Bangor University

DOCTOR OF PHILOSOPHY

How familiarity influences attention and visual working memory for faces and other complex stimuli

Jackson, Margaret Cecilia

Award date:
2005

Awarding institution:
Bangor University

How Familiarity Influences Attention and Visual Working Memory for Faces and Other Complex Stimuli

Margaret Cecilia Jackson, BSc, MSc

This thesis is submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy, completed in the School of Psychology, University of Wales, Bangor.

ACKNOWLEDGEMENTS

Foremost, I thank Jane Raymond for being a supportive and inspiring supervisor. She has given me the opportunity and encouragement to achieve my goals and more. Thanks also to my other committee members, Paul Downing and Oliver Turnbull, who provided useful and insightful feedback along the way. I am grateful to Romina Palermo and an anonymous reviewer for their helpful comments on a version of Chapter 3 that is now in press in Perception and Psychophysics. I would also like to say thanks to the girls in the Consumer Lab – Nikki, Carys, Helena, Flis, and Jen – for providing some well-needed sanity and insanity now and again, and to Steph for the early days. My brother Andrew has been a great and patient proofreader – pints are definitely on me. I am grateful of the support and love from my mum and dad. They have given me some of the most important skills to achieve what I want from life. Finally, I thank Simon for his remarkable love, wisdom, and cheer throughout. This is for us.

TABLE OF CONTENTS

ABSTRACT

This thesis investigated how familiarity can ease visual processing and promote what Jacoby and Dallas (1981) called the *fluency heuristic*. I conducted a series of nine experiments that investigated the influence of familiarity on attention and visual working memory (WM). Attentional blink (AB) and visual search paradigms were employed to measure attention; a change detection task was employed to measure visual WM capacity. Familiarity effects were examined by comparing unfamiliar versus famous faces, and Hanzi experts versus Hanzi novices. Hanzi are Chinese characters, and represented complex, non-face objects. Whereas unfamiliar faces showed a significant AB effect, famous faces did not. Search for famous faces was significantly more efficient than search for unfamiliar faces; this familiarity benefit was replicated with Hanzi. Change detection was significantly more accurate and WM capacity estimates were larger for famous than unfamiliar faces, suggesting that familiarity increased visual WM maintenance effectiveness; this familiarity benefit was replicated with Hanzi. When faces and Hanzi were inverted, the familiarity advantage was removed from the visual WM task and remained only for faces in the visual search task. In summary, (upright) familiar stimuli required less attention for processing, were more efficiently encoded, and were more effectively maintained in visual WM than unfamiliar stimuli. I propose that the familiarity advantage reflects better *within-item integration* (binding between features and their configurations), perhaps enabled by enhanced neural synchrony, and supported by long-term memory representations. I present a new model – *fluency-by-integration* – to illustrate how familiarity might promote fluent visual processing and reduce the burden on attention and visual WM resources.

PART 1

AN INTRODUCTION TO OBJECT RECOGNITION, FACE RECOGNITION, AND EFFECTS OF FAMILIARITY

CHAPTER 1

INTRODUCTION

Simon is leaving for work in the morning. He locks the front door with the front door key. He unlocks his car with the car key. At work, he greets his boss in the corridor and hands her the report she requested. At lunchtime he encounters his neighbour in the bakery and asks how his holiday was. After work Simon goes to the supermarket to buy his girlfriend some of her favourite chocolates. He completed all these tasks with apparent ease.

Simon's car and house keys are visually similar and lie next to each other on his key-ring. Yet he has never mistakenly tried to unlock his car with the front door key. If Simon had not recognised his boss in the corridor he might have handed the report to the wrong person. If he mistook his neighbour for someone else in the bakery he might have looked foolish asking about a holiday that did not exist. Had he not been able to easily distinguish the correct chocolates from several cluttered shelves containing numerous brands of chocolate, he might have spent a long time in the supermarket, arrived late home for dinner, and received a frosty welcome from his girlfriend.

The ability to recognise the visual content of our environment, such as objects and faces, and discriminate between one form and another, is fundamental to normal human cognition. Visual familiarity with a stimulus, achieved through learning and exposure, allows us to deal quickly, efficiently, and accurately with incoming information. This has been termed the *fluency heuristic* (Jacoby & Dallas, 1981).

The objective of this thesis is to investigate how such fluent processing is achieved. Specifically, how are the perceptual and cognitive processes that underpin visual recognition different for familiar compared to unfamiliar stimuli? Faces and complex non-face objects (Chinese characters known as *Hanzi*) were used as stimuli. Faces were considered useful for three reasons. First, face recognition is fundamental to normal human cognition and social communication. Second, there is a large body of research into face recognition processes on which assumptions and hypotheses can be built. Third, it is easy to select unfamiliar and familiar face images for experimental comparison. I did not intend my research to focus soley on the effect of familiarity on face recognition, therefore, Hanzi were used to determine whether effects of familiarity on face recognition could apply to complex non-face objects. Familiarity effects were examined by comparing Hanzi recognition among Hanzi experts versus Hanzi novices.

This chapter reviews theories of object recognition and face recognition, and presents what is currently known about the effect of familiarity on perceptual processes. It has been suggested that because of the extreme social relevance of face recognition, humans might have evolved special perceptual mechanisms dedicated to (Kanwisher, 2000; Moscovitch & Moscovitch, 2000; Tanaka & Sengco, 1997), or at least specially suited for (Gauthier, Skudlarski, Gore, & Anderson, 2000; Gauthier & Tarr, 2002; Rossion, Gauthier, Goffaux, Tarr, & Crommelinck, 2002) face processing that are distinct from mechanisms used to recognise non-face objects. Due to such arguments, I review object recognition and face recognition theories independently. Distinctions between object and face recognition are briefly outlined below.

There are relatively few studies of the effect of familiarity on higher-level cognitive processes such as attention and visual working memory (WM). Attention is important for recognition in that, when directed to a stimulus, it can enhance perceptual processing (known as *perceptual enhancement*) (e.g., Kastner & Ungerleider, 2000; Raymond, O'Donnell, & Tipper, 1998). Attention is also required for combining perceptual information about different properties of a stimulus into a single percept (known as *binding*) (e.g., Luck & Beach, 1998). Visual WM is a complex system that enables us to maintain and manipulate visual information for a few seconds when original sensory stimuli are no longer present. It is essential for creating constant and coherent visual percepts in the presence of eye, head, and object motion that can interrupt direct sensory stimulation.

This thesis is presented in four parts. *Part 1* (Chapter 1, presented currently) provides a general review of the object and face recognition literature. *Part 2* (Chapters 2-4) focuses on familiarity and attention. Chapter 2 provides a selective review of attention theory and presents existing evidence for familiarity effects on attentional processes. In Chapters 3 and 4, two experimental paradigms – *attentional blink* (AB) and *visual search* – are used to address the effect of familiarity on attention. Chapter 3 presents a series of AB experiments (Experiments 1-3) using unfamiliar and famous faces. Chapter 4 presents four visual search experiments: Experiment 4 compared upright unfamiliar and famous faces; Experiment 5 compared inverted unfamiliar and famous faces; Experiment 6a used upright complex non-face stimuli (Hanzi) and compared expert versus novice performance; Experiment 6b used inverted Hanzi and compared expert versus novice performance. *Part 3 (*Chapters 5-6) focuses on familiarity and visual WM. Chapter 5 provides a review of WM theory. The distinction between visual, verbal, and spatial WM is outlined to provide clarity and context, but I focus on *visual* WM. To the best of my knowledge, the effect of familiarity on visual WM has not been investigated to date. Therefore, evidence for familiarity effects on verbal and spatial WM is used to draw predictions about visual WM. Chapter 6 presents four experiments designed to investigate familiarity effects on visual WM: Experiment 7 compared upright unfamiliar and famous faces; Experiment 8 compared inverted unfamiliar and famous faces; Experiments 9a and 9b used upright and inverted Hanzi respectively and compared expert and novice performance. All experiments in Chapter 6 used a change detection paradigm, outlined in the General Methods section.[1] Finally, *Part 4* (Chapter 7) provides a general discussion of the findings from all attention and visual WM experiments and returns to the concept of fluency.

Object Versus Face Recognition

The idea that the perceptual processing of faces might make use of a special set of mechanisms is rooted in three lines of evidence. The first, derived from human brain imaging studies, is that a specific brain area, the *fusiform face area* (FFA), is selectively activated by face stimuli and not by stimuli from other categories such as flowers, houses, and cars (Grill-Spector, Knouf, & Kanwisher, 2004; Kanwisher, McDermott, & Chun, 1997; McCarthy, Puce, Gore, & Allison, 1997). Such studies are supported by event-related potential (ERP) and magnetoencephalography (MEG) studies that showed face-specific neuronal activation occurred in the occipitotemporal

[1] Note that visual search and visual WM experiments were part of several large within-subjects studies as the former paradigm was used to set experimental design parameters in the latter. The same groups of participants were involved in the following linked within-subjects experiments: Experiments 4 and 7, Experiments 5 and 8, and Experiments 6a, 6b, 9a, and 9b.

region 170 ms after a face was presented, labelled the N170 and M170 components respectively (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Jeffreys, 1996; Liu, Harris, & Kanwisher, 2002; Liu, Higuchi, Marantz, & Kanwisher, 2000). The second line of evidence comes from neuropsychological investigation of patients with different discrete brain lesions. Studies have revealed a double dissociation between impairments in face recognition (*prosopagnosia*) and object recognition (*object agnosia*) – that is, object and face recognition processes typically remain intact for prosopagnosics and agnosics respectively (De Renzi, 1989; Moscovitch, Winocur, & Behrmann, 1997). The third line of evidence comes from behavioural studies that showed face perception was dramatically more impaired by inversion of the image than non-face objects (Farah, Tanaka, & Drain, 1995; Tanaka & Farah, 1993). Because inversion has considerably greater impact on the perception of faces than other classes of stimuli, this has been taken to indicate that face processing involves a unique orientation-dependent mechanism.

Evidence is provided that counters the above claims of face-specific mechanisms, however. First, studies using bird and car experts and novel objects for which expertise was developed (*Greebles*; described in more detail later) also demonstrated significant FFA activation when these stimuli were presented (Gauthier et al., 2000; Gauthier & Tarr, 2002; Rossion et al., 2002). Gauthier and colleagues suggested that the FFA might be a *centre for expertise* that is activated by any stimulus class for which expertise has been developed rather than specifically activated by faces. In the above studies, face activation remained higher than that for birds among bird experts and cars among car experts, and was attributed to the fact that face recognition experience is gained over a sufficiently longer period of time than experience with any other object category.

Second, prosopagnosia does not always produce a human face-specific deficit. A number of studies have reported prosopagnosics with impairments for discriminating exemplars of animals of the same class, such as cows (see Ellis & Young, 1989, and Farah, 1990 for reviews). Third, an inversion effect for (human) body images (Reed, Stone, Bozova, & Tanaka, 2003) and dogs (for dog experts: Diamond & Carey, 1986) has also been reported, suggesting that inversion effects might be a property of a wider range of non-face stimuli for which expertise has been developed.

Whether or not face processing involves a special perceptual mechanism is tangential to this project and I will set this issue aside. The key issues addressed in the remainder of this chapter are how objects and faces are recognised and whether familiarity affects perceptual processes subserving object and face recognition.

Object Recognition

For clarity, the word *object* will be used to denote anything tangible that exists within our environment that is not a human face.

How do we recognise objects? This question has interested many cognitive psychologists for many decades and investigations have produced three perceptual routes to object recognition: *featural*, *configural*, and *holistic*. Featural (or *local*) processes are thought to involve narrow-scale analysis of relatively small, distinctive structural and surface features of an object that are represented independently from one another. Structural features, known as *primitives*, include 2-dimensional (2-D) lines, vertices, and edges (Edelman, 1998; Tarr & Bulthoff, 1995), and more complex 3-dimensional (3-D) parts (Marr, 1982; Marr & Nishihara, 1978). Surface features include colour and texture. Configural processes are thought to involve the analysis of spatial relationships between structural parts. Holistic (or *global*) processes are thought to involve the least amount of visual detail and rely on more abstract information such as outline shape or silhouette. For example, in recognising a dog, featural analysis might focus on the colour and texture of its fur, and the shape of its ears; configural analysis might focus on the distance between its ears and the ratio of body length to leg length; holistic analysis might focus on the silhouette of its head, body, and tail. (See Peterson and Rhodes, 2003 for various reviews.)

Object recognition need not follow one of these routes exclusively, however.

Biederman (1987) proposed a model of object recognition that can be interpreted as a combination of both featural and configural processes. His *Recognition-by-Components (RBC)* theory proposes that object representations are based on 36 basic components called *geons*. Geons are 3-D shapes such as cylinders, cones, blocks, and wedges. For example, a dog might have four cylinders for legs, one cylinder for a tail, and a cone for a head. Information about the shape of each geon and its spatial relationship with other geons is combined to form a more complex structural description. An advantage of RBC theory is that a small number of geons and spatial relations can be used to describe a vast number of different objects. A major limitation is that it is not refined enough to discriminate between highly similar objects. Many objects that belong to the same object category share similar geons: because a labrador and a border collie are physically similar, accurate discrimination between the two might require more fine-grained analysis of structural shape (e.g., length of the nose) and surface features such as fur colour, length, and texture.

RBC theory implies that object recognition proceeds in a sequential fashion from featural to configural analysis. In contrast, processing of larger scale information, such as global form, has been shown to occur more rapidly than local processing (Kimchi, 1992; Love, Rouder, & Wisniewski, 1999; Navon, 1977) and suggests a coarse-to-fine object recognition procedure that Navon termed *global precedence*. Before one can recognise a border collie for example, one must first recognise that it is a dog.

Regardless of whether local or global processes are engaged first, any theory of object recognition that involves such step-by-step processes might seem effortful and time-consuming. *Image-based* theory, also known as *template matching* theory (Tarr & Bulthoff, 1998), offers a more efficient route to object recognition. According to this theory, 2-D image representations of 3-D objects are stored in long-term memory (LTM) and object recognition involves matching a representation of an actual image to a stored representation of that image. Image-based theory implies that all three routes to recognition – featural, configural, and holistic – proceed in parallel. An advantage of this theory is that recognition can proceed rapidly if an exact match exists in LTM. A major limitation is that any variation in an object's properties, such as its colour when in shade or its silhouette when viewed from a particular angle, might impair recognition. As a simple example, the letter "**A**" can be presented in various different font styles – A, ***A***, A – and in various different orientations (turn this page on its side or upside down). Yet, despite variations in form and orientation, we still recognise it as the letter "A". It has been suggested that to allow for variations in object properties multiple templates of an object in all its possible forms are stored (Tarr, 1995; Ullman, 1989). Others have proposed that changes in object form can be compensated for by various *normalisation* processes, such as mental rotation and view extrapolation, that require only a small number of object representations to generalise to a larger number of representations (e.g., Poggio & Edelman, 1990; Tarr & Pinker, 1989).

Most objects have an orientation in which they are most frequently perceived, usually what we think of as an upright position. This has been termed the *canonical* orientation and is the optimum for object recognition. There has been substantial debate on whether divergence from the canonical orientation (e.g., a sideways or upside down object view) affects recognition ability (Pinker, 1984). Object recognition that is impaired by divergence from the canonical orientation is known as *orientation-dependent* recognition, while object recognition that is unaffected by changes in orientation is known as *orientation-independent* recognition. The presence or absence of such orientation effects is commonly used to differentiate between alternative routes to object recognition. In general, featural processing accounts predict that changes in orientation will have little impact on object recognition as long as features or parts that are diagnostic for recognition are still visible (Biederman & Cooper, 1991); configural theory predicts that divergence from a canonical orientation will slow processing of spatial relations because the "top" of the object must be found before further analysis can continue (Diamond & Carey, 1986; McMullen & Jolicoeur, 1992); global processing accounts predict that changes in an object's orientation will impair recognition if a representation of the object in that orientation is not stored in LTM (Poggio & Edelman, 1990; Riesenhuber & Poggio, 1999).

Traditionally, different routes to object recognition have been neurologically dissociated by a left hemisphere (LH) versus right hemisphere (RH) distinction. Many studies have identified a LH bias for processing featural information and a RH bias for processing global (and configural) information (Heinze, Hinrichs, Scholz, Burchert, & Mangun, 1998; Martinez et al., 1997; Yamaguchi, Yamagata, & Kobayashi, 2000). Such studies typically presented hierarchical letter stimuli, known as *Navon* letters (Figure 1), in the left visual field (LVF) or right visual field (RVF). Participants were required to direct their attention to either the small, local letters (e.g., "E") or the global letter form (e.g., "H") and name the attended letter as quickly and accurately as possible. Performance was better for local letters in the RVF (LH) and for global letters in the LVF (RH). Clinical studies provide further support for this distinction. For example, Delis, Robertson, and Efron (1986) showed that patients with LH damage made more errors recalling local elements of hierarchical letter stimuli whereas patients with RH damage made more errors recalling global elements.

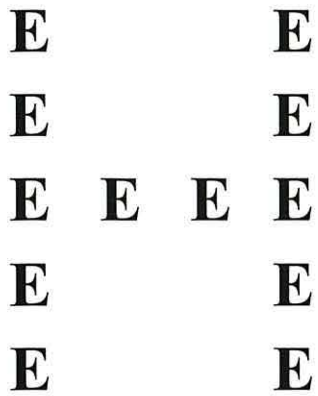


Figure 1. Example of a *Navon* figure. The global letter is "H"; the local letter is "E".

An extrapolation of orientation and hemispheric investigations suggests that featural processes are orientation-independent and are localised in the LH, whereas holistic/configural processes are orientation-dependent and are localised in the RH.

Which of these routes, or combination of routes, is most effective for object recognition?

The general conclusion is emerging that each route is used at different stages of object processing. Rosch, Mervis, Gray, Johnson, and Boyesbraem (1976) proposed a taxonomy of levels of object recognition that ranges from *superordinate* (animal), to *basic-level* (dog), to *subordinate* (border collie). Only a handful of studies have investigated the processes underlying different stages of object recognition. Hamm and McMullen (1998) suggested that superordinate and basic-level recognition is orientation-independent and subordinate recognition is orientation-dependent. Participants in their study verified whether superordinate, basic-level, and subordinate names matched objects presented at varying degrees from the upright orientation. While reaction times (RTs) in the superordinate and basic-level conditions were not significantly affected by changes in orientation, RTs in the subordinate condition were significantly slower for objects rotated from their canonical orientation.

Marsolek (1999) proposed that basic-level processes are located in the LH while subordinate processes are located in the RH. He presented objects in the LVF (RH) or RVF (LH), followed by a repeated presentation of the same object (e.g., grand piano – grand piano) or by presentation of a different exemplar from the same object category (e.g., grand piano – upright piano). Participants were required to name the second object (the *to-be-named* object) according to its subordinate level category. In the example provided here, the correct answer is "piano".[2] He found that naming was more accurate when the second object was exactly the same as the first object than when it was a different exemplar from the same category as the first, but this effect was only found for LVF presentation (RH). Naming accuracy in the two conditions did not differ in the RVF (LH). Marsolek interpreted the naming advantage for the same versus different exemplars presented in the LVF (RH) to reflect subordinate-level processes because the stored perceptual information supporting naming was able to distinguish between same and different object exemplars. Equivalent naming accuracy for the same versus different exemplars in the RVF (LH) was interpreted to reflect basic-level processes that were unable to distinguish between category exemplars.

Taken together, the two studies outlined above are consistent with the dichotomy that orientation-independent processes are located in the LH and

[2] This procedure is called a *perceptual* or *repetition priming* paradigm and is explained in more detail in the section on the effect of familiarity on object recognition.

orientation-dependent processes are located in the RH. Based on their results it could be predicted that superordinate and basic-level recognition, served by orientation-independent processes located in the LH, involve analysis of featural information, while subordinate recognition, served by orientation-dependent processes located in the RH, involves analysis of global/configural information.

In direct contrast to this prediction, however, Collin and McMullen (2005) suggested that featural processes best support subordinate discrimination while global processes best support basic-level categorisation. They asked participants to verify subordinate-, basic-, and superordinate-level names of object images that were presented in three different spatial frequency conditions: low, high, and unfiltered (see Figure 2 for an example). It has been proposed that low spatial frequencies (LSF) emphasise global/configural object attributes whereas high spatial frequencies (HSF) emphasise featural object details (Fink et al., 1997). Unfiltered images emphasise both global and local properties and acted as a control condition in this study. Their results revealed a significantly higher number of subordinate-level naming errors in the LSF (global) condition that HSF (local) and unfiltered conditions.

