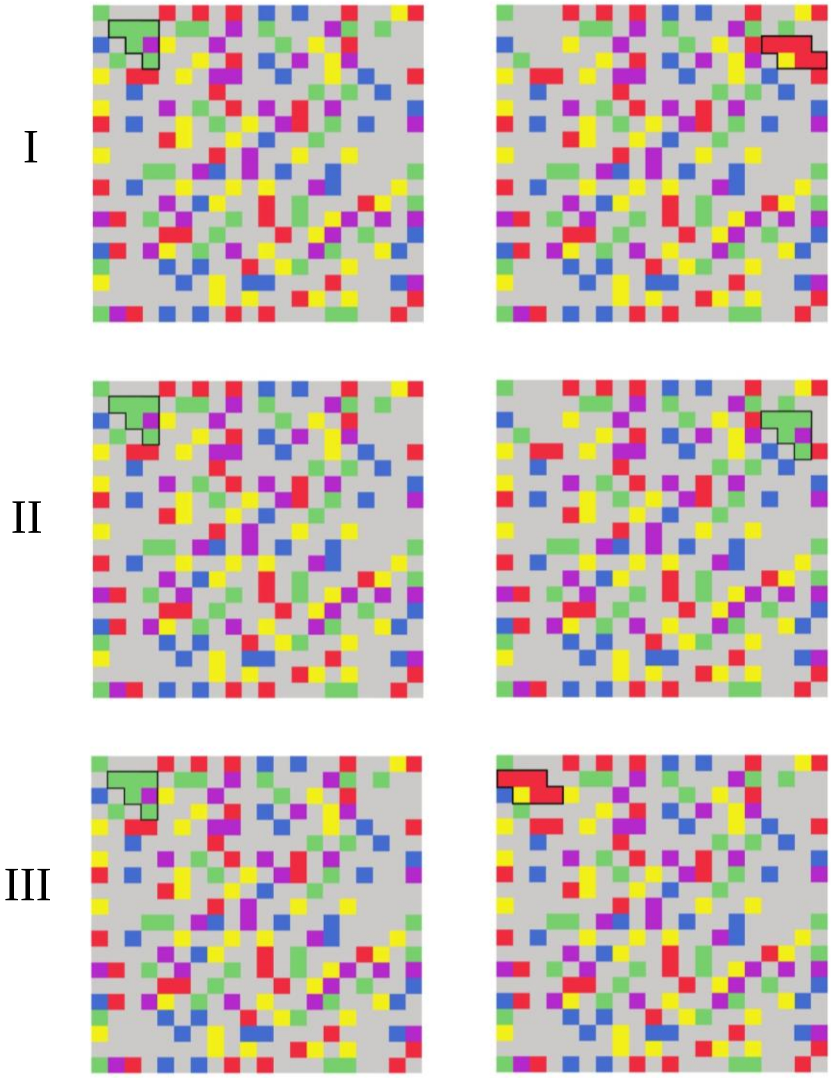

Figure 6: Examples of stimuli used in Experiment 5. All display the common background (X) with unique features outlined in black (this outline was not present during the experiment). Panel I is an example of the Content and Location condition (Left: A top left, Right: B top right); Panel II shows the Location condition (Left: A top left, Right: A top right); Panel III shows the Content condition (Left: A top left, Right: B top left)

was constrained such that A and B were added to the X background either top left or top right and C and D were added to the Y background either bottom left or bottom right. The stimuli were presented centrally on a 17 inch monitor with standard pixel height and width 576 x 576, subtending an approximate visual angle of 22.5° x 22.5°. DirectRT software was used to control the presentation of the stimuli on a PC. The area of the screen surrounding the checkerboard stimuli was a mid-grey, equidistant in lightness between the background greys of stimulus X and Y. A black border separated the checkerboard from the remainder of the screen. The individual squares within the checkerboard were not separated with any border.

Design and Procedure

As noted above, all participants were exposed to a single pair of checkerboards before completing a same/different discrimination test with the exposed stimuli and an equivalent nonexposed pair. For half the participants, the exposed checkerboards used the X background (and thus the Y background was used for the novel test stimuli) and for the remainder the Y background was used for the exposed stimuli. In the Content and Location condition, the unique features differed in both location and content (e.g., A top left, B top right, see Figure 6I). In the Location only condition, a single unique feature was used but appeared in two different locations (e.g., A top left, A top right, see Figure 6II). For the Content only condition, two unique features were used but they appeared at the same location (e.g., A and B top left, see Figure 6III). Critically, the assignment of stimuli to conditions was counterbalanced such that each of the unique features (A-D), each of the possible locations (top left, top right, bottom left, bottom right), and each of the two background patterns (X and Y) were used equally often in each of the three stimulus conditions (and equally often in the exposed and novel stimuli). Thus, across subjects, there were no differences in the particular unique features, locations, or backgrounds that could have artefactually contributed to the discrimination performance across conditions.

At the start of the experiment, participants were sat approximately 60cm from the computer screen and presented with a set of standardised instructions:

> *You will be exposed to a set of checkerboards. Pay attention to the stimuli, any stimulus differences will be useful later in the experiment. During exposure, please press the space bar to proceed from one trial to the next. If there are any questions please ask the experimenter now. If there are no questions press the space bar to begin.*

During the exposure phase, each stimulus was presented 60 times, for 470ms each trial. The two exposed stimuli were presented in strict alternation (e.g., AX, BX, AX, BX,….). Each stimulus presentation was followed by a blank grey screen, during which participants made their space bar presses. Regardless of a space bar press, the following trial was initiated 2000ms after the offset of previous stimulus.

Following the completion of the exposure phase, a second set of instructions was displayed in the same manner as the first. Participants were informed that they would be presented with a succession of pairs of checkerboards, one stimulus at a time. They were told to press the "Z" key if the two stimuli appeared the same and the "/" key if the stimuli appeared different. This instruction remained on screen throughout the test period. On every discrimination trial, the first stimulus was presented for 800ms, followed by a blank screen for 550ms before the presentation of the second stimulus for 800ms. A white square was displayed at the interval between trials; this remained on screen for 1400ms after a response had been made, the next trial then commenced.

At test, there were two types of stimuli; exposed and novel. The exposed stimuli had been displayed in the exposure phase of the experiment. The novel stimuli were unfamiliar for participants. Thus, there were four types of test trial: *different-exposed*, in which both the exposed stimuli were presented (e.g., AX→BX, or BX→AX), *different-novel,* in which two different novel stimuli were presented (e.g., CY→DY, or DY→CY), *same-exposed,* in which two copies of one of the exposed stimuli were presented (e.g., AX→AX, or BX→BX), and *same-novel*, in which two copies of one of the novel stimuli were presented (e.g., CY→CY, or DY→DY). There were 160 test trials, divided into 4 blocks of 40 trials. Each block contained 10 trials of each type (half starting with one of the stimuli from a relevant pair, and half starting with the other). Within each block, trial order was randomised. Between blocks, participants were able to pause before continuing.

Statistical Analysis

The data were examined in terms of proportion of correct same/different judgments (as has been typical with previous experiments of this type) as well as with a signal detection analysis. Sensitivity scores, *d'*, for each participant were calculated by treating hits as the proportion of correct responses given on different trials and false alarms as the proportion of incorrect responses to same trials (i.e., respond "different" when the two images were actually the same). Factorial ANOVA procedures were used to assess the output of both the proportion correct and *d'* data. A significance level of p <.05 was set for all analyses.

As noted above, Bayesian analyses were also conducted as a means of assessing the strength of empirical support for the hypothesis that two conditions do not differ. The calculation of the Bayes factor requires the specification of an effect size for the alternate hypothesis. The suggested default for this is that the manipulation will produce a difference of one standard deviation between the treatment and control means (as was implemented in Chapter 2). While the beneficial effect of exposure on perceptual learning effect is well established, it is difficult to justify which particular demonstration or demonstrations of perceptual learning should be used to set the expected effect size for the current studies. Therefore, in the analyses reported here, I based the specified effect size on the default as suggested by Rouder *et al.* (2009). While there are no published algorithms for factorial

ANOVA procedures, the key theoretical questions in the current chapter can generally be reduced to t-tests equivalent to the comparisons between two groups or conditions, in which case paired or unpaired Bayes t-tests were performed as appropriate. As in Chapter 2, results are treated as either supporting the null or alternative (or neither) by adopting the convention suggested by Jeffreys (1961) and recommended by Rouder *et al*. (2009).

3.2.2 Results and Discussion

Figure 7 Panel A shows the test data as mean proportion of correct responses on the same and different trials, as a function of stimulus type (Content and Location, Content only, and Location only) and exposure condition (Exposed/Novel). It is apparent that accuracy was generally higher on the same trials than that on different trials. Furthermore, overall accuracy appears to be greater for exposed compared to novel stimuli. Critically, neither of these factors were influenced by condition type. That is, learning appeared to be unaffected by whether the stimuli differed in either the content alone or location alone of the unique features (or both). A mixed ANOVA confirmed main effects of trial type (Same correct vs. Different correct) and stimulus type (Exposed vs. Novel) $F(1,33) = 40.36$, $p < .001$, MSE = 0.012, and $F(1,33) = 12.61$, $p = .001$, MSE = 0.024, respectively, but no main effect was observed due to condition (Content and Location, Content only, and Location only), $F < 1$. There was also an interaction between trial type (Same vs. Different) and stimuli type (Exposed vs. Novel), $F(1,33) = 15.38$, $p < .001$, MSE = 0.011. Simple effects analysis revealed that the performance on different trials was better for exposed stimuli compared to novel stimuli, $F(1,33) = 15.39$, $p < .001$, MSE = 0.030, B_{01} =.012, but no reliable difference was observed between exposed and novel stimuli on same trials, $F(1,33) = 2.01$, $p = .165$, MSE = 0.005, B_{01} = 2.837. Moreover, simple effects analysis confirmed that there was superior performance for exposed than novel stimuli (on the different trials) in all three stimulus conditions: Content and Location, $F(1,33) = 6.90$, $p = .013$, MSE = 0.005, B_{01} =

.511; Content only, $F(1,33) = 4.41$, $p = .045$, MSE = 0.005, $B_{01} = 1.051$; and Location only, $F(1,33) = 4.27$, $p = .048$, MSE = 0.005, $B_{01} = .329$. No other interactions were observed (*Fs* < 1).

Panel B of Figure 7 shows the mean *d'* sensitivity scores as a function of stimulus type (Exposed/Novel) and condition (Content and Location, Content only, and Location only). ANOVA confirmed a main effect of stimulus type $F(1,33) = 8.935$, $p =.005$, MSE = .894, but no other main effect for condition or any interaction between condition and stimulus type was observed (*Fs* < 1). Before moving to the implications of these results, it is worth noting that the observation here that performance was generally better on same than different trials is entirely consistent with previous investigations using these types of stimuli (e.g., Lavis & Mitchell, 2006; Mitchell, Kadib, *et al*., 2008; Mitchell, Nash, *et al*., 2008). Presumably, this effect of trial type represents a bias to report that the two stimuli presented on each test trial were the same which might be attributable to how difficult the stimuli are to discriminate as the bulk of them comprise the same common background (Lavis, *et al*., 2011). Moreover, my observation that the effects of exposure were restricted to the different test trial types is also consistent with previous observations.

Perhaps the most theoretically important feature of the current results was that discrimination performance was superior for exposed as compared to novel stimuli, a finding that was also supported by the sensitivity scores. Moreover, that this exposure effect was equally large regardless of whether the stimuli involved differed in both the content and

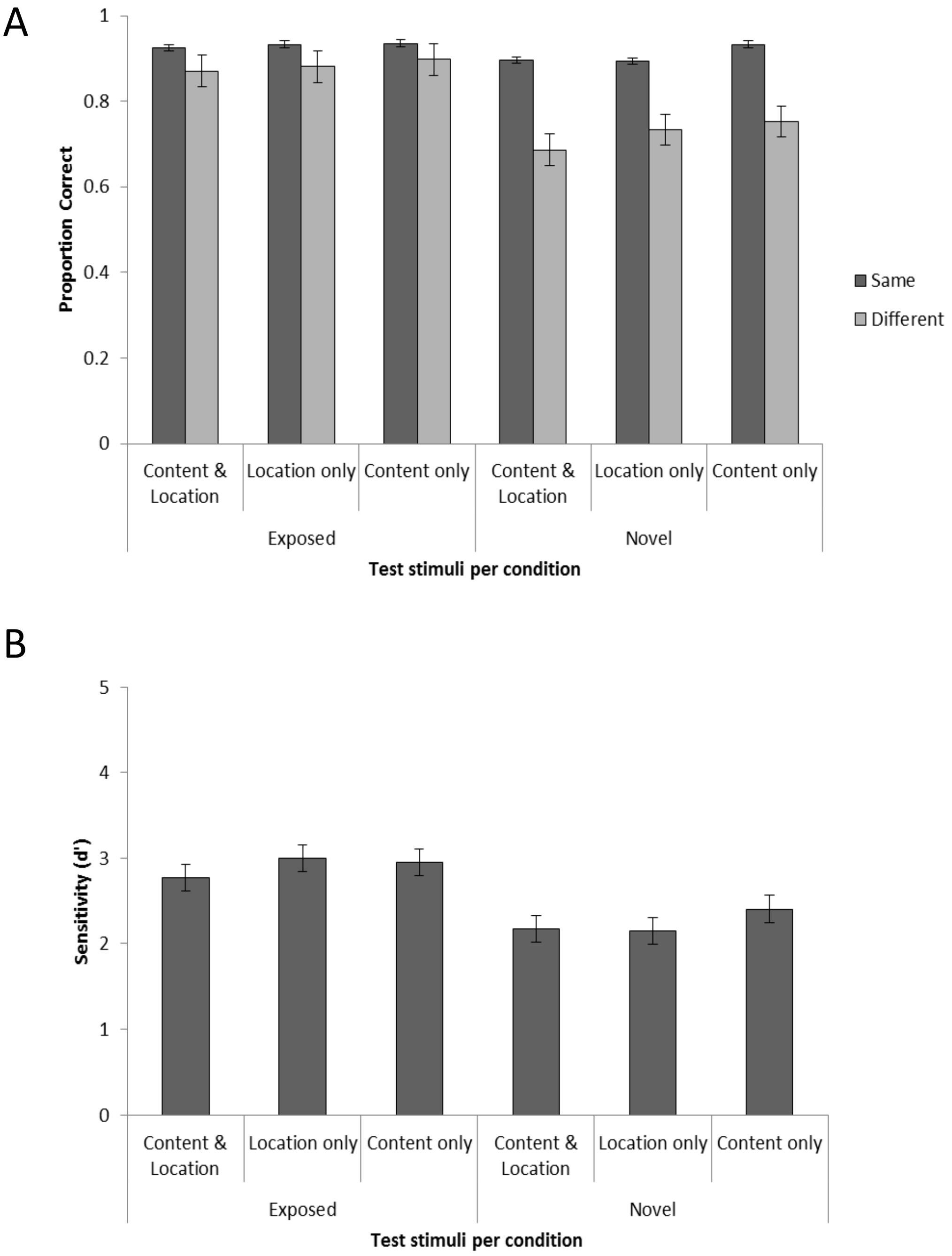

Figure 7: Panel A shows the test data from Experiment 5 as mean proportion correct (with SEM). Performance at chance level is 50%. Data are presented as a function of trial type (Same or Different), stimulus type (Exposed or Novel), and condition (Content and Location, Location only, Content only). Panel B shows mean (with SEM) sensitivity scores (*d'*) for each stimulus type and condition.

location of the unique features, or in either the content or location of these features alone. This suggests that differences in either location alone, or content alone, are sufficient to support perceptual learning. That said, as both unique features appeared in the same place in the Content only condition, then performance here could also be due to focusing on that one place. Regardless, the fact that all three groups showed equivalent exposure learning effects implies that perceptual learning in the Content and Location condition might have multiple, redundant, determinants. This makes it difficult to unambiguously attribute learning to either a change in the processing of the content of the unique features (as has typically been done) or of their location. As I will now turn to an alternative means of investigating the contribution of location and content I will defer further consideration of this issue to the General Discussion (section 5.2.2).

3.3 Experiment 6

Experiment 6 examined whether learning based on exposure to stimuli that differed in the content and location of unique features would transfer to stimuli for which only one of the content or location was maintained from the exposure phase. Figure 8 shows examples of the stimuli used, and Table 5 summarizes the design. All subjects were exposed to stimuli which differed in terms of the content and location of the unique features (see Figure 8-I): For example, one unique feature (A) was added to the top left of the background (X) and a second unique feature (B) was added to the top right. At test, all subjects were tested with these exposed stimuli (the Exposed condition). All subjects were also tested with stimuli containing two novel unique features, C and D, that were presented at the same location as A and B had appeared in during the exposure phase (the Location-Same condition: Figure 8 Panel II). If the content of unique features is critical to what is learnt during exposure, then performance in the Exposed condition will be superior to that in the Location-Same

Table 5

Design of Experiment 6

Group	Condition	Exposure	Test
Content-Same	Exposed	AX/BX	AX/BX
	Location-Same		C*X/D*X
	Location- Different (Content Same)		A*X/B*X
Content-Different	Exposed	AX/BX	AX/BX
	Location-Same		C*X/D*X
	Location-Different (Content Different)		EX/FX

Note: A-E represent unique features while X and Y represent common background checkerboards (see Experiment 6 methods for details and Figure 6 and 8 for examples). Stimuli C*X/D*X comprise new unique features but in the same location that A and B were presented (Figure 8 Panel II). Stimuli A*X/B*X comprise the same unique features but in a new location (Figure 8 Panel III). Stimuli EX/FX comprise the new unique features in a new location (Figure 8 Panel IV).

condition. However, if learning where to look for differences between stimuli is sufficient to support exposure effects, then performance in these two conditions will be equivalent. Finally, all subjects were tested with stimuli where the unique features appeared in different places from that in which A and B were presented during the exposure phase (the Location-Different condition). For half of the participants, these test stimuli used the previously exposed unique features (A and B) but in a different location (see Panel III of Figure 8 – Location-Different–Content-Same), while for the remainder of the participants, the test stimuli used novel unique features (see Panel IV of Figure 8 – Location-Different–Content-Different). For the purposes of description, the participants receiving the Location-Different–Content-Same condition will be referred to as the Content-Same group, while the participants receiving the Location-Different–Content-Different condition will be referred to as the Content-Different group. If the exposure to the unique features supports perceptual learning, regardless of location, then performance in the Location-Different–Content-Same condition will be superior to performance in the Location-Different–Content-Different condition. However, if exposure effects are entirely determined by learning where to look for differences, with no independent contribution of the content of those differences, then discrimination performance in these conditions will be equivalent.

3.3.1 Method

Participants

Participants consisted of 24 undergraduate students, between the ages of 18-24. They were recruited from the School of Psychology at Cardiff and participated in return for course credit.