Figure 2. Example of an unfiltered image, a high-pass filtered image (high spatial frequency) and a low-pass filtered image (low spatial frequency), reproduced from Collin and McMullen (2005).

It is difficult to disentangle these conflicting results. More research is required to better illuminate the role of local and global processes in basic-level and subordinate object recognition.

To summarise, three hypothetical routes to object recognition have been outlined. Featural/local processes, controlled by the LH, rely on fine-grained analysis of structure and surface properties that are represented independently of one another. Changes in orientation do not affect recognition as long as features necessary for recognition remain visible. Configural processes, controlled by the RH, rely on intermediate-level analysis of object parts with specific emphasis on the spatial relation between parts. Changes in orientation impair recognition because configural analyses cannot proceed efficiently until a normalised, canonical view has been established. Holistic/global processes, controlled by the RH, rely on broad-scale analysis of shape outline. Changes in orientation impair recognition if a representation of an object in that view is not stored in LTM. It has been suggested that featural processing supports subordinate recognition and configural/holistic processing supports basic-level recognition. Evidence for this is mixed, however, and no firm conclusions have been reached at present.

The Effect of Familiarity on Object Recognition

The benefit of familiarity on object processing appears to be rapidly acquired. Compared to items presented for the first time, a single repetition of an object has been shown to improve the accuracy and speed of its identification in a succeeding encounter, without effortful or explicit recall of the initial encounter (Treisman, 1992; Tulving & Schacter, 1990). This effect, known as perceptual or repetition priming, has been shown to occur even when the observer was unaware of the first presentation of the stimulus (*subliminal priming*) (Bar & Biederman, 1998; Marcel, 1983a, 1983b).

The literature on perceptual priming is extensive and it is beyond the bounds of this thesis to present a thorough theoretical discussion of this topic. Nevertheless, several key properties of perceptual priming are relevant. First, perceptual priming effects are sensitive to repetition frequency: increasing the number of object repetitions improves object recognition and naming accuracy (Brown, Jones, & Mitchell, 1996; Wiggs, Martin, & Sunderland, 1997). Second, effects can persist over time: some facilitation of object naming has been shown to remain for 1 week (Musen & Treisman, 1990) to 48 weeks (Cave, 1997) after the initial presentation. Third, repetition priming typically leads to decreased activation of stimulus-specific neurons, an effect known as *repetition suppression* (Schacter & Buckner, 1998).

Several studies have shown that repeated exposure to a stimulus over a short period of time led to a decline in neuronal response activity (Baylis & Rolls, 1987; Li, Miller, & Desimone, 1993; Miller, Li, & Desimone, 1991; Xiang & Brown, 1998). It has been suggested that this reduction in neuronal firing rates reflects the *sharpening* of an object's representation in the visual cortex, resulting in a sparse and selectively tuned neuronal network (Desimone, 1996) that recruits only key neurons necessary for successful object recognition.

Some researchers have proposed that perceptual priming reflects the first step in the development of visual familiarity (Poldrack, Desmond, Glover, & Gabrieli, 1998; Poldrack, Selco, Field, & Cohen, 1999; Reber, Gitelman, Parrish, & Mesulam, 2005) but one piece of evidence suggests otherwise: perceptual priming effects are most pronounced when familiarity with the repeated object already exists prior to experimentation. For example, several studies found that repetition had a greater benefit for subsequent recognition of real objects than nonsense objects (Gruber & Muller, 2005; Vuilleumier, Henson, Driver, & Dolan, 2002), and of familiar stimuli (i.e., famous faces) than unfamiliar stimuli (Henson, Shallice, Gorno-Tempini, & Dolan, 2002; Schweinberger, Pickering, Burton, & Kaufmann, 2002). Furthermore, whereas repetition suppression (considered to be linked to the perceptual priming effect) has been found for repeated real and familiar objects, *repetition enhancement* – an increase in neural activity – has been reported for meaningless and novel stimuli (Henson, Shallice, & Dolan, 2000). Henson and colleagues proposed that repetition enhancement for a novel object reflects the creation of a new cortical network that is used to represent that particular object. Support for the possibility that repetition enhancement reflects object learning is provided by studies that reported higher neuronal firing rates for learned than unlearned images (Kobatake, Wang, & Tanaka, 1998; Miyashita, Date, & Okuno, 1993; Sakai & Miyashita, 1994). In addition, Holscher, Rolls, and Xiang (2003) showed that perirhinal cortex activity, considered to reflect long-term familiarity for visual objects, increased as object exposure increased.

Mixed evidence for increased and decreased neuronal activation for repeated objects muddies the view that visual familiarity is acquired quickly and simply through basic stimulus repetition. In reality, repeated exposure to the same 2-D image of an object is rare. Familiarity is more often achieved via rich and varied exposure to objects in three dimensions, under different lighting conditions, across different viewpoints, during motion, during part occlusion, and in the context of other objects and certain actions.

What are the processing consequences of more naturally developed visual familiarity that has been acquired over longer periods of time? The term *visual expertise* is often used to describe such a type of familiarity. Gauthier and Tarr (1997) defined visual expertise as the ability to recognise an object at its subordinate level as fast as it can be recognised at its basic-level category. For example, using a name verification task Tanaka and Taylor (1991) found that bird experts were as fast to recognise a picture as a "robin" as they were to recognise it as a "bird", whereas novices were slower to recognise a picture as a "robin" than a "bird". They found the same pattern of results for dog experts versus dog novices.

Studies of expert object recognition have revealed two main consequences of visual expertise: (1) increased reliance on configural information (which I will call *configural precedence*), and (2) increased activation of the FFA.

Diamond and Carey (1986) presented one of the first studies that implicated configural precedence in expert object recognition. Based on the assumption that inversion effects signal the use of configural processes, they presented upright and inverted pictures of breeds of dog to dog experts and novices, and measured the effect of inversion on recognition. They reported a significant reduction in recognition accuracy for inverted than upright dogs among experts but no significant inversion impairment for novices. In an imaginative study, Gauthier and Tarr (1997) trained participants to recognise a homogeneous set of artificial stimuli at a subordinate level. These stimuli, called *Greebles*, were given an identity by assigning them a gender category and a name. (See Figure 3 for an example of Greeble stimuli.) They compared trained (expert) versus untrained (novice) participants on a name verification task for whole Greebles versus isolated Greeble parts that were presented upright or inverted. They found that experts were significantly better at recognising whole than parts of upright Greebles whereas novices showed equivalent performance in both conditions. The benefit of familiarity for recognising wholes over parts (termed the *whole/part advantage*) was interpreted to reflect greater reliance on configural processing for intact upright stimuli for which expertise has been developed. When Greebles were inverted, the whole/part advantage for experts disappeared, a result considered indicative of the disruptive effect of inversion on configural processing.

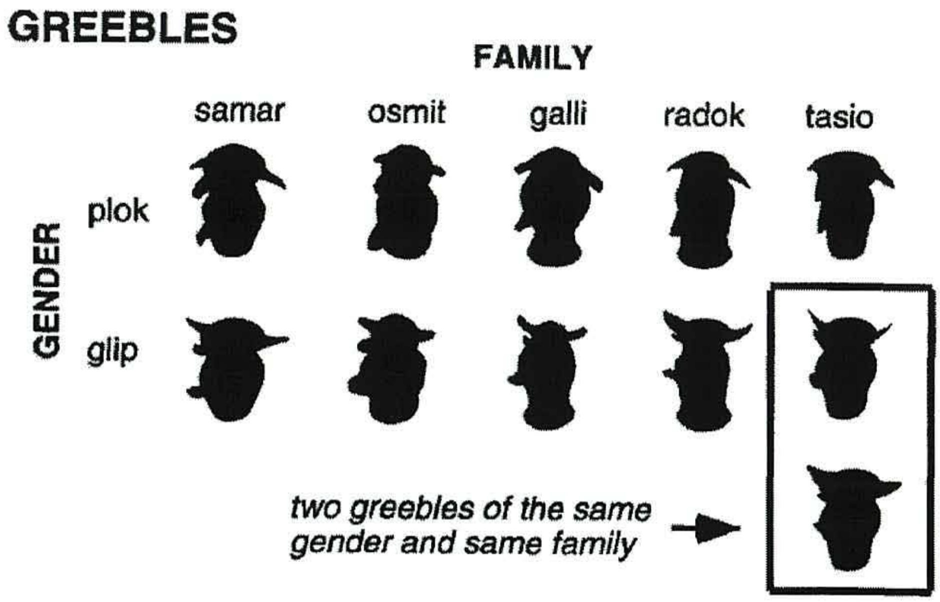


Figure 3. Example of Greebles, reproduced from Gauthier and Tarr (1997).

Neurophysiological evidence of configural precedence for familiar objects is provided by a study of single-cell activity in monkeys. Baker, Behrmann, and Olson (2003) found that neurons in inferior temporal cortex (IT), a region considered important for object recognition (Desimone & Ungerleider, 1989), responded more selectively for whole than for parts of learned objects compared to novel objects.

A second consequence of expert object recognition is increased activation of the FFA, a region considered specialised for face or expert visual processing. These effects were outlined in the previous section that compared object and face recognition but they are briefly recapped here. Using Greeble, car, and bird experts, Gauthier and colleagues consistently showed that FFA activation was greater for experts than novices (Gauthier et al., 2000; Gauthier, Tarr, Anderson, Skudlarski, &

Gore, 1999; Gauthier & Tarr, 2002). Convergent with this, the N170 ERP component – considered to reflect the structural encoding of faces and non-face objects for which expertise has been developed – responded more strongly in Greeble, dog, and bird experts than in novices of such stimuli (Rossion et al., 2002; Rossion, Kung, & Tarr, 2004; Tanaka & Curran, 2001). FFA and N170 activation has long been considered specific to faces, but the results of the above studies led some researchers to consider that each signified expert visual recognition in general (e.g., Rossion et al., 2004; Tarr & Gauthier, 2000).

Face Recognition

Faces are unarguably an extraordinary class of visual stimulus. We all have one. From birth onwards we typically encounter them every day. We recognise individuals almost exclusively by their face. And we retain the ability to recognise people close to us, such as family and friends, despite significant changes in facial appearance that can result from child-to-adult development and aging.

How do we recognise faces? Three different routes to object recognition were outlined in the previous section: featural, configural, and holistic. These terms are also used in the face processing literature. There is consensus that, as for objects, featural processing of faces involves structural parts of the face that are analysed independently of each another, such as the eyes, nose, and mouth, and surface features such as eye and skin colour. Similarly, it is agreed that configural processes rely on the analysis of spatial relations between structural facial features. There is little consensus, however, on what constitutes holistic processing of faces.

Tanaka and Farah (1993) offer a view of holistic processing that is most similar to the object recognition account. They proposed that faces are processed as unparsed wholes in which there is no internal part structure. For others, the concept of holistic processing is tightly interwoven with that of configural processing. Bartlett, Searcy, and Abdi (2003) proposed that holistic processing involves the analysis of configural information that spans the entire face, including internal components such as the eyes, nose, and mouth, and external components such as chin line and hair. Moscovitch et al. (1997) suggested that holistic processing involves the analysis of configurations that span the internal face region only. For Schyns and Gosselin (2003), the term holistic simply means that configural information from several separate regions (whether spatially linked or not) is diagnostic for accurate recognition. Diamond and Carey (1986) introduced the terms *first-order relational information* to describe the overall arrangement of parts spanning the whole face, and *second-order relational information* to describe more localised relations between features. For clarity I will use the term configural rather than holistic, and, where necessary, I will use the terms *broad-span* and *narrow-span* to refer to first- and second-order relational information respectively.

A key question that has fuelled significant debate among face recognition researchers over the past thirty years is: which of these routes – featural or configural – is most effective for face recognition?

In the previous section on object recognition, evidence was outlined that suggested featural processes were most effective for subordinate discrimination (e.g., border collie) whereas holistic processes were most effective for basic-level (e.g., dog) and superordinate (e.g., animal) categorisation. At which level does face recognition proceed? To recognise that a face is a face requires basic-level categorisation, just like recognising that a dog is a dog. However, subordinate recognition is different for objects than for faces. Subordinate object recognition is said to involve within-category discrimination, for example between one breed of dog (border collie) and another (labrador). But this does not identify the individual dog, only the breed of dog. The equivalent level of recognition for faces might involve the distinction between males and females or Caucasians and Afro-Carribeans. Although we sometimes need to attend to properties of gender and race, typically we *identify* a face on a more refined individual level: "this is my neighbour".

Therefore, a further level of recognition is required to adequately describe how we recognise faces. To the best of my knowledge, no term exists at present so I will call this *micro-subordinate* recognition, and in the following section the term *face recognition* will implicitly assume this level of person identification. As a side issue, note that micro-subordinate recognition is not specific to faces: a border collie can be identified as the neighbour's border collie; a Volkswagon can be identified as my Volkswagon. Processes that subserve micro-subordinate object recognition do not appear to have been investigated, so the question of whether featural or configural routes are most effective for face recognition must be addressed directly from the face recognition literature.

The relative contribution of featural and configural processes to face recognition has been widely debated and continues to be so. At one extreme, some researchers have proposed that face recognition is primarily based on independent feature analysis. At the other extreme, it has been suggested that face recognition is primarily based on broad-scale configural analysis. An intermediary view is that faces are recognised using a combination of featural and configural analyses, known as the *dual-route hypothesis*. Each of these three theories is reviewed in turn.

Featural Face Processing

When we describe a face we often use feature descriptions. Examine the face in Figure 4 below and imagine describing it to someone else: he has a rounded nose, a full mouth, large eyes, and short hair.

Figure 4. A face.

Perhaps as a consequence of this, many applied approaches to face recognition focus on facial features. For example, the police force often uses Identikit or Photofit procedures to help eyewitnesses recognise the criminal's face, an approach that involves constructing a face resembling that of the criminal on a feature-by-feature basis. However, only a handful of studies provide empirical evidence that featural processing plays a primary role in face recognition.

Macho and Leder (1998) presented two faces, Target A and Target B, side by side for 8000 ms. Participants were instructed to examine each target face closely and determine whether a subsequently presented single test face was more similar to Target A or Target B. The test face could differ from either of the target faces in its local features (nose width and mouth size) and/or a configural property (distance between the eyes). A varied range and combination of featural/configural changes were applied. They found that similarity decisions were sensitive to local feature differences independently of a change in the other feature, and concluded that face processing involves independent feature analysis.

Leder and Bruce (2000) provided results that can be similarly interpreted. During a study session, their participants learned to attach names to two sets of six faces, a *local* set and a *local+relational* set. The local set had the same shaped features in the same spatial relationship, but differed in respect of local properties that did not affect configuration (hair, eye, and mouth colour). The local+relational set had the same range of hair, eye, and mouth features as the local set, but they were displaced vertically and horizontally within the same facial outline. The recognition test involved matching each face to its learned name. The results showed that recognition performance with the local+relational face set was not significantly better than the local face set, suggesting that featural information alone was sufficient for effective face recognition.

Configural Face Processing

Compared to featural processing, there is substantially more evidence to suggest that configural processing is the primary contributor to face recognition. Face inversion effects are frequently used to diagnose the presence or absence of configural processing. Just as significant inversion effects on object recognition are taken as evidence for global processing, significantly impaired recognition of inverted versus upright faces is similarly interpreted as evidence for configural processing of upright faces.

Several studies have suggested that upright face recognition relies on configural, orientation-dependent processes, whereas inverted face recognition relies on featural, orientation-independent processing, engaged when configural information is unavailable. Tanaka and Sengco (1997) created two configurations of a face, one with eyes close together and one with eyes far apart. After subjects studied faces presented in one of the two configurations they were tested for their recognition of specific features in three different test conditions: isolated features (i.e., features removed from the face entirely), features shown in a new face configuration, and features as shown in the old face configuration. Performance was highest when features were presented in the old face configuration and poorest when they were presented in isolation. This result is similar to the whole/part advantage reported in

Greeble studies. In contrast, when test stimuli were inverted the advantage for whole faces over isolated features disappeared. This finding favours the view that configural information is important for normal, upright face recognition.

Tanaka and Farah (1993) reached a similar conclusion. They taught participants to match names to either whole faces or to isolated face parts and then tested their ability to name upright and inverted whole faces in a subsequent session. Whereas performance was better for upright than inverted faces when whole faces had been learned, this inversion effect was absent when isolated face parts had been learned.

Using a different approach, Young, Hellawell, and Hay (1987) asked participants to make recognition judgements of the top half of face composites (faces created by merging the top half of one face with the bottom half of another face) that were either aligned or misaligned horizontally. They found that, when faces were upright, naming was slower for aligned than misaligned face halves, known as the *composite effect*. This benefit of misalignment is thought to occur because configural processing of the composite is disrupted, allowing the top half to be processed without interference from the bottom half. When inverted faces were presented, however, there was no benefit of misalignment, suggesting that configural processes were not engaged in inverted face recognition (see also Carey & Diamond, 1994; Hole, 1994).

Studies of left/right visual field effects lend support for the role of configural processing in face recognition: a robust RH advantage for face recognition has been found. For example, normal upright faces were identified more rapidly and accurately when they were presented in the LVF than in the RVF (Hillger & Koenig, 1991; Rhodes, 1993). Consistent with the proposal that RH activity is associated with global object processing, dominant RH activity for faces is interpreted as dominance for configural face processing. Hillger and Koenig found that when inverted faces were presented to the LVF and RVF, the RH advantage seen for upright faces was eliminated.

Neuroimaging and clinical evidence lends further support for RH, configural dominance in upright face processing. Functional magnetic resonance imaging (fMRI) studies of normal individuals have shown that face recognition is mediated predominantly by a right-lateralised neural network (Haxby et al., 1996; Kim et al., 1999; McCarthy et al., 1997; McDermott, Buckner, Petersen, Kelley, & Sanders, 1999; Young, Hay, & McWeeny, 1985). The majority of patients with prosopagnosia have specific damage to the right FFA (De Renzi, 1989; Sergent & Signoret, 1992). Prosopagnosia confined mainly to the RH has been interpreted as a specific deficit in configural processing, as featural processing mechanisms appear to remain intact. For example, Farah, Wilson, Drain, and Tanaka (1995) measured prosopagnosic patient *LH*'s ability to recognise upright compared to inverted faces. Patient *LH* demonstrated opposite results to that obtained in studies using healthy participants: his recognition of inverted faces was superior to his recognition of upright faces. These findings were interpreted to reflect his ability to use featural information to recognise inverted faces and his inability to use configural information to recognise upright faces.

The finding of a double dissociation between prosopagnosia and object agnosia lends further support to the distinction between configural upright and featural inverted face processing. Moscovitch and colleagues reported a series of studies with an individual with object agnosia, patient *CK*. While *CK* showed impairments in object recognition, his ability to recognise upright faces was intact. In contrast, his ability to recognise inverted faces was significantly poorer than controls (Moscovitch & Moscovitch, 2000; Moscovitch et al., 1997).

Finally, ERP studies demonstrated that the N170 component was significantly larger in the right than the left hemisphere (e.g., Sagiv & Bentin, 2001). An increased and delayed N170 amplitude has also been shown for inverted relative to upright faces, an effect considered to reflect the enhanced difficulty of processing faces without the availability of configural information (Rossion et al., 1999; Rossion et al., 2000). As Sagiv and Bentin pointed out, however, the assumption that an enhanced N170 amplitude reflects greater difficulty in face processing is at odds with fMRI data that showed inverted faces produced less activity in the FFA than upright faces (Kanwisher, Tong, & Nakayama, 1998).

Dual-Route to Face Recognition

A dual-route to face recognition has been proposed that involves both featural and configural processes. Some researchers defined this dual-route hypothesis on the basis of configural processes for upright faces and featural processes for inverted faces, as discussed above. Others suggested that both featural and configural processes are used for upright face recognition, but that configural analysis is the hallmark of efficient recognition, hence the sensitivity of face recognition to inversion (Bartlett et al., 2003; Rhodes, Brake, & Atkinson, 1993). The majority of configural processing advocates based their view on inversion effects, although Rakover and Teucher (1997) found inversion effects for isolated facial features (forehead, eyes, nose, mouth, and chin). This evidence casts doubt on the use of inversion effects to diagnose exclusivity of configural processing in whole upright faces.

ERP data supports the dual-route hypothesis. Based on their finding that the N170 component is significantly elicited by isolated facial features and face contours in addition to whole intact faces, Bentin et al. (1996) suggested that face processing is modulated both by featural and configural mechanisms.

Moscovitch and Moscovitch (2000) proposed an *interactive activation* model in which face recognition involves the exchange of featural and configural information. Their idea is based on findings that agnosic patient *CK*'s object recognition deficit lies not with the identification of object parts *per se*, but in appreciating their relation to one another and integrating them into a whole percept (e.g., Behrmann, Moscovitch, & Winocur, 1994). The interactive activation model posits that when configural information is not readily available, as in inverted faces, featural information is analysed and translated into configural information, which is then used to recognise a face. Indeed, in contrast to reports of a RH bias for face processing, several functional imaging studies have shown bilateral fusiform activation elicited by faces (e.g., Henson et al., 2003; Katanoda, Yoshikawa, & Sugishita, 2000; McCarthy et al., 1997; see Kampf, Babkoff, & Nachson, 2005 for a review).

It is possible that the diversity of evidence for featural and/or configural routes to face recognition arises from the variety of tasks employed and the nature of the face stimuli used. The majority of studies typically involved matching, naming (learned face-name associations), or passive viewing tasks and used unfamiliar (or recently learned) faces. There are three key limitations to assessing mechanisms underlying face recognition in these ways. First, matching unfamiliar faces might have little to do with face processing *per se* but is merely a perceptual exercise in matching general pictorial information. Second, studies that involve naming unfamiliar faces could be considered a measure of explicit short-term memory that is bound to the context of a learning session. As in perceptual priming studies, they tell us little about natural long-term face recognition processes. In addition, tasks with and without a naming component might differentially affect laterality results: LH activation has been found for face naming and RH activation has been found for perceptual analysis of faces (e.g., Tsukiura et al., 2002). Finally, ERP measures of brain activation from passive viewing of faces appear to provide information on how we recognise that a face is a face, i.e., basic-level recognition, but might tell us little about how we identify faces at a micro-subordinate level. Several reports that N170 activation was unaffected by face familiarity support this possibility (Bentin & Deouell, 2000; Eimer, 2000a, 2000b).

Familiar faces that have been learned through long-term natural exposure contain a larger amount of additional related information that can be activated on visual presentation than unfamiliar or artificially learned faces. For example, we typically know concrete things such as the person's name, occupation, marital status, where he or she lives, and also abstract things such as his or her political and social beliefs (all examples of *semantic* information). Bruce and Young (1986) presented a model that incorporates the rich and varied information we might use to recognise familiar faces. Their model contains four stages to face recognition. Stage 1 involves the use of featural and configural information to create a structural representation of the face. Stage 2 involves comparisons between the structural representation and stored representations in LTM known as *face recognition units* (FRUs). If a match is found, a sense of familiarity arises at this stage. Activation of a matching FRU then allows Stage 3 to proceed, in which *person identity nodes* (PINs) are accessed. PINs contain specific semantic information about the individual that has been stored in the past. The final stage involves access to the person's name, considered possible only through PIN activation.

Bruce and Young's (1986) model implies that to fully recognise a person his/her name must be retrieved via detailed semantic information. Yet you might be able to think of many instances where you are visually familiar with a person and not know his/her name or many details about him/her: for example, a neighbour that you greet in the morning, a waiter in your favourite restaurant, the clerk in your local post office. Although widely cited, their model tells little of how perceptual and cognitive processes involved in visual recognition are different for unfamiliar and familiar faces.

The Effect of Familiarity on Face Recognition

Look at the faces in Figure 5 below. Subjectively, your perception of the face on the right might feel different from your perception of the other two faces. You might even feel different looking at the face in the middle compared to the face on the left. The face on the right is the current Prime Minister of Britain – Tony Blair – a person highly visually familiar to many. You have most likely never seen the face on the left before. The face in the middle might seem a little familiar to you: it was presented on page 26.

Figure 5. These three faces illustrate different levels of familiarity. The face on the left is *novel*, presented for the first time here. The face in the centre was *recently seen* on page 26. The face on the right is a famous face – Tony Blair (British Prime Minister at the time of writing).

The majority of people we interact with are highly familiar to us: family, partner, friends, and colleagues. Television and cinema also provide us with a vast number of people that we find highly recognisable: actors, presenters, newsreaders, and politicians. It is less frequent, and probably less important, that we recognise less familiar people such as retail assistants, bank clerks, and the average person on the street. Understanding how we recognise highly familiar faces is valuable to understanding how visual processes might aid effective social communication. Yet surprisingly little research has directly addressed whether perceptual processes differ for recognition of highly familiar versus less familiar faces. Existing evidence is varied and disorganised.

For clarity, the term *unfamiliar* is used to describe faces that have been seen once before or recently learned over multiple exposures in an experimental setting, and the term *familiar* is used to describe faces that have been naturally learned over longer periods of time (such as famous or personally familiar faces). A face presented once only for the very first time is described as a *novel* face.

As discussed in the previous section, studies of expert object recognition suggested that, compared to recognition of unfamiliar objects, recognition of highly familiar objects relies more heavily on configural processes, diagnosed by greater inversion effects, RH dominance, and greater use of low spatial frequency (global) information than high spatial frequency (local) information. Expert object recognition has also been found to result in greater FFA activation.

Familiarity effects on face recognition are generally inconsistent with what might be predicted from the object recognition literature, however: there is little evidence to suggest greater reliance on configural processing and little evidence for greater activity in the FFA for familiar compared to unfamiliar faces.

To examine the relative roles for featural and configural processing in familiar and unfamiliar face recognition, Collishaw and Hole (2000) examined the extent of recognition impairment produced by blurring, inversion, and scrambling (and combinations of these modifications, not discussed here). Participants stated whether a face was famous or non-famous (task 1), or learned or unlearned (task 2), as quickly and as accurately as possible. The authors reasoned that if configural information were more important for familiar than unfamiliar face recognition, inversion and/or scrambling would affect recognition of familiar faces more than unfamiliar faces. Based on the assumption that featural processes dominate the recognition of unfamiliar faces, blurring was expected to impair recognition of unfamiliar faces more than familiar faces due to the disruption of fine-grained local features. Results were compared to recognition of un-manipulated control faces presented in a block of trials prior to the manipulation trials. Recognition of both familiar and unfamiliar faces was impaired in inverted, scrambled, and blurred conditions. Although the impact of manipulation was overall smaller for unfamiliar faces, the authors interpret similar manipulation effects in each condition to reflect equivalent use of featural and configural processes for familiar and unfamiliar face recognition.

In stark contrast to the prediction that familiar face recognition relies heavily on configural processes, Sergent (1986, cited in Bruce, 1989) found that familiar faces were recognised more efficiently when fine-grained local information was available than when it was removed. Using an approach similar to that adopted in studies of object recognition, she presented high and low spatial frequency (HSF, LSF) images of familiar faces (participants' work colleagues) to right or left visual fields and measured naming speed. Results showed that HSF images showed a LH advantage while LSF images showed a RH advantage, consistent with suggestions from object recognition studies that fine-grained features are processed more effectively in the left hemisphere and coarse-grained configural (or global) properties are processed more effectively in the right hemisphere. The key finding was that participants named HSF faces significantly faster than LSF faces. This is consistent with the notion from traditional object recognition literature that analysis of fine details supports subordinate (or in this case, micro-subordinate) recognition, but inconsistent with findings that expert object recognition promotes greater use of configural information. Sergent did not conduct this experiment with unfamiliar faces, therefore our understanding of how recognition of familiar faces might differ from unfamiliar faces according to the type of information provided by different spatial frequencies is limited.

Investigations into face familiarity effects on FFA activation have found no difference between familiar and unfamiliar faces (Gorno-Tempini & Price, 2001; Gorno-Tempini et al., 2000). Like the N170 component, the FFA might be specifically suited for structural encoding of faces – i.e., used for basic-level recognition that a face is a face – rather than for discrimination between different face identities. But there is little evidence to support this possibility: prosopagnosic patients do not demonstrate impaired discrimination between faces and non-face objects; counter-intuitively, larger FFA activation for *unfamiliar* than famous faces has been found (Rossion, Schiltz, & Crommelinck, 2003). Rossion et al. interpreted their results to reflect the role of the FFA in discriminating between unfamiliar and familiar faces rather than between different individual identities.

Setting aside comparisons with object recognition, several differences between familiar and unfamiliar face recognition processes have been demonstrated. Studies have shown that, compared to recognition of unfamiliar faces, recognition of familiar faces (1) is reliant on internal rather than external face regions, (2) is more robust to image quality degradation, (3) is less affected by changes in viewing angle (i.e., rotation on a vertical axis), (4) is based on a wider neural network region, and (5) involves greater bilateral hemispheric cooperation.

Evidence that familiar face recognition is more reliant on internal (i.e., eyes, nose, and mouth) than external (i.e., head outline, hair, ears, and chin) face regions was established over 20 years ago (e.g., Ellis, Shepherd, & Davies, 1979; Young, Hay, McWeeny, Flude, & Ellis, 1985). Recent investigations continue to support this finding. Clutterbuck and Johnston (2002) presented complete images of unfamiliar, moderately familiar, or highly familiar faces simultaneously with images of internal and external face regions. Participants decided if whole and regional images depicted the same individual (match) or not (non-match). They found that correct whole/internal matches were made significantly faster than correct whole/external matches for highly familiar than moderately familiar or unfamiliar faces. Correct whole/internal matches were also significantly faster than correct whole/external matches for moderately familiar than unfamiliar faces, indicating a graded shift in the use of external features for unfamiliar face processing to internal features for highly familiar face processing. Similarly, Bonner, Burton, and Bruce (2003) found that as unfamiliar faces were learned over a period of three days, whole/internal match decision accuracy significantly improved while whole/external performance remained constant. The internal face region bias for familiar faces is considered to reflect the use of stable, unchangeable, learned components (and their relations) to guide recognition. External regions, such as hair, are considered unstable, changeable characteristics that provide less reliable information on which to base a recognition judgement, but which are relatively distinctive and allow for early recognition judgements to be made (Ellis et al., 1979).

Studies using poor quality video surveillance footage to test face recognition have found that recognition performance was superior for personally familiar faces than unfamiliar faces seen prior to the test, despite degraded image quality (Bruce, Henderson, Newman, & Burton, 2001; Burton, Wilson, Cowan, & Bruce, 1999).

Using a perceptual priming procedure combined with ERP measures, Schweinberger, Pickering, Jentzsch, Burton, and Kaufmann (2002) showed that familiar face recognition was not impaired by changes in the rotation angle of the face image (i.e., frontal, three-quarter, and profile views). When both first (the prime) and second (to-be-named) images were of the same famous person in the same view, they found a negative waveform that peaked at about 250 ms post-stimulus. Termed the N250r, this component was still found when the first and second images were different views of the same face, albeit a smaller effect than when the same view was repeated. Schweinberger and colleagues concluded that famous faces are represented in a view-independent, abstract manner and suggested that the N250r component to some extent reflects access to stored representations of familiar faces. Relative to familiar faces, unfamiliar faces have produced smaller or negligible N250r effects using this procedure (e.g., Begleiter, Porjesz, & Wang, 1995; Pfutze, Sommer, &

Schweinberger, 2002; Schweinberger, Pfutze, & Sommer, 1995) and were not tested in this study.

In a similar vein, Pourtois, Schwartz, Seghier, Lazeyras, and Vuilleumier (2005) measured perceptual priming effects for familiar and unfamiliar faces using fMRI. The second face presented was always a different view of the first face. During scanning, participants made gender judgements to famous and unfamiliar faces and were told that familiarity was irrelevant to the task. Counter to the suggestion that famous faces are robust to changes in viewpoint, they found no priming effect, i.e., no reduction in activation, for repeated famous or unfamiliar faces in left or right FFA regions.

There is evidence that recognition of familiar faces involves a wider neural network than recognition of unfamiliar faces. Leveroni et al. (2000) compared recognition for familiar and unfamiliar faces. In an encoding session, participants were instructed to memorise familiar and unfamiliar faces for a subsequent identification task (unfamiliar faces therefore became recently learned faces). The identification task involved old/new judgements referring specifically to whether a face was presented in the encoding session. While recognition accuracy and speed for both familiar and unfamiliar faces was equivalent, there was evidence of wider spread bilateral activation involving prefrontal, lateral temporal, and mesial temporal (hippocampal and parahippocampal) regions for familiar faces. Rather than indicating differential perceptual processes, however, such findings were interpreted to reflect greater use of semantic information (e.g., name and occupation) used to aid recognition of familiar faces. ERP studies reported greater negativity 400 ms post-stimulus (N400) and greater positivity 600 ms post-stimulus (P600) for familiar than unfamiliar faces (Bentin & Deouell, 2000; Eimer, 2000a, 2000b; Schweinberger, Pickering, Burton et al., 2002), and such data were similarly interpreted to reflect activation of semantic information.

While many studies of face recognition using unfamiliar faces provide evidence for a right hemisphere bias, investigations of the effect of familiarity on face recognition have found a bilateral advantage for familiar over unfamiliar faces. Mohr, Landgrebe, and Schweinberger (2002) presented famous and unfamiliar faces in the LVF, RVF, and bilaterally to both visual fields. Participants made fame judgements (famous or non-famous). They found that ability to judge a famous face as famous was significantly better when faces were presented bilaterally, whereas they found no difference in unfamiliar face judgement accuracy between any of the conditions. Replicated by Schweinberger, Baird, Blumler, Kaufmann, and Mohr (2003), these results were interpreted to reflect better inter-hemispheric cooperation for recognition of familiar faces.

In summary, there is little evidence that recognition of familiar faces involves greater reliance on configural information or increased FFA activation. Greater reliance on information contained internally and better resistance to image quality degradation suggests that familiar face representations are more stable and robust than unfamiliar face representations. The presence of a bilateral advantage for familiar faces could be interpreted in light of the dual-route hypothesis of face recognition. Perhaps all faces use featural information processed in the left hemisphere and configural information processed in the right hemisphere, but integration of featural and configural processes is better for familiar faces.

PART 2

THE EFFECT OF FAMILIARITY ON ATTENTION

CHAPTER TWO

THEORETICAL REVIEW OF ATTENTION, THE ATTENTIONAL BLINK EFFECT, AND THE ROLE OF FAMILIARITY

This chapter serves two purposes. First, it provides an overview of attention theory, focusing specifically on limits in attentional capacity and the attentional blink (AB) – the paradigm used in Experiments 1-3. Second, it presents current evidence for the impact of familiarity on attention.

The word *attention* can have several different connotations. Medin, Ross, and Markman (2005) provide a useful description: when reading a book, attention can mean *concentration*; when listening to a particular conversation in a crowded room, attention can mean *selection*; when talking about being able to attend to a limited number of things at one time, we are referring to *limits in capacity*; performing a task effortlessly with little attention is known as *automaticity*. An additional role of attention is *rejection*, or *inhibition*, of irrelevant information. For example, when selectively listening to a particular conversation in a crowded room, we must filter out other conversations to prevent interference.

Most relevant to the current study are attentional limits and automaticity. Evidence for limits in attention are reviewed below. The subsequent section – the effect of familiarity on attention – reviews evidence for our ability to recognise visual stimuli in a seemingly automatic manner, using minimal attentional resource.

Limits of Attention

It has been argued that a limited capacity attentional reource exists in the form of a *central pool* of attention from which portions are allocated to relevant, attention-demanding tasks (e.g., Broadbent, 1958; Kahneman, 1973). If insufficient attention is available, impairments in perception can occur. The detrimental impact on perceptual processing when little or no attention is available has been demonstrated by use of *dual-task* (or *divided-attention*) paradigms. In such paradigms, attention is absorbed by a central task and performance on a concurrent (or rapidly succeeding) secondary task is measured. If there is insufficient residual attention from the first task, performance on the secondary task will suffer. This is known as a *dual-task cost.*

There are two main schools of thought regarding how dual-task costs occur. The *bottleneck* model of attention posits that the central attentional pool is prevented from overload by means of a selective filter that sifts through incoming information (Broadbent, 1958). This filter admits only relevant information at the exclusion of other types of information, as in the selection and rejection processes outlined above. Eventually, the amount of information being selected and rejected reaches a maximum; this is the point at which a bottleneck in attentional resource is reached and processing of subsequently presented stimuli is impaired. Alternatively, the *attentional capacity* theory proposes that the inability to perform two tasks at the same time results from general depletion of the central pool of attentional resource (Kahneman, 1973). This theory assumes that information can be processed until attentional resources are fully drained.

Attentional limits have been demonstrated in several ways and have coined a number of different effects – *repetition blindness*, *change blindness*, and *inattentional blindness*. One of the most notable and widely studied demonstrations of dual-task costs is the *attentional blink* (AB) effect. The AB effect is reviewed in depth below, followed by a brief description of the other effects.

Attentional Blink (AB) Effect

The AB effect describes impaired report of a second target, *T2*, when it is presented in close temporal succession to a first target, *T1* (Raymond, Shapiro, & Arnell, 1992). It was analogised to an eye blink because the large portion of attentional resource allocated to T1 processing appears to render T2 "invisible" due to lack of available attention, as if the observer closed his/her eyes at the time of its presentation.

In the original demonstration of the AB, Raymond et al. (1992) presented letters in *rapid serial visual presentation* (RSVP). Each letter, chosen at random from the alphabet, was black in colour and presented in the centre of a computer screen at an approximate rate of 11 letters per second. T1 and T2 were embedded in an RSVP stream of letters (known as *distractors*): T1 was a white letter, and T2 was the letter X. T1 always preceded T2. The crucial manipulation was the serial position of T2 relative to T1 (known as the T2 *lag*). Figure 6a illustrates a typical trial. In a dual-task condition, participants were required to report the identity of the white letter and state the presence or absence of the letter X. In a single-task control condition, participants ignored T1 and only responded to T2. In the dual-task condition, Raymond and colleagues found that detection of T2 was significantly impaired when it was presented between 160 ms and 480 ms after T1 – the AB effect. In contrast, they found no T2 processing deficit in the single-task condition. Figure 6b illustrates their results. They suggested that the AB effect is caused by suppression of T2 processing, and that this suppression is initiated by attentional events related to T1 processing.

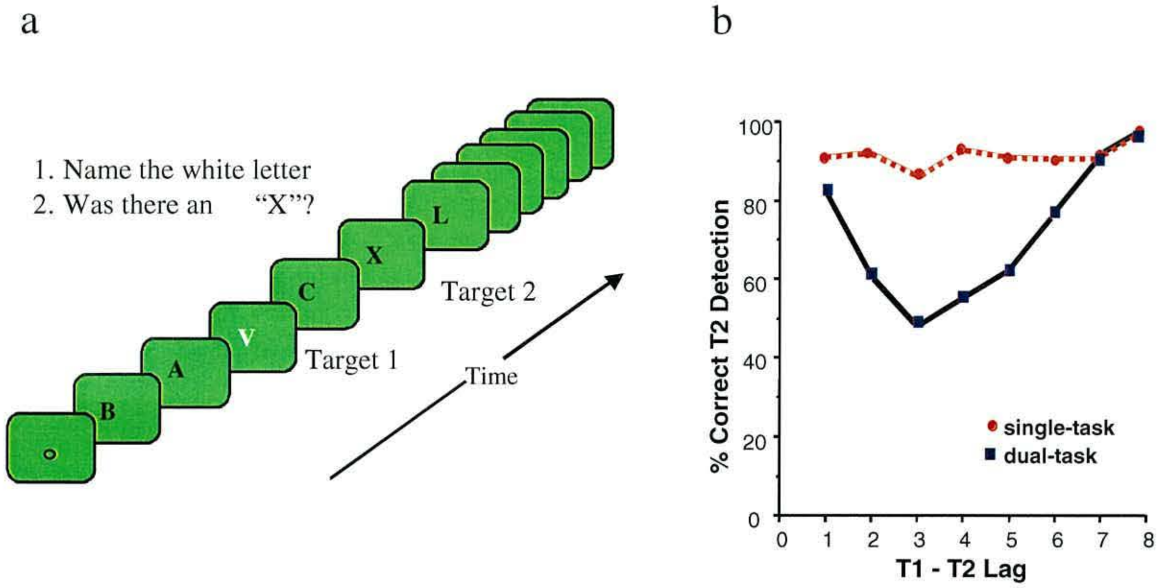


Figure 6. (a) Example of an AB trial as used in the original study by Raymond et al. (1992). The T1 task was to name the white letter and the T2 task was to state whether the letter X was present or absent from the RSVP stream. (b) Example of an AB effect as found in Raymond et al. When participants were required to perform both T1 and T2 tasks (dual task condition), detection of T2 was impaired when it followed in close temporal succession to T1 (short T1-T2 lags). This T2 deficit is known as the AB effect. When only the T1 task was performed (single-task condition), the AB effect was absent.