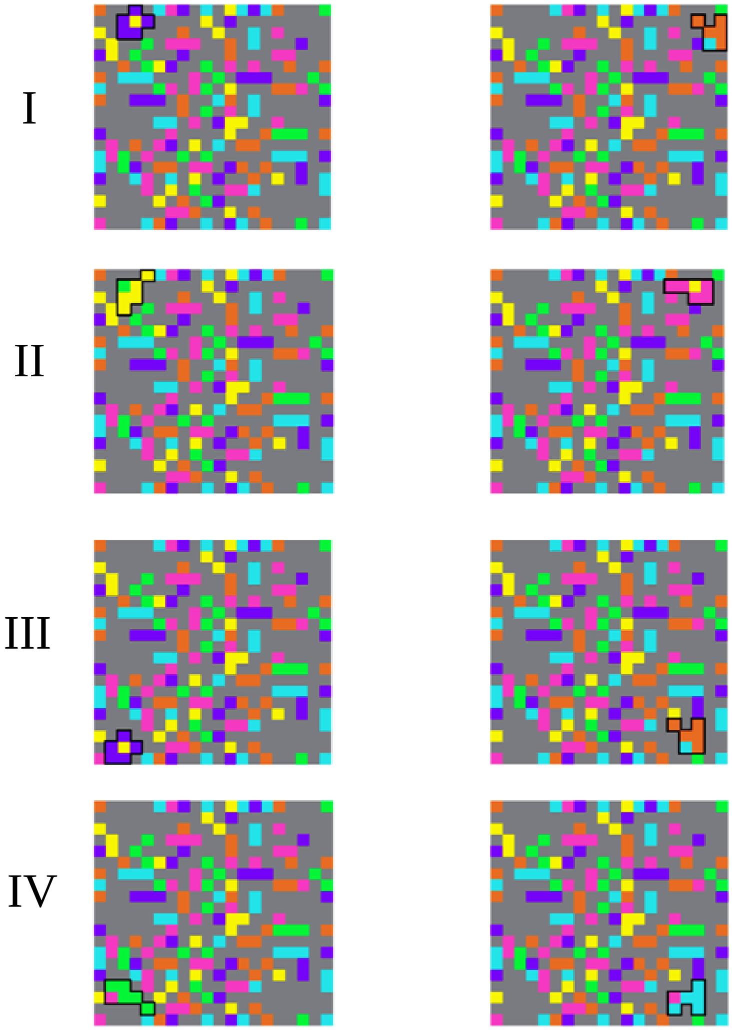

Figure 8: Examples of stimuli used in Experiments 6 and 7. Panel I represents the Exposed condition (e.g., AX and BX). The remaining checkerboards are examples of the transfer tests after exposure to AX and BX. Panel II shows stimuli with new unique features at the same location as the unique features used in exposure (Location-Same). Panel III shows stimuli with the same unique features, but in different locations (Location-Different–Content-Same). Panel IV shows stimuli with new unique features at a new location (Location-Different–Content-Different). Each of the checkerboards features a unique feature; displayed on the common element X, highlighted by a lack border to distinguish it from the common element (this outline was not present during the experiments).

Apparatus and Stimuli

Stimuli consisted of 20 x 20 colour checkerboards, created in the same way as in Experiment 6, save that the unique features could be added in six different locations (roughly top, middle, and bottom on right or left). There were two distinct common backgrounds (X and Y), and two sets of distinct unique features (A-F and G-L: comparison of Panel I of Figure 6 with Panel I of Figure 8 illustrates these differences). Thus, there were two entirely separate sets of stimuli allowing each subject to run through the basic exposure and the test procedure twice (see below).

Design and Procedure

The key test phase involved participants making same/different judgements under multiple conditions which might undermine any transfer of perceptual learning from the exposure to the test phase. Thus, all participants were given a run through the exposure and test procedure similar to the Content and Location condition from Experiment 5 (the only difference being that the Novel condition comprised only novel unique features and locations but the same common background as the exposed stimuli). Each of the two background patterns (X/Y), and each of the unique feature sets (A-F/G-L) were used equally often across participants. The feature sets were assigned to conditions in pairs (e.g., A/B, C/D, E/F) such that each pair was used equally often as the exposed or novel pair (with one pair from the set not being used).

In this basic phase the application of the unique stimuli to the backgrounds was constrained such that A/B (or G/H) were applied to the top left or right of the background, C/D (or I/J) were applied to the middle left or right, and E/F (or K/L) applied to the bottom left or right. Participants showed the typical advantage for discriminating exposed over novel stimuli in this phase[9]. Thus, the results will not be considered further.

Following this, all participants received a second phase of intermixed exposure in which two checkerboards were presented in alternation 60 times each. This used the set of stimuli that had not been seen in the practice run (that is, new unique features, common background, and exposure locations). As noted in Table 5, after the exposure phase participants were given a same-different discrimination task. Half the participants received test trials in three conditions; with the same stimuli as in the exposure phase (the Exposed condition, see Figure 6 Panel I), with new unique features placed in the same positions as the unique features from the exposure phase (the Location-Same condition, see Figure 8 Panel II), and with the same unique features from the exposure phase but in a new location (the Location-Different–Content-Same condition, see Figure 8 Panel III). The other half of the participants also received test trials in which the location of the critical stimuli was different, but in this case the content was also different (the Location-Different–Content-Different Condition, see Figure 8 Panel IV) – these participants also received the Exposed and Location-Same conditions as described above. The unique feature set and common background not used in the practice phase were used here. For participants in the Content-Same group, the feature sets were assigned to conditions in pairs (e.g., A/B, C/D, E/F) such that each pair was used equally often in the Exposed and Location-Same/Location-Different

---

[9] It is worth noting that the two phases, one with two types of test trial (Exposed/Novel) and one with three types of test trial (Exposed/Location-Same/Location-Different) were originally planned to be counterbalanced such that each occurred first or second equally often. However, due to a human error, the basic phase was always run first. However, the results of this, and the following, experiment indicated that the transfer trials did not compromise the basic exposure effect (which was the concern that led me to include the basic condition in the first place). Thus the ordering of the tests in the current experiment seems unlikely to have compromised the interpretation of the current results.

conditions (with a third pair not presented for each participant). For participants in the Content-Different group, the feature sets were assigned such that each pair was used equally often in the exposed, Location-Same, and Location-Different conditions. The location of the stimuli was constrained such that the general region of the background where novel stimuli appeared in the practice phase was used as the location for the Exposed condition, and the stimuli in the Location-Different conditions were presented at the unused set of locations from the practice phase. Therefore, across participants, the assignment of stimuli was counterbalanced such that each of the unique features (A-L), each of the possible locations (top, middle, and bottom on right or left), and each of the two background patterns (X and Y) was used equally often for all conditions.

The test phase comprised three blocks of 30 trials each. Within each block there were 10 trials from each of the test conditions (5 same and 5 different). The order of trials was randomised within a block and participants were allowed to rest between blocks. Participants followed the same instructions as in Experiment 5 and all other details (e.g., stimulus timings) not explicitly mentioned here were the same as in Experiment 5.

3.3.2 Results and Discussion

Figure 9 Panel A shows the test data as mean proportion of correct responses for the same-different responses for the three test conditions (Exposed, Location-Same, and Location-Different), as a factor of group (Content-Same on the left, Content-Different on the right). Inspection of the figure suggests that performance was equivalent in the Exposed and Location-Same conditions, and both of these were superior to the Location-Different conditions. There is little suggestion of any difference between the Content-Same and Content-Different groups for any of the test conditions. In addition, performance was generally higher on same than different trials, and differences between conditions were larger on different trials.

The test data was analysed with a mixed ANOVA with a between subject factor of group (Content-Same or Content-Different), and within-subject factors of test condition (Exposed, Location-Same, or Location-Different) and test trial type (Same or Different). There were main effects of test trial type, $F(1,22) = 16.17$, $p = .001$, MSE = 0.054, test condition, $F(2,44) = 42.33$, $p < .001$, MSE = 0.028, and an interaction between them, $F(2,44) = 21.56$, $p < .001$, MSE = 0.036. Simple effects analyses of the interaction revealed significant effects of test condition on the different trials for Exposed vs. Location-Different, $F(1,22) = 39.79$, $p < .001$, MSE = 0.156, $B_{01} <.001$, and Location-Same vs. Location-Different, $F(1,22) = 40.91$, $p < .001$, MSE = 0.135, $B_{01} <.001$. However, there was no significant effect for Exposed vs. Location-Same, $F < 1$, $B_{01} = 5.333$ on different trials. There were also some differences between conditions on the same trials (Exposed vs. Location same, $F < 1$, $B_{01} = 5.382$, Exposed vs. Location-Different, $F(1,22) = 4.56$, $p = .044$, MSE = 0.018, $B_{01} = .774$, Location-Same vs. Location-Different, $F(1,22) = 3.49$, $p = .075$, MSE = 0.015, $B_{01} = .1.388$). That is, discrimination performance was equivalent for the exposed stimuli and for stimuli that had novel unique features appearing in the same place as the unique features of the exposed stimuli. Discrimination performance in both these conditions was superior to the test stimuli that had unique features in a different place to that of the exposed stimuli. There was no main effect of group, nor any interaction involving this factor, largest $F(1,22) = 1.32$, $p = .236$, MSE = 0.054, for the trial type by group interaction. This is not particularly surprising because the test trials for the Exposed and Location-Same conditions were the same for Content-Same and Content-Different groups (albeit that the accompanying test trials were different, and so the Content-Same and Content-Different groups differed in the number times that the exposed unique features appeared in the test phase). Critically, there was no hint of a significant difference between the Content-Same and Content-Different groups for the Location Different condition ($F < 1$ for both same $B_{01} =$

4.511 and different trials B_{01} = 5.742), despite the fact that in the Content-Same group the stimuli tested had the same unique features as the Exposed condition, while for the Content-Different group the unique features were novel.

Panel B of Figure 9 displays the mean *d'* sensitivity scores as a factor of group (Content Same and Content Different) and test condition (Exposed, Location-Same, or Location-Different). Analysis of the *d'* scores showed a main effect of test condition, $F(2,44) = 37.84$, $p < .001$, MSE = .941. Simple effects analyses revealed that the Exposed and Location-Same conditions did not differ from each other (Exposed vs. Location Same, $F < 1$, B_{01} = 5.229) and that both of these conditions resulted in higher *d'* scored than the Location-Different conditions (Exposed vs. Location-Different, $F(1,22) = 50.90$, $p < .001$, MSE = 2.265, B_{01} <.001, Location-Same vs. Location-Different, $F(1,22) = 56.78$, $p < .001$, MSE = 1.721, B_{01} <.001). There was no main effect of group, nor any interaction involving this factor, largest $F(1,22) = 2.36$, $p = .106$, MSE = .459, for the test condition by group interaction. Like the standard analysis there was no hint of a difference between the Content-Same and Content-Different groups for the Location Different condition ($F < 1$, B_{01}= 6.120).

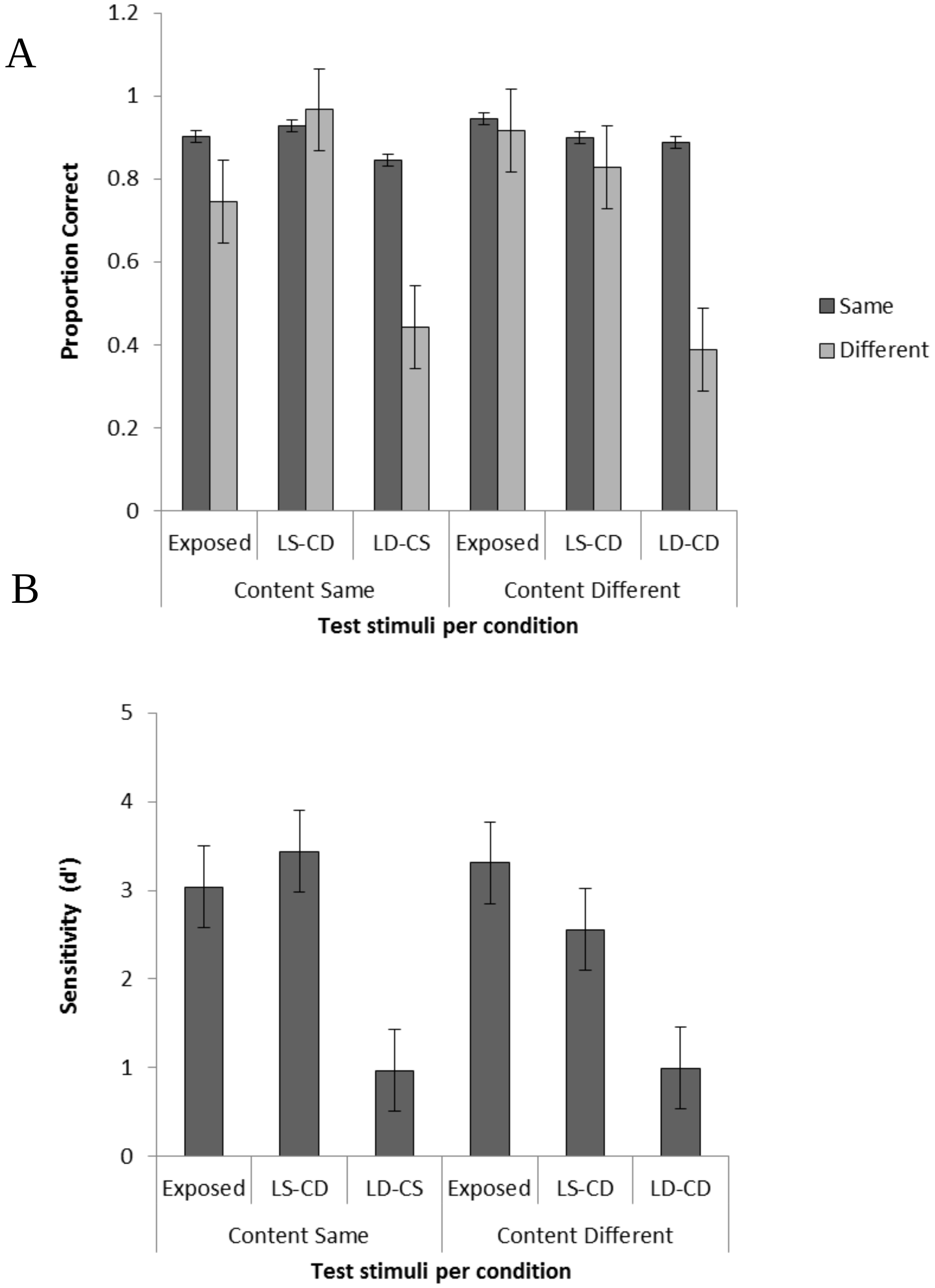

Figure 9: Panel A shows the test data from Experiment 6 as mean proportion correct (with SEM). Performance at chance level is 50%. Data are displayed as a function of test trial type (Same or Different), test condition (Exposed, Location-Same (LS-CD), or Location-Different (LD-CS/LD-CD), and training group (Content-Same or Content-Different). Panel B shows mean (with SEM) sensitivity scores (*d'*) for each test condition and training group.

In summary, discrimination between stimuli that had novel unique features was equivalent to discrimination between exposed stimuli, as long as the novel features appeared at the same location as the unique features that had been present in the exposed stimuli. Moreover, discrimination between stimuli that had the exposed unique features at a different location than that at which they appeared during initial exposure was no better than was discrimination of entirely novel stimuli. That is, the improvement in discrimination produced by exposure transferred entirely to novel content at the exposed location, but not at all to the exposed content at a novel location. These conclusions are unaffected by whether the data was examined as proportion correct or *d'*, and the interpretation supported by standard null hypothesis testing was bolstered by Bayes factor analyses indicating that the absence of significant differences between critical conditions. That is, Bayesian analysis between the critical comparisons; Exposed and Location-Same conditions, and the Content-Same and Content-Different groups for the Location-Different condition, reflects genuine evidence in favour of true absences of any effects.

3.4 Experiment 7

While the results of Experiment 6 are certainly consistent with the idea that perceptual learning with the current stimuli is entirely determined by learning where to look for the critical differences, rather than learning about what those differences are, there are two aspects of that experiment that might have led to an underestimation of learning about the content of the unique features. Firstly, the comparison between transfer to exposed features at a new location and a totally novel control was between-subject, while the examination of transfer based on location was within-subject. To the extent that between-subject comparisons are less powerful than within-subject comparisons then Experiment 6 might have underestimated the former effect. Secondly, the fact that test trials examining content- and location-based transfer were intermixed puts these two effects into direct competition, as any tendency to attend to a particular location would reduce the ability to detect the exposed features when they appeared at a different location[10]. Moreover, for two thirds of the test trials, the critical unique features (either exposed or novel) appeared at the same locations used for the exposed stimuli during exposure, and only one third of the test trials used new locations. The preponderance of trials using this exposed location might have further enhanced any tendency for participants to focus on location to the exclusion of content. While the fact that there was no hint of content-based transfer and there was excellent location-based transfer leaves the relative importance of the two effects in no doubt, the complete absence of content-based learning with these stimuli remains to be established.

The design of Experiment 7 (see Table 6) addressed these issues by examining location and content based transfer in separate groups of participants. All participants received two exposure and test runs, in both of these participants were exposed in an alternating fashion to a pair of checkerboards with unique features which differed in both location and content (as in Panel I of Figure 8). In one run (Exposed conditions) all participants received a same/different test phase with these exposed stimuli (Exposed), as well as with stimuli that had different unique features in different locations on the same common background (Exposed-Control). For participants in group Content, the other run (Transfer conditions) involved a same/different test phase with stimuli that retained the same unique feature content as seen in the exposure phase, but moved to a different location (Transfer). There were also test trials with stimuli that had novel unique features in different locations on the same common background (Transfer-Control). For participants in group location, the transfer conditions involved a same/different test phase with stimuli that changed the unique feature content from the exposure phase, but retained the location (Transfer). There were also test trials with stimuli that had novel unique features in different

---

[10] A similar issue is present in the test trials of Experiment 3 of Wang *et al*. (2012).

locations on the same common background (Transfer-Control). In short, the experiment comprised a within-subject manipulation of whether the test stimuli had been exposed in any fashion (Exposed & Transfer vs. Exposed-Control & Transfer-Control) and a within-subject manipulation of whether the test stimuli were exactly the same as in the exposure phase or not (Exposed vs. Transfer). Whether the Transfer conditions maintained the content or location of the exposed unique features (Content vs. Location) was assessed in separate groups. By assessing the transfer of learning based on content and location in separate participants and sessions, this design avoided the direct competition between attending to location and content that may have been present in previous studies.

If the results from Experiment 6 are reliable then in the Location group the difference in performance on the discrimination task between the Exposed and Exposed-Control conditions should be the same as that between the Transfer and Transfer-Control conditions. In contrast, the Content group should only show a difference in discrimination between the Exposed and Exposed-Control conditions, but not show any difference between the Transfer and Transfer-control conditions. That is, there should be transfer based on location, but not the content, of unique features.

Table 6

Design of Experiment 7

Group	Condition	Exposure	Test
Content	Exposed	AX/BX	AX/BX
	Exposed-Control		CX/DX
	Transfer	GY/HY	G*Y/H*Y
	Transfer-Control		KY/LY
Location	Exposed	AX/BX	AX/BX
	Exposed-Control		CX/DX
	Transfer	GY/HY	I*Y/J*Y
	Transfer-Control		KY/LY

Note: A-L represent unique features, while X and Y represent common background checkerboards (see Experiment 7 methods for details and Figure 8 for examples). Transfer stimuli (G*Y/H*Y) comprise the same unique features but in a new location (Figure 8 Panel III). Transfer stimuli (I*Y/J*Y) comprise new unique features but in the same location that G and H were presented (Figure 8 Panel II).

3.4.1 Method

Participants Apparatus and Stimuli

Participants consisted of 48 Undergraduate students, between the ages of 18 and 25, recruited from the School of Psychology at Cardiff University. They received course credit in return for their participation. Stimuli consisted of 20 x 20 colour checkerboards created as in Experiment 2 that were presented using the same equipment as described previously.

Design and Procedure

All participants were given two runs through an exposure and test sequence. Participants followed the same instructions as those in Experiment 5 during both runs. The basic exposure and test procedures timings were as outlined in Experiments 5 and 6, so only the differences are noted here.

In each run, participants were exposed to a pair of checkerboards that shared a common background (X in one run and Y in the other), and were distinguished by unique features that differed in both content and location. As outlined in Table 6, during the Exposed condition the exposure and test run and the same/different discrimination task consisted of trials with exactly the same stimuli as in exposure, plus novel controls (these used the same common background, but had new unique features presented at a new location). During the Transfer condition, the same/different discrimination task consisted of trials with stimuli that shared some aspect of the exposed stimuli, plus novel controls (these used the same common background, but had new unique features presented at a new location). For half of the participants (i.e., the Content group), the transfer test stimuli retained the content of the unique features from exposure, but moved them to a new location. For the remaining participants (i.e., the Location group), the transfer test stimuli retained the location of the unique features, but changed the content. The test phase for each of the two runs comprised two blocks of 40 trials each. Within each block, there were 10 trials from each of the test conditions (5 same and 5 different). Within each block trial order was randomised. Between blocks participants were able to pause before continuing by pressing the spacebar (there was also an opportunity to pause mid-way through each block).