In further experiments outlined in the same paper, Raymond et al. (1992) found that an item had to immediately follow T1 in order to produce the AB effect. No AB effect was found when a blank screen followed T1 presentation. They suggested that the item immediately following T1 (the T1 + 1 item; often labelled the *T1 mask*) interfered with T1 processing, and that this interference was necessary to initiate attentional suppression of T2. Subsequent studies found that the AB effect was also dependent on an item immediately following T2 – the *T2 mask* (Giesbrecht & Di Lollo, 1998), and suggested that, under conditions of limited attention, T2 is susceptible to interference from succeeding stimuli.

In an earlier paper, Reeves and Sperling (1986) suggested that attention is allocated episodically to each item in the RSVP stream, and that target identification involves the opening and closing of an *attentional gate*. When T1 is detected, the gate opens to enable T1 processing and remains open until processing is complete. Due to its close temporal proximity to T1, the T1 mask can enter through the attentional gate, share the attentional resource, and is processed along with T1. Based on this idea, Raymond et al. (1992) proposed a *shut and lock* attentional gate model to explain the AB effect. They suggested that potential confusion between T1 and the T1 mask initiates an attentional suppression episode that serves to minimse further confusion from subsequent items in the RSVP stream. This suppression hypothetically locks the attentional gate, thus disabling T2 processing, until the period of confusion ends. Raymond and colleagues likened this suppressive mechanism to the bottleneck model of attention, and proposed that the attentional gate behaves as a filter that effectively rejects, or locks out, potentially confusing stimuli, and improves selection of stimuli relevant to the task.

A recent electrophysiological study by Gross et al. (2004) underscores the role of the T1 mask in producing a T2 deficit. They used MEG to study the interaction, or communication, among neural structures involved in the operation of attention during a standard AB task. In their study, communication among brain regions was identified via *neural synchronisation* (see Singer, 1999 for a review). Different neurons discharge activity at different rates, or frequencies, measured in Hertz (Hz); repetitive neuronal activity at a particular frequency is known as *oscillation*. Neural synchronisation occurs when one group, or *assembly*, of neurons in one part of the brain oscillates rhythmically in phase with another neural assembly in another part of the brain. Gross et al. defined five regions of interest that were subjected to synchronisation analysis: occipital, cingulum, bilateral frontal, bilateral temporal, and bilateral posterior parietal areas. On trials in which an AB was absent (i.e., T2 was successfully detected), they found that the T1 mask elicited a reduced neural network response (*desynchronisation*) followed by a strong synchronisation to T2. In contrast, on trials in which an AB was present (i.e., T2 was not reported), desynchronisation to the T1 mask and synchronisation to T2 was significantly smaller than on AB absent trials. Gross and colleagues concluded that desynchronisation reflects suppression and synchronisation reflects facilitation, and interpreted their results to suggest that successful suppression of the T1 mask is required in order to produce successful detection of T2.

A critical question is: what stage of processing is impaired by limited attention? Information processing can be divided into early perceptual stages that involve analysis of physical stimulus properties, and later post-perceptual stages that involve analysis of meaning and transfer of information into short-term memory (STM). Across the wide range of AB studies there has been great debate about the *locus* of the AB effect: while original proposals cited an early perceptual locus, there is mounting evidence that postperceptual processes play an important role.

Perceptual Accounts of the AB

The original Raymond et al. (1992) study proposed an early selection account of the AB in which the attentional gate filters information according to physical stimulus feature similarities between targets and distractors. In confirmation of this, Raymond, Shapiro, and Arnell (1995) found that a significantly larger AB was produced when targets and distractors were featurally similar to one another than when they were different. In addition, they found that categorically similar targets and distractors did not produce a larger AB effect than categorically dissimilar targets and distractors, suggesting that postperceptual categorisation processes are unaffected by limited attention.

Early perceptual T2 deficit accounts came into question, however, when evidence was found that the meaning of T2 could be processed during the AB period. Luck, Vogel, and Shapiro (1996) recorded ERPs during an AB task and specifically measured the N400 component, a negative waveform that is considered to reflect semantic processing (Bentin & Deouell, 2000; Eimer, 2000a, 2000b; Schweinberger, Pickering, Burton et al., 2002). In their task, T1 was a row of identical digits and T2 was a word presented in red font colour. Distractors were random sequences of letters. Each trial began with the presentation of a "context" word that either matched semantically with T2 (e.g., razor-shave) or mismatched semantically (e.g., wheel-jewel). Participants were required to identify the T1 digit and state whether T2 matched the context word or not. While Luck and colleagues found impaired T2 performance at lag 3, this failure of conscious report was not accompanied by a reduction in the N400 amplitude at lag 3. They concluded that the presence of an N400 during the AB period reflected that T2 was processed to a postperceptual semantic level, a finding that argues against a perceptual locus of the AB.

In a study that combined AB and *semantic priming* paradigms, Shapiro, Driver, Ward, and Sorensen (1997) provided further evidence for the survival of semantic information during the AB period. Semantic priming occurs when an initial (prime) word significantly facilitates processing of a subsequent (probe) word, due to a semantic link between prime and probe. They presented three targets in RSVP – T1, T2 (the prime), and T3 (the probe). On half of the trials T2 and T3 were related semantically (e.g., doctor-nurse) and on the other half of trials they were unrelated semantically (e.g., table-nurse). They found that T2 significantly primed T3, even when T2 was presented during the AB period. Martens, Wolters, and van Raamsdonk (2002) conducted a similar study and supported this finding. Further electrophysiological evidence for the dissociation between the AB effect and N400 activity is provided by a study by Rolke, Heil, Streb, and Hennighausen (2001). Using a similar procedure as Shapiro, Driver et al. they found N400 activation whether an AB effect was present or not, consistent with Luck et al.'s (1996) findings.

Postperceptual Accounts of the AB

Two influential theories of a postperceptual AB have been proposed: the *interference* model and the *two-stage* model.

Shapiro, Raymond, and Arnell's (1994) *interference* model proposes that confusion between T1, T2, and their masks arises once these items have been transferred to and are held in a visual short-term memory (VSTM) store, rather than at a featural processing stage as proposed by Raymond et al. (1992). The interference model assumes that T1, T2, and their masks are processed to varying degrees and compete with each other for attentional resources. T1 always appears before T2, thus it receives preferential weighting of attentional resource over T2. Once perceived, T1 information must be securely transfered, or *consolidated*, into VSTM if it is to be correctly reported at the end of a trial; so must T2. Competition for retrieval of T1 and T2 from VSTM is thought to occur when they have been consolidated in close temporal succession to one another; insufficient time has elapsed to *flush* irrelevant items, such as T1 and T2 masks, from the VSTM store, thus producing an AB. Based on an idea by Duncan and Humphreys (1989), Shapiro et al. proposed that a long lag between T1 and T2 does not produce an AB because sufficient time has been allowed for VSTM to be flushed without further demands being placed on it.

Chun & Potter (1995) introduced the *two-stage model* of attention to account for the AB effect. In their view, in Stage 1, a perceptual and semantic representation of every item presented in RSVP is created. Each representation is transient and unstable and cannot serve as the basis for subsequent report. Unless a representation is selected for further processing (e.g., due to its target status) it will be subject to rapid decay when interference from subsequently presented stimuli disrupts processing. Because every item is processed to this level, the first stage is proposed to have unlimited capacity. If an item has been selected from Stage 1, it can proceed to Stage 2. This second stage involves the sequential consolidation of selected items into a more durable representation that is less susceptible to decay and interference, rendering the target reportable at the end of a trial. Chun and Potter proposed that Stage 2 is limited in capacity and that no item is processed beyond Stage 1 until consolidation of items that reached Stage 2 is complete. This second stage has been likened to the concept of working memory (WM), a complex mechanism that enables very short-term, or *online*, maintenance and manipulation of relevant information. (WM theory is discussed in more detail in Chapter 5.) The two-stage model proposes that the AB effect is elicited by impaired consolidation of T2 due to the occupation of this process by T1. Chun and Potter described this impairment as a *second-stage bottleneck.*

Note the subtle difference between the interference and two-stage models. The interference model implies that T2 undergoes consolidation and gains access to VSTM (WM) and that retrieval from WM is impaired due to interference from T1, the T1 mask, and the T2 mask. The two-stage model implies that T2 cannot enter the consolidation phase until T1 consolidation is complete, and does not make any specific inferences regarding the role of T1 or T2 masks.

It is also clear that the interference and two-stage models contain some overlap. Shapiro, Arnell, and Raymond (1997) attempted to draw some common ground in their *unified* model of the AB. This model states that sufficient attention to T1 is required for it to reach a reportable level of awareness; T2 cannot therefore be consolidated into a VSTM/WM storage buffer until T1 processing is complete, leaving T2 vulnerable to decay from other competing stimuli; and if further demands are placed on the attentional system (e.g., by rapid response requirements) then response-selection factors will additionally decrease T2 accuracy. The key difference between the two-stage model and the unified theory is that the former implies that the T2 deficit is caused by a bottleneck in limited capacity WM consolidation processes, whereas the latter implies that general attentional capacity limits constrain T2 processing at the consolidation stage.

Vogel, Luck, and Shapiro (1998) provided more direct evidence for a postperceptual locus of the AB that is linked to WM. They applied ERP measures to determine the first stage at which processing is suppressed during the AB, and examined four ERP waveforms: P1, N1, N400, and P3. Early perceptual processes have been shown to elicit the P1 and N1 waveforms, considered to reflect sensory analysis and visual discrimination respectively (Hillyard & Picton, 1987; Vogel & Luck, 2000, cited in Luck, Woodman, & Vogel, 2000). As mentioned previously, the N400 waveform is considered to reflect semantic processing. The P3 component is considered to reflect the updating of WM (e.g., Donchin, 1981; Donchin & Coles, 1988). While P1, N1, and N400 amplitudes were unchanged during the AB period, Vogel and colleagues found that the P3 waveform was completely suppressed during the AB. They interpreted this finding to suggest that the AB might be located before or during the consolidation of a stable representation of T2 into WM. They proposed a theory of the AB that incorporates some aspects of previous models and introduces a new concept of *conceptual STM (CSTM)*. According to their account, all items in the RSVP stream are fully identified to a conceptual, semantic level and stored in CSTM. Items in this store are unavailable for report, prone to decay and interference, and become durable and reportable only when consolidated into a visual WM store. Attention is required for this consolidation process and items are transferred from CSTM to visual WM based on their degree of match to stored representations of target stimuli. Similar to the unified model of the AB, Vogel et al. concluded that the

T2 deficit is caused by engagement of attention during consolidation of T1 into visual WM.

In a later ERP study, Vogel and Luck (2002) clarified the AB deficit to reflect delayed consolidation rather than absent consolidation of the T2 item. They found that when T2 was the last item in the RSVP stream (unmasked), the P3 component was not suppressed in amplitude but occurred approximately 104 ms later on lag 3 trials than on lag 7 trials, despite the absence of an AB effect. They suggested that the postponement of T2 consolidation produces an AB because items following T2 are given the opportunity to interfere with T2 processing before its consolidation is complete.

Having outlined theories of the cause of the AB, a particularly interesting characteristic of the effect is also worth noting. Two types of AB function are commonly reported: (1) a monotonic, linear increase from lag 1 (i.e., when T2 immediately follows T1) onwards, and (2) a U-shaped function where performance at lag 1 is high (known as *lag-1 sparing*), dips for up to 500ms, and recovers at later lags (as in the traditional AB effect illustrated in Figure 6b). Lag-1 sparing has been attributed to the attentional gate hypothesis outlined above with reference to the T1 mask (Raymond et al., 1992; Reeves & Sperling, 1986): when T2 immediately follows T1 it is essentially the T1 mask; it enters through the attentional gate with T1 and enjoys a shared attentional resource.

Lag-1 sparing does not always occur, however. Visser, Bischof, and Di Lollo (1999) conducted a comprehensive review of the lag-1 sparing effect and found that equal numbers of studies found lag-1 sparing as those that did not, and that certain conditions were necessary to produce the effect. Lag-1 sparing only appears to occur under conditions where T1 and T2 enjoy similar properties. Studies have shown an absence of lag-1 sparing when T1 and T2 shared different properties, for example when T2 was presented in a different spatial location or when T2 was a different object category and required a different response to T1. The absence of lag-1 sparing suggests that T2 is unable to proceed through the attentional gate with T1 and share attentional resource. Visser et al. concluded that the presence or absence of lag-1 sparing is an indication of how efficiently and effectively the attentional system can reconfigure itself to cope with the changing demands of processing T2 when it is presented directly after T1. When T1 and T2 share similar properties, reconfiguration is efficient and lag-1 sparing is present; when T1 and T2 share different properties, reconfiguration is effortful and lag-1 sparing is absent.

Other Demonstrations of Attentional Limits

Three other notable effects of limited attention on object processing have been described in the attention literature: repetition blindness, change blindness, and inattentional blindness.

Repetition blindness is a term used to describe a situation where observers are unable to report a second presentation of the same object when it is repeated in close succession to the first presentation (Kanwisher, 1987). Experimental measures of repetition blindness typically involve the presentation of stimuli in RSVP in the same spatial location with one item (*R1*) repeated (*R2*). In two of Kanwisher's original experiments, words were presented in RSVP. Observers were required to report the word that had been repeated, or read aloud a sequence of words as they appeared in a sentence. When R2 was repeated in close temporal succession to R1, observers were less able to report its repetition or neglected to read it out loud the second time it appeared in a sentence. Observer reports suggested that they were unaware of the repetition. When the temporal distance between R1 and R2 was increased, repetition was detected more often.

Repetition blindness is similar to the AB effect: both occur only when R1 (or T1) is attended, and both reflect similar temporal lag-dependent effects. But whereas attentional deficits linked to WM consolidation have been proposed to account for the AB effect, repetition blindness has been interpreted as a failure in *token individuation* for the second occurrence of the repeated item (Kanwisher, 1987). Kanwisher distinguished between *types* – basic-level representations of a class of stimuli that can be activated (e.g., a dog), and *tokens* – specific individual instances of a class exemplar (e.g. that border collie), encoded in time and space and retrieved from *episodic memory* (memory for a specific event). She proposed that all items in an RSVP stream are represented in type form. However, only items to which attentional resources are favourably allocated reach a token individuation level of representation, and this level is required in order for successful detection and report. When R1 is repeated, tokenisation for R2 is thought to fail due to an interpretation of R2 as residual activation from R1. When R1 and R2 are emphasised as distinct, and when spatial or temporal separation is emphasised, repetition blindness is reduced (Arnell, Shapiro, & Sorensen, 1999). (See Chun, 1997 for further distinctions between the AB and repetition blindness.)

The term change blindness is typically used to describe an event in which a physical change in the environment is not detected due to lack of attention (Rensink, O'Regan, & Clark, 1997). For example, if we do not pay attention to the traffic lights looming in front of us, we are unlikely to notice them change from green to red. Often, a change that can seem very obvious when attention is allocated to it can be missed when attention is busy elsewhere. Experimental measures of the effect typically involve a *change detection* task. Traditional change blindness experiments involve alternating brief presentation of two complex scenes that differ in one visual property, for example a dog changes to a cat, and back to a dog again, and so on (this is known as a *flicker* paradigm). Only when attention is directed to the changing object can the change be detected. A modification of this change detection design involves brief single presentation of two images, each of which is followed by a mask (e.g., Buttle & Raymond, 2003). Because the experimental conditions between change detection and AB paradigms are similar – presentation of the first and second images is brief; the temporal interval between each image is short; the task requires attention to both targets – the inability to detect a change between two items could also be interpreted as an AB effect.

The final effect of limited attention outlined here is inattentional blindness. Inattentional blindness refers to instances where, in the absence of attention, an observer does not perceive (i.e., is functionally blind to) a highly visible stimulus presented at or within a few degrees of fixation (Mack & Rock, 1998). The effect can be likened to driving a car while using a mobile phone: if the majority of attention is temporarily allocated to the speaker on the other end of the phone, the driver is more likely to miss a van pulling out in front even if his/her eyes were looking in the direction of the van. You might hear the phrase "I don't understand, I just didn't see it" after a collision on the road. Behavioural measures of inattentional blindness typically involve the presentation of an unexpected object in close proximity to a central object with which the observer is attentionally engaged. Observers are asked whether they saw anything on the computer screen other than the central object. An inattentional blindness effect exists when observers are unable to report the unexpected object.

Attentional capacity limits using the AB and other paradigms have been demonstrated for a variety of stimuli, such as letters, digits, words, objects (e.g., Raymond, 2003) and scenes (e.g., Marois, Yi, & Chun, 2004). Are attentional requirements therefore the same, i.e., *fixed*, for the processing of all types of stimuli? There is growing evidence that attentional capacity is not fixed: the level of *familiarity* with a particular stimulus exemplar appears to modulate attentional demand.

The Effect of Familiarity on Attention

It was suggested several decades ago that the more skilled we become at performing a certain task, the less attention we need to allocate to that task. The term automaticity was introduced to describe the ability to perform a task without attention (e.g., Schneider & Shiffrin, 1977). Everyday examples include operating a car, reading, and typing. The literature on automaticity is extensive and a detailed review is beyond the scope of this thesis. Two key findings are briefly presented to illustrate automatic processing theory. Using a dual-task procedure Shaffer (1975) showed that highly skilled typists could type at nearly normal speed and accuracy while reciting nursery rhymes and shadowing (repeating spoken input). The *Stroop effect* (Stroop, 1935) is also often used as a classic example of dual-task interference: when required to read aloud the colour of the ink in which a word was printed, participants' responses were slowed when the word was an incompatible colour word, e.g., response to the word "red" printed in green ink was slower than response to the word "dog" in green ink (correct response is "green"). The Stroop effect is considered to reflect the automaticity of word processing: reading the word and computing the word's meaning cannot be suppressed despite instructions to ignore it, and thus interferes with naming the ink colour.

Critics of automaticity theory argued that although tasks such as these were believed to reflect automatic processing (i.e., proceeded in the absence of attention), they might have in fact used attentional resource in a particular way. Perhaps the ability to type and recite concurrently is a reflection of task switching strategies in which attention is directed to one task to the other and back again in rapid succession (known as *central switching*; Broadbent, 1982; Shaffer, 1975; Welford, 1980). The automaticity interpretation of the Stroop effect has also been criticised. For example, Besner, Stoltz, and Boutilier (1997) found that the Stroop effect was significantly reduced when only one letter in a colour word was coloured than when all letters were coloured. Based on this finding, they concluded that the interpretation of automatic word reading as the cause of Stroop interference is exaggerated. If words are read automatically, both conditions should have produced equivalent Stroop effect magnitudes (but see also Driver & Tipper, 1989; Mari-Beffa, Estevez, & Danziger, 2000). An alternative, more cautious account of the effects of practice and learning was proposed: *less* attention is required for stimulus processing, as opposed to *no* attention.

If familiar stimuli require less attention for processing, what is the effect of familiarity on tasks in which attentional resources are taxed? Surprisingly, this question has received little investigation to date. A handful of studies have employed the AB, change blindness, and inattentional blindness paradigms to address this issue. Attenuation of these effects by the use of highly familiar stimuli has consistently been found.

Shapiro, Caldwell, and Sorensen (1997) compared the ability to detect one's own name with that of other names and common nouns in a typical AB paradigm using a full RSVP stream. Across four experiments, they presented common nouns and names as T1, T2, and distractors in varying combinations, and presented participants' own name as T1 or T2. While they found an AB effect for common nouns and names as T2 in the majority of T1 and distractor conditions, no AB effect was found when a participant's own name was presented as T2 among other names or nouns. Shapiro and colleagues proposed that the elevated salience of one's own name effectively protected it against interference from competing stimuli at the consolidation stage of processing in WM. This finding concurs with Moray's (1959) seminal dichotic listening study in which a participant's own name was the only additional auditory stimulus that could be perceived during an attention-demanding task in which different auditory information was presented to each ear. (See also Wood and Cowan, 1995.)

An effect of familiarity on change blindness has been demonstrated for faces stimuli. Buttle and Raymond (2003) found that the ability to detect a change in identity between two briefly presented masked face images was enhanced when the change involved a famous face (i.e., unfamiliar to famous, famous to unfamiliar, or famous to famous changes). Even when observers were given rich background information about the unfamiliar faces and participated in naming and information retrieval sessions before the experiment began, the change detection advantage was only found for famous faces and was absent for the recently learned but non-famous faces. They concluded that highly familiar faces are processed more efficiently and demand less attentional resource than unfamiliar faces. It could also be argued that famous faces captured attention more readily than unfamiliar faces in their study, thus facilitating detection of a change. Indeed, reductions in change blindness have been reported when objects in a scene or features of an object automatically grab attention, and when attention is directed toward the relevant object before potential change of that object occurs (Lamme, 2003).

Using the inattentional blindness paradigm, Mack and Rock (1998) explored whether the surprise presentation of one's own name would be detected while attention was allocated to a central visual discrimination task (participants determined which axis of a cross was longer). They found that one's own name was detected significantly more often than other names and common nouns, and concluded that the perceptual salience and personal importance of one's own name captured attention.

This effect could also be interpreted to reflect a modest attentional resource requirement for processing one's own name, available as a residual from the central task.

Tong and Nakayama (1999) used a different approach to explore the effect of familiarity on visual processing. They conducted a series of visual search experiments in which participants searched for an unfamiliar target face or their own (highly familiar) target face in displays of between one and six unfamiliar distractor faces. In different conditions, target and distractor faces were presented upright, inverted, in three-quarter view, or in profile view. They found that search for one's own face was consistently more efficient than search for an unfamiliar face when faces were upright, inverted, and in profile, even after hundreds of presentations of the unfamiliar target face. They introduced the term *robust representation* to describe a visual stimulus for which we have an extreme level of familiarity. Properties of robust representations were defined as follows: they may "(1) mediate rapid asymptotic visual processing, (2) require extensive visual experience to develop, (3) contain some abstract or view-invariant information, (4) facilitate a variety of visual and decisional processes across tasks and contexts, and (5) demand less attentional resources" (Tong & Nakayama, pp. 1017).[3] Tong and Nakayama's (1999) study has

---

[3] Tong and Nakayama (1999) also conducted an experiment comparing the effect of unfamiliar versus familiar *distractors* on search for an unfamiliar or familiar target face (Experiment 3 in their study). In contrast to the range of different unfamiliar distractor identities used in their experiments discussed above, distractors in this particular experiment were multiple views of the same face. They found that search for an unfamiliar face among familiar own face distractors was more efficient than search for one's own face among unfamiliar distractors and concluded that familiar face distractors are easier to reject than unfamiliar face distractors. This experiment is not discussed in the main text here because repeated use of different views of the same face might have confounded the results: if familiar faces are more robust to changes in viewpoint than unfamiliar faces, then it is possible that multiple views of one's own face as familiar distractors were easily identified as the same person, and thus reduced the perceived set size, relative to multiple views of a stranger's face as unfamiliar distractors, which might

one major limitation. The presence of one's own face among unfamiliar faces might have produced a *pop-out* effect: due to its status as the only familiar face, one's own face might have captured attention and stood out from the display of unfamiliar faces, making it easier to find. This potential for pop-out was absent when an unfamiliar face was displayed with other unfamiliar faces and might explain why search was less efficient. They did not present a familiar target face among familiar distractors (indeed this would have been impossible given that each familiar face was the participant's own). If such a design were implemented, the roles of familiarity and pop-out in visual search efficiency could be better examined. The pop-out argument can also be applied to Shapiro, Caldwell et al.'s (1997) study. Perhaps one's own name "escaped" the blink because it was the only self-relevant word in the RSVP stream.[4]

Evidence for the effect of familiarity on attention can be summarised as follows: one's own name did not produce an AB effect and was not susceptible to inattentional blindness; a change between two faces was better detected when the change involved a famous face; search for a personally familiar face was more efficient than an unfamiliar face. These results suggest that familiarity can affect attention, facilitate encoding, and perhaps influence WM processes. Evidence for the

---

have been perceived as different individuals. A reduction in the perceived set size would result in more efficient search.

[4] Note that the term pop-out is often used to describe the *pre-attentive* perception of a single item that is distinctly dissimilar from its surrounding distractors. Such pre-attentive processing is traditionally defined by search speed that remains constant when the number of distractors increases (a flat search slope). Tong and Nakayama (1999) found that search rate for a familiar face among unfamiliar distractors increased as set size increased, suggesting that pop-out, in the traditional sense of the term, was not evident. I therefore use the term pop-out to describe how a single item might be more easily distinguished from a display of dissimilar items (compared to a display of similar items) and facilitate search by easing the burden on attentional resource rather than removing the need for attention altogether.

effect of familiarity on attention is incomplete, however, and the studies described above raise five questions to be addressed. (1) How do highly complex stimuli, such as familiar faces, affect the AB? (2) Is the lack of AB for one's own name and efficiency of visual search for one's own face due to visual familiarity, or do these results reflect the salience of self-relevant stimuli? (3) Do familiar stimuli more readily *attract* attention or *require less* attention for processing? (4) Are attentional resource requirements for names and faces special in some way due to their connection with person identification? (5) If visual WM is implicated in the AB, what is the direct effect of familiarity on visual WM processes?

Chapters 3 and 4 report six experiments designed to address the first four questions. In Experiments 1-3 (Chapter 3), detection of an unfamiliar versus familiar T2 face was measured in a traditional AB paradigm. This is the first time that familiarity effects in the AB have been explored using faces, and addressed question 1. In these and all subsequent face experiments, familiar faces were famous faces (i.e., were not self-relevant), addressing question 2. Question 3 was addressed by examining the role of pop-out in AB Experiments 1-3. This was achieved by manipulating the familiarity relationship between the T2 face and distractor faces: in Experiment 1, an unfamiliar T2 face was presented among unfamiliar distractor faces (no pop-out); in Experiment 2, a famous T2 face was presented among unfamiliar distractor faces (potential pop-out); in Experiment 3, an unfamiliar or famous T2 face was presented among famous distactors (potential pop-out in the former condition; no pop-out in the latter condition).

Experiments 4-6 (Chapter 4) employed a visual search task based on Tong and Nakayama's (1999) study. In order to control for pop-out (and address question 3), target and distractor stimuli were either all unfamiliar or all familiar in any search display. In Experiments 4 and 5, unfamiliar and famous faces were compared. In Experiments 6a and 6b, complex non-face stimuli (Hanzi) were used and addressed whether any difference in attentional demand for familiar versus unfamiliar faces could similarly apply to that for familiar versus unfamiliar non-face stimuli (addressing question 4).

Question 5 is addressed in Part 3 (Chapter 6) of this thesis: performance on a change detection task designed to measure visual WM capacity was compared for unfamiliar versus familiar faces (Experiments 7 and 8) and Hanzi (Experiments 9a and 9b).

One final, overarching question that is applied to all experiments reported in this thesis is this: if familiarity modulates attention and visual WM processes, how is this achieved? To address this question, the nature of visual representations and visual recognition processes for familiar versus unfamiliar stimuli are discussed throughout. Experiments 5, 6b, 8, and 9b used inverted stimuli to further address this issue and investigate the relative contribution of featural and configural processes to familiar and unfamiliar item recognition.

CHAPTER THREE

THE EFFECT OF FAMILIARITY ON THE ATTENTIONAL BLINK FOR FACES: EXPERIMENTS 1-3

Previous studies using one's own name or face and famous faces suggested that familiar stimuli either more readily capture attention or require less attention for processing than unfamiliar stimuli. The aim of this chapter was to examine the effect of familiarity on attentional demand more closely. To do this, I presented unfamiliar and familiar (famous) faces in a conventional AB paradigm.

Before I could address whether familiarity influences attentional demand for faces it was important to establish whether an AB effect could be obtained with faces at all: does face identification *in general* require attention? Experiment 1 used unfamiliar faces and served this purpose. Colleagues have informally reported difficulty in eliciting an AB for faces and the only published paper that has addressed this issue reports no AB effect (Awh et al., 2004).

As discussed in Chapter 1, the extreme social relevance of faces led many researchers to consider that a unique perceptual mechanism has specifically evolved for face processing that is distinct from object processing (evidenced by superior activation of the FFA, a neurological dissociation between object and face recognition, and particularly dramatic inversion effects for faces compared to non-face objects).

Similarly, some have argued that a unique attentional mechanism has specifically evolved for face processing. Before presenting Experiments 1-3, literature addressing this issue is reviewed. Unless stated otherwise, all studies discussed below assumed that optimum (upright) face processing is specifically reliant on configural information whereas processing of non-face objects and inverted faces is reliant on featural information.[5]

Does Face Processing Involve a Specifically Evolved Attentional Mechanism?

Three different views on whether face processes involve a specifically evolved attentional mechanism have been expressed which I term the *no-attention*, *special-attention*, and *default* hypotheses.

The No-Attention Hypothesis

The no-attention hypothesis states that face processing is automatic, obligatory, and requires no attention (e.g., Farah, 1996; Farah, Wilson et al., 1995). This view is plausible because humans are unarguably expert face analysers and it has been proposed that little or no attention is needed when processing stimuli for which expertise has been developed (e.g., Schneider & Shiffrin, 1977). Support for this hypothesis comes from reports that irrelevant face distractors presented either side of a central attention-demanding name-categorization task exerted significant interference on central task performance, even though attention was directed away from the distractors (Lavie, Ro, & Russell, 2003; Young, Ellis, Flude, McWeeny, & Hay, 1986). These studies demonstrated that when attention was (presumed)

[5] An article that includes all three experiments reported in this chapter has been accepted for publication [Jackson, M. C. & Raymond, J. E. (in press). The role of attention and familiarity in face identification. *Perception & Psychophysics*]. Experiments 1 and 2 were presented as a poster at the 2003 Vision Sciences (VSS) conference, Florida [Jackson, M. C. & Raymond, J. E. (2003). Familiarity effects on face recognition in the attentional blink. *Journal of Vision*, *3*(9), 817a].

unavailable for face recognition, identity information could still be processed to a level that caused interference with an on-going task.

The Special-Attention Hypothesis

The special-attention hypothesis states that optimal face processing requires access to a face-specific attentional resource that is dedicated to configural processing and is separate from a featural attentional resource (Awh et al., 2004; Palermo & Rhodes, 2002).

Palermo and Rhodes (2002) reported a series of dual task experiments in which participants were required to match the identity of two peripherally presented faces (primary task) while concurrently encoding face features (e.g., eyes) in a centrally presented face (secondary task). Recognition for the encoded features was subsequently tested by presenting them in the whole face they were learned in, or as isolated parts. Their idea, based on a finding by Tanaka and Farah (1993), was that if features were encoded as an integral part of the broader face configuration (as opposed to independently from other features), then feature recognition would be better with whole faces than with isolated face parts (the whole/part advantage). To investigate the notion of a special configural attentional channel, peripheral faces were presented in either upright or inverted orientations. Palermo and Rhodes found that performance on the central task was worse when upright peripheral faces were presented. They proposed that upright peripheral faces consumed configural attentional resource, impaired configural processing of the central face, and thus eliminated the whole/part advantage. Inverted peripheral faces, assumed to engage a featural attentional resource, were thought to have produced little inteference with configural processing of the central face and thus revealed the whole/part advantage.

Awh et al. (2004) supported the notion of a special configural attention resource for faces. They reported greater impairment in T2 performance when T1 and T2 involved configural processing (e.g., when T1 and T2 were both faces) than when attention was divided between configural (face) and featural (non-face) targets. Their study is reviewed in more detail in the discussion section at the end of Experiment 1.

Directly contradicting the view that optimal face processing requires access to a configural attention resource, but consistent with the notion of separate featural and configural attention channels, Boutet, Gentes-Hawn, and Chaudhuri (2002) suggested that configural processes do not require attention whereas featural processes do. Using the composite effect paradigm, participants viewed whole faces and houses presented in transparency at the same spatial location. In one condition they were asked to attend to faces (and ignore houses) and in another to attend to houses (and ignore faces). They then made recognition judgements of the top half of face composites that were aligned or misaligned horizontally with the lower half. As mentioned previously, the composite effect reflects impaired recognition of the top half of a face when face halves are aligned. This impairment is considered to reflect automatic configural processing that introduces interference from the lower face half (Carey & Diamond, 1994; Hole, 1994; Young et al., 1987). Boutet et al. found that the size of the composite effect was equivalent with and without full attention to faces during the initial session and concluded that configural face processing does not require attention. To the extent that configural processing is seen as a special property of face processing (Farah, Wilson et al., 1995), Boutet and colleagues' view can also be seen as a version of the no-attention hypothesis.

The Default Hypothesis: Faces are Like Objects

Both the no-attention and the special-attention views can be contrasted with a third possibility, the default hypothesis. The default hypothesis states that attention is needed to process faces in the same way as is needed for any other complex stimuli (Downing, Liu, & Kanwisher, 2001; Wojciulik, Kanwisher, & Driver, 1998). It is supported by observations that activation in face-specific brain areas is modulated by the degree of attentional allocation to faces (Downing et al., 2001; Pessoa, McKenna, Gutierrez, & Ungerleider, 2002; Vuilleumier, Armony, Driver, & Dolan, 2001; Wojciulik et al., 1998). In addition, Tong and Nakayama's (1999) demonstration that the rate of visual search for faces slowed as set size increased (i.e., search became more effortful) indicates that attention is needed for face processing. Other visual search studies with face stimuli have found similar results (Brown, Huey, & Findlay, 1997; Kuehn & Jolicoeur, 1994; Nothdurft, 1993).

The default view has parsimony, a feature lacking in the suggestion of separate attentional mechanisms for configural and featural processes. Indeed, the special-attention view is broadly inconsistent with numerous cross-modal attention studies indicating a single pool of attentional resource for all sensory modalities tested, whether visual, auditory, or tactile (e.g., Arnell & Jolicoeur, 1999; Jolicoeur, 1999; Pashler, 1998; Soto-Faraco et al., 2002).

To summarise, three views of the role of attention in face processing are (1) attention is not needed, (2) a special configural attention mechanism is needed, or (3) attention is needed as for any other stimulus. Empirical literature indicates that attention can enhance performance on face perception tasks, yet attention is not always necessary for the extraction of information from a face. When attentional resource has been allocated to a face, other face stimuli appear to compete more heavily for this resource than non-face stimuli. Put this way, the role of attention in face processing does not seem particularly special and these summarising statements apply well to other stimuli. For example, object perception is degraded without attention (Mack & Rock, 1998); information about an object can be extracted without directing attention to it (Tipper & Cranston, 1985); dual-task costs are greater when target stimuli for each task are of the same rather than a different stimulus class (Kanwisher & Potter, 1990).

Experimental Outline

In Experiment 1, all faces in the RSVP stream (T2 and distractors) were unfamiliar. In Experiment 2, distractors remained unfamiliar but T2 was a famous face. Experiment 3 investigated the possibility of pop-out effects that could be caused by a difference in familiarity between T2 and its distractors: distractors were famous faces, and T2 faces were either unfamiliar (potential pop-out, as in Experiment 2) or famous (no pop-out, as in Experiment 1). In all experiments the T1 task involved a shape judgement regarding texture elements in a pattern (described in more detail in

General Methods). This task was considered to involve featural processing and allowed specific examination of the special-attention hypothesis. The T2 task involved the detection of a pre-specified target face. Because successful detection of T2 required the ability to identify that individual as the target, the T2 task was considered a face identification task.[6]

The presence of an AB effect is interpreted to reflect that significant attentional resource is required to detect T2; the absence of an AB effect is interpreted to reflect that little attentional resource is required to detect T2. Different results can be expected from the different hypotheses outlined above. The no-attention hypothesis would predict that no AB should be found for either unfamiliar or familiar faces because face processing in general is automatic. The special-attention hypothesis would predict that no AB should be found for either unfamiliar or familiar faces because the featural T1 task does not consume configural attentional resource required for face identification. The default hypothesis would predict that a significant AB should be found for unfamiliar faces and that no AB should be found for familiar faces, just like other non-face stimuli such as names (as in Shapiro, Caldwell et al.'s 1997 AB study where common names and nouns produced an AB but one's own name did not).

Each experiment shared a similar methodology so a General Methods section is provided below.

---

[6] Note that *implicit* identification based on *visual* information (e.g., "was a particular face present or absent?"), rather than *explicit* identification based on *verbal* name retrieval (e.g., "whose face was presented?"), was involved in these and all subsequent experiments in this thesis.

General Methods

Participants

Participants recruited from the University of Wales Bangor Student and Community Subject Panels participated in exchange for course credits or money. All were white European adults who reported normal or corrected to normal vision and were naïve to the purpose of their experiment. Informed consent was obtained prior to participation. No participant completed more than one experiment.

Apparatus

Stimuli were displayed on a 22-inch Mitsubishi DiamondPro 2060u monitor (32-bit true colour; resolution 1280 x 1024 pixels) and generated by E-Prime software (Version 1.0; Schneider, Eschman, & Zuccolotto, 2002) using a Dell computer. Responses were recorded via the computer keyboard. A chin rest stabilised participants' head position and ensured the display was always viewed binocularly from a distance of 70cm. Testing was conducted in a small room with low ambient illumination.

Stimuli

Each trial consisted of 15 items presented successively in RSVP at the same central screen location. One of these, T1, was a computer-generated greyscale, abstract, elliptical pattern composed of either 20 small circles or 25 small squares, each element having a grey value randomly selected (with replacement) from ten levels (Figure 7a). Ten exemplars of each pattern type (circles/squares) were used in each experimental block of trials. A T1 item was presented on every trial and its serial position in the RSVP stream was randomly selected as 3, 4, or 5. Half of the trials featured a circles pattern and the other half featured a squares pattern, randomised.

All other items were faces. Each was a greyscale image of a Caucasian adult with hair present. As far as possible, neutral faces seen in frontal view were selected. None wore glasses or sported facial hair. Faces were either unfamiliar or familiar (famous). Unfamiliar faces were selected from the Psychological Image Collection at Stirling (PICS). Famous faces were selected from Google Image web search results by using famous names as search terms. In order to mirror natural diversity, faces were not matched on dimensions such as attractiveness or distinctiveness. Luminance and contrast values of face images were not manipulated because only reasonably high quality images were used and no obvious, systematic differences on these dimensions between famous and unfamiliar faces were apparent. All images (face and non-face) were displayed against a grey rectangular uniform background that subtended a visual angle of approximately 2.9° x 3.4°. Each face within the background subtended an average of 2.8° x 3.3°. Minor variations in face/head size reflected natural individual differences. All unfamiliar and famous face images are provided in Appendix A.

Design

In each experimental block, T2 was the same adult male and was presented on half of all trials. When presented, it always appeared after T1 at a lag of between 1 (i.e., the first image after T1) and 8 (i.e., the eighth image after T1). A face always followed T1. A minimum of two items followed T2 to complete the RSVP series. On T2 absent trials, a T2-filler item was presented in its place. This was one of eight randomly selected different adult male faces matched in apparent age to the T2 face. The item following T2 (T2 mask) was randomly selected from eight other adult males, also matched in apparent age to the T2 face. All other faces (distractors) varied in apparent age; half were female and half were male. Each factorial combination of T1 serial position, T2 lag, and T2 presence/absence was presented on an equal number of trials, in a pseudo random order. Each block was composed of 192 trials. Collapsing across T1 serial position, T2 was presented 12 times at each lag.

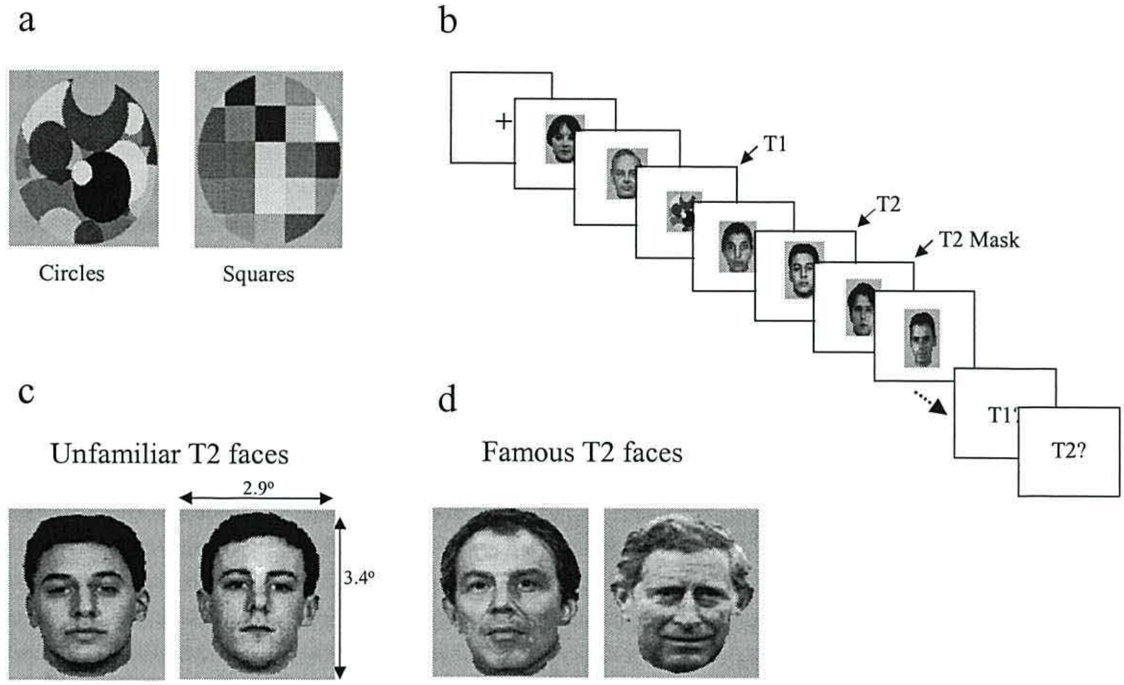


Figure 7. (a) Example of T1 items: "Circles" and "Squares". (b) Example RSVP sequence illustration with T1 at serial position 3 and T2 at lag 2. Each item was presented for 85 ms with no inter-stimulus interval. (c) Images of the unfamiliar T2 faces used in Experiments 1 and 3. (d) Images of the famous T2 faces (Tony Blair and Prince Charles) used in Experiments 2 and 3.