The presentation order of the exposure and transfer runs were counterbalanced so that half the participants were given the Transfer run first and the other half of the participants were given the Exposure run first. Within these groups, each of the two background patterns (X/Y) and each of the unique feature sets (A-F/G-L) were used equally often in the exposure and transfer runs. For the exposure run, the feature sets were assigned to conditions in pairs (e.g., A/B, C/D, E/F), such that each pair was used equally often as the exposed or novel pair (with one pair from the set not being used for each participant). In the Location group, the remaining feature sets were assigned to conditions in pairs (e.g., G/H, I/J, K/L) such that each pair was used equally often as the transfer-exposed, transfer-test, or transfer-control stimuli. In the Content group, these remaining features were assigned such that each pair was used equally often as transfer-exposed or transfer-control stimuli (with one pair from the set not being used for each participant). The locations at which the unique features appeared were assigned such that each set of locations (top, middle, bottom, on left and right) were used equally often across participants for the exposed condition, with the exposed-control stimuli appearing equally frequently at one of the other two locations. The exposure phase of the transfer run stimuli always appeared at the locations not used in the exposure run. During the test phase, in the Content group, the transfer, and transfer-control stimuli appeared in the other two locations with equal frequency (thus for half the subjects the transfer-control stimuli appeared where the exposed stimuli were placed and for the other half the transfer-control stimuli appeared where the exposed-control stimuli were placed). In the test phase, in the Location group, the transfer-control stimuli appeared equally often in either the location where the exposed, or exposed-control stimuli were placed. Therefore, across participants, the assignment of stimuli to condition ensured that each of the common backgrounds (X or Y), each of the unique features (A-L), and each of the possible locations (top, middle, and bottom on right or left) was used equally often for all conditions. Moreover, the assignment of locations was constrained such that for half of the subjects attending to the same location in each of the exposure and transfer runs would assist performance in the second run, and for half of the subjects it would hinder performance in the second run[11].

3.4.2 Results

Figure 10 displays the proportion of correct responses as a factor of group (Content on the left, Location on the right), stimulus type (Exposed/Transfer vs. Control), and test trial type (Same/Different). As has been seen previously, performance was generally better on same than different trials, with differences between conditions carried largely by the different trials. Turning first to the Content group, performance was greater in the Exposed than Exposed-Control condition, but there was little or no difference between the Transfer and Transfer-Control conditions. In contrast, for the Location group, the difference between Exposed and Exposed-Control was equivalent to the difference between Transfer and Transfer-Control conditions.

This data was initially subjected to a mixed ANOVA with a between subjects factor of group (Content or Location), and within subject factors of test trial type (Same or Different), transfer condition (Exposed or Transfer), and exposure treatment (Exposed/Transfer vs. Control). Consistent with the description of the results above, there was a 3-way interaction between group, Transfer condition, and Exposure condition, $F(1,46) = 6.79$, $p = .012$, MSE = 0.052, indicating that the relative size of the Exposed vs. Exposed-Control and Transfer vs. Transfer-Control differences was influenced by whether the transfer

---

[11] An initial analysis of the data indicated that there were in fact no carry-over effects of this type, and thus test order was not included in the reported analyses.

conditions were content- or location-based. There was also a significant 4-way interaction, $F(1,46) = 4.72$, $p = .035$, MSE = 0.050, which is consistent with the 3-way interaction being driven by performance on the different trials. In order to explore the different effects of content or location-based transfer indicated by these interactions, separate 3-way ANOVAs were performed for each of group Content and group Location[12].

Taking first the Content group, the most theoretically important results were the significant interactions between Transfer condition and Exposure condition, $F(1,23) = 10.27$, $p = .004$, MSE = 0.046 (which demonstrates that discrimination was better in the Exposed condition rather than Transfer Conditions), and the interaction between test trial type, Transfer condition and Exposure condition, $F(1,23) = 11.55$, $p = .002$, MSE = 0.048 (which suggests that the previous interaction was largely carried by the different trials)[13]. Simple effects analyses of the three-way interaction revealed that there was a difference between Exposed and Exposed-Control for different trials, $F(1,23) = 23.35$, $p < .001$, MSE = 0.007, $B_{01} < .001$, but not for same trials, $F < 1$, $B_{01} = 5.958$. There were no differences between Transfer and Transfer-Control on either same, $F < 1$, $B_{01} = 5.688$, or different trials, $F < 1$, B_{01}

---

[12] The remainder of the 4-way ANOVA was as follows. As with the previous experiments, performance was generally better on same than different trials, $F(1,46) = 57.13$, $p < .001$, MSE = 0.087. There was also a significant effect of Exposure treatment whereby novel control stimuli were discriminated less well than exposed/transferred stimuli, $F(1,46) = 12.50$, $p < .001$, MSE = 0.047. These two factors interacted with one another, $F(1,46) = 14.72$, $p < .001$, MSE = 0.045. There was also a significant 3-way interaction between test trial type, Transfer condition, and Exposure condition, $F(1,46) = 6.56$, $p = .014$, MSE = 0.050. No other main effects or interactions were significant, largest $F(1,46) = 2.72$, $p = .106$, MSE = 0.052, for the interaction between Transfer condition, and Exposure condition.

[13] The remainder of the ANOVA revealed significant effects of Exposure, $F(1,23) = 7.68$, $p = .011$, MSE = 0.061, test trial type, $F(1,23) = 48.11$, $p < .001$, MSE = 0.062, and an interaction between them, $F(1,23) = 7.06$, $p = .014$, MSE = 0.063, (reflecting the usual advantage for "same" trials, and the fact it interacted with whether stimuli were exposed or novel). The main effect of Transfer condition did not reach conventional levels of significance, $F(1,23) = 3.56$, $p = .072$, MSE = 0.040.

= 6.207. That is, discrimination between stimuli that shared their unique features with exposed stimuli but with these unique features appearing at a new location, was no better than with stimuli that had entirely novel unique features.

Turning to the Location group, the key results here were that there was a significant effect of Exposure, $F(1,23) = 8.21$, $p = .009$, MSE = 0.034, but that there was no significant effect of Transfer condition, $F < 1$, and critically no interaction between Transfer condition and Exposure, $F < 1$, nor any other significant interaction involving Transfer condition (largest $F(1,23) = 1.12$, $p = .291$, MSE = 0.034, for the interaction between Transfer condition and test trial type)[14]. That is, discrimination of novel control stimuli was worse overall than for the Exposed/Transfer conditions combined, and there was no difference in discrimination performance between Exposed and Transfer conditions. In order to match the analysis performed on the Content group I also examined the simple effects for the 3-way interaction (even though this was not significant here): The difference between Exposed and Exposed-Control for different trials approached standard levels of significance, $F(1,23) = 4.03$, $p = .057$, MSE = 0.004, $B_{01} = 1.041$, but not for same trials, $F(1,23) = 1.57$, $p = .223$, MSE = 0.001, $B_{01} = 3.002$. The difference between Transfer and Transfer-Control for different trials approached standard levels of significance, $F(1,23) = 3.06$, $p = .094$, MSE = 0.008, $B_{01} = 1.561$, but not for same trials, $F(1,23) = 1.29$, $p = .268$, MSE = 0.001, $B_{01} = 3.405$. That is, discrimination between Exposed stimuli was entirely equivalent to that with stimuli that had novel unique features which appeared in the same location to those of the exposed stimuli (albeit that discrimination in both of these conditions was numerically smaller than that in the Content group).

---

[14] The remainder of the ANOVA revealed a significant effect of test trial type, $F(1,23) = 18.17$, $p < .001$, MSE = 0.112, and an interaction between test trial type and Exposure type, $F(1,23) = 8.65$, $p = .007$, MSE = 0.027, which again reflects an advantage for “same” trials, and the fact it interacted with whether stimuli were exposed or novel.

Analysis of *d'* suggested a similar pattern of results to the standard analysis as displayed in Panel B of Figure 10. Like the proportion analysis, there was a 3-way interaction between group, Transfer condition, and Exposure condition, $F(1,46) = 10.94$, $p = .007$, MSE = 1.395, indicating that sensitivity scores followed a similar trend to the previous analysis. Turning first to the Content group, there was a significant interaction between Transfer condition and Exposure condition, $F(1,23) = 13.04$, $p = .004$, MSE = 1.239. Simple effects analyses of the interaction revealed that there was a difference between Exposed and Exposed-Control, $F(1,23) = 19.41$, $p < .001$, MSE = 0.147, $B_{01} = .008$. There were no differences between Transfer and Transfer-Control, *F*s < 1, $B_{01} = 5.204$. That is, discrimination between stimuli that shared their unique features with exposed stimuli but with these unique features appearing at a new location, was no better than with stimuli that had entirely novel unique features. The remainder of the ANOVA revealed a significant effect of exposure $F(1,23) = 11.082$, $p = .003$, MSE = 1.959, and that the effect of Transfer condition approached standard levels of significance, $F(1,23) = 3.85$, $p = .062$, MSE = 1.331.

For the Location group, the key result here was a significant effect of Exposure, $F(1,23) = 11.70$, $p = .002$, MSE = .944, but no effect of Transfer condition, $F < 1$. Most critically, there was no significant interaction between Exposure condition and Transfer condition, $F < 1$. In order to match the analysis performed on the Content group, in both the standard and sensitivity analysis, I also examined the simple effects interaction (even though this was not significant here): The difference between Exposed and Exposed-Control approached standard levels of significance, $F(1,23) = 3.02$, $p = .095$, MSE = 0.138, $B_{01} = 1.582$, while the difference between Transfer and Transfer-Control reached standard levels of significance, $F(1,23) = 5.82$, $p = .024$, MSE = 0.070, $B_{01} = .512$. Again, discrimination between exposed stimuli was equivalent to that with stimuli that had novel unique features which appeared in the same location to those of the exposed stimuli.

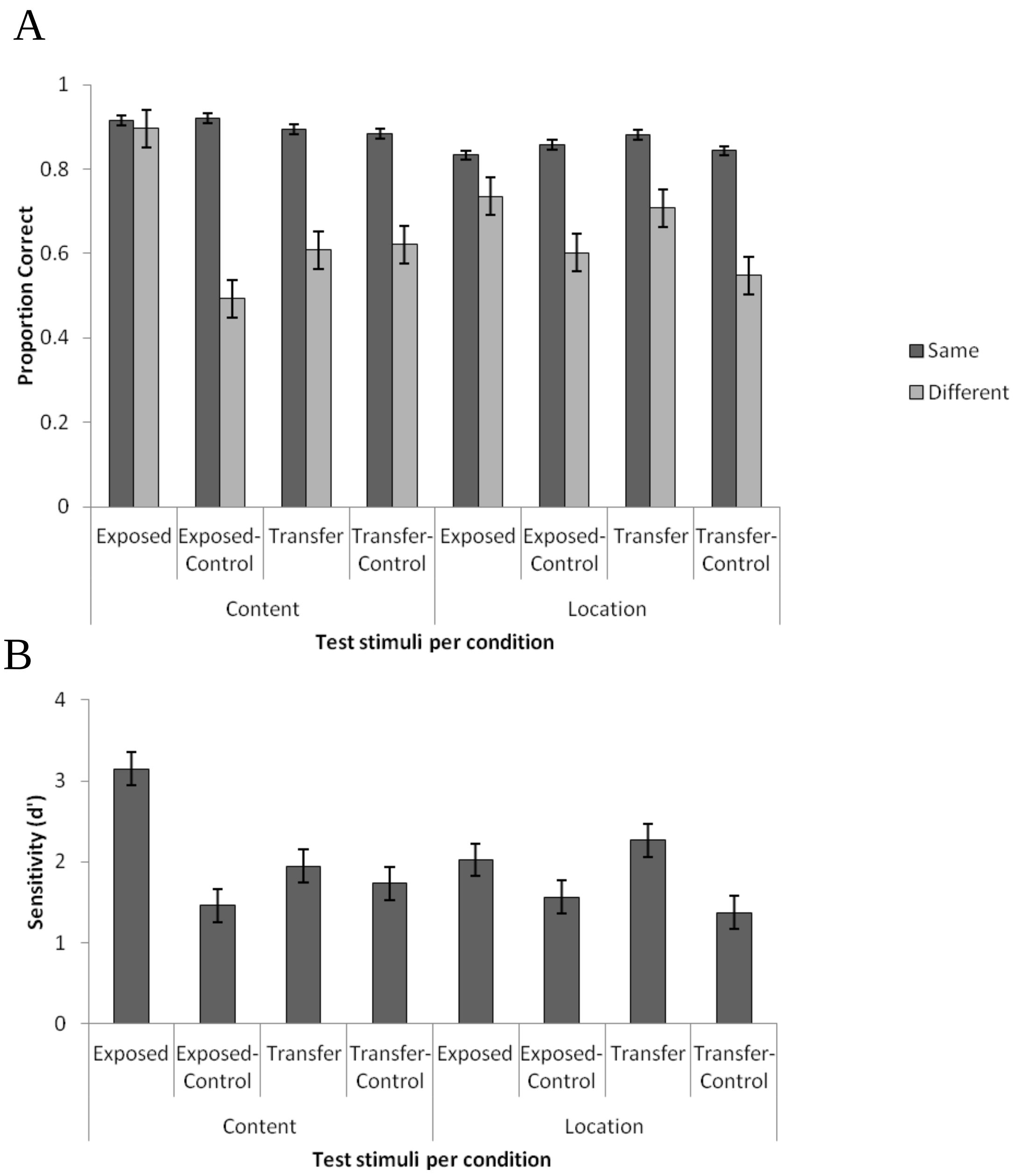

Figure 10: Panel A shows the test data from Experiment 7 as mean proportion correct (with SEM). Performance at chance level is 50%. The data are organised by Transfer group (Content group on the left, Location group on the right), and are presented as a function of test trial type (Same or Different), test condition (Exposed or Transfer), and Exposure condition (Exposed/Transfer or Exposed-Control/Transfer-Control). Panel B displays mean (with SEM) sensitivity scores (*d'*) for both Transfer groups presented as a function of Test and Exposure conditions.

3.4.3 Discussion

In summary, after training with stimuli that differed in both the content and location of the unique features, performance on the transfer test was determined by whether this involved the exposed content at a new location (Content group), or new content that was presented at the same location (Location group). Performance with exposed stimuli was superior to performance due to content-based transfer, with no evidence of any difference between the Transfer and Transfer-Control conditions for the Content group. As in Experiment 6, Bayes factor analyses supported the view that the lack of a significant difference genuinely supports the idea that there was no content-based transfer. In contrast, performance due to location-based transfer was no different from performance with the exposed stimuli. While the simple effects analyses of the Transfer and Transfer-Control conditions for the Location group offer only equivocal support for the presence of a location-based transfer (as these were significant for the *d'* analysis but only approached standard significance levels for the proportion correct analysis), it should be remembered that there was no difference in the size of the exposure and location-based transfer effects in this experiment (if anything, the transfer effects were bigger), and that in Experiment 6 very reliable location based transfer effects were observed.

Therefore, the discrimination phase of Experiment 7 replicated the key findings from Experiment 6. That is, exposure-dependant improvements in discrimination ability transferred to new test stimuli when novel unique features of the to-be-discriminated stimuli appeared in the same location as the unique features of the exposed stimuli. However, when the to-be-discriminated stimuli maintained the same unique features, but presented them at a new location, there was no transfer of the exposure-dependant improvements in discrimination.

3.5 Chapter Discussion

The three experiments reported here examined the ability to discriminate between checkerboard stimuli made similar by placing unique features on a common background. In Experiment 5, exposure to stimuli that could be distinguished on the basis of either the content or location of the unique features resulted in improvements in discrimination performance over novel stimuli that were equivalent to those produced by exposure to stimuli that differed in both the content and location of the unique features. In Experiments 6 and 7, the improvement in discrimination performance produced by exposure to stimuli that differed in the content and location of the unique features transferred entirely to stimuli that had new unique features in the same location as the unique features of the exposed stimuli. In contrast, there was no suggestion of any transfer of exposure-produced improvement in discrimination performance when the unique features of the exposed stimuli were moved to a different location. The fact that the content of unique features was unable to support any transfer of exposure learning, but that the location of those unique features supported complete transfer of learning is entirely consistent with the improvements in discrimination on the basis of exposure being due entirely to participants learning where to look (at least with the type of stimuli and procedures examined here).

The findings of these experiments seem to link well with the issues of face recognition, and the differences in the features of a face which are utilised by viewers. Briefly put, recognition is faster and more accurate on the basis of the internal features rather than external features for familiar faces (Ellis *et al*., 1979; Young *et al*., 1985). Moreover, the ability to accurately identify internal features increases with increasing familiarity (e.g., Osborne & Stevenage, 2008), and there is evidence that experience or familiarisation with a face promotes better recognition through these internal features (e.g., O’Donnell & Bruce, 2001; see also Clutterbuck & Johnston 2002, 2005). Thus, this fits with the notion that attending to the crucial location of the diagnostic features is a product of experience with a stimulus, and that this is due to experience increasing the salience of the most diagnostic aspect of a stimulus. In the case of faces this is thought to be the internal features (Fletcher, Butavicius, & Lee, 2008).

What is critical for theoretical accounts of perceptual learning is the demonstration from the current experiments that it is *only* learning about the location of unique features that matters for checkerboards constructed here. As noted in the introduction, the idea that exposure-produced improvements in discrimination depend on learning about where the critical differences in stimuli might appear, is problematic for all theoretical accounts of perceptual learning that are based on mechanisms involving the content of the exposed stimuli. Obviously, the idea that exposure effects with one type of stimulus is potentially subject to artefacts due to spatial attention (as was seen here), does not mean that content-based mechanisms do not contribute to perceptual learning at all. Indeed few, if any, studies of perceptual learning in non-human animals would admit explanation in terms of deliberate allocation of spatial attention, especially as most such studies have used stimuli such as flavours that cannot be discriminated on location alone (e.g., Blair & Hall, 2003; Dwyer & Honey, 2007; Symonds & Hall, 1995). Moreover, as will be considered in detail in the General Discussion (see section 5.2.2), not all human-based studies are subject to these attentional confounds. Briefly put, not all stimuli used to investigate perceptual learning in humans will afford the possibility for location-based learning observed with these checkerboards. As such, the general implications of these current results for theoretical accounts of perceptual learning in humans, and the implications for prior studies that used directly comparable stimuli will be considered fully in the General Discussion (section 5.2.2).

Chapter 4: Multiple synthesised views

4.0 Abstract

Recent evidence suggests that training using synthesised faces created from a single view and presented at multiple yaw rotations can aid recognition (Liu *et al*, 2009). Chapter 4 reports the ability of commercially available photogrammetric software, (i.e., FaceGen), to produce a similar advantage in a recognition task containing a line-up of faces (an array). All experiments gave participants exposure to target individuals – some of whom were presented along with synthetically produced versions of the individual at different rotations. Subsequently, participants were asked to identify the target from an array of nonexposed foils. The general aim of these experiments was to reveal how recognition could be supported by synthesised views of a face. Experiment 8 found that targets were best identified within an array of front view faces, when they had been previously exposed within a schedule consisting of multiple photographs of different views of a target compared to the same views using computer generated stimuli and a brief presentation of a 30° photograph. Experiment 9 found that computer generated stimuli facilitated the best learning of a target, at a subordinate angle (30°), compared to the controls (Brief exposure and FaceGen 30° exposure), and that this learning was equivalent to exposing photographs of a target displayed at multiple yaw rotations. In Experiment 10, using the same test procedure, repeated presentation of the front view of a face yielded similar performance to that of the synthesised faces, which were again both better than a brief exposure. However, the lack of difference between repeated exposure and the generated stimuli may indicate a limit to the amount of additional information which is conveyed using this software (Experiment 10). These findings strengthen the claim that identifying a target can be improved using multiple synthesised views generated from a single front view of a face, and suggest that this improvement may be affected by the quality of synthesised material.