Procedure

Each RSVP trial (illustrated in Figure 7b) was initiated by pressing the space bar. A central fixation cross appeared for 1000 ms immediately followed by a 15-item RSVP series. Each item, presented at the same location as the fixation cross, was presented for 85 ms with no inter-stimulus interval (ISI). Participants were required to identify the abstract image (T1) as a circles or squares pattern (using key C labelled "circles" and key N labelled "squares"), and report whether the pre-specified target face (T2) was present or not (using key K labelled "yes" and key S labelled "no"). Both responses were un-speeded and no feedback was provided.

Before beginning each experimental block, participants completed four short practice blocks designed to deliver a consistent amount of pre-test exposure to the T2 face used in the succeeding experimental block, and to ensure familiarity with the two tasks. Practice blocks were like experimental blocks in all respects, with the following exceptions. First, performance feedback on both T1 and T2 tasks was provided at the end of every trial. Second, the presentation duration of each RSVP item was reduced in each successive practice block, beginning with 400 ms, reducing to 300 ms, then 200 ms, and finishing with 85 ms (the value used in the experimental blocks). Third, each practice block had only 12 trials (50% T2 present). Within each practice block T2 was presented at lags 2, 5, and 7, and appeared twice at each lag. This yielded a total of 24 exposures to T2 in the practice session. Before the practice session began, the T2 face was presented in the centre of the screen with instructions to examine the face carefully before proceeding. The T2 face was presented once more before the experimental block as a reminder. Participants tended to examine the face for approximately 5-10 seconds before initiating the first trial of a block.

Data Analysis

Data regarding performance on the T2 task was only analyzed if the T1 response had been correct on that trial. T2 false alarm (FA) rate, i.e., the percentage of "present" responses when T2 was absent, was calculated for each participant and if this value exceeded 20% that participant's data was excluded. I also excluded data from participants whose mean T2 hit rate (percentage of "present" responses when T2 was present) exceeded 98%. The number of people excluded on these bases varied across experiments and is detailed in each relevant section.

For each remaining participant, I calculated a post-AB baseline level of T2 detection by averaging the hit rates obtained at lags 6, 7, and 8. This measure was based on prior studies that reported the AB effect was typically over when the interval between T1 and T2 exceeded approximately 500 ms. A repeated-measures analysis of variance (ANOVA) using T2 hit rates obtained at lags 6, 7, and 8 (conducted separately for each experiment) showed a non-significant effect of lag in all cases, and justified the use of this method for establishing baseline.

In each experiment, I conducted repeated-measures ANOVA on T2 hit rates using lag (1, 2, 3, 4, 5, and baseline) as a within-factor to test for overall effects of lag on T2 performance. Planned post-hoc comparisons used paired-samples *t*-tests (with Bonferroni corrections where applicable) and compared T2 performance at short versus long lags.[7] Where a significant AB effect was evident, I used a within-subjects contrast to determine the presence or absence of lag-1 sparing: a significant linear effect indicated lag-1 sparing was absent; a significant cubic effect indicated lag-1 sparing was present. False alarm rates did not differ as a function of T1-T2 lag in any condition across all experiments. This indicates that hit rates, rather than *dprime* values (*d'*), are an adequate measure of performance in this context.[8] I also repeated all analyses with *d'* values, however, and found similar results in all cases. Alpha levels were set at .05 in these and all subsequent experiments in this thesis.

[7] In these and all subsequent experiments in this thesis, deviations of sphericity within the data (Mauchly's test, repeated-measures ANOVA) were corrected for by using the Huynh-Feldt statistic.
[8] Note, however, that analysis of false alarms as a function of lag is a crude estimate of guess rates because on target absent trials there was no way of knowing at which lag a participant mistakenly thought T2 was present.

Experiment 1: Unfamiliar Faces and the AB

Two hypotheses of the attentional demand on face processing are directly assessed in this experiment: the no-attention view and the special-attention view. To re-cap, if unfamiliar face identification requires attention, a significant dual-task cost in detecting a specific T2 face presented at short, but not long, T1-T2 lags (an AB effect) is expected. Conversely, if unfamiliar face identification places few demands on attentional resources expended by the T1 item, the difference between short and long lags is expected to be non-significant (no AB effect).

To test the notion of a special attentional resource dedicated to configural processing and needed for identifying upright faces (Awh et al., 2004; Palermo & Rhodes, 2002), I used a T1 task that was considered to require featural processing. Based on the assumption that face identification predominantly involves configural processes, the special-attention view would predict no AB for a T2 face when a featural T1 task is used. If attentional demands on face processing are not special in this way, and if face processing does require attention, an AB effect is predicted.

Two different unfamiliar T2 faces (Figure 7c) were each tested in a single session with two different participant groups. T2 masks, T2 fillers, and distractor faces were also unfamiliar. The methods were as outlined in the General Methods with the following exceptions.

Methods

Participants

Twenty-six British participants (14 females, 12 males; mean age 22 years) were randomly assigned to one of two face groups. Data from four participants were excluded due to excessively high FA rates, leaving 13 people in one group and 9 in the other.

Results

First, the effect of target face was examined by conducting a mixed design repeated-measures ANOVA on T2 hit rates with face (face 1 and face 2) as a between factor and lag (1, 2, 3, 4, 5, and baseline) as a within factor. This revealed a non-significant main effect of face and a non-significant face by lag interaction. On this basis, and because I was not concerned with attentional requirements for specific individual faces, data from both faces were combined for further analyses.

T1 Performance

The mean percent correct T1 score was 96.3% (*SE* = 0.6%). In this and all subsequent experiments reported in this chapter all participants performed at 85% or better on the T1 task and an ANOVA on T1 performance with lag as a within factor showed a non-significant main effect of lag.

T2 Performance

The mean T2 FA rate was 8.8% (SE = 1.4%). T2 percent correct scores on target present trials (hit rate) as function of lag are shown in Figure 8. An ANOVA on these data using lag (1, 2, 3, 4, 5, and baseline) as a within factor revealed a significant main effect of lag, $F(5, 105) = 2.44$, $p < .05$, indicating the presence of an AB effect. Without considering lag 1 performance, mean T2 detection reached a minimum of 66.1% at lag 2, a value significantly below baseline (76.3%), $t(21) = 3.40$, $p < .05$. This supports the claim of an AB effect. Performance at lag 1 (M = 65.0%; SE = 6.2) was comparable to that at lag 2, indicating the absence of lag-1 sparing. This is supported by a within-subjects contrast analysis that showed the lag effect to be significantly linear in nature, $F(1, 21) = 6.36$, $p < .05$.

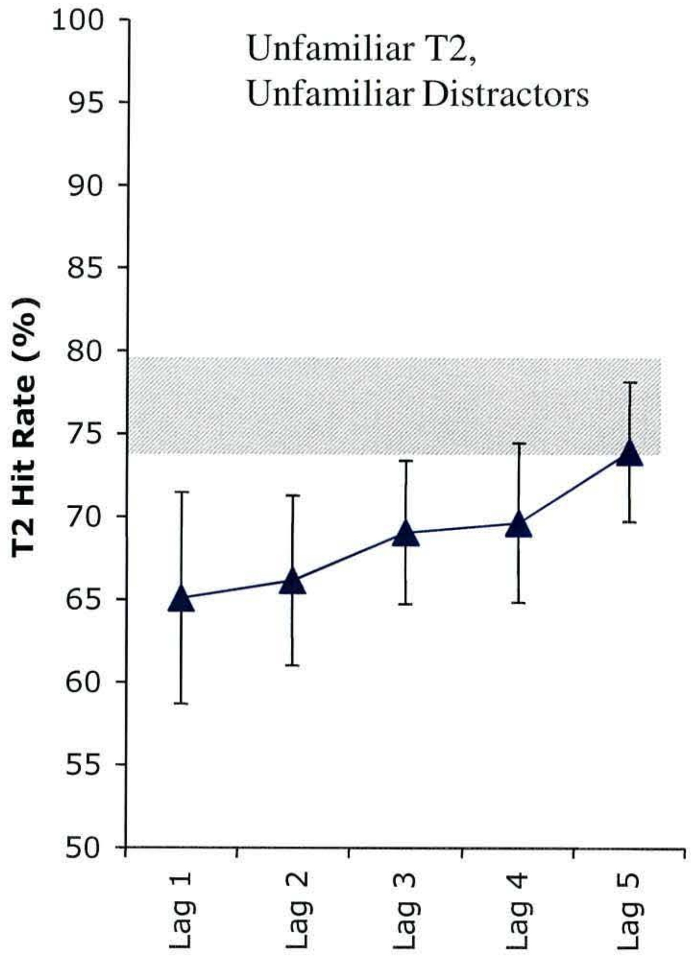


Figure 8. Mean percent hit rate as a function of lag for detection of unfamiliar T2 faces among unfamiliar distractor faces, Experiment 1. Bars represent ± 1 standard error and the shaded area represents the mean hit rate at baseline (the mean of lags 6-8) ± 1 standard error.

Discussion

Experiment 1 produced two main findings. First, identification of unfamiliar faces does require attention: a significant AB effect was found. Second, this AB effect for unfamiliar faces was produced when the T1 task was featural (shape discrimination) and the T2 task was, or at least was assumed to be, configural (face detection). These results contrast with the suggestion that face processing is automatic

(Lavie et al., 2003) and are incompatible with the reported lack of AB for a T2 face using a featural T1 task (Awh et al., 2004). I deal with each issue separately.

The presence of an AB for unfamiliar faces indicates that attention is needed for face identification and does not support Lavie and colleagues' (2003) findings. In their study, participants were required to search for a famous target name among a list of centrally presented words and state whether the target name belonged to the occupation category of pop star or politician. Famous face distractors, either congruent or incongruent with the target category, were presented on either side of the central task as irrelevant flankers. They manipulated the perceptual load of the central task and expected a high load to leave no attention for processing distractor faces. Despite removal of attentional resource from the flanker faces, they reported interference effects from these flankers on the name categorisation task. Lavie et al. interpreted this finding to reflect automatic, or obligatory, processing of faces. They fell short of proposing that faces are processed with no attention, however, and stated instead that little attention is required for automatic face processing.

One important feature of Lavie et al.'s (2003) study is the use of familiar (famous) faces. As mentioned previously, high stimulus familiarity was found to reduce the need for attention (Buttle & Raymond, 2003; Shapiro, Caldwell et al., 1997; Tong & Nakayama, 1999); this might account for their findings. Experiment 2 of the current study specifically explored the effect of familiarity on attention with the expectation that if famous faces were used as T2 items an AB effect would be absent.

The results of Experiment 1 also run contrary to Awh et al.'s (2004) reported lack of AB for T2 faces with a featural T1 task. The T1 task used in the current experiment is featural, yet a robust AB effect was produced. The discrepancy may lie in the operational definition of the AB effect. Awh et al. defined an AB as a significant difference between T2 performance rates obtained in dual-task trials with that obtained when the T1 task was to be ignored (single-task trials). This makes sense if there is no effect of lag for the single-task condition (indicating that T1 was effectively ignored). Yet, in their critical T1-digit/T2-face experiment for which they report no AB effect, there was a significant effect of lag in both single- and dual-task conditions. This sheds doubt on their claim that an AB effect was absent. If an AB effect is defined as a significant difference in performance for long versus short lags, then an AB effect for faces is evident in their dual-task results. This observation weakens the basis for their proposal of a special attention mechanism for configural face processing.

These criticisms aside, it is possible that the current texture T1 task did employ configural processes and that this accounted for the presence of an AB effect for unfamiliar faces. This is unlikely for two reasons. First, lag-1 sparing was not found. Recall that the term lag-1 sparing refers to the finding of no obvious perceptual deficit for stimuli presented at lag 1 in an AB procedure that produces large deficits for the same stimuli presented at lag 2. Lag-1 sparing is generally found to occur when T1 and T2 tasks do not require a stimulus category switch (Visser et al., 1999). If T1 stimuli used here were somehow face-like, or used the same processes as a face identification task (i.e., configural processing), then lag-1 sparing should have been evident. This was not the case. Second, in this and all subsequent experiments, I found no systematic effects of T1 type (circles versus squares) on T2 performance, precluding the argument that the T1 circles image might have been more face-like than the T1 squares image and contributed especially to the AB finding. The AB effect for faces found here does not appear to depend on prior engagement of face processing mechanisms and these results support the more parsimonious default view that the attentional resource needed to process faces is not particularly different from that required to process any other stimulus. Existing, non-stimulus-specific accounts of the AB effect (e.g., Shapiro, Arnell et al., 1997) appear adequate.

Finding an AB effect for unfamiliar faces was important in two regards: it provides the first evidence that an AB effect for faces can be achieved, and sets a benchmark to which familiar faces can be compared.

Experiment 2: Famous Faces and the AB

Experiment 2 aimed to explore the effect of familiarity on face identification within the AB. If identification of familiar faces requires little or no attention, no AB effect is expected.

To investigate this, I presented two famous British males as T2 targets – Tony Blair (UK Prime Minister at the time of study) and Charles Windsor, the Prince of Wales (Figure 7d). Two groups of participants were tested: a British (GB) group and an Other European (OE) group (Europeans excluding UK and Republic of Ireland citizens). British participants were expected to be highly familiar with the T2 faces and show no AB effect, whereas Group OE were expected to be less familiar and produce an AB effect. Note that Tong and Nakayama (1999) claimed that the benefits of familiarity for attention can only come about with extreme familiarity and cannot be produced even when participants receive thousands of exposures to a face. Similarly, in Shapiro, Caldwell et al.'s (1997) study, common, familiar names were used as T2 stimuli, but only when the name was the participant's own did it fail to produce an AB. In the current experiment, the OE group were expected to recognize Tony Blair and Prince Charles but they were not expected to have the same level of extreme familiarity that British citizens have. The between-group experimental design allowed control over the contribution of specific stimulus artefacts (such as luminance or contrast) driving T2 detection. The methods were as outlined in the General Methods with the following exceptions.

Methods

Participants

Participants recruited for the GB group were UK citizens who had spent at least the past 5 years living in the UK. Participants recruited for the OE group were all born and raised in continental Europe and had been living in the UK for less than one year. They were citizens from Italy, Spain, Germany, Greece, France, and Norway. One OE participant was excluded because of a high FA rate. Nine GB participants and five OE participants had mean T2 hit rates greater than 98%, rendering their data un-interpretable due to ceiling performance. Excluding these, data was obtained from 16 GB participants (9 females, 7 males; mean age 22 yrs) and 12 OE participants (7 females, 5 males; mean age 22 yrs).

Stimuli

The T2 items, an image of Tony Blair and an image of Prince Charles, were greyscale photographs similar to the other faces stimuli (as described in the General Methods). As in Experiment 1, all T2 masks and T2 fillers depicted men matched in apparent age to each T2 face. These and all distractors were unfamiliar faces.

Procedure

Each participant was tested in two blocks, one block for each T2 face (counterbalanced). Before the experiment began, each participant rated their familiarity with the written names of each T2 stimulus, along with 24 other famous and non-famous names that were used to fill out this task (the scale ranged from 0-5; 0 = no name recognition, 5 = high familiarity). Participants were also asked to name each T2 face on completion of both experimental blocks.

Results

Familiarity Ratings

Familiarity ratings for the T2 names provided by Group GB ($M = 4.8$, $SE = 0.2$) were only marginally higher, and not significantly so, than those provided by Group OE ($M = 4.5$, $SE = 0.2$). On completion of the experiment, all GB participants correctly named each T2 face whereas three OE participants made naming errors. These OE participants were still included in the analyses.

T1 Performance

Mean percent correct T1 scores were 97.2% (*SE* = 0.5) and 97.3% (*SE* = 0.8) for groups GB and OE respectively, a non-significant group difference.

T2 Performance

Data from both T2 faces were combined.[9] The mean FA rates for groups GB and OE were 6.8% (*SE* = 0.9) and 5.0% (*SE* = 1.0) respectively, a non-significant group difference. The mean T2 hit rate for each group is plotted as a function of lag in Figure 9. There are two points to note. First, no AB effect is found for Group GB (Figure 9a), a finding that contrasts with the AB effect observed for a similar group of participants using an unfamiliar T2 face in Experiment 1. Second, an AB effect is clearly evident in the data from Group OE (Figure 9b), even though they saw the same faces as Group GB. A mixed-design ANOVA with group (GB, OE) as a between factor and lag (1, 2, 3, 4, 5, and baseline) as a within factor confirmed that the interaction of lag by group was significant, $F(5, 130) = 2.76, p < .05$. While the main effect of lag was non-significant for Group GB ($F < 1$), the lag effect was highly significant for Group OE, $F(5, 55) = 5.09, p < .01$, and significantly cubic in nature indicating lag-1 sparing, $F(1, 11) = 17.90, p < .01$. For group GB, the performance minimum seen at lag 3 was high (91.3%) and did not differ from baseline performance (91.4%). For Group OE, a performance minimum of 78.7% was

[9] In contrast to similar performance for the two unfamiliar T2 faces used in Experiment 1, significant performance differences between Tony Blair and Prince Charles were revealed in Experiment 2 for both GB and OE groups. However, Group GB showed no AB effect for either Tony Blair or Prince Charles, and Group OE showed a significant AB effect for Tony Blair and a marginal AB effect for Prince Charles. Differences in performance between the two target faces therefore appeared to reflect overall task difficulty rather than attentional requirements.

observed at lag 2 and was marginally significantly different from baseline (86.5%), $t(11) = 1.90, p = .08$.

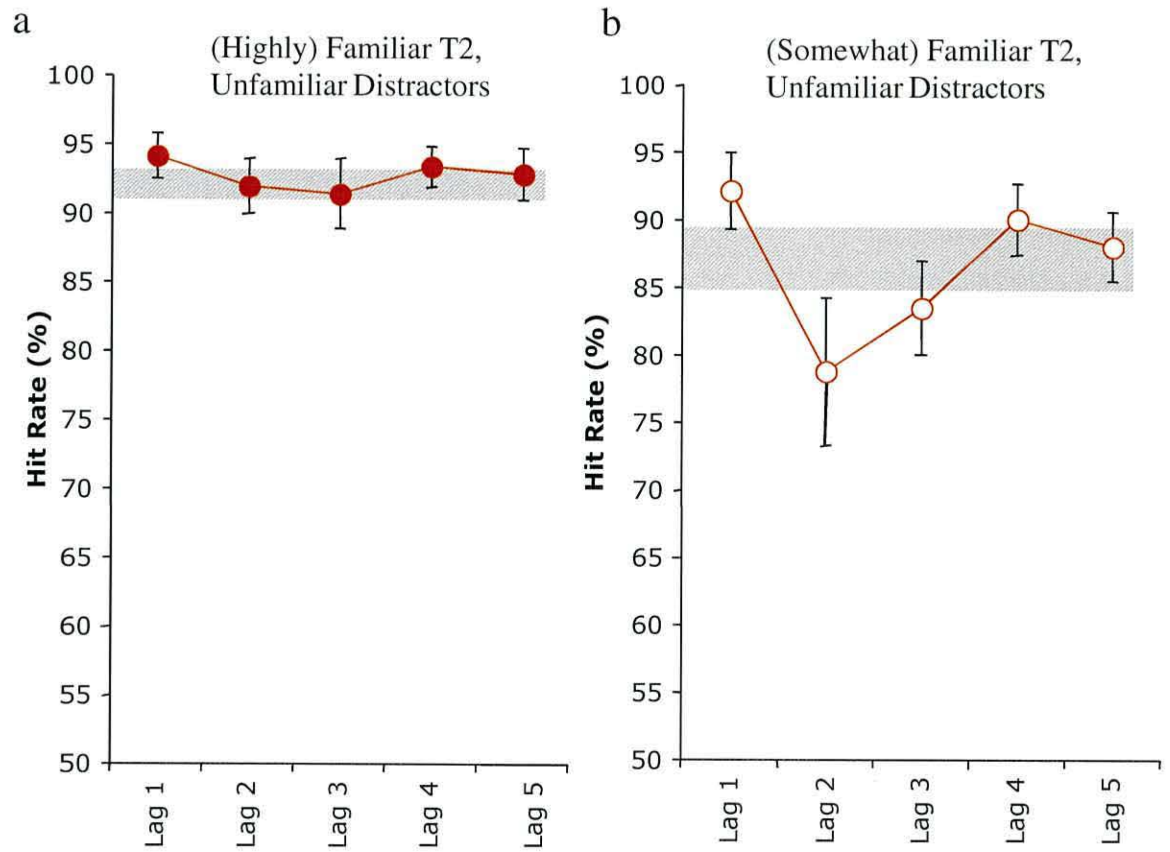


Figure 9. Mean percent hit rate as a function of lag in Experiment 2 for: (a) detection of highly familiar T2 faces among unfamiliar distractor faces, Group GB (British); (b) detection of somewhat familiar T2 faces among unfamiliar distractor faces, Group OE (Other Europeans). Bars represent ± 1 standard error and the shaded area represents the mean hit rate at baseline (the mean of lags 6-8) ± 1 standard error.