4.1 Introduction

The experiments reported in Chapter 2 suggest that the effectiveness of comparison-based perceptual learning training with faces is limited, and that any enhanced discrimination gained from exposure fails to transfer to nonexposed faces. That is, comparison between relatively few faces is not an effective means of enhancing face recognition, at least not at a practical level. Returning to these forensic concerns, there remains a need to improve face recognition based on a limited number of images (e.g., a single view of an individual). While Chapter 2 gave exposure to one view of a face, it has been found that exposing multiple views of a face can facilitate improved recognition (Chen & Liu, 2009). In this chapter, ways of enhancing recognition are explored that involve generating different views of the person such that these can be used in training. Three experiments used an engineering technique in which multiple views of a face are generated (see section 1.2.3) and exposed during familiarisation in order to improve subsequent recognition.

Exposures to multiple views are thought to facilitate recognition because they help create an internal representation of the 3D structure of a face such that it is possible to generalise to novel views of a face (see Hole & Bourne, 2010). It has been argued that view point dependence is a function of familiarity (Jiang, Blanz, & O'Toole, 2009), such that familiarity with a face can support generalisation to novel viewing contexts (Jiang, Blanz, & O'Toole, 2007). Indeed, laboratory research suggests that seeing an unfamiliar face in one view does not generalise well to other views (Bruce, 1982), and that the view dependent nature of unfamiliar faces is well established (Longmore *et al.*, 2008; Troje & Bülthoff, 1996).

These findings regarding unfamiliar faces should give us concern regarding police investigations that utilise mug shots of suspects that typically consist of a very few images. Suspects are generally unfamiliar and typically presented in limited views and so the chances for individuals to avoid detection are high. One way in which this could be overcome is by utilising technology that allows a single image of a face to be modelled onto an average face. This method then allows for generation of differential face views based upon this modelled face (e.g., Blanz & Vetter, 1999).

To date, a small number of studies have utilised this synthesised face generation as a means of improving recognition accuracy. For example, Bailenson *et al.* (2004) found high identification for cybernetic busts, albeit not as high as original photographs, and that these type of virtual stimuli are processed in a similar manner to that of face photographs. Liu *et al.* (2009) found that experimental conditions involving learning a photograph alongside synthesised views (constructed based on the original photograph) consistently produced higher accuracy compared to that in the control conditions across experiments. While these results do indeed suggest that computer generated views might assist in learning to recognise a face, they are based on bespoke software, and the effects might be restricted to these generation mechanisms.

The main aim of the current experiments was to use a commercially available software program to examine whether generated views of a face can be used to improve face recognition in general. The generated views were created using SI FaceGen Modeller, a commercially available software program. All experiments involved a period of exposure whereby a target was presented before a recognition task. The way in which the generated faces were used to support recognition was varied over the experiments. The basic task was similar to that used in Chapter 2, in that participants were asked to select a target from a line-up of faces (arrays). Further to this, participants were asked to give an indication of how confident they felt about their decision.

4.2 Experiment 8

The main purpose of Experiment 8 was to examine the potential benefit of exposing participants to synthesised target faces using commercially available software (i.e., FaceGen). These synthesised faces were used to create multiple yaw rotations from a single front-view input. Figure 11 shows examples of the stimuli used, and Table 7 summarises the design.

In the FaceGen condition, participants were shown a target face at 30° left of full face, plus synthetically generated views of the target at other yaw rotations. In the Brief 30° photo condition, only the target face at 30° yaw was presented, while in the Photo condition all yaw rotations were real photographs of the target individual. These three within-subject conditions were all tested by examining the ability to select the targets from arrays of front-facing novel foil faces. If the opportunity to view synthesised multiple rotations of a target does allow for accurate recognition, as suggested by Liu *et al,* (2009), then it is expected that the multiple views in the FaceGen condition should allow for better recognition compared to the Brief 30° photo conditions. Moreover, if the generated faces are entirely equivalent to veridical photographs then performance in the FaceGen and Photo conditions should be equivalent.

4.2.1 Method

Participants:

Eighteen students, aged between 18 and 29, were recruited from Cardiff University and completed the experiment in return for payment of £2. All had normal or corrected-to-normal vision.

Table 7

Design of Experiment 8

Condition	Training sequence	Test Arrays
Brief 30°	+, +, 30°L, +, +, +	Select face from an array of five faces
FaceGen	90°L, 60°L, *30°L, 30°R, 60°R, 90°R	
Photo	90°L, 60°L, 30°L, 30°R, 60°R, 90°R	

Note: 30°-90° represent the different angles that the target face was exposed at, L (Left) or R (Right) represents the direction of the face when looking at the screen, and + symbolises a fixation cross that was presented in the brief condition in place of exposure at different yaw rotations. Within the FaceGen training sequence * (e.g., *30°) indicates that unlike the other exposures this view was a photograph and not a synthesised view. During training exposure, each stimulus was present for 2s with a 1s inter-stimulus interval (ISI). Test arrays required recognition of the presented target to the appropriate stimulus in a choice array consisting of five faces (i.e., 4 foils plus the target individual). See Experiment 8 methods for details and Figure 11 for stimuli examples.

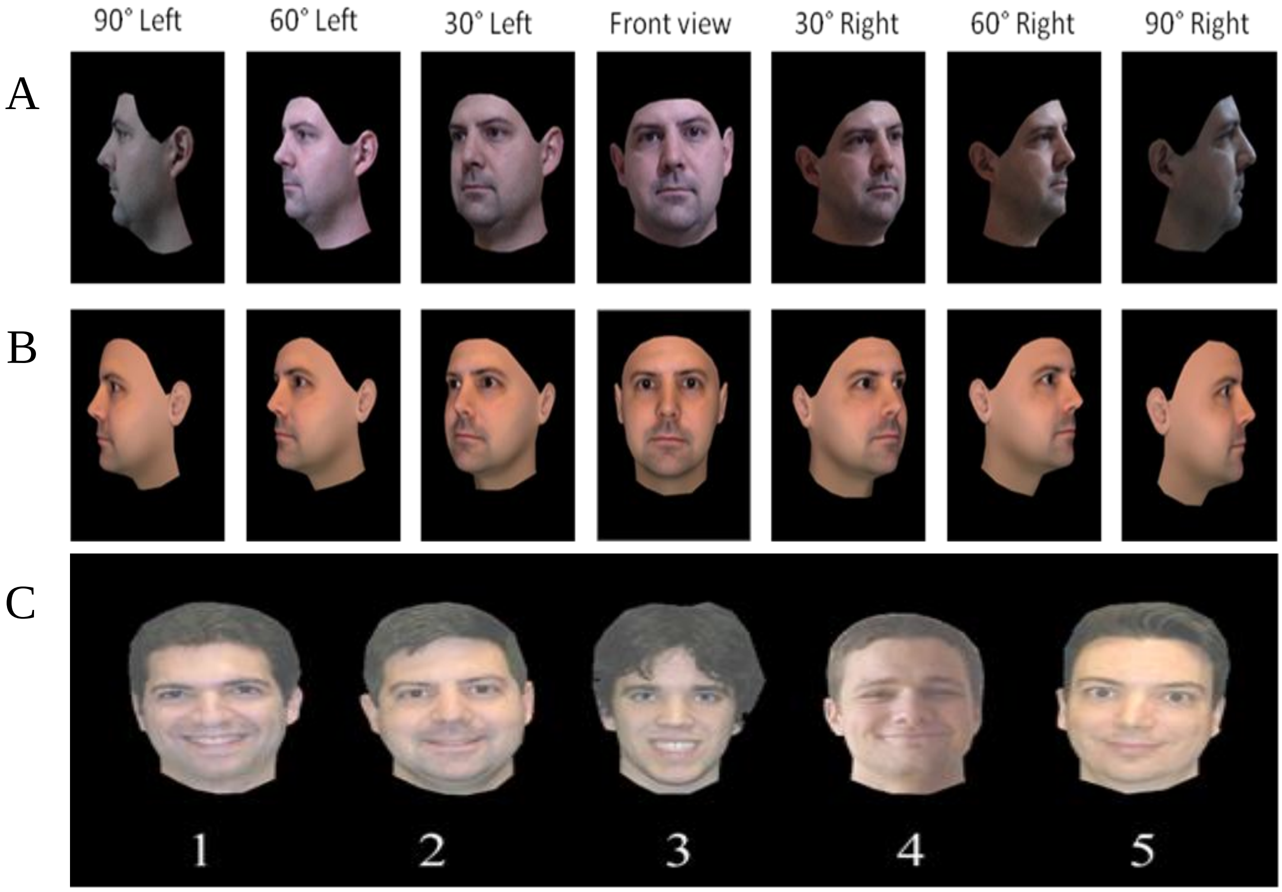

Figure 11: Examples of stimuli used in Experiments 8-10. Panel A represents the photo stimuli used in Experiment 8-9. Panel B represents the counterpart computer generated examples used in Experiments 8-10. All computer generated examples were created from a single front view input. Panel C represents the front view arrays displayed in Experiment 8.

Stimuli: Cropping, Arrays, and Face Generating

Twelve photographs were chosen from the Multi-Pie database (Gross, Matthews, Cohn, Kanade, & Baker, 2008) to become target individuals. Each target face was between the ages of 19 and 45 years old and selected avoiding the presence of non-face cues (e.g., glasses and facial hair). Along with the front-view photo, six other photos of the target were chosen. These 6 photos panned from the left-side view to a right-side view with 30° increments between each photo (see Figure 11, Panel A). Test arrays were made by cropping photographs of the target individual and 4 other individuals (foils) that were not used as target faces. Each array was homogenised so that the sizes of each array were identical (800 × 267 pixels, subtending an approximate visual angle of 29.9° × 10.9°).

Furthermore, each array displayed faces with a happy expression and was subject to a contrast change to minimise recognition through non-face cues such as skin tone (see Figure 11, Panel C). Synthesised faces were generated using SI FaceGen Modeller 3.1 (developed by Singular Inversions, Toronto, Canada). Using a single front-view photo, the software was able to synthesise a computational representation of a face on a rotatable 3D model of the image. This is achieved by placing landmarks upon the key features of the face (e.g., corners of the mouth, jaw line, eyes). The software was used to synthesise still 2D shots of the face as if taken at a variety of angles. Counterpart computer generated versions of the photo targets and foils were created, and 2D stills were taken akin to the angles of the photo stimuli (See Figure 11, Panel B). Each photo and synthesised face image was cropped to remove hair, using Adobe Photoshop 6™, and displayed on a black background on screen at 600 × 463 pixels, subtending to an approximate visual angle of 23.4° × 18.4°. The stimuli were presented centrally on a 17 inch monitor. A custom programme ran using DirectRT software was used to control the presentation of the stimuli on a PC. Responses were registered using a computer keyboard. For the experiment, the numbers 1-5 at the top of the keyboard were used to select a face matching the number depicting each face in the test array.

Design, Procedure and Counterbalance

Participants completed three conditions (Brief 30°, FaceGen, Photo), as part of a within-subjects design. The FaceGen condition consisted of two-second exposures to a photograph of the target face at 30° left orientation, plus multiple synthesised views of the target individual presented at 90° left, 60° left, 30° right, 60° right and 90° right. The Photo condition, gave the same two-second exposures as the FaceGen condition, but used original photographs of the face at all orientations. Presentations in these conditions always ran from 90° left to 90° right, returned, and were then repeated, such that the last presentation before test was a right-facing 90° profile shot. The Brief 30° condition displayed a target, for two seconds, at a left 30° degree angle (this was time matched to the multi-view conditions, such that a fixation cross replaced the target faces, and only the photo target remained). To repeat, all conditions presented a target photo at the left 30° orientation within the exposure sequences, so differences between conditions must relate to the effects of the additional real or generated views.

Participants were sat approximately 70cm from the computer screen. A brief practice trial familiarised them with order of the exposure and test phases. During the exposure phase, face presentations were separated by a 1s inter-stimulus interval. Participants were then given an array of faces which were displayed for 10s and asked to make their identification during this period. Confidence was then measured using a button response to a 7-point Likert scale (1: “Not at all confident”, 7: “Extremely confident”).

Each condition was tested on four different face stimuli within each experimental condition. All conditions were counterbalanced such that each condition was presented first, second, or third equally often, and this was rotated such that every condition was placed in every presentation order across the counterbalance. Similarly, every face appeared within each condition across the counterbalance.

Data Analysis:

The measures used in all the experiments in this chapter are the same as those described within Chapter 2, that is, accuracy of recognition and CA score. Unlike in the experiments reported in Chapter 2, the pattern of results observed with the CA scores does not always mirror the performance of accuracy, thus it is reported throughout this chapter.

4.2.2 Results

Figure 12 shows the test data as percentage correct for each condition (FaceGen, Photo, and Brief 30°). It is evident that exposures to targets in the Photo condition facilitated the highest mean recognition of a target face, while exposures given in the FaceGen condition seems to be no better than the Brief 30° control. A within subject ANOVA observed a main effect of condition, $F(2, 34) = 5.43$, $p = .009$, MSE = 0.338; planned comparisons revealed that there were significant improvements in recognition accuracy for faces exposed in the Photo condition compared to training with faces in either the FaceGen or Brief 30° (control) condition, $F(1, 17) = 9.56$, $p = .007$, MSE = 1.125, and $F(1, 17) = 9.37$, $p = 0.07$, MSE =0.889, respectively. There were no differences observed between the FaceGen and the control condition $F < 1$. As can be seen from Figure 12 Panel B a similar pattern of results are observed for CA scores. ANOVA of CA scores confirms the trend observed in accuracy. That is, a main effect of condition, $F(2, 34) = 5.57$, $p = .008$, MSE = 3.504. Further analysis revealed higher CA scores for the Photo condition compared to both the FaceGen, $F(1, 17) = 7.701$, $p = .013$, MSE = 6.493, and Brief 30° (control) conditions, $F(1, 17) = 8.765$, $p = .009$, MSE = 7.544, but no differences observed between the FaceGen and the control condition $F < 1$.

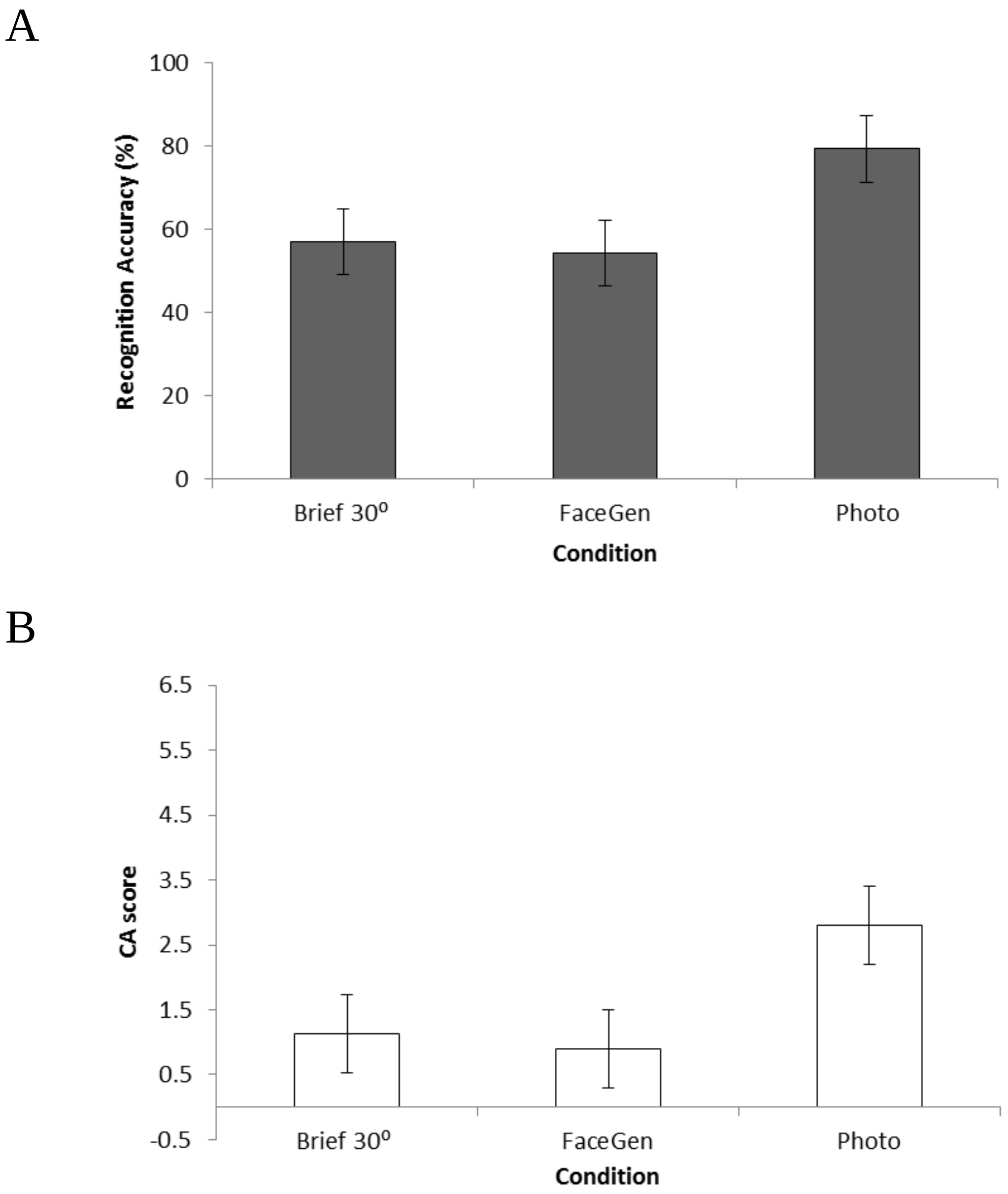

Figure 12: Panel A displays recognition accuracy as a percentage for each condition (with SEM) as a function of condition (Brief 30°, FaceGen, and Photo). Performance at chance level is 20%. Panel B displays mean CA scores (with SEM) for each condition (Brief 30°, FaceGen, and Photo).

4.2.3 Discussion

The results of the current experiment suggest that the Photo condition represents the optimal level of performance for these stimuli. Indeed, the level of performance is relatively high for a test which required generalisation to a novel viewpoint. For the FaceGen stimuli, performance was no better than that of the Brief 30° control. It is possible that these results reflect the merits of the human visual system which alone extracts as much information from a single image as the currently used modeller. However, the failure of the computer generated condition to support an advantage despite the extra views included may represent a limit of the particular modeller (i.e., FaceGen), rather than a more general limit on the amount of information contained within a single image. The current modeller is simply wrapping the texture map computed from a front view image around a pre-defined model. The fact that the image angle at test is the same as the image used to create the synthesised faces could be problematic. That is, it could be a demonstration of what details the modeller is not including from the face photograph (this is discussed in more detail in Chapter 5). This seems particularly likely given that it has been reported previously that creating multiple views of face from a single front view image and testing recognition of that image can produce improved recognition (Liu *et al*, 2009).

4.3 Experiment 9

Experiment 8 did not observe any benefit of multiple synthesised views when testing the front view of a face. However, this may underestimate the influence of such synthesised viewpoints, because Experiment 8 did not involve the presentation of a generated view at the test orientation. This occurred because the SI FaceGen Modeller requires at least a single front-view input to model a face. Thus, when using front facing test faces, it is not possible to use this software to produce a generated image at the test orientation that is independent of the input image. Therefore Experiment 9 investigated if the advantage of exposing multiple views transferred to a novel angle not based on the original image (Table 8 summarizes the design). Participants were given an exposure phase similar to that of the previous experiment, yet the recognition test required that participants selected faces from an array displayed at 30° angles (see Figure 13). In addition, a FaceGen 30° condition presented a single synthesised face at the test angle to examine the possibility that it is only the generated view at the test angle that matters. Under these circumstances, it is now possible to provide a generated image at the test angle. Thus, it allows for an assessment of whether generated views can support enhanced recognition when they include presentation of a generated image at the to-be-tested angle.

Table 8

Design of Experiment 9

Condition	Training sequence	Test Arrays
FaceGen 30°	+, +, 30°L, *0°, +, +, +	Select face from an array of five faces
Brief front	+, +, +, 0°, +, +, +	
FaceGen multiple	90°L, 60°L, 30°L, *0°, 30°R, 60°R, 90°R	
Photo	90°L, 60°L, + , 0°, 30°R, 60°R, 90°R	

Note: 0°-90° represent the different angles that the target face was exposed at, L (Left) or R (Right) represents the direction of the face when looking at the screen, and + symbolises a fixation cross that was presented in the Brief condition in place of exposure at different yaw rotations. Within the FaceGen training sequence * (e.g., *0°) indicates that unlike the other exposures this view was a photograph and not a synthesised view. Test arrays required recognition of the target individual from five faces in a choice array consisting of 30° photos including the target and 4 foils (See Experiment 9 Method for details and Figure 13 for an example of the Test array).

4.3.1 Method

Participants:

Twenty four students, aged 18-24, were recruited from Cardiff University and completed the experiment in return for payment of £2. None of the participants had taken part in Experiment 8.

Stimuli

The stimuli consisted of the 8 target faces taken from same set of cropped and synthesised faces for the previous experiments in this chapter. Target faces used in the training phase for this experiment were front-view photos. For this experiment, a new set of arrays were constructed, using the same method as previous experiments, but displaying the target individual and foils at a 30° angle (facing left when looking at the screen) within the array (see Figure 13). Again, a contrast change was applied to the array - all other stimuli details, such as the sizes of the targets and arrays, were identical to that of Experiment 8.

Design and Procedure

Each participant completed four conditions (FaceGen 30°, Brief front, FaceGen multiple, and Photo) as part of a within-subjects design. Each condition began by displaying a front-view photo of a target individual. The FaceGen multiple condition consisted of exposure to multiple computer generated views that were presented either side of a front-view photograph of the target face. These were presented at 90° left, 60° left, 30° left, a photograph of target in a front-view, 30° right, 60° right, and 90° right. The Photo condition gave the same exposures as the FaceGen condition, but used photos at all orientations with the exception that the 30° left facing Photo was removed and replaced with a fixation cross in order to prevent direct matching to the images used in the test array.

Figure 13: 30° test arrays used in Experiments 9-10. Participants were given a 10s presentation of an array during which they were asked to choose the number that corresponded to the target seen during exposure.

Presentations in the multiple exposure conditions (FaceGen multiple and Photo) always ran from 90° left to 90° right and returned, such that, the last presentation before test was a left-facing 90° profile view. The Brief condition gave two exposures to a front-view photo that was time matched to the multiple conditions, such that a fixation cross replaced the angled target displays, and only the front-view photo target remained. The FaceGen 30° condition displayed the front view photos consistent with previous conditions and a FaceGen target at a left 30° degree angle. This was the same angle as test arrays and thus this condition examined whether only the presentation of generated image at the to-be-tested angle was required to improve recognition. This condition also utilised fixation crosses like the Brief control condition (Table 8 summarizes the design). All other procedural details remained the same as Experiment 8. The assignment of faces to condition, and the order in which the exposure conditions were presented was counterbalanced as in Experiment 8.

4.3.2 Results and Discussion

Figure 14 displays recognition accuracy as percentage correct for each condition (FaceGen 30°, Brief front, FaceGen multiple and Photo). Exposure to the FaceGen multiple condition resulted in the best recognition of a target face compared to the brief controls (FaceGen 30°, Brief front). In addition, the FaceGen multiple condition seemed to produce equivalent learning to the Photo condition, and there was little or no difference apparent between the two control conditions. The within-subjects ANOVA failed to observe a main effect of condition, $F(3, 69) = 2.28$, $p = .087$, MSE = 0.148. However, planned comparisons revealed that higher accuracy was observed for the FaceGen multiple condition compared to the Brief front exposure, $F(1,23) = 9.47$, $p = .005$, MSE = 0.510 and the FaceGen 30° exposure conditions, $F(1,23) = 4.28$, $p = .049$, MSE = 0.510. No other differences were observed between the conditions (largest, $F(1,23) = 2.76$, $p = .110$, MSE = 0.375, between Photo and FaceGen 30° condition).

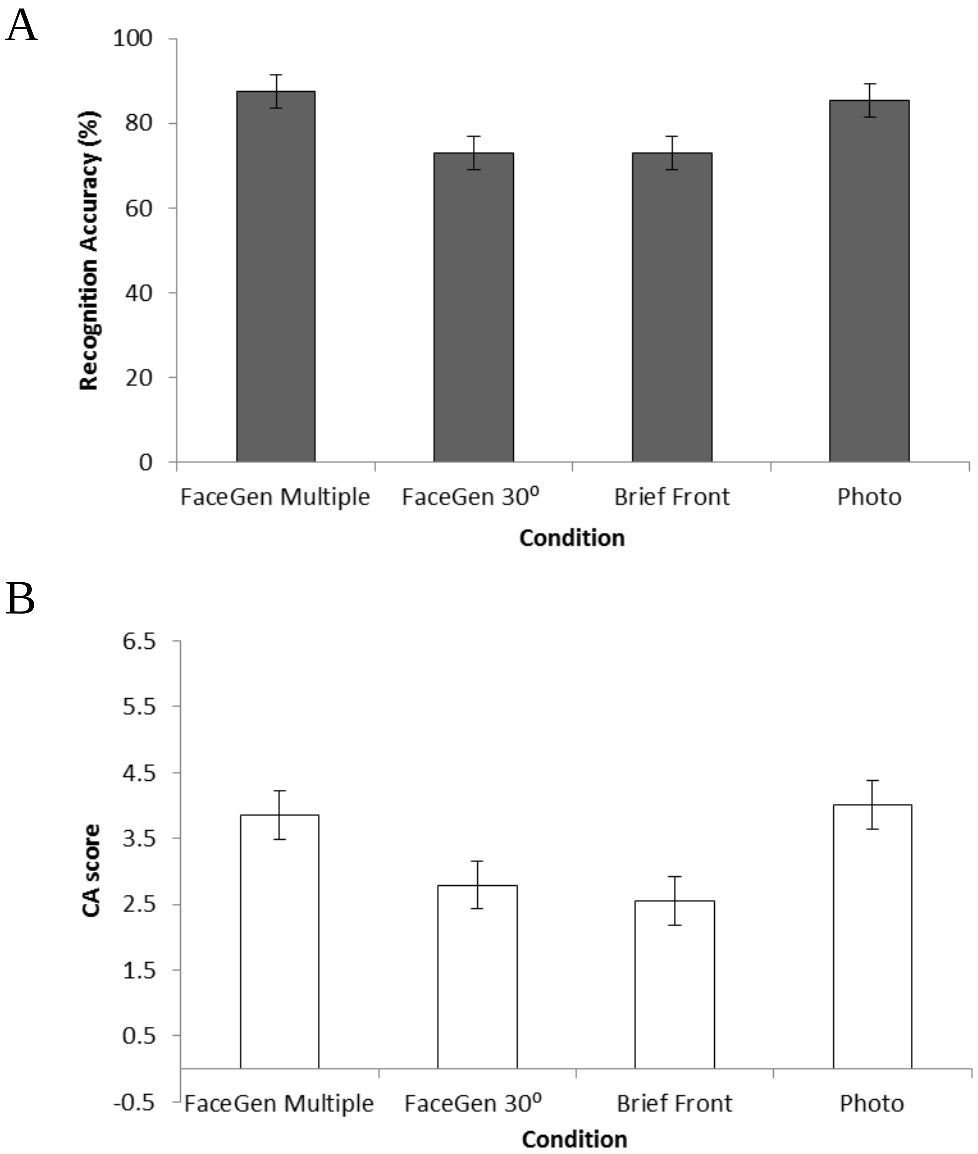

Figure 14: Panel A is recognition accuracy (with SEM) as a percentage for each condition (FaceGen multiple, FaceGen 30°, Brief front, and Photo). Performance at chance level is 20%. Panel B is CA score (with SEM) as a function of the same conditions (FaceGen multiple, FaceGen 30°, Brief front, and Photo).

Panel B of Figure 14 displays a similar trend to that of the accuracy data. However in this instance the Photo condition had higher CA score compared to that of the FaceGen multiple condition. Analysis of the CA scores revealed a main effect of condition, $F(3,69) = 3.30$, $p = .025$, MSE =13.044. Planned comparisons revealed that training in the Photo exposure was better than the Brief front exposure $F(1,23) = 5.77$, $p = .025$, MSE = 51.042, and similarly, but not quite significantly, compared to the FaceGen 30° exposure condition, $F(1,23) = 4.18$, $p = .052$, MSE = 35.042. The remainder of the CA scores were consistent with the accuracy data. That is, like accuracy, there was no significant difference between the FaceGen multiple and Photo conditions $F < 1$. In addition, the FaceGen multiple condition was had a better CA score compared to the Brief exposure $F(1,23) = 8.115$, $p = .014$, MSE = 41.344, and the FaceGen 30° exposure conditions, $F(1,23) = 5.20$, $p = .032$, MSE = 27.094.

Recognition performance in both the Photo and the FaceGen multiple conditions were more accurate compared to the controls. That is, training exposing either photographic images or synthesised images at multiple yaw rotations facilitated better recognition than controls. Perhaps most importantly was the equivalent learning of the FaceGen multiple condition compared to performance of the Photo condition. This suggests that any important features for recognition of the photographs are being replicated by the photogrammetric software, despite the rather impoverished nature of the FaceGen stimuli. Moreover, both conditions facilitated greater learning, which transferred to a novel test angle, compared to the FaceGen 30° exposure and the Brief front exposure controls. Albeit that the FaceGen 30° exposure may have required less transfer than the Photo condition, because the transfer in the FaceGen 30° condition gave exposure to the test angle, but with 30° generated faces. That said, certain conditions differed in the total number of face images presented, so poor performance in the control conditions may have been a product of the relatively small number of exposures compared to the amount given in the multiple view conditions (i.e., FaceGen multiple & Photo), rather than any advantage gained from multiple angles *per se*. Thus, the advantage may reflect the number of exposures in the multiple view conditions compared to the controls (Brief front and FaceGen 30°), rather than any change in representation related to the type of exposure.

4.4 Experiment 10

In Experiment 9, training using synthesised views resulted in superior learning compared to the controls. However there are several differences between the controls and multiple view conditions. Firstly, the total number of exposures rather than the type of exposure may have resulted in better performance of the multiple conditions compared to the control. Secondly, the difference in the interval between the last face exposure and test phase may have affected performance. Experiment 10 assessed how exposure to a repeated single viewpoint compares to how well the multiple generated views can aid recognition. The Photo condition from the previous experiment was replaced with a repeated presentation of a photograph of a target in the front-view (Repeat front condition). The multiple presentations in a single view enable an assessment of whether the effect observed in the previous experiment, and that of Liu *et al*. (2009), was a product of the amount of exposure given.

In order to match the interval between the first and last exposure images, and the presentation of the test array all conditions started and ended with a front-view image of the target face. While this controls for differences in the exposure-test interval, the added presentation of front view face at the end of each exposure condition may provide enough recently presented information to allow participants to generalise to a novel view point displayed on test, thus creating a ceiling effect based on recency. Such potential effects of recency were examined by manipulating the exposure test interval by the addition of a distractor face between the target exposure and test trials. If the presentation of a front view target does create a recency effect then the distractor face should allow enough interruption to assess changes in the representation of a target following each exposure. Such post list delays have been demonstrated to reduce recency effects for face stimuli (e.g., Kerr, Avons, & Ward, 1999).

If, as suggested, recognition following repeated exposures is based on image or view specific codes (Longmore *et al.*, 2008); then it follows that the FaceGen exposure condition should facilitate better recognition compared to the Repeat and Brief training conditions because the multiple angles should allow a better representation of the face which can generalise when testing recognition at a novel viewpoint (Table 9 summarises the design).

4.4.1 Method

Participants

Sixty-one participants, aged 18-24, completed the experiment in return for course credit. All participants were recruited from School of Psychology at Cardiff University. None of the participants had participated previously described in this chapter. Twenty-four students completed the part of the experiment without the distractor, while 27 participants completed the half with the distractor face.

Stimuli

All faces were taken from the same set of cropped and computer generated faces used in the previous experiments in this chapter plus the female distractor faces taken from the same database. Distractor faces were photographs displaying a different gender (i.e., female) that were cropped and presented in a fashion identical to the exposed photo stimulus.

Table 9

Design of Experiment 10

Condition	Training sequence	Distractor Face	30° Test Arrays
Brief front	0°, +, +, +, +, +, +	No	Select face from an array of 5 faces
Repeat front	0°, 0°, 0°, 0°, 0°, 0°, 0°		
FaceGen multiple	*0°, 90°L, 60°L, 30°L, 30°R, 60°R, 90°R		
Brief front	0°, +, +, +, +, +, +	Yes	Select face from an array of 5 faces
Repeat front	0°, 0°, 0°, 0°, 0°, 0°, 0°		
FaceGen multiple	*0°, 90°L, 60°L, 30°L, 30°R, 60°R, 90°R		

Note: 0°-90° represent the different angles that the target face was exposed at, L (Left) or R (Right) represents the direction of the face when looking at the screen, and + symbolises a fixation cross that was presented in the Brief condition in place of exposure at different yaw rotations. Within the FaceGen training sequence * (e.g., *0°) indicates that unlike the other exposures this view was a photograph and not a synthesised view. Distractor faces were a female face that was unrelated to the target face (See Experiment 10 Method for details and Figure 13 for an example of the Test array).

Design and procedure

A within-subjects design gave participants three different exposures (Brief front, Repeat front and FaceGen multiple). The Brief front condition was identical to the one in Experiment 9. Similarly, the FaceGen multiple condition gave exposure to multiple computer generated views like Experiment 9. That is, the exposures were identical except that exposures at 30° left are followed by 30° right. Presentations in the FaceGen condition always ran from 90° left to 90° right, and returned. The Repeat front condition gave the same repeated exposure to the front-view photograph of a target throughout. This was time matched to correspond with the length of presentation time in the FaceGen multiple condition. All exposures began by displaying a front view photograph of a target individual and ended with a presentation front view (0°) of the target. Those who participated in the distractor conditions saw the same exposure sequences as described above (see Table 9), with the only modification being that the final exposure of a target was followed by a presentation a novel distractor face at the same view angle (0°). These distractor faces were displayed for 2s followed by a 1s ISI and then presentation of the test procedure. Distractor faces were counterbalanced such that each face was presented in every distractor condition equally often. All other details including procedure and counterbalancing were identical to Experiment 9.

4.4.2 Results

Figure 15 Panel A displays percentage of correct responses as a factor of condition (FaceGen multiple, Brief front, Repeat front) and distractor (Distractor, No Distractor). Performance was generally better in the FaceGen multiple and Repeat front conditions than in the Brief front, but there was apparently little effect of distractors. A mixed ANOVA with the within subject factor of condition (FaceGen multiple, Brief front, Repeat front) and the between subjects factor of distractor (Distractor, No Distractor) found a main effect of exposure condition on accuracy $F(2, 98) = 3.78$, $p = .026$, MSE = 0.041, but no other main

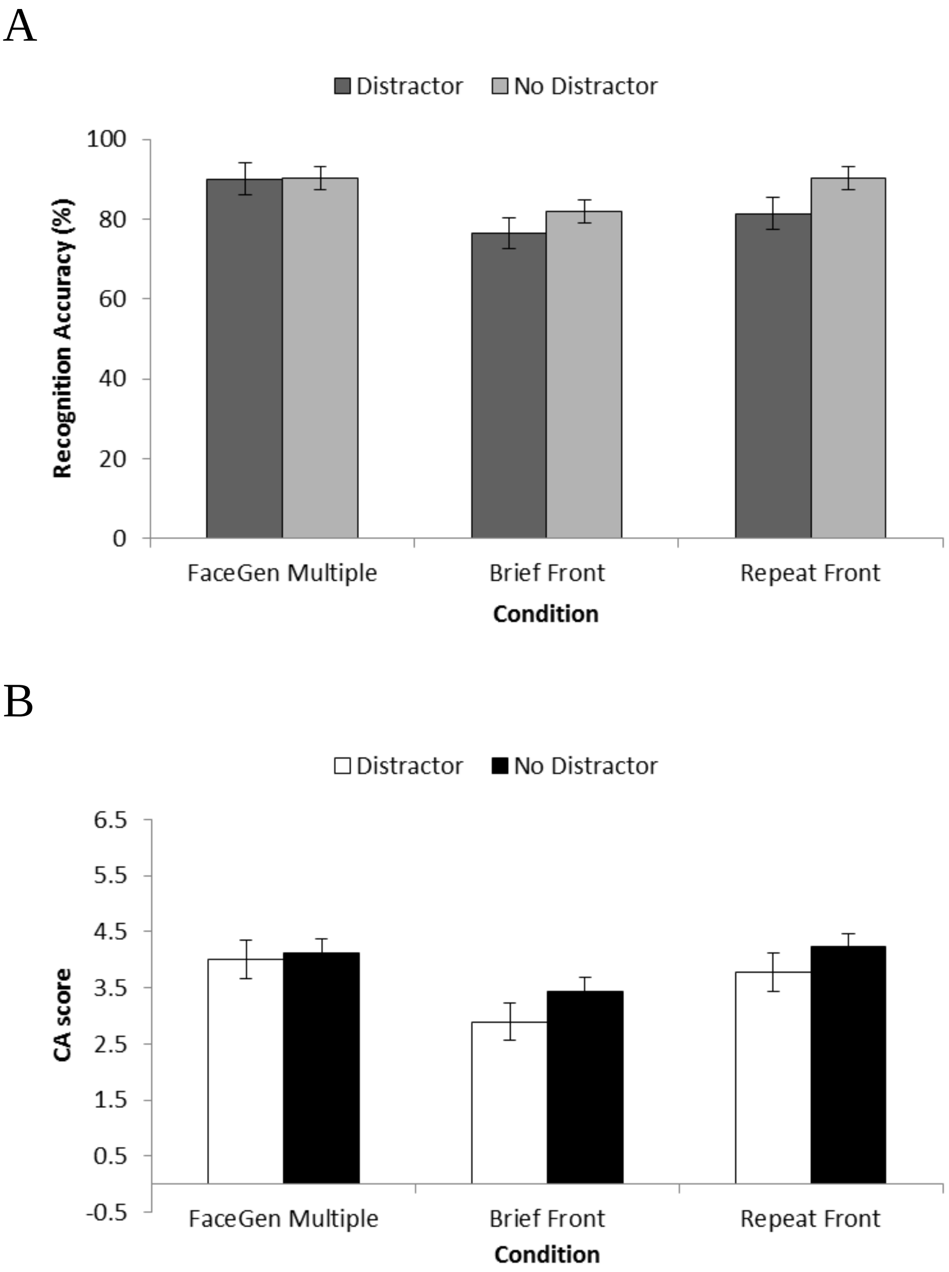

Figure 15: Panel A is recognition accuracy as a percentage correct (with SEM) for each condition (FaceGen multiple, Brief front, and Repeat front) organised as a function of distractor (Distractor and No Distractor). Performance at chance level is 20%. Panel B is CA scores (with SEM) for each condition (FaceGen multiple, Brief front, and Repeat front) organised as a function of distractor (Distractor and No Distractor).

effects or interactions (largest *F*(1,49) = 2.44, *p* = .125, MSE = 0.036, for the effect of Distractor). Pairwise analysis suggested that Brief Front exposure produced lower recognition scores than the FaceGen multiple condition *F*(1, 49) = 6.896, *p* = .011, MSE = 0.088, but no other differences were observed (largest *F*(1,49) = 2.24, *p* = .140, MSE = 0.057, between Brief front and Repeat front exposure conditions).