Discussion

In contrast to the significant AB effect found for unfamiliar T2 faces in Experiment 1, I found no AB effect for highly familiar T2 faces (Group GB).

Participants for whom the T2 faces were *somewhat* familiar (Group OE) did show a significant AB effect. This concurs with previous suggestions that the burden on attentional resources is only substantially lifted for highly familiar stimuli (Tong & Nakayama, 1999; Buttle & Raymond, 2003).

An important feature of the experimental design used here was that I used the same T2 faces for both groups of participants; only the presumed familiarity of the participants with those images differed. Thus I can be confident that group differences in T2 detection were due to familiarity and not stimulus artefacts. The pattern of AB effects reported here (AB absent for the GB group; AB present for the OE group) suggests that, despite similar subjective ratings of familiarity for both groups, visual experience with specific faces determines the amount of attentional resource needed for rapid identification.

There is one notable difference between the AB effect for unfamiliar T2 faces (Experiment 1) and that for somewhat familiar T2 faces (Group OE, Experiment 2). Whereas lag-1 sparing was absent in Experiment 1, it was present in Group OE. Experiments 1 and 2 comprised identical T1-T2 category switches (i.e., the attentional system had to reconfigure from abstract image to face image), so the presence of lag-1 sparing in Group OE is surprising. It is possible that an additional moderator of lag-1 sparing exists, connected to the familiarity of the T2 item immediately following T1. Some degree of familiarity with T2 in Group OE might have allowed it to be rapidly or distinctively encoded so that when attention was allocated to T1 the representation of T2 shared this allocation, was not confused with that of T1, and was easily detected. When T2 (at lag 1) was an unfamiliar face, encoding might have been more time-consuming, or resulted in a less robust, less distinct representation, which reduced the likelihood of conscious awareness for it.

Experiment 3: The Effect of Pop-Out on T2 Performance

Although the results of Experiment 2 suggest that high familiarity provides protection from the consequences of the AB, an alternative interpretation is that the stimulus conditions (for Group GB) allowed T2 to pop out of the RSVP stream whereas those in Experiment 1 did not. Perhaps a familiar T2 face among unfamiliar distractor faces was easy to detect because of its unique status as the only familiar face presented, rather than because of familiarity *per se*. Indeed, Barnard, Scott, Taylor, May, and Knightley (2004) found that distractors that were distinct from targets within an RSVP stream caused less interference with target identification and produced a smaller AB effect than distractors that were similar to targets. This form of pop-out by distinction might be analogous to the phenomenon of *novel pop-out*, defined by more accurate localisation of a singularly novel item within a display of familiar (repeated) items than localisation of targets within a display of homogenous items (i.e., all novel or all familiar) (Johnston, Hawley, Plewe, Elliot, & DeWitt, 1990). Although the presence of an AB effect in Group OE – having used a (somewhat) familiar T2 among unfamiliar distractors – suggests that pop-out is unlikely to account for the lack of AB in Group GB, it was important to examine this possibility in more detail.

To determine whether the distinctiveness of the T2 item could account for the absence of an AB effect for familiar faces in Group GB, in Experiment 3 I replaced unfamiliar distractors with familiar (famous) face distractors in the RSVP stream and presented either unfamiliar or famous faces as T2 items. An unfamiliar T2 face among famous distractors created a pop-out condition where T2 was unique in its lack of familiarity. A famous T2 face among famous distractors created an RSVP configuration where pop-out was absent.

If familiarity reduces the burden on attentional demand, I predicted that no AB effect would be found when T2 was famous and an AB effect would be present when T2 was unfamiliar. Alternatively, if pop-out captures attention then I predicted that an AB effect would be found when T2 was famous and no AB effect would be present when T2 was unfamiliar.

To ensure high familiarity with the T2 faces, I only used British participants. One group was asked to detect unfamiliar T2 faces and another group was asked to detect famous T2 faces. The T2 faces were those used in Experiments 1 and 2. The methods were as outlined in the General Methods with the following exceptions.

Methods

Participants

All participants were UK citizens and had spent at least the past five years living in the UK. They were randomly assigned to one of two groups: Group U (unfamiliar T2) or Group F (famous T2). Unexpectedly, a large number of participants in Group U had extremely high false alarm rates. In 14 cases (39%) this was greater than 20% so data from these participants was excluded. High FA rates might have resulted from inducement of a generalised, false sense of recognition for the unfamiliar T2 faces, influenced by the high familiarity of the distractor faces. Only 5 participants (19%) from Group F were excluded on this basis, a percentage comparable to that found in Experiment 1 (15%). After exclusions, data were analysed from 22 (12 females, 10 males; mean age 22 yrs) and 21 (14 females, 7 males; mean age 23 yrs) participants in Groups U and F respectively.

Stimuli and Procedure

To maintain consistency across experiments, T2 images were the two unfamiliar T2 faces used in Experiment 1 and the two famous T2 faces used in Experiment 2. Of the famous distractor faces, half were male and half were female. Five were British politicians, six were from the British Royal Family (politicians and royalty were never used as T2 masks), and the remainder were a mixture of actors, singers, sports stars, and models considered internationally famous. Within each condition the experiment was split into two blocks, one for each face (counterbalanced). (See Appendix A for face stimuli.)

On completion of the experiment, each participant rated T2 and distractor faces for familiarity. Each face used in the study was presented in the centre of a computer screen and participants were required to make a familiarity judgement based on a scale of 0-5 (0 indicating no recognition of the face and 5 indicating high familiarity). Participants in Group F were also asked to name each T2 face as a further check of recognition.

Results

Familiarity Ratings

Group U familiarity ratings for the unfamiliar T2 faces ($M = 1.91$, $SE = 0.20$) were significantly lower than Group F ratings for the famous T2 faces ($M = 4.81$, $SE = 0.07$), $U = 0.50$, $p < .01$, as expected. Group U familiarity ratings for the famous distractors ($M = 3.56$, $SE = 0.12$) were not significantly different from those in Group F ($M = 3.44$, $SE = 0.15$). All participants in Group F correctly named each famous T2 face.

T1 Performance

The mean percent correct T1 score in Group U was 96.4% ($SE = 0.7$), a value not statistically different from that in Group F ($M = 96.4\%$; $SE = 0.6$).

T2 Performance

As in Experiments 1 and 2, data from both T2 faces in each group were combined. The FA rates were 8.5% ($SE = 1.0$) and 8.2% ($SE = 1.0$) in Groups U and F respectively, and did not differ statistically. Mean T2 hit rates for each group are plotted as a function of lag in Figure 10. For Group U (Figure 10a), a repeated-measures ANOVA with lag (1, 2, 3, 4, 5, and baseline) as a within factor revealed a significant main effect of lag, $F(5, 105) = 3.11$, $p < .05$, indicating the presence of an AB effect for unfamiliar T2 faces. Performance reached a minimum of 73.1% at lag 2 and improved with longer lags to reach a baseline value of 82.0%. The difference between these values was significant, $t(21) = 3.31$, $p < .05$, thereby confirming a clear

AB effect. Unlike the AB function found in Experiment 1, performance of Group U at lag 1 was not significantly different from baseline ($p > .1$). This suggests that lag-1 sparing was present. A within-subjects contrast analysis, however, revealed that the main effect of lag was significantly linear in nature, $F(1, 21) = 5.19, p < .05$, suggesting that lag-1 sparing, if present, was minimal.

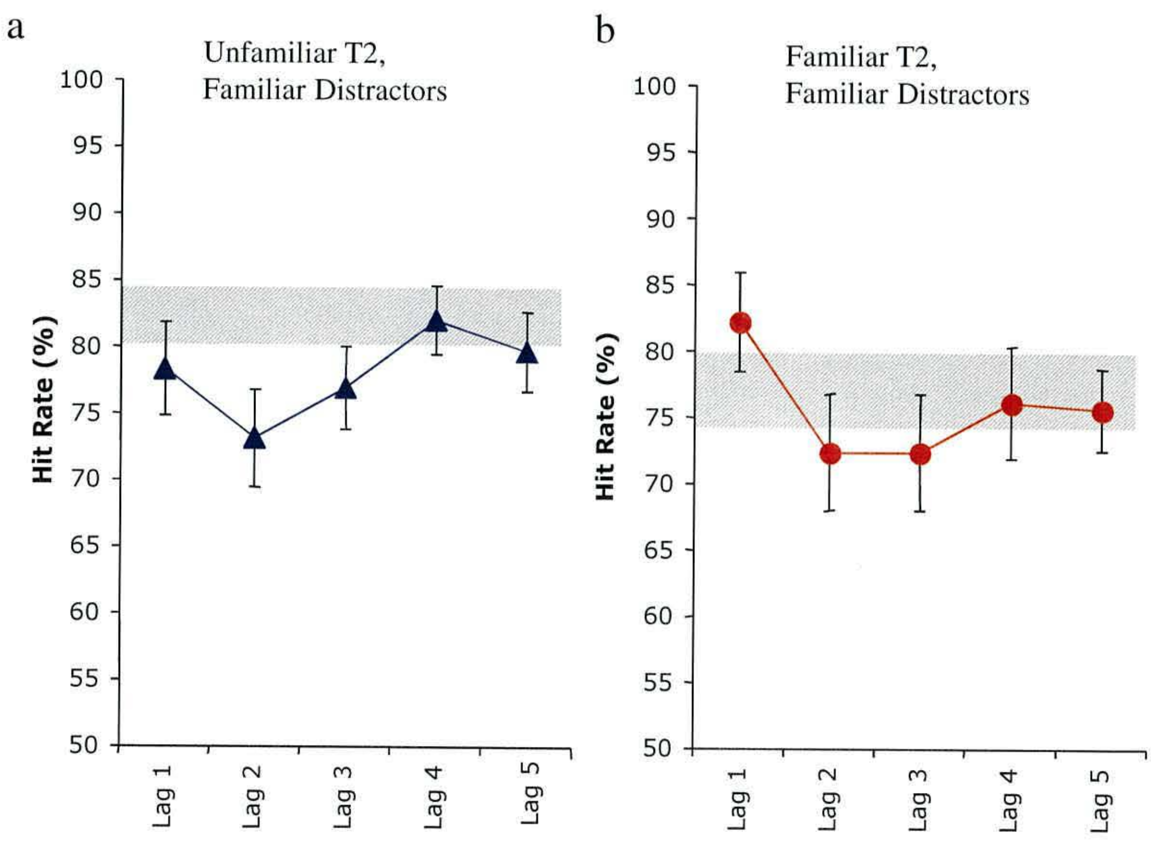


Figure 10. Mean percent hit rate as a function of lag in Experiment 3 for: (a) detection of unfamiliar T2 faces among highly familiar distractor faces (Group U); and (b) detection of highly familiar T2 faces among highly familiar distractor faces (Group F). Bars represent ± 1 standard error and the shaded area represents the mean hit rate at baseline (the mean of lags 6-8) ± 1 standard error.

In contrast to Group U, the detection of familiar T2 faces in a stream of other familiar faces (Group F, Figure 10b) showed no AB effect. Statistically the main effect of lag was significant for this group, $F(5, 100) = 2.75$, $p < .05$, but this result was influenced by performance at lag 1 (82.2%) which was marginally significantly *higher* than baseline (77.0%), $t(20) = 1.93$, $p = .07$. The lag-1 advantage for familiar faces was similar to that found in Group OE (Experiment 2; familiar T2), strengthening the suggestion that familiarity allows rapid access through a hypothetical attentional gate opened by T1. With lag 1 data excluded from Group F, there was a non-significant main effect of lag, $F(4, 80) = 1.11$, $p > .3$. To make sure that I was not simply eliminating a true AB effect in Group F by removing lag 1 data, I excluded lag 1 data from all other conditions that had revealed a significant AB and re-ran the analyses. The previous results were not affected: with lag 1 excluded, the main effect of lag (and therefore the AB effect) remained statistically significant in Experiment 1 [$F(4, 84) = 2.94$, $p < .05$], Experiment 2 (Group OE) [$F(4, 44) = 4.00$, $p < .01$], and Group U in this experiment [$F(4, 84) = 4.59$, $p < .01$].

In Group F, performance reached a minimum of 72.3% at lag 3 but this value was not significantly different from baseline, providing further support for the absence of an AB effect.

One notable feature of the results from Group F is that overall performance on the T2 task was markedly reduced compared to Group GB in Experiment 2 (compare Figure 10b with Figure 9a). This reduction in performance might have been caused by a general increase in task difficulty produced when a familiar T2 face had to be distinguished from equally familiar distractors. Performance in Group F was not lag-dependent, therefore the drop in overall performance does not appear to have an attentional basis. Note that performance was also slightly depressed when an unfamiliar T2 had to be distinguished from unfamiliar compared to familiar distractors. In this case performance did have an attentional basis.

Familiarity and pop-out effects were further examined by drawing comparisons across Experiments 1, 2, and 3. A mixed-design repeated-measures ANOVA with T2 type (unfamiliar, familiar) and RSVP configuration (pop-out, no pop-out) as between factors and lag (1, 2, 3, 4, 5, and baseline) as a within factor was conducted on hit rates obtained from each participant. In support of my proposal that T2 familiarity modulated the AB, a significant interaction between lag and T2 type, $F(5, 385) = 3.14$, $p = .01$ was found, coupled with a non-significant interaction of lag by RSVP configuration, $F < 1$. Critically, the triple interaction of lag by T2 type by RSVP configuration was non-significant, $F(5, 385) = 1.41$, $p > .1$. It is clear that familiarity, not pop-out, modulated attentional demand for face processing in these experiments.

Similar cross-experimental analyses conducted on T1 percent correct data and FA rates showed that, in all cases, main and interaction effects were non-significant, indicating that neither processing demands of the T1 task nor guess rates on the T2 task can account for the effect of T2 familiarity on the AB.

Discussion

When unfamiliar T2 faces were presented among familiar distractors (Group U) a significant AB effect was found that was largely similar to that obtained for unfamiliar T2 faces presented among unfamiliar distractors (Experiment 1). When familiar T2 faces were presented among familiar distractors (Group F) there was no evidence of an AB effect, a result that mirrored the lack of AB for familiar T2 faces presented among unfamiliar distractors (Group GB; Experiment 2). Table 1 summarises the predicted and actual AB effect results in relation to familiarity versus pop-out predictions.

Table 1. Predicted and actual attentional blink results relative to the pop-out and familiarity accounts of the AB for each condition: U-U (unfamiliar T2, unfamiliar distractors; Experiment 1); F-U (GB) (familiar T2, unfamiliar distractors; Experiment 2); U-F (unfamiliar T2, familiar distractors; Experiment 3), and F-F (familiar T2, familiar distractors; Experiment 3).

Condition	U-U	F-U	U-F	F-F
Experiment	1	2	3	3
Pop-out prediction	AB	No AB	No AB	AB
Familiarity prediction	**AB**	**No AB**	**AB**	**No AB**
Result	**AB**	**No AB**	**AB**	**No AB**

These findings indicate that, regardless of distractor type, when T2 was familiar it appeared protected from the AB effect, but when T2 was unfamiliar an AB effect was found. Uniqueness of the T2 item relative to distractor faces, on a familiarity dimension, cannot account for why detection of familiar faces was unperturbed by an immediate prior task.

Chapter Discussion

In three experiments, participants were required to discriminate the texture of an abstract pattern (T1) and detect the presence of a specific face (T2), embedded in a series of rapidly presented distractor faces. In Experiment 1, all faces (both T2 and distractor) were unfamiliar, i.e., unknown to participants prior to the experiment, and an AB effect was found. In Experiment 2, the distractor faces remained unfamiliar but the T2 faces used were highly familiar to one group of participants and only somewhat familiar to another group. No AB effect was observed when T2 was highly familiar but an AB was produced when T2 was somewhat familiar. Because the same T2 faces were used for both groups, the AB effect for somewhat familiar faces provides strong evidence against the argument that T2 stimulus artefacts could explain the lack of AB for highly familiar faces. Experiment 3 examined whether the potential for T2 to pop-out of the RSVP display, due to its status as the only familiar face, could account for the lack of AB observed for highly familiar faces in Experiment 2. In Experiment 3, all distractor faces were highly familiar faces and T2 faces were either unfamiliar or highly familiar. I found an AB effect for unfamiliar T2 faces and no AB effect for highly familiar T2 faces, replicating the results obtained from Experiments 1 and 2, and illustrating that pop-out (i.e., attentional capture) cannot account for the absence of an AB effect.

The pattern of results across these experiments supports two main conclusions. First, unfamiliar faces require attention if their identity is to gain access to awareness. The results also provide no compelling reason to suppose that the attentional resource needed for attending faces is qualitatively different from that needed to attend any other stimuli. Second, high familiarity reduces the amount of attentional resource required for successful face identification. Each point is discussed separately.

Faces and Attention

The above results refute the proposal that face identification *per se* requires no attention (e.g., Farah, Wilson et al., 1995). Clearly, finding an AB effect for unfamiliar and somewhat familiar faces is evidence of an attentional requirement for the conscious awareness of face identification. Studies have suggested that face information (especially emotional expression information) subliminally presented (Dimberg, Thunberg, & Elmehed, 2000) or presented outside the focus of attention (Eastwood, Smilek, & Merikle, 2001), can implicitly influence behaviour and cause emotion-specific brain activation (Anderson, Christoff, Panitz, De Rosa, & Gabrieli, 2003; Vuilleumier et al., 2001; Whalen et al., 1998; but see also Holmes, Vuilleumier, & Eimer, 2003; Pessoa, Kastner, & Ungerleider, 2002; Pessoa, McKenna et al., 2002, for contrasting results). In contrast, evidence that face *identification* can modify behaviour without explicit recognition has only been obtained using famous faces (Buttle & Raymond, 2003; Jenkins, Lavie, & Driver, 2003; Lavie et al., 2003). As my findings suggest, such phenomena probably stem from stimulus familiarity and are unlikely to be indicative of general face processing mechanisms.

In addition, my results do not support the notion of a special attentional mechanism dedicated either to face or configural processing (Awh et al. 2004; Jenkins et al., 2003; Palermo & Rhodes, 2002). The special-attention view of face processing posits that configural processing, thought to be required for optimal face identification, requires a separate attentional resource from that used for featural processing of non-face stimuli. Based on results from AB experiments using digits, letters, unfamiliar faces, and greebles, Awh et al. concluded that only when identification of T1 used configural processing (and hence depleted a hypothetical configural attention resource) would an AB effect be observed for face T2 stimuli. Contrary to this, using a featural T1 task I found clear evidence of an AB effect for unfamiliar and somewhat familiar faces. This suggests that explicit face identification utilises a similar attentional resource as that required for processing non-face, featural stimuli.

As discussed in Experiment 1, the discrepancy between the current results and those of Awh et al. (2004) might arise from differences in how an attentional blink is operationally defined than in actual findings. They found a significant main effect of lag (in their Experiment 5) using a featural T1 task (digit identification) and a configural T2 task (face identification). This finding would traditionally indicate an AB effect. Yet they conclude an AB effect was absent for T2 faces due to lack of difference between single-task (ignore T1) and dual-task performance. The presence of a significant lag effect in the single task confounds the use of this comparison.

Their results are therefore at odds with the proposal of a special configural attention resource for faces. It lacks parsimony to propose a special attentional mechanism for faces, or face-like, stimuli, and the results of Experiments 1-3 provide empirical evidence against this notion.