CA scores displayed in Figure 15 Panel B indicated a similar pattern of results to the percentage correct analysis. There was a main effect of condition, *F*(2, 98) = 5.06, *p* = .008, MSE = 2.571, but no other main effects or interactions (largest *F*(1,49) = 1.20, *p* = .278, MSE = 4.106, for the effect of Distractor). Pairwise analysis suggested that Brief front exposure produced lower CA scores than either the FaceGen multiple condition, *F*(1, 49) = 7.00, *p* = .011, MSE = 5.996, or the Repeat front condition, *F*(1, 49) = 6.912, *p* = .011, MSE = 4.237, but no difference was observed between FaceGen multiple and Repeat front exposure conditions F < 1.

4.4.3 Discussion

The results of Experiment 10 demonstrates that exposure to either multiple synthesised views (i.e., FaceGen multiple) or a repeated single view, produces an advantage over briefly presenting a face, and that recognition of face is similar for these two conditions. That is, both the repeated presentations of a single front view (i.e., Repeat front) or as multiple artificially generated yaw rotations (i.e., FaceGen multiple) supported improved recognition, but did not differ from each other. In addition, the results indicate that recognition performance was not influenced by the placement of a distractor, albeit that this may well be due to a lack of power to detect the relevant interaction as Figure 10 suggests some trend towards the differences between exposure conditions being larger when there was a distractor face.

The current experiment shows that repeated exposure to a single front-view of the face supported recognition of that face at a novel viewing angle to the same degree as did presenting views at multiple viewing angles that were generated by a commercially available modeller. Given that presenting multiple views based on actual pictures should produce superior performance when compared to a single viewpoint, the similarity of these two conditions suggests that the current modeller is not producing images that are equivalent to actual photographs. It is possible that this may be due to the modeller not capturing all the information stored within a photograph. Alternatively, there may be a limit to the amount of information that can be extracted from a single view, and that both the modeller and the unaided visual system are extracting the same amount of relevant detail (a more detailed discussion of this is undertaken in sections 5.2.3 and 5.3.3). That said, it has been shown that accurate recognition from repeated presentation to a single view diminishes the further a face is rotated from a learned view (e.g., Longmore *et al.*, 2008). Therefore, it may be that the test phase of the current experiments was favourable to the repeated photo condition, given that the difference between exposure and test angles was minimal (i.e., 0°-30°). As even the rather limited modeller examined here supported recognition at least as well as did repeated exposure alone, and it did so under conditions that may have favoured the repeated single view conditions, there is at least prima-facie evidence that using photogrammetric software could be a viable means with which to improve unfamiliar face recognition (see General Discussion sections 5.2.3 and 5.3.3).

4.5 Chapter Discussion

The experiments presented in this chapter examined the potentially beneficial effects of synthesising faces from a single front view image. Three experiments reported here examined the benefit of such method, using a commercially available modeller to produce additional training images, and this was tested using a line-up recognition test procedure. In

Experiment 8, no benefit of having seen the synthesised stimuli (i.e., the FaceGen condition) was observed relative to the control condition, while, as expected, exposure to the photographic images displayed at multiple yaw rotations facilitated the improved accuracy in face recognition. The findings of Experiment 9, however, suggest that multiple views of synthesised faces can aid accurate recognition equal to that of the multiple still photos of the target individual when multiple generated views included one at the to-be-tested angle of presentation. Experiment 10, however, showed no advantage for multiple synthesised images over simply repeating a single view of a target, but performance in both conditions was better than that of the control (i.e., Brief exposure).

The fact that recognition can be supported in any way on the basis of photogrammetric images, but that this support did not exceed repeated exposure to a single image, leaves open two general possibilities (these will be discussed further in the General Discussion: see sections 5.2.3 and 5.3.3). Firstly, the SI FaceGen software provides impoverished stimuli, and thus does not support perfect transfer because it loses information when making the synthesised images. As such, even relatively impoverished stimuli may be of some assistance, but the level of support derived from such images may be limited. Secondly, that there is a fundamental limit on the overlap between the information common to different views of a face – and that the human visual system is just as good at extracting this common information as is any computer program following multiple exposures to an individual. Obviously both possible effects may well be in operation to some extent.

Regardless of these caveats, the results reported here provide additional evidence that people exposed to multiple computer generated views of a face were more accurate in recognising it in a subsequent test than if they were only exposed to the original photographic image upon which the generated views were based. Similar results have also been obtained by Liu *et al*. (2009). Comparisons between these studies and their implications will be considered more fully in the General Discussion (see section 5.3.3).

Chapter 5: General Discussion

5.1 Summary of results

Ten experiments used a range of methods to support recognition of an unfamiliar faces. These examined: exposure training on face recognition using a perceptual learning framework, the role of location and content of unique features within complex visual stimuli, and the effect of exposing multiple synthesised faces on recognition.

5.1.1 Exposure training with faces

Chapter 2 reported four experiments examining the effect of comparison and similarity for faces familiarised using techniques inspired by theoretical analyses of perceptual learning. In Experiments 1 and 2, exposure to a target face along with similar comparators of the same gender (either morphed or not) produced no benefit when selecting the target from an array of novel foils, relative to target faces that were exposed with either dissimilar comparators or no comparators at all. In Experiment 3, increasing the number of comparators during the exposure phase also produced no improvement in recognition. Experiment 4 found no advantage when manipulating the amount of facial information given during exposure. The implications of these findings will be discussed in section 5.2.1.

5.1.2 Location and content of unique features

Chapter 3 reported the effects of manipulating unique feature placement on perceptual learning with checkerboard stimuli. In Experiment 5, exposure to stimuli that could be distinguished on the basis of either the content or location of the unique features resulted in improvements in discrimination performance over novel stimuli that were equivalent to those produced by exposure to stimuli that differed in both the content and location of the unique features. In Experiments 6 and 7, the improvement in discrimination performance produced by exposure to stimuli that differed in the content and location of the unique features transferred entirely to stimuli that had new unique features in the same location as the unique features of the exposed stimuli. In contrast, there was no suggestion of any transfer of exposure-produced improvement in discrimination performance when the unique features of the exposed stimuli were moved to a different location. The implications of these results will be discussed in sections 5.2.2 and 5.3.2.

5.1.3 Multiple synthesised views

Chapter 4 examined the effect of exposure to multiple views of a face that were generated from a single photo of an individual. In Experiment 8, there was no advantage of synthesised stimuli over brief exposure. Multiple photographic views of a target facilitated the best recognition performance for arrays displaying the front view of face. In Experiment 9, using a recognition task displaying foils at the three-quarter view (30°), synthesised stimuli produced equal learning to that of the photo condition and both of facilitated better recognition performance than relevant control conditions. In Experiment 10, using the same test procedure, repeated presentation of the front view of a face yielded similar performance to that of the synthesised faces that were again both better than a brief exposure. These findings are discussed in detail within section 5.2.3.

5.2 Implications of current findings

The implications of the current findings relate to the three main issues outlined in the rationale (section 1.1). That is, the implications are discussed in terms of: exposure training, location and content of unique features, and the use of multiple synthesised faces

5.2.1 Exposure training with faces

Mundy *et al*. (2007) reported that comparison during exposure training facilitated performance when the faces to be discriminated at test were the same as those presented during the exposure phase (see also, Dwyer *et al*, 2009; Dwyer *et al*, 2011; Mundy *et al.*, 2009). Similar effects have also been demonstrated using identical twins (e.g., Stevenage, 1998), in which case the enhancement in discrimination transferred to new images of the same twins (see also, Robbins & McKone, 2003). In this light, the experiments presented in Chapter 2 suggest that these facilitatory effects of comparison will be strongest when discriminating between a target and exposed comparators, and may not even extend to novel test foils at all. This is consistent with the findings of Blair and Hall (2003) wherein rats were exposed to two compound flavours AX and BX on an intermixed schedule (AX, BX, AX, BX…) and a third compound CX on a blocked schedule, before a taste aversion was established to AX. This aversion did not generalise to BX at test, demonstrating good discrimination between the intermixed stimuli. However, the aversion did generalise to the test stimulus CX which had been presented in a blocked fashion, suggesting that there was no improvement in discrimination performance between an exposed target and similar stimuli that were not exposed in alternation.

Against this background, and the current results, the fact that Dwyer and Vladeanu (2009) found that recognition of a target against novel test foils was improved by the presence of similar comparators requires explanation. As noted in the introduction (section 1.4), the stimulus set used by Dwyer and Vladeanu consisted of artificial generated faces, whereas my stimuli consisted of face photographs. These generated stimuli may have belonged to a distinctive or restricted portion of face-space. In this light, the difference between these sets of results could be attributed to way in which the stimulus set of Dwyer and Vladeanu were generated using a computer programme which models each face upon an average. That is, each face produced from the modelling software is framed upon an average model of a collection of input faces used to build the model[15]. It follows from this that any face produced by the model will be defined using a constrained amount of dimensions (see also section 5.2.3 for more discussion on the software). If we again consider the idea of the face-space framework, the dimensions which span this space are assumed to encode physical or abstract attributes that render different faces discriminable from one another (Hancock, Burton, & Bruce, 1996). Moreover, if we consider estimates of the amount of dimensions as between 15-22 (Lewis, 2004), then reducing these dimensions may have facilitated a viewing strategy whereby differentiation highlighted the crucial differences on a small number of dimensions allowed for easier identification. That is, the similar condition allowed a more focused approach on the perceptually relevant information with which to identify an individual. In short, the fact that Dwyer and Valdeanu used the same methods to produce both the comparison faces and the test foils is problematic. This method combined with the notion that the stimulus set was very distinctive suggests that comparison faces and test foils may well have been so similar that they were essentially examining a situation where there was effectively no discrimination from novel test foils.

Conversely, the experiments reported in Chapter 2 using real faces, where the dimensions are less restricted, failed to produce any similarity advantage. The complex nature of face stimuli may be the reason for this, and so considering them as simply a collection of common and unique elements (as is common in the perceptual learning literature) may not be beneficial considering the multitude of ways faces can differ or be similar.

It should be noted that Mundy *et al.* (2007) and Dwyer *et al.* (2009) attributed the effects of their intermixed exposure procedures to an unsupervised learning process which

---

[15] Note that SI FaceGen uses a similar approach to that of Blanz and Vetter (1999, see also Vetter & Poggio 1997; O'Toole *et al.*,1994 for a detailed explanation of how some of these models are constructed, and a detailed discussion on how they work).

afforded comparison driven stimulus differentiation, which in turn improved discrimination of the exposed faces. In the experiments reported in Chapter 2, names were assigned to the target faces. Attaching labels of this type to a stimulus increases the range of distinctive features it possesses (Hall, 2008; see also, James, 1890). That is, the acquisition of labels becomes another set of unique features, which can support supervised learning processes leading to acquired distinctiveness if the labels differ (or acquired equivalence, if they are the same). Thus, it could be argued that learning in the studies from Chapter 2 was not unsupervised and that the presence of the name labels informed subjects of the distinctions they were required to learn. For example attaching the name Bob to a face may invoke a process whereby a participant closely inspects what is special about Bob compared to Not Bob. This is almost certainly true, as the experimental subjects were instructed to learn the target faces. However, comparison should have aided discrimination regardless of whether the learning was supervised or not. It should be noted, however, that the results of Chapter 2 found no evidence of improvements based upon comparison, and so the question of whether supervised and unsupervised learning was involved is entirely moot.

5.2.2 Location and Content of unique features

In one sense, the idea that attention to particular regions of these checkerboard stimuli is key to performance is unsurprising. Indeed, Wang and Mitchell (2011) have clearly demonstrated that participants look to the location where exposed unique features appear – even when those features are absent on a given trial. Moreover, in Experiment 3 of Wang *et al*. (2012) this tendency to look at the location where the exposed features appeared was maintained even when novel features appeared in those places. However, while this provides evidence that participants have learnt the location of the unique features of the exposed stimuli, examining gaze direction in this manner does not assess whether they have genuinely learnt nothing about the content of those features at all. Wang *et al*. (2012) also observed that discrimination accuracy was higher during test when unique features appeared in the location of the trained unique features, regardless of whether those test features had been exposed or were novel. This is certainly consistent with the idea that subjects learn more about location than content, but, again, it does not directly assess whether there was no content-based learning at all. One reason for this is that there was no analysis in Wang *et al*. (2012) of whether the absence of a significant effect of content exposure genuinely supports the absence of such an effect (such as using the Bayes factor analysis as described here). More importantly, in Wang *et al*. (2012) participants received a single test-phase involving trials where the exposed features A/B appear in the trained locations or in a new location, while novel features C/D appear either in the trained location for A/B or in a new location. Thus, successful performance on A/B same location trials would effectively reinforce any tendency to attend to this location (and thus support good performance when C/D appears in the same place). But, by reinforcing the tendency to look in a particular location, this combined test does not offer an uncontaminated assessment of whether learning about content (i.e. the A/B features themselves) could support enhanced discrimination at all. In essence, this design puts the tendency to respond based on location in opposition to any tendency to respond based on content. Importantly, my experiments demonstrate that it is *only* learning about the location of unique features that matters for discriminating checkerboards constructed in the fashion used here. The results of Experiment 6 here, and Experiment 3 of Wang *et al*. (2012), are consistent with this possibility, and the results of Experiment 7 confirm this, even when content- and location-based performance is not directly opposed.

The first experiments using colour checkerboards, of the type used in Chapter 3, were reported by Lavis and Mitchell (2006). Experiments 1A and 1B, of Lavis and Mitchell, simply showed that intermixed exposure was better than blocked exposure for promoting subsequent discrimination, and thus do not help to distinguish between the different accounts of perceptual learning. Thus, the possibility of an attentional artefact is of little importance (a similar analysis can be applied to the experiments reported by Mitchell, Nash *et al*. (2008) who demonstrated that trial spacing cannot explain the superiority for intermixed over blocked exposure). In Experiment 2A, participants were exposed to three pairs of stimuli (AX/BX and CX/DX, each exposed in alternation, while EX/FX were exposed in blocks) and Experiment 2B used a similar design, save that two pairs were exposed in blocks and one was exposed in alternation. Following this exposure participants were tested for their discrimination within pairs (e.g., AX vs. BX, or EX vs. BX) or between pairs (e.g., AX vs. CX, or AX vs. EX). Discrimination involving only blocked stimuli was less accurate than discrimination involving a stimulus exposed in alternation regardless of whether the discrimination involved between or within pair comparisons. On the face of it, the facility with which between pair discriminations were made is inconsistent with accounts based on mutual inhibition (e.g., McLaren & Mackintosh, 2000) and thus seems to favour an explanation in terms of intermixed exposure enhancing the salience of the unique features (which was exactly the analysis made by Lavis & Mitchell, 2006). However, because the unique features remained in the same place for within- and between-pair tests, if participants had simply learnt to look to the locations where the unique features of intermixed stimuli appeared, then the success of between-pair discriminations can be explained without recourse to changes in feature salience. Similarly, Mitchell, Kadib *et al*. (2008), report that after exposure to AX/BX, discrimination was equivalently good for AX/X as it was for BY/Y (i.e., exposure effects generalised to a new common background – Experiment 2). Again, the fact that the unique features remained in the same place regardless of what background was used on test means that a response strategy based on simply looking at the locations where differences occurred during exposure could entirely explain the observed data without recourse to a change in the salience of the unique features. Finally, Lavis *et al*. (2011)

reported that exposure to the unique features alone facilitates discrimination (Experiment 2). But again, the additional unique feature alone exposures maintained their location, and so the influence of these exposures can also be explained purely by an attention to location mechanism. The pattern across all these studies is largely the same – a transfer of exposure learning to test performance that appears to be informative by being apparently inconsistent with one or other theoretical account of perceptual learning appears not to be theoretically decisive at all because the transfer of exposure effects can be explained simply in terms of where participants chose to look. Thus, none of the studies examined here require explanation in terms of a modification of feature salience (however that modification might be supposed to occur).

Now, one obvious rejoinder to the contention that looking in a particular place obviates the necessity for theoretical accounts of changes in feature salience (or indeed any other account of improved discrimination performance) is to speculate that where something appears in a complex visual stimulus should be considered as a feature of that stimulus. Considered in this way, the data reported here become a demonstration that where something appears is the critically important feature. While this suggestion certainly merits consideration, it does not fully address the critique described above. Firstly, the idea that location is a feature directly challenges the interpretation of studies examining the transfer of learning from one situation to another - if the key feature is location, then this remains constant despite changes in things like the background or the comparison stimuli, and so no real transfer is being examined at all. Secondly, it is not the location that distinguishes the stimuli (e.g., all of them have a top left), but the fact that there is a difference in the content that appears at that location between two stimuli. Thus, attending to a location is not to attend to the distinctive aspects of a complex stimulus at all. But perhaps most critically, even if location is considered as a feature then this characterisation of the stimuli still does not address the possibility that looking at a particular location after exposure is the result of a strategic choice on behalf of the participants, rather than being due to their attention being drawn to some particularly salient feature.

The potential for strategic choices to influence the performance has long been identified as a challenge for those interested in examining the effects of mere exposure on perceptual learning in humans (e.g., McLaren & Mackintosh, 2000). Although not manipulated in any experiments reported here, the fact that some procedural details could contribute to a learning strategy involving location has been highlighted. In my studies, and many others, participants are instructed to look for differences between stimuli during the exposure phase. Assuming that they follow these instructions, then when they discover a way of distinguishing the critical stimuli (such as looking in a particular place) then this behaviour will be implicitly reinforced by the success of achieving the task that has been set for them (Mackintosh, 2009). While recognising the potential for people to deploy attention in this sort of strategic manner, Lavis *et al*. (2011) downplay the importance of this possibility by suggesting that this account does not explain how different exposure schedules influence the ability to detect the location of distinctive features. However, during alternating exposure, the critical difference between stimuli is present on every trial, and any possible difference that was identified by deliberate search can thus be checked at will. For blocked exposure, only the single transition trial affords the opportunity to directly check whether a feature really does discriminate two stimuli. Thus, there is an obvious mechanism whereby stimulus scheduling could influence the effectiveness of strategic processes.

5.2.3 Multiple synthesised views

The results of several studies have indicated that accurate recognition of an unfamiliar face is dependent on the view at which it is presented (e.g., Hill *et al.*, 1997; Krouse, 1981; O’Toole, Edelman, & Bülthoff, 1998). Recognition is best when presented in the three-quarter view above that of a front facing view, both of which, are greater than the profile view (e.g., Bruce *et al*., 1987; Hill & Bruce 1996; Liu & Chaudhuri, 2002). It is thought that the superior 3D information provided by the three-quarter views includes more structural information (Hole & Bourne, 2010). In light of this, a few experiments have utilised a method of stimulus creation that allows multiple views from even a single 2D photographic input. The experiments of Chapter 4 suggest that using this method to create multiple views of a face from a single input can aid recognition equivalent to that of exposing the original photos of an unfamiliar face. This is consistent with the findings of Bailenson *et al.* (2004) that virtual busts, created using photogrammetric software, can facilitate recognition close to the level of recognition produced by real photographic stimuli presented during training.

Against this background, and the current results, the findings of Liu *et al.* (2009) suggest that that participants exposed to a combination of original photos in one view and synthesised faces in a different view improved recognition (see Experiment 1), and that displaying the different angles of faces sequentially was better than a single presentation of a front view image (Experiment 2). Similarly, the experimental findings of Chapter 4, using low-cost photogrammetric software (FaceGen), suggest that similar effect can contribute to the recognition of a familiarised target individual in a task requiring participants to identify a target individual from a line-up. Taken together, these results indicate the effectiveness of utilising this method to overcome the viewpoint dependence associated with unfamiliar faces. According to Liu *et al.* (2009), the ability to synthesise an angle close to that of a test view can help bridge the gap between original photo and test image. This is consistent with evidence from the object and face recognition literature that indicates that is it easier to generalise based on exposures to multiple views rather than a single view (e.g., Edelman & Bülthoff, 1992; Hill *et al.*, 1997). The fact that recognition can be supported on the basis of low-cost synthesised images suggests that even relatively impoverished stimuli may be of some assistance.

On the other hand, there are several caveats that may limit the utility of synthesising extra views of a face. That is, the experiments of Chapter 4 failed to produce an advantage for the synthesised stimuli when the faces were tested in the front-view (especially in the presence of real photographs). However, as previously outlined, this likely occurred because it is not possible to use this software to produce a generated image independent of the input image at the front-view orientation. Further to this, when a single front-view face was repeatedly presented, it produced equal recognition to that of the synthesised images. This might be taken to suggest that the level of support derived from such images may be limited. Indeed, the advantage of exposure conditions because of the "extra views provided by the synthesised images" has been acknowledged by Liu *et al.* (2009 p. 996); although in their demonstration this was not investigated. As discussed in Chapter 4, it is not clear whether the equivalence of multiple generated views to a repeated single veridical image is due to a fundamental limit in the amount of information that is common to different views (and thus the modeller and the human visual system are equivalently good at extracting this common information) or whether the FaceGen software used here was simply not optimal in extracting all the overlapping information. While nothing in the current experiments, or those of Liu and his colleagues, directly supports either option, the fact that FaceGen produces images that are clearly less than photographic quality and lack accurate textural information (at least) would certainly be consistent with the idea that superior modelling software could afford even better support for face learning across different viewing angles. That is, the limited support gained from synthesised views may be due to the weaknesses of the software used in Chapter 4 rather than a fundamental limit on the information common to different views of a face.

As outlined in the General Discussion of Chapter 2 (section 5.2.1), each face produced from the modelling software is framed upon an average model of a collection of input faces used to build the model; faces which are modelled upon an “average” are derived from a limited pool that may lack sufficient variation and detail to aid the most accurate recognition. The commercial photogrammetric software used may not have included enough information within the pool of synthesised views. This may mean that not enough information is being conveyed within a synthesised face. What type of information is lacking in the synthesised face view is apparent when comparing the synthesised faces to their photo counterparts. Most notably is the contrast in surface pigmentation between images (see Figure 11). If, as Longmore *et al.* (2008) suggest, familiar faces that have been seen many times allow for extraction of structural codes, then the period of familiarisation should have allowed extraction of some of these codes. These codes are thought to include information similar to that of object recognition e.g., three-dimensional shape, and surface pigmentation (Marr & Nishihara, 1978). Thus, repeated exposure to a front-view photo in the current studies may have allowed some of this information, particularly surface pigmentation, to be conveyed, but it is unlikely much of the shape information was gained. Conversely, the synthesised images displayed at different angles may have allowed more three-dimensional shape information, but lacked the surface pigmentation that would be afforded by real photos or a better modeller.

5.3 Future Directions

5.3.1 Exposure training with faces

Returning to the forensic considerations with which I began. Proving our identity with some form of photographic identification is rapidly becoming a vital aspect of our daily routine, despite the fact that there is clear evidence that people are ineffective and error-prone at this task (Kemp *et al*., 1997; Logie *et al*., 1987). It is of particular importance that ways to overcome this deficiency are explored. Unfortunately, despite evidence of strong perceptual learning based improvements in a wide range of real-life situations such as viewing X-ray images (Sowden, Davies, & Roling, 2000), wine tasting (Solomon, 1997) and chick sexing (Biederman & Shiffrar, 1987), my current results suggest any direct forensic application of comparison-based perceptual learning for face recognition is limited. That is, despite its theoretical plausibility, comparison through a process of perceptual learning does not improve the learning of new faces in a practically tractable fashion. That said, one heartening implication of the current results is that repeated brief exposures to a face does support at least some ability to reliably identify an individual from their photograph.

The current studies appear to conclusively rule out the contribution of comparison to the development of face familiarity, at least within the early stages of acquiring familiarity (c.f. Stacey *et al*., 2005). Thus, future work in this area must look to alternative means as a way to support recognition of unfamiliar individuals. This has been considered in Chapter 4 with respect to generated views to overcome the view dependence associated with recognition of unfamiliar faces. Alternatively, another option would be return to familiarity development to offer some suggestions. For example, recent evidence has suggested that familiarity can develop for learnt (originally unfamiliar) faces with repeated experience. Comparing recognition while manipulating the degree of this experience, using different schedules, may help to identify the most robust and forensically applicable schedule.

5.3.2 Location and Content of unique features

If many studies based on the type of checkerboard stimuli used here are fatally compromised by the possibility of a strategic direction of attention to location, then where does this leave the investigation of perceptual learning in humans? Perhaps most critically, the basic schedule effects underpinning many analyses of perceptual learning are present in stimuli that are not open to strategic spatial attention (e.g., flavours: Dwyer *et al.*, 2004; Mundy *et al.*, 2006). Moreover, other visual stimuli, such as morphed faces (e.g., Dwyer *et al.*, 2011; Mundy *et al.*, 2007), and checkerboards (e.g., McLaren, 1997; Wills *et al.*, 2004) or icon arrays (De Zilva & Mitchell, 2012) with probabilistically defined features, have no single defining feature at a constant location, and thus strategic attention to particular locations cannot play a role[16]. In addition, the fact that exposure to the common element alone improves subsequent discrimination (Mundy *et al.*, 2007; Wang & Mitchell, 2011) cannot be explained by subjects learning to attend to the critical location during the exposure phase, as there is no indication at this point of what the critical location might be. The corpus of unconfounded experimental studies might be reduced by placing the current style of checkerboard to one side, but many studies using other stimulus types remain. There are also studies (including with checkerboards as used here) where the results cannot be explained by strategic attention to location. Thus, the results reported here do not require the wholesale questioning of theoretical accounts of perceptual learning in humans.

That said, the impact of the current studies should not be underestimated. The contribution of strategic allocation of attention to particular regions of stimulus space, independent of any change in the representation properties or salience of the features that occur at that space, has formerly been cited as a logically possible confound (Mackintosh, 2009). The current experiments explicitly demonstrate that such content-independent mechanisms can entirely explain exposure-dependant improvements in discrimination performance in one commonly used type of visual stimulus. This both raises questions about the theoretical interpretation of all other studies using the same type of stimulus, and provides

---

[16] That is not to say that participants could not approach these tasks in a strategic manner, but that mere attention to a particular area will not suffice to reliably distinguish the stimuli.

concrete evidence for the contribution of implicitly reinforced attentional mechanisms that must be considered in all studies of human perceptual learning.

One obvious direction to follow on from the current results would involve re-running the confounded transfer experiments. For example, re-examining the studies that used the now invalidated cues, by examining perceptual learning with stimuli that negate any potential confound of where to look. Thus, the relative salience of the unique features can be accurately assessed in order to advance the understanding of perceptual learning.

5.3.3 Multiple synthesised views

The fact that a variety of computer-based face modelling approaches support the generation of multiple views from even a single 2D photographic input raises the question of whether this technology could be used to support human face recognition. While these studies provide some evidence for the idea that computer generated faces might support human recognition, they provide little or no detail of the mechanisms involved, and of how this process might be optimised for applied use. For example, people clearly have some ability to generalise from a single photographic image to a different view of the target face, so the if generated views support recognition beyond the level possible with multiple exposures to a single image then implies that the models must (in some circumstances at least) perform the transformation over viewing angles more effectively than an unaided person. Identifying where the advantage for the generated models lie will provide information about how unaided human recognition works, as well as providing information about how best to provide artificial support.

Therefore, examining a number of face generation approaches to ascertain which affords the greatest support to unaided human recognition is an obvious potential future direction for this research. As yet, only two commercially available Photogrammetric software products have been tested, e.g., FaceGen (Chapter 4) and 3DMeNow (see, Bailenson *et al.*, 2004). However, one of the main restrictions of photogrammetric software is the poor contrast of the human face surface (see, Galantucci, Percoco, & Di Gioia, 2009). While the other demonstrations of this effect have used custom software (Liu *et al.*, 2009), the modification of existing models or even the development of new ones (based on more laser scans of original faces – thus providing a better average for which an image could be modelled upon) would allow for a closer examination of what types of generated images best support human recognition. The quality of already published generated images based on the Basel Face Model (Paysan, Knothe, Amberg, Romdhani, & Vetter, 2009) suggests that it offers one option for improved modelling of face texture while using a system that still operates via a projection onto a general 3D norm representation. Alternatively, direct application of PCA processes could support the development of direct projections from a front-view to multiple angled views without the intermediate step of morphing onto a 3D norm. Comparing these two classes of opproach would address both the theoretical issue (what aspect of human recognition are being supported by the computer generated images) and the practical one (how is the generation technology best applied to support human recognition).

5.4 Concluding remarks

In sum, the results reported within Chapter 2 of this thesis suggest that an exposure schedule based on perceptual learning is not an effective means of facilitating recognition of an originally unfamiliar individual. Moreover, I have identified that in some instances of apparent perceptual learning with complex visual stimuli, the mechanisms behind any improvement test performance are based on strategic attention to particular locations, rather than driven by changes in the perception of the content of stimulus features (Chapter 3). However, this does not exclude any processes of perceptual learning in learning to recognise a face. Indeed, taken together, the results of Chapter 2 and 4 indicate that experience with a face does improve subsequent recognition above that of a brief encounter. Moreover, artificial support in the form of computer generated views offers the promise of improving the ability to recognise a face from a single photograph beyond the level that could be achieved by the human visual system alone.

Appendices

Appendix 1. Mean performance accuracy from Chapter 2 for number of comparator (3A) compared to baseline performance (3B)

Number of comparators (3A) vs. baseline (3B)	Result of t-test	Test of variance
0 comparators	*t*(31) = -4.365, *p*< .001	Unequal *F*=5.895 *p*= .019
1 comparator	*t*(36) = -3.636, *p*= .001	Unequal *F*= 6.483 *p*= .014
4 comparators	*t*(32) = -3.475, *p*= .001	Unequal *F*= 6.942 *p*= .011
16 Comparators	*t*(50) = -4.925, *p*< .001	Equal *F*= 2.024 *p*= .161

References

Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature, 387*(6631), 401-406.

Angeli, A., Bruce, V., & Ellis, H. D. (1999). Getting familiar with faces. *Perception* 28, Suppl. 115.

Bahrick, H. P., Bahrick, P. O., & Wittlinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General, 104*(1), 54-75.

Bailenson, J. N., Beall, A. C., & Blascovich, J. (2003). Using Virtual Heads for Person Identification: An Empirical Study Comparing Photographs to Photogrammetrically-Generated Models. *Journal of Forensic Identification, 53*(6), 722-728.

Bailenson, J. N., Beall, A. C., Blascovich, J., & Rex, C. (2004). Examining virtual busts: Are photogrammetrically generated head models effective for person identification? *Presence: Teleoperators and Virtual Environments, 13*(4), 416-427.

Bartlett, J. C., & Searcy, J. (1993). Inversion and Configuration of Faces. *Cognitive Psychology, 25*(3), 281-316.

Beale, J. M., & Keil, F. C. (1995). Categorical effects in the perception of faces. *Cognition, 57*(3), 217-239.

Biederman, I., & Shiffrar, M. M. (1987). Sexing Day-Old Chicks: A Case Study and Expert Systems Analysis of a Difficult Perceptual-Learning Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*(4), 640-645.

Bindemann, M., Avetisyan, M., & Blackwell, K. A. (2010). Finding Needles in Haystacks: Identity Mismatch Frequency and Facial Identity Verification. *Journal of Experimental Psychology: Applied, 16*(4), 378-386.

Bindemann, M., & Sandford, A. (2011). Me, myself, and I: Different recognition rates for three photo-IDs of the same person. *Perception, 40*(5), 625-627.

Blair, C. A. J., & Hall, G. (2003). Perceptual Learning in Flavor Aversion: Evidence for Learned Changes in Stimulus Effectiveness. *Journal of Experimental Psychology: Animal Behavior Processes, 29*(1), 39-48.

Blair, C. A. J., Wilkinson, A., & Hall, G. (2004). Assessments of changes in the effective salience of stimulus elements as a result of stimulus preexposure. *Journal of Experimental Psychology: Animal Behavior Processes*, *30*(4), 317-324.

Blanz, V., & Vetter, T. (1999). A Morphable Model for the Synthesis of 3D Faces. *Proceedings of Siggraph'99.* 187-194.

Blanz, V., & Vetter, T. (2003). Face recognition based on fitting a 3D morphable model. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 25*(9), 1063-1074.

Bonner, L., Burton, A. M., & Bruce, V. (2003). Getting to know you: How we learn new faces. *Visual Cognition, 10*(5), 527-536.

Bower, Gordon H., & Karlin, Martin B. (1974). Depth of processing pictures of faces and recognition memory. *Journal of Experimental Psychology October, 103*(4), 751-757.

Bruce, V. (1982). Changing faces: visual and non-visual coding processes in face recognition. *The British Journal of Psychology, 73*(Pt 1), 105-116.

Bruce, V., & Burton, A. M. (2002). Learning new faces. In T. Poggio & M. Fahle (Eds.), *Perceptual learning* (pp. 317-334). Cambridge, MA: MIT Press.

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P. J. B., Burton, A. M., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied, 5*(4), 339-360.

Bruce, V., Henderson, Z., Newman, C., & Burton, A. M. (2001). Matching Identities of Familiar and Unfamiliar Faces Caught on CCTV Images. *Journal of Experimental Psychology: Applied, 7*(3), 207-218.

Bruce, V., & Valentine, T. (1988). When a nod's as good as a wink: The role of dynamic information in face recognition. In M. M. Gruneberg, P. E. Morris & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues.* (Vol. 1) (pp. 169-174). Oxford, England: John Wiley & Sons.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology,* (77), 305-327.

Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: the power of an expertise framework. *Trends in Cognitive Sciences, 10*(4), 159-166.

Burton, A. M. (1994). Learning new faces in an interactive activation and competition model. *Visual Cognition, 1*(2-3), 313-348.

Burton, A. M., Bruce, V., & Hancock, P. J. B. (1999). From pixels to people: A model of familiar face recognition. *Cognitive Science, 23*(1), 1-31.

Burton, A. M., Bruce, V., & Johnston, R. A. (1990). Understanding face recognition with an interactive activation model. *The British Journal of Psychology, 81*(3), 361-380.

Buttle, H., & Raymond, J. E. (2003). High familiarity enhances visual change detection for face stimuli. *Perception and Psychophysics, 65*(8), 1296-1306.

Calder, A. J., Young, A. W., Perrett, D. I., Etcoff, N. L., & Rowland, D. (1996). Categorical perception of morphed facial expressions. *Visual Cognition, 3*(2), 81-117.

Campbell, R., Coleman, M., Walker, J., Benson, P. J., Wallace, S., Michelotti, J., & Baron-Cohen, S. (1999). When does the inner-face advantage in familiar face recognition arise and why? *Visual Cognition, 6*(2), 197-216.

Carey, S. (1992). Becoming a face expert. *Philosophical transactions of the Royal Society of London. Series B: Biological sciences, 335*(1273), 95-102.

Chen, W., & Liu, C. H. (2009). Transfer between pose and expression training in face recognition. *Vision Research, 49*(3), 368-373.

Chung, M. S., & Thomson, D. M. (1995). Development of face recognition. *The British Journal of Psychology, 86*(1), 55-87.

Clutterbuck, R., & Johnston, R. A. (2002). Exploring levels of face familiarity by using an indirect face-matching measure. *Perception, 31*(8), 985-994.

Clutterbuck, R., & Johnston, R. A. (2005). Demonstrating how unfamiliar faces become familiar using a face matching task. *European Journal of Cognitive Psychology, 17*(1), 97-116.

Cole, S. A. (2001). *Suspect identities: A history of fingerprinting and criminal identification*. Cambridge MA: Harvard University Press.

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), 94-107.

Davies, G., & Valentine, T. (2007). Facial composites: Their utility and psychological research. In R. C. L. Lindsay, D. Ross, J. Read & M. Toglia (Eds.), *Handbook of Eyewitness Psychology: Memory for people* (pp. 59-83). Mahwah, New Jersey: Lawrence Erlbaum.

Davies, G., Van Der Willik, P., & Morrison, L. J. (2000). Facial composite production: A comparison of mechanical and computer-driven systems. *Journal of Applied Psychology, 85*(1), 119-124.

De Zilva, D., & Mitchell, C. J. (2012). Effects of exposure on discrimination of similar stimuli and on memory for their unique and common features. *Quarterly Journal of Experimental Psychology, 65*(6), 1123-1138.

Diamond, R., & Carey, S. (1986). Why Faces Are and Are Not Special. An Effect of Expertise. *Journal of Experimental Psychology: General, 115*(2), 107-117.

Dwyer, D. M., Bennett, C. H., & Mackintosh, N. J. (2001). Evidence for inhibitory associations between the unique elements of two compound flavours. *Quarterly Journal of Experimental Psychology Section B: Comparative and Physiological Psychology, 50*(2), 97-107.

Dwyer, D. M., Hodder, K. I., & Honey, R. C. (2004). Perceptual learning in humans: Roles of preexposure schedule, feedback, and discrimination assay. *Quarterly Journal of Experimental Psychology Section B: Comparative and Physiological Psychology, 57*(3), 245-259.

Dwyer, D. M., & Honey, R. C. (2007). The Effects of Habituation Training on Compound Conditioning Are Not Reversed by an Associative Activation Treatment. *Journal of Experimental Psychology: Animal Behavior Processes*, *33*(2), 185-190.

Dwyer, D. M., & Mackintosh, N. J. (2002). Alternating exposure to two compound flavors creates inhibitory associations between their unique features. *Animal Learning and Behavior, 30*(3), 201-207.

Dwyer, D. M., Mundy, M. E., & Honey, R. C. (2011). The role of stimulus comparison in human perceptual learning: Effects of distractor placement. *Journal of Experimental Psychology: Animal Behavior Processes, 37*(3), 300-307.

Dwyer, D. M., Mundy, M. E., Vladeanu, M., & Honey, R. C. (2009). Perceptual learning and acquired face familiarity: Evidence from inversion, use of internal features, and generalization between viewpoints. *Visual Cognition, 17*(3), 334-355.

Dwyer, D. M., & Vladeanu, M. (2009). Perceptual learning in face processing: Comparison facilitates face recognition. *Quarterly Journal of Experimental Psychology, 62*(10), 2055-2067.

Edelman, S., & Bulthoff, H. H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. *Vision Research*, *32*(12), 2385-2400.

Ellis, H. D., Davies, G. M., & Shepherd, J. W. (1978). A critical examination of the Photofit System for recalling faces. *Ergonomics, 21*(4), 297-307.

Ellis, H. D., Jeeves, M. A., Newcombe, F., & Young, A. (1986). *Aspects of Face Processing*. Dordrecht: Martinus Nijhoff.

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception, 8*(4), 431-439.

Fahle, M., Edelman, S., & Poggio, T. (1995). Fast perceptual learning in hyperacuity. *Vision Research, 35*(21), 3003-3013.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "Special" about Face Perception? *Psychological Review, 105*(3), 482-498.

Field, A. P. (2009). *Discovering statistics using SPSS (and sex and drugs and rock 'n' roll)*. 3rd edition. London: Sage.

Fleishman, J. J., Buckley, M. L., & Klosinsky, M. J. (1976). Judged attractiveness in recognition memory of women's faces. *Perceptual and motor skills, 43*(3 I), 709-710.

Fletcher, K. I., Butavicius, M. A., & Lee, M. D. (2008). Attention to internal face features in unfamiliar face matching. *British Journal of Psychology, 99*(3), 379-394.

Frowd, C. D., Carson, D., Ness, H., Richardson, J., Morrison, L., McLanaghan, S., & Hancock, P. (2005). A forensically valid comparison of facial composite systems. *Psychology, Crime and Law, 11*(1), 33-52.

Galantucci, L. M., Percoco, G., & Di Gioia, E. (2010) Low cost 3D face scanning based on landmarks and photogrammetry: A new tool for a surface diagnosis in orthodontics. *Lecture Notes in Electrical Engineering, 52,* 93-106.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3*(2), 191-197.

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

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

Gibson, E. J. (1963). Perceptual learning. *Annual Review of Psychology, 14*, 29-56.

Gibson, E. J. (1969). *Principles of perceptual learning and development*. New York: Appleton-Century-Crofts.

Gibson, E. J., & Walk, R. D. (1956). The effect of prolonged exposure to visually presented patterns on learning to discriminate them. *Journal of Comparative and Physiological Psychology, 49*(3), 239-242.

Gold, J., Bennett, P. J., & Sekuler, A. B. (1999a). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research, 39*(21), 3537-3560.

Gold, J., Bennett, P. J., & Sekuler, A. B. (1999b). Signal but not noise changes with perceptual learning. *Nature, 402*(6758), 176-178.

Gold, J., Sekuler, A. B., & Bennett, P. J. (2004). Characterizing perceptual learning with external noise. *Cognitive Science, 28*(2), 167-207.

Goldstone, R. L. (1994). Influences of Categorization on Perceptual Discrimination. *Journal of Experimental Psychology: General, 123*(2), 178-200.

Goldstone, R. L. (1995). Effects of categorization on color perception. *Psychological Science*, 298-304.

Goldstone, R. L. (1996). Isolated and interrelated concepts. *Memory and Cognition, 24*(5), 608-628.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49*, 585-612.

Goren, C. C., Sarty, M., & Wu, P. Y. K. (1975). Visual following and pattern discrimination of face like stimuli by newborn infants. *Pediatrics, 56*(4), 544-549.

Gross, R., Matthews, I., Cohn, J. F., Kanade, T., & Baker, S. (2008). Multi-PIE. *IEEE International Conference on Automatic Face and Gesture Recognition, 28*(5), 807-813.

Hall, G. (1991). Perceptual and associative learning. Oxford, UK: Clarendon Press/Oxford University Press.

Hall, G. (2003). Learned changes in the sensitivity of stimulus representations: Associative and nonassociative mechanisms. *Quarterly Journal of Experimental Psychology Section B: Comparative and Physiological Psychology, 56 B*(1), 43-55.

Hall, G. (2008). Perceptual learning. *Learing and Memory: a Comprehensive Reference. Vol.1: Learning Theory and Behavior*, 103-121.

Hall, G., Mitchell, C., Graham, S., & Lavis, Y. (2003). Acquired Equivalence and Distinctiveness in Human Discrimination Learning: Evidence for Associative Mediation. *Journal of Experimental Psychology: General, 132*(2), 266-276.

Hancock, P., Bruce, V., & Burton, A. Mike. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences, 4*(9), 330-337.

Hancock, P., Burton, A. M., & Bruce, V. (1996). Face processing: Human perception and principal components analysis. *Memory and Cognition, 24*(1), 26-40.

Hill, H., & Bruce, V. (1996). Effects of Lighting on the Perception of Facial Surfaces. *Journal of Experimental Psychology: Human Perception and Performance, 22*(4), 986-1004.

Hill, H., Schyns, P. G., & Akamatsu, S. (1997). Information and viewpoint dependence in face recognition. *Cognition, 62*(2), 201-222.

Hole, G., & Bourne, V. (2010). *Face Processing: Psychological, Neuropsychological and Applied Perspectives*. New York: Oxford University Press.

Honey, R. C., Bateson, P., & Horn, G. (1994). The role of stimulus comparison in perceptual learning: an investigation with the domestic chick. *Quarterly Journal of Experimental Psychology Section B: Comparative and Physiological Psychology, 47*(1), 83-103.

Honey, R. C., & Hall, G. (1989). Acquired Equivalence and Distinctiveness of Cues. *Journal of Experimental Psychology: Animal Behavior Processes, 15*(4), 338-346.

Honey, R. C., & Hall, G. (1991). Acquired equivalence and distinctiveness of cues using a sensory-preconditioning procedure. *Quarterly Journal of Experimental Psychology Section B: Comparative and Physiological Psychology, 43*(2), 121-135.

Howell, D. C. (2006).*Statistical methods for psychology*. 6th edition. Belmont, CA: Thomson.

James, W. (1890). *The principles of psychology*. New York, NY, US: Holt and Co.

Jeffreys, H. (1961). *Theory of Probability*. Oxford, UK: Clarendon.

Jeter, P. E., Dosher, B. A., Liu, S. H., & Lu, Z. L. (2010). Specificity of perceptual learning increases with increased training. *Vision Research, 50*(19), 1928-1940.

Jiang, F., Blanz, V., & O'Toole, A. J. (2007). The role of familiarity in three-dimensional view-transferability of face identity adaptation. *Vision Research, 47*(4), 525-531.

Jiang, F., Blanz, V., & O'Toole, A. J. (2009). Three-dimensional information in face representations revealed by identity aftereffects. *Psychological Science, 20*(3), 318-325.

Johnson, M. H. (2011). Face Perception: a Developmental Perspective. In A. J. Calder, G. Rhodes, M. H. Johnson & J. V. Haxby (Eds.), *The Oxford Handbook of Face Perception* (pp. 4-14). New York: Oxford University Press.

Johnston, R. J., & Edmonds, A. J. (2009). Familiar and unfamiliar face recognition: A review. *Memory, 17*(5), 577-596.

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., & 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.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological Science, 18*(12), 1084-1089.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Gibson, A., Smith, M., Pascalis, O. (2005). Three-month-olds, but not newborns, prefer own-race faces. *Developmental Science, 8*(6), F31-F36.

Kemp, R. (1996). Perception and recognition of normal and negative faces: The role of shape from shading and pigmentation cues. *Perception, 25*(1), 37-52.

Kemp, R., Towell, N., & Pike, G. (1997). When Seeing should not be Believing: Photographs, Credit Cards and Fraud. *Applied Cognitive Psychology, 11*(3), 211-222.

Kerr, J. R., Avons, S. E., & Ward, G. (1999). The effect of retention interval on serial position curves for item recognition of visual patterns and faces. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *25*, 1475-1494.

Krouse, F. L. (1981). Effects of pose, pose change, and delay on face recognition performance. *Journal of Applied Psychology, 66*(5), 651-654.

Lavis, Y., Kadib, R., Mitchell, C., & Hall, G. (2011). Memory for, and Salience of, the Unique Features of Similar Stimuli in Perceptual Learning. *Journal of Experimental Psychology: Animal Behavior Processes, 37*(2), 211-219.

Lavis, Y., & Mitchell, C. (2006). Effects of preexposure on stimulus discrimination: An investigation of the mechanisms responsible for human perceptual learning. *Quarterly Journal of Experimental Psychology, 59*(12), 2083-2101.

Lavrakas, P. J., Buri, J. R., & Mayzner, M. S. (1976). A perspective on the recognition of other race faces. *Perception and Psychophysics, 20*(6), 475-481.

Lawrence, D. H. (1949). Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus. *Journal of Experimental Psychology, 39*(6), 770-784.

Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience, 4*(1), 89-94.

Levin, D. T., & Beale, J. M. (2000). Categorical perception occurs in newly learned faces, other-race faces, and inverted faces. *Perception and Psychophysics, 62*(2), 386-401.

Lewis, M. B. (2004). Face-space-R: Towards a unified account of face recognition. *Visual Cognition, 11*(1), 29-69.

Light, L. L., Kayra-Stuart, F., & Hollander, S. (1979). Recognition memory for typical and unusual faces. *Journal of Experimental Psychology: Human Learning and Memory, 5*(3), 212-228.

Lindsay, R. C. L, Mansour, J. K., Bertrand, M. I., Kalmet, N., & Melsom, E. I. (2011). Face Recognition in Eyewitness Memory. In A. J. Calder, G. Rhodes, M. H. Johnson & J.

V. Haxby (Eds.), *The Oxford Handbook of Face Perception* (pp. 307-328). New York: Oxford University Press.

Liu, C. H., Chai, X., Shan, S., Honma, M., & Osada, Y. (2009). Synthesized views can improve face recognition. *Applied Cognitive Psychology, 23*(7), 987-998.

Liu, C. H., & Ward, J. (2006). The use of 3D information in face recognition. *Vision Research, 46*(6-7), 768-773.

Logie, R. H., Baddeley, A. D., & Woodhead, M. M. (1987). Face recognition, pose and ecological validity. *Applied Cognitive Psychology, 1*(1), 53-69.

Longmore, C. A., Liu, C. H., & Young, A. W. (2008). Learning Faces From Photographs. *Journal of Experimental Psychology: Human Perception and Performance, 34*(1), 77-100.

Lubow, R. E. (1973). Latent inhibition. *Psychological Bulletin, 79*(6), 398-407.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review, 82*(4), 276-298.

Mackintosh, N. J. (2009). Varieties of perceptual learning. *Learning and Behavior, 37*(2), 119-125.

Mackintosh, N. J., Kaye, H., & Bennett, C. H. (1991). Perceptual learning in flavour aversion conditioning. *The Quarterly Journal of Experimental Psychology Section B: Comparative and Physiological Psychology, 43*(3), 297-322.

Malpass, R. S. (1981a). Effective size and defendant bias in eyewitness identification lineups. *Law and Human Behavior, 5*(4), 299-309.

Malpass, R. S. (1981b). Training in face recognition. In G. M. Davies, H. D. Ellis & J. W. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 271-285). London: Academic Press.

Malpass, R. S., Lavigueur, H., & Weldon, D. E. (1973). Verbal and visual training in face recognition. *Perception and Psychophysics, 14*(2), 285-292.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London. Series B. Biological sciences, 200*(1140), 269-294.

McKone, E. (2008). Configural Processing and Face Viewpoint. *Journal of Experimental Psychology: Human Perception and Performance, 34*(2), 310-327.

McKone, E., & Robbins, R. (2011). Are Faces Special? In A. J. Calder, G. Rhodes, M. H. Johnson & J. V. Haxby (Eds.), *The Oxford Handbook of Face Perception* (pp. 150-176). New York: Oxford University Press.

McLaren, I. P. L. (1997). Categorization and perceptual learning: An analogue of the face inversion effect. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 50*(2), 257-273.

McLaren, I. P. L., Kaye, H., & Mackintosh, N. J. (1989). An associative theory of the representation of stimuli: Applications to perceptual learning and latent inhibition. *Parallel Distributed Processing: Implications for Psychology and Neurobiology*, 102-130.

McLaren, I. P. L., & Mackintosh, N. J. (2000). An elemental model of associative learning: I. Latent inhibition and perceptual learning. *Animal Learning and Behavior, 28*(3), 211-246.

McLaren, I. P. L., & Mackintosh, N. J. (2002). Associative learning and elemental representation: II. Generalization and discrimination. *Animal Learning and Behavior, 30*(3), 177-200.

Megreya, A. M., & Bindemann, M. (2009). Revisiting the processing of internal and external features of unfamiliar faces: The headscarf effect. *Perception, 38*(12), 1831-1848.

Megreya, A. M., Bindemann, M., & Havard, C. (2011). Sex differences in unfamiliar face identification: Evidence from matching tasks. *Acta Psychologica, 137*(1), 83-89.

Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, and Computers, 36*(4), 630-633.

Mitchell, C., Kadib, R., Nash, S., Lavis, Y., & Hall, G. (2008). Analysis of the Role of Associative Inhibition in Perceptual Learning by Means of the Same-Different Task. *Journal of Experimental Psychology: Animal Behavior Processes, 34*(4), 475-485.

Mitchell, C., Nash, S., & Hall, G. (2008). The Intermixed-Blocked Effect in Human Perceptual Learning Is Not the Consequence of Trial Spacing. *Journal of Experimental Psychology: Learning Memory and Cognition, 34*(1), 237-242.

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception, 31(1),* 553-566.

Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review, 98*(2), 164-181.

Mundy, M. E., Dwyer, D. M., & Honey, R. C. (2006). Inhibitory associations contribute to perceptual learning in humans. *Journal of Experimental Psychology: Animal Behavior Processes, 32*(2), 178-184.

Mundy, M. E., Honey, R. C., Downing, P. E., Wise, R. G., Graham, K. S., & Dwyer, D. M. (2009). Material-independent and material-specific activation in functional MRI after perceptual learning. *NeuroReport, 20*(16), 1397-1401.

Mundy, M. E., Honey, R. C., & Dwyer, D. M. (2007). Simultaneous Presentation of Similar Stimuli Produces Perceptual Learning in Human Picture Processing. *Journal of Experimental Psychology: Animal Behavior Processes, 33*(2), 124-138.

Natu, V., & O’Toole, A. J. (2011). The neural processing of familiar and unfamiliar faces: A review and synopsis. *British Journal of Psychology, 102*(4), 726-747.

Nosofsky, R. M. (1986). Attention, Similarity, and the Identification-Categorization Relationship. *Journal of Experimental Psychology: General, 115*(1), 39-57.

O'Donnell, C., & Bruce, V. (2001). Familiarisation with faces selectively enhances sensitivity to changes made to the eyes. *Perception, 30*(6), 755-764.

O'Toole, A. J. (2011). Cognitive and Computational Approaches to Face Recognition. In A. J. Calder, G. Rhodes, M. H. Johnston & J. V. Haxby (Eds.), *The Oxford Handbook of Face Perception* (pp. 16-30). New York: Oxford University Press.

O'Toole, A. J., Deffenbacher, K. A., Valentin, D., & Abdi, H. (1994). Structural aspects of face recognition and the other-race effect. *Memory and Cognition, 22*(2), 208-224.

O'Toole, A. J., Edelman, S., & Bülthoff, H. H. (1998). Stimulus-specific effects in face recognition over changes in viewpoint. *Vision Research, 38*(15-16), 2351-2363.

O'Toole, A. J., Peterson, J., & Deffenbacher, K. A. (1996). An 'other-race effect' for categorizing faces by sex. *Perception, 25*(6), 669-676.

O'Toole, A. J., Vetter, T., & Blanz, V. (1999). Three-dimensional shape and two-dimensional surface reflectance contributions to face recognition: An application of three-dimensional morphing. *Vision Research, 39*(18), 3145-3155.

O'Toole, A. J., Vetter, T., Troje, N. F., & Bülthoff, H. H. (1997). Sex classification is better with three-dimensional head structure than with image intensity information. *Perception, 26*(1), 75-84.

Osborne, C. D., & Stevenage, S. V. (2008). Internal feature saliency as a marker of familiarity and configural processing. *Visual Cognition, 16*(1), 23-43.

Paysan, P., Knothe, R., Amberg, B., Romdhani, S., & Vetter, T. (2009). A 3D face model for pose and illumination invariant face recognition. *IEEE International Conference on Advanced Video and Signal Based Surveillance,* 296-301.

Pike, G. E., Kemp, R. I., Towell, N. A., & Phillips, K. C. (1997). Recognizing moving faces: The relative contribution of motion and perspective view information. *Visual Cognition, 4*(4), 409-437.

Poggio, T., Fahle, M., & Edelman, S. (1992). Fast perceptual learning in visual hyperacuity. *Science, 256*, 1018-1021.

Ramon, M., Busigny, T., & Rossion, B. (2010). Impaired holistic processing of unfamiliar individual faces in acquired prosopagnosia. *Neuropsychologia, 48*(4), 933-944.

Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition, 47*(1), 25-57.

Roark, D. A., O'Toole, A. J., Abdi, H., & Barrett, S. E. (2006). Learning the moves: The effect of familiarity and facial motion on person recognition across large changes in viewing format. *Perception, 35*(6), 761-773.

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.

Ross, D. A., Hancock, P. J. B., & Lewis, M. B. (2010). Changing faces: Direction is important. *Visual Cognition, 18*(1), 67-81.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin and Review, 16*(2), 225-237.

Shapiro, P. N., & Penrod, S. (1986). Meta-Analysis of Facial Identification Studies. *Psychological Bulletin, 100*(2), 139-156.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84*(2), 127-190.

Shiu, L. P., & Pashler, H. (1992). Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Perception and Psychophysics, 52*(5), 582-588.

Solomon, G. E. A. (1997). Conceptual Change and Wine Expertise. *Journal of the Learning Sciences, 6*(1), 41-60.

Sowden, P. T., Davies, I. R. L., & Roling, P. (2000). Perceptual learning of the detection of features in X-ray images: A functional role for improvements in adults' visual sensitivity? *Journal of Experimental Psychology: Human Perception and Performance, 26*(1), 379-390.

Stacey, P. C., Walker, S., & Underwood, J. (2005). Face processing and familiarity: Evidence from eye-movement data. *British Journal of Psychology, 96,*407-422.

Stevenage, S. V. (1998). Which twin are you? A demonstration of induced categorical perception of identical twin faces. *British Journal of Psychology, 89*(1), 39-57.

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

Symonds, M., & Hall, G. (1995). Perceptual learning in flavor aversion conditioning: Roles of stimulus comparison and latent inhibition of common stimulus elements. *Learning and Motivation, 26*(2), 203-219.

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

Tanaka, J., & Gauthier, I. (1997). Expertise in Object and Face Recognition. In R. L. Goldstone, D. L. Medin & P.G. Schyns (Eds.), *Mechanisms of perceptual learning,* Vol.36, (pp. 83-125). San Diego, CA: Academic Press.

Troje, N. F., & Bülthoff, H. H. (1996). Face recognition under varying poses: The role of texture and shape. *Vision Research, 36*(12), 1761-1771.

Turati, C., Bulf, H., & Simion, F. (2008). Newborns' face recognition over changes in viewpoint. *Cognition, 106*(3), 1300-1321.

Turati, C., Cassia, V. M., Simion, F., & Leo, I. (2006). Newborns' face recognition: Role of inner and outer facial features. *Child Development, 77*(2), 297-311.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 43*(2), 161-204.

Valentine, T., & Bruce, V. (1986). The effect of race, inversion and encoding activity upon face recognition. *Acta Psychologica, 61*(3), 259-273.

Valentine, T., Chiroro, P., & Dixon, R. (1995). An account of the own-race bias and the contact hypothesis on a "face-space" model of face recognition. In T. Valentine (Ed.), *Cognitive and computational aspects of face recognition: Explorations in face space* (pp. 69-94). London: Routledge.

Valentine, T., & Endo, M. (1992). Towards an Exemplar Model of Face Processing: The Effects of Race and Distinctiveness. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 44*(4), 671-703.

Vetter, T., & Poggio, T. (1997). Linear object classes and image synthesis from a single example image. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 19*(7), 733-742.

Vetter, T., & Walker, M. (2011). Computer-Generated Images in Face Perception. In A. J. Calder, G. Rhodes, M. H. Johnston & J. V. Haxby (Eds.), *The Oxford Handbook of Face Perception* (pp. 388-399). New York: Oxford University Press.

Wang, T., Lavis, Y., Hall, G., & Mitchell, C. J. (2012). Location and salience of unique features in human perceptual learning. *Journal of Experimental Psychology: Animal Behavior Processes, 38,* 407-418.

Wang, T., & Mitchell, C. J. (2011). Attention and Relative Novelty in Human Perceptual Learning. *Journal of Experimental Psychology: Animal Behavior Processes, 37*(4), 436-445.

Wills, A. J., Suret, M., & McLaren, I. P. L. (2004). The role of category structure in determining the effects of stimulus preexposure on categorization accuracy. *Quarterly Journal of Experimental Psychology Section B: Comparative and Physiological Psychology, 57*(1), 79-88.

Woodhead, M. M., Baddeley, A. D., & Simmonds, D. C. V. (1979). On training people to recognize faces. *Ergonomics, 22*(3), 333-343.

Xiao, L. Q., Zhang, J. Y., Wang, R., Klein, S. A., Levi, D. M., & Yu, C. (2008). Complete Transfer of Perceptual Learning across Retinal Locations Enabled by Double Training. *Current Biology, 18*(24), 1922-1926.

Yin, R. K. (1969). Looking at upide-down faces. *Journal of Experimental Psychology, 81*(1), 141-145.

Young, A. , Hay, D. C., McWeeny, K. H., Flude, B. M., & Ellis, A. W. (1985). Matching familiar and unfamiliar faces on internal and external features. *Perception, 14*(6), 737-746.

Young, A., Hay, D. C., & Ellis, A. W. (1985). The faces that launched a thousand slips: everyday difficulties and errors in recognizing people. *The British Journal of Psychology*, *76*(4), 495-523.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, *16*(6), 747-759.