Instead, the default view of attentional allocation for face processing is supported. This states that perceptual and cognitive processes needed for the identification of faces are susceptible to limits in attention, as are non-face stimuli. I did not measure AB effects for non-face stimuli in this study so I cannot directly compare, either qualitatively or quantitatively, the attention needed for face versus non-face object identification. Nevertheless, the AB function for unfamiliar faces (Experiment 1) is qualitatively consistent with AB functions obtained previously using non-face objects (e.g., Raymond, 2003). Parsimony thus eliminates any need to posit a special attentional mechanism for faces on the basis of the current data.

Familiarity, Face Processing, and Attention

The second conclusion drawn from these results is that the processing of familiar faces appears to require little attention. This is consistent with several previous findings: irrelevant famous face distractors caused interference effects on a difficult attention-demanding central name categorisation task (Jenkins et al., 2003; Lavie et al., 2003); performance on a change detection task was better with famous faces than with unfamiliar faces (Buttle & Raymond, 2003); visual search for one's own face was more efficient than search for a recently learned stranger's face (Tong & Nakayama, 1999). The current findings are also consistent with reports of attentional advantages for one's own name (Arnell et al., 1999; Mack & Rock, 1998; Shapiro, Caldwell et al., 1997). The Shapiro et al. study – in which one's own name as T2 did not produce an AB effect – is comparable to the current study and the similarity in findings suggests that highly familiar faces "escaped" the AB because of their high familiarity, not their stimulus class. Consistent with this, a brain imaging study that used faces and buildings showed fame-specific activations in the medial temporal gyrus that were unaffected by stimulus category (Gorno-Tempini & Price, 2001).

The ability of familiar face identification processes to proceed with little attention also resolves some of the seemingly conflicting results about the role of attention in face processing. It suggests that empirical results obtained with famous faces cannot be used to make general statements about the role of attention in face processing.

Why are highly familiar faces protected from the AB effect? According to several theories of the AB, if a T2 stimulus is to gain access to awareness and be reportable at the end of an RSVP trial, its representation must be consolidated into WM (Chun & Potter, 1995; Shapiro, Arnell et al., 1997). The current finding that highly familiar faces escaped the AB suggests two possibilities. First, highly familiar faces might be perceptually processed, or *encoded*, faster than less familiar faces, making their representations less susceptible to interference from the immediately succeeding stimulus (the T2 mask) and more likely to gain access to WM. Tong and Nakayama (1999) proposed that each familiar stimulus representation is stored as a *compact visual code* that is rapidly processed. This is similar to the idea that increased familiarity with a stimulus results in the creation of a sparse neuronal network that only contains information diagnostic for identification (Desimone, 1996), and which might be used rather like a short-cut. In this sense, consolidation of perceptual face information into WM might be more efficient for familiar than unfamiliar stimuli.

The notion of a compact visual code for processing familiar faces could be analogous to broad-scale configural processes that rely more upon spatial relations between large regions of a face than detailed featural information contained within each region. Are familiar faces processed more configurally than unfamiliar faces? Studies of expert object recognition suggested that, as familiarity with a visual object increases, processing shifts from a featural mode to a configural mode. In contrast, there is no evidence to suggest that familiarity with *faces* alters processing in this manner (Collishaw & Hole, 2000). In addition, such a hypothetical advantage in perceptual processing is not supported empirically as yet by brain imaging studies. As mentioned in Chapter 1, N170 waveform activity, thought to be indicative of early structural encoding of faces (Bentin et al., 1996), is not speeded or diminished by familiarity with a face (e.g., Bentin & Deouell, 2000), an effect that might be expected if familiarity enhanced compaction of structural face information.

A second possibility is that the familiarity advantage is located *in* WM rather than during transfer into WM. That is, significant perceptual experience with a stimulus might allow *attention-friendly maintenance* of a distinct, durable representation within WM, once entry has been gained. It has been suggested that the persistence of representations in WM is enhanced by strong long-term memory

(LTM) representations (Cowan, 2001; Ruchkin, Grafman, Cameron, & Berndt, 2003). Given that highly familiar faces are more richly encoded in LTM than unfamiliar faces, they might be more robustly or distinctively maintained in WM and enjoy better protection from decay or interference. Evidence for rich encoding of familiar faces is supported by brain imaging studies that showed famous faces, unlike unfamiliar faces, activated regions of the posterior cingulate, including the retrosplenial cortex, areas associated with episodic memory and emotional salience (e.g., Shah et al., 2001), and areas of the left anterior middle temporal gyrus (Gorno-Tempini & Price, 2001) associated with semantic processing and categorization (e.g., Devlin et al., 2002). Support for this late, postperceptual familiarity influence is provided by studies that showed an effect of face familiarity on later ERP components: as mentioned earlier, enhanced negativity of the N400 waveform has been reported for familiar compared to unfamiliar faces (e.g., Bentin & Deouell, 2000; Eimer, 2000a, 2000b), an effect considered to reflect semantic activity involved in the identification of familiar faces.

The results of Experiments 1-3 cannot distinguish between the alternatives of an early perceptual versus late postperceptual influence of familiarity. Visual search and visual WM experiments, reported in Chapters 4 and 6 respectively, aimed to address these issues more directly.

To summarise, the results of Experiments 1-3 indicate that successful report of unfamiliar face identies requires attentional resource, probably in much the same way as other non-face stimuli. I find no support for the notion that a dedicated configural attentional channel mediates face identification. The key finding was that high familiarity dramatically eased the burden on attentional resource required for the report of face identies. Highly familiar stimuli, such as famous faces, might be protected from AB effects because they are efficiently encoded and consolidated into WM. It is also possible that famous faces benefit from superior maintenance in WM, aided by enduring, stable, and highly resilient representations supported by LTM.

CHAPTER 4

THE EFFECT OF FAMILIARITY ON VISUAL SEARCH FOR FACES AND COMPLEX NON-FACE OBJECTS: EXPERIMENTS 4-6

The experiments in the previous chapter demonstrated that familiarity reduces the burden on attentional resources required for face identification: an AB effect was present for unfamiliar faces but absent for (highly) familiar faces. Two possibilities for this familiarity advantage were outlined.

One possibility is that early perceptual encoding of familiar faces might be particularly efficient due to the use of a compact visual code that only activates neurons diagnostic for identification and enables swift and unperturbed transfer into WM. The notion of a compact visual code is akin to that of configural processing, e.g., a compact code might exclusively process a relatively small number of spatial relations between features in a face: distances between the eyes, eyes and nose, nose and mouth, and eyes and mouth.

A second possible account of the lack of AB for familiar faces is that later postperceptual processes involved in familiar face recognition might be particularly efficient. These might include superior maintenance of familiar face representations in WM once access has been gained. I suggested that the maintenance of familiar face representations might be more durable and resistant to interference from competing stimuli than unfamiliar face representations, due to support from LTM.

In this chapter, I report a series of visual search experiments that aimed to investigate the first proposal of whether familiar faces are more efficiently encoded than unfamiliar faces. In Chapter 6 I report a series of visual WM experiments that aimed to directly investigate the second proposal of a postperceptual locus for the familiarity advantage.

Although as yet there are no neurophysiological data to support the notion of an early perceptual advantage for familiar faces, Tong and Nakayama's (1999) visual search study showed that perceptual encoding of familiar faces was enhanced relative to unfamiliar faces: search for one's own face was more efficient than search for an unfamiliar face among unfamiliar distractors. As noted in the previous chapter, one limitation of their findings is that the search advantage for a familiar face might have been caused by a pop-out effect in which one's own target face was relatively easy to find because it was the only familiar face in the display. Unfamiliar target faces presented among other unfamiliar faces would not have enjoyed this distinction.

This chapter reports a series of visual search experiments that used faces (Experiments 4 and 5) and complex non-face symbols (Experiments 6a and 6b). It is beyond the scope of this thesis to provide an in-depth review of visual search theory, but some relevant aspects are outlined.

The visual search paradigm is considered a useful measure of perceptual encoding efficiency. A typical visual search task involves speeded search for a predefined target among a display of distractor items. Reaction times (RTs) to state whether a target item is present or absent are measured. Typically (in a display of relatively homogeneous items), RTs increase linearly as the number of items in the display (set size) increases, considered by many to indicate a serial search process. Theories of serial search propose that each item is checked against an internal representation of the target item until the target is found (at which point search is self-terminated and a target present response is made), or until all items have been checked and the target is considered absent (known as exhaustive search). The time taken to complete the search increases as set size increases; target present trials are typically responded to faster than target absent trials. Treisman and Gelade (1980) offered the slope of the function relating RT to set size as a measure of search efficiency. The logic is that if the search rate per item is slow, then adding more distractors to the display will markedly affect the amount of time it takes to find the target (or determine that it is absent), and result in a steep search slope. If the search rate per item is fast, then adding more distractors will have less impact on the time needed to complete the search task, and a shallower search slope is produced. (See Duncan & Humphreys, 1989; Moore & Wolfe, 2001; Sternberg, 1966; Townsend, 1990; Townsend & Fific, 2004; Woodman & Luck, 2003, for various reviews of visual search processes.)

The visual search task is also considered a useful measure of attentional requirements for visual processing. Attention is required to reject distractor items and select a target item, and, according to the serial search model, is allocated to each item in a sequential fashion (Treisman & Gelade, 1980). The less attention that is required for distractor rejection and target selection, the faster each item will be processed in turn, resulting in more efficient search.[10] The aims of these experiments were fourfold. First, I wanted to confirm whether perceptual encoding is more efficient for familiar than unfamiliar faces and could thus account for at least part of the attentional advantage found in Chapter 3. Crucially, I controlled for potential pop-

[10] The serial model of visual search is most appropriate for my data. Another type of search process has been evidenced, known as *parallel* search, in which the time needed to complete the search is independent of the number of items in the display. Parallel search is typically characterised by a flat search slope. Theories of parallel search propose that attention is allocated to, and information is accumulated from, most or all items in a display simultaneously (e.g., Humphreys & Muller, 1993, cited in Moore & Wolfe, 2001).

out effects by presenting familiar target faces among familiar distractors, and unfamiliar target faces among unfamiliar distractors (Experiment 4).

Second, I wanted to examine more closely whether increased use of configural information could explain the familiarity advantage for faces. Whereas upright faces were used in Experiment 4, inverted faces were used in Experiment 5. Based on the assumption that inversion disrupts configural processing, inverting the faces allowed me to assess the relative contribution of featural and configural processes to unfamiliar and familiar face identification.

Third, I wanted to determine whether familiarity effects on encoding efficiency for faces could also apply to complex non-face objects. To this aim, visual search efficiency for Chinese characters, called Hanzi, was compared between Hanzi experts and Hanzi novices (Experiment 6a). Inversion effects for Hanzi were also investigated (Experiment 6b).

Finally, visual search rates obtained in the above experiments were used to set encoding duration parameters in the series of visual WM experiments reported in Chapter 6. Each visual search experiment reported here was paired with an equivalent visual WM experiment in which the same participants were tested under the same familiarity and orientation conditions. For example, participants who completed the visual search task for unfamiliar and familiar upright faces completed the visual WM task for unfamiliar and familiar upright faces. This within-subjects design allowed a degree of control over individual differences in encoding speed, attentional limits, and visual WM capacity.

Face and Hanzi experiments are reported in two separate sections below, labelled accordingly.

Visual Search for Faces: Experiments 4 and 5

Experiment 4 examined visual search for upright unfamiliar and familiar faces. Experiment 5 examined visual search for inverted unfamiliar and familiar faces. A different group of participants completed each experiment. Specific hypotheses are presented at the start of each experiment section. Similar procedures were used in both experiments, therefore a General Methods section is provided below.

General Methods

Apparatus

Stimuli were displayed on a 22-inch Mitsubishi DiamondPro 2060u monitor (32-bit true colour; resolution 1280 x 1024 pixels) and generated by E-Prime software (Version 1.0; Schneider et al., 2002). Viewing distance was 70 cm.

Stimuli

Faces were greyscale images of Caucasian adult males with hair present. As far as possible, neutral faces seen in frontal view were selected. None wore glasses or sported facial hair. A set of 30 unfamiliar and 30 familiar (famous) faces were used.

Unfamiliar faces were selected from the Psychological Image Collection at Stirling (PICS). Famous faces were selected from Google Image web search results by using famous names as search terms. In order to mirror natural diversity, faces were not matched on dimensions such as attractiveness or distinctiveness. Luminance and contrast values of face images were not manipulated because only reasonably high quality images were used and no obvious, systematic differences on these dimensions between famous and unfamiliar faces were apparent. Faces subtended $2.8^{\circ} \times 3.3^{\circ}$ and were displayed at random locations within a white region (approximately $17.0^{\circ} \times 14.4^{\circ}$), with the constraint that each was separated by at least 2.9° on the horizontal axis and 3.4° on the vertical axis, centre to centre. All unfamiliar and famous face images are provided in Appendix A.

Design

Participants searched for a target face among a display of distractor faces. Sets of 12 target and 18 distractor faces were used. Although Tong and Nakayama (1999) suggested that even hundreds of exposures to an unfamiliar face were insufficient to yield a search efficiency that was comparable to that for highly familiar faces, the use of a number of different target and distractor faces served here to minimise any potential learning effects that might have occurred during the experimental session. The target face for a particular trial was randomly selected from the set of 12 and presented at the start of that trial. Distractor faces were randomly selected from the set of 18. Set sizes 4, 7, and 10 were used for upright faces in Experiment 4; set sizes 2, 4, and 6 were used for inverted faces in Experiment 5. No face appeared more than once in any given display. There were 72 trials per set size (50% target present), yielding 216 trials in total per session. Set size and target present/absent conditions were pseudorandomised. Participants in each experiment completed two sessions, one with unfamiliar faces and the other with famous faces. Each session was completed on a different day and session order was counterbalanced.

Procedure

Each search trial began with a 1000 ms central fixation cross, followed by a target face presented for 1000 ms in the centre of the screen. After a 500 ms blank interval the faces search display was presented until the participant reported, as quickly and accurately as possible, whether the target face was present or not (using key I labelled "yes" and key E labelled "no"). Figure 11 illustrates a typical trial. A short practice session was provided before the main experiment began.

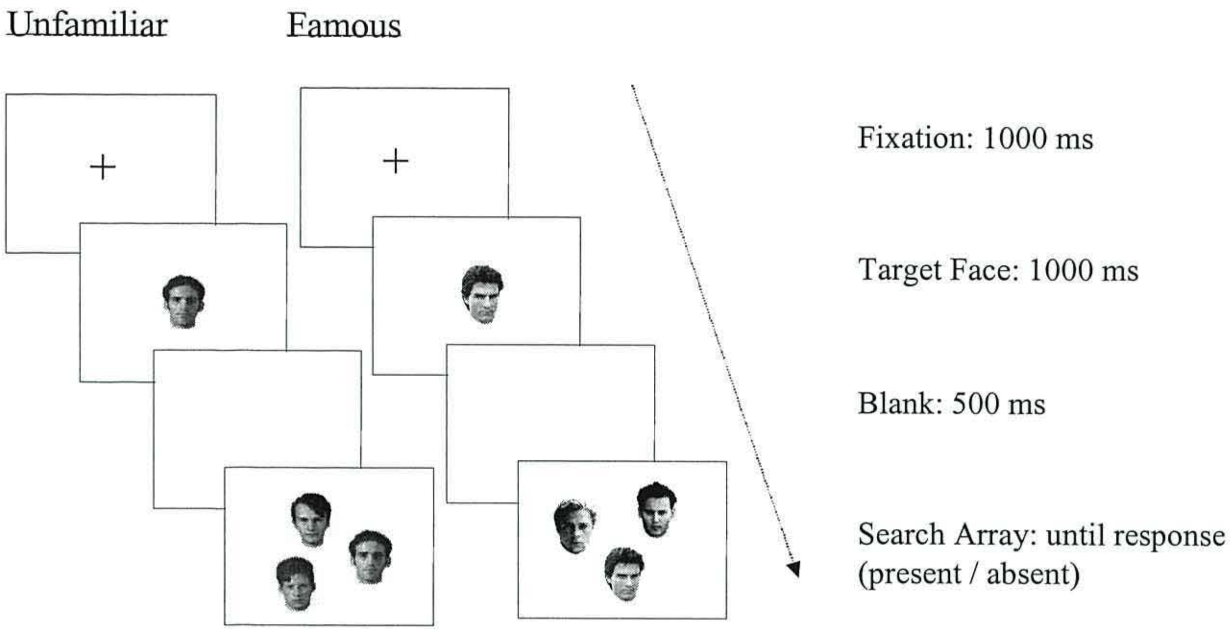


Figure 11. Example of the visual search trial procedure. Unfamiliar and famous faces were used in Experiments 4 (upright) and 5 (inverted). Hanzi (complex non-face objects) were used in Experiments 6a (upright) and 6b (inverted) in which Hanzi expert versus Hanzi novice participant groups were compared.

As mentioned previously, participants in the visual search experiments also completed an equivalent visual WM experiment (reported in Chapter 6). After both visual search and visual WM study sessions were completed, participants rated the familiarity of each of the 30 famous faces used across studies (subsets of 11 famous and 11 unfamiliar faces used in the visual search experiments were used in the visual WM experiments). As ratings are pertinent to the current visual search experiments, the procedure and results are reported here. Face images were presented upright in a paper questionnaire. Participants rated familiarity on a scale of 1-5, where 1 indicated no recognition of the face and 5 indicated high familiarity. Participants were also asked to name each famous face as a further check of recognition. Ratings confirmed that participants in both experiments were highly familiar with the famous faces

(Experiment 4: $M = 4.23$, $SE = 0.09$; Experiment 5: $M = 4.04$, $SE = 0.13$). Participants were also able to name the faces relatively accurately (Experiment 4: $M = 73\%$, $SE = 3$; Experiment 5: $M = 71\%$, $SE = 3$). Although naming accuracy might appear lower than expected of highly familiar faces, anecdotal evidence indicated that some participants did clearly recognise the faces but experienced problems generating the corresponding names. Familiarity ratings and naming accuracy did not significantly differ between experiment groups.

Data Analysis

Analyses of RT data were only conducted on trials in which the search response was correct. Accuracy data is presented in each experimental section. Data from target present and target absent trials were combined to produce an overall measure of efficiency in both selecting a target and rejecting distractors.[11] To exclude outliers and minimise anticipatory results, RTs two standard deviations above or below each participant's mean RT in each familiarity and set size condition were removed. The percentage of trial exclusions from each set size and familiarity condition in each faces experiment can be seen in Tables B1 and B2 in Appendix B.

In both upright and inverted face experiments, the effect of session order (unfamiliar-familiar; familiar-unfamiliar) was calculated using search slope values and was found to be non-significant.

[11] Note that in all visual search experiments target present trials were performed significantly faster than target absent trials, consistent with a serial model of visual search.

Experiment 4: Upright Unfamiliar and Famous Faces in Visual Search

The aim of this experiment was to determine whether search for a familiar face is more efficient than search for an unfamiliar face while controlling for a potential pop-out effect. In one session, all target and distractor faces were unfamiliar. In another session, all target and distractor faces were famous. If familiarity increases encoding efficiency with little drain on attentional resource, search was expected to be faster and reveal a shallower search slope for famous versus unfamiliar faces.

Methods

Participants.

Twenty-four participants (16 females, 8 males; mean age 22 years) volunteered to participate in exchange for course credits or monetary remuneration. All reported normal or corrected to normal vision and were naïve to the purpose of their experiment. Informed consent was obtained prior to participation.

Results and Discussion

The results support the notion that familiar faces are encoded more efficiently than unfamiliar faces, even with the potential for pop-out removed. Figure 12a shows RT data for unfamiliar and famous upright faces as a function of set size. It is clear that search for famous faces was faster overall and did not slow as rapidly with increasing set size as search for unfamiliar faces. A repeated-measures ANOVA with familiarity (unfamiliar, famous) and set size (4, 7, and 10) as within factors revealed a significant main effect of familiarity, $F(1, 23) = 21.00, p < .001$ and a significant interaction between familiarity and set size, $F(2, 46) = 5.10$, $p < .05$. A comparison of the slopes for functions relating RT to set size revealed a significantly shallower search slope for famous than unfamiliar faces, $t(23) = 2.45$, $p < .05$. Search was also more accurate for famous (M = 97%, SE = 0.00) than unfamiliar (M = 95%, SE = 0.01) faces and a paired samples *t*-test confirmed that this difference was significant, $t(23) = 3.78$, $p < .01$.

These data are consistent with Tong and Nakayama's (1999) proposal that familiar faces are encoded more efficiently, and require less attention for processing, than unfamiliar faces.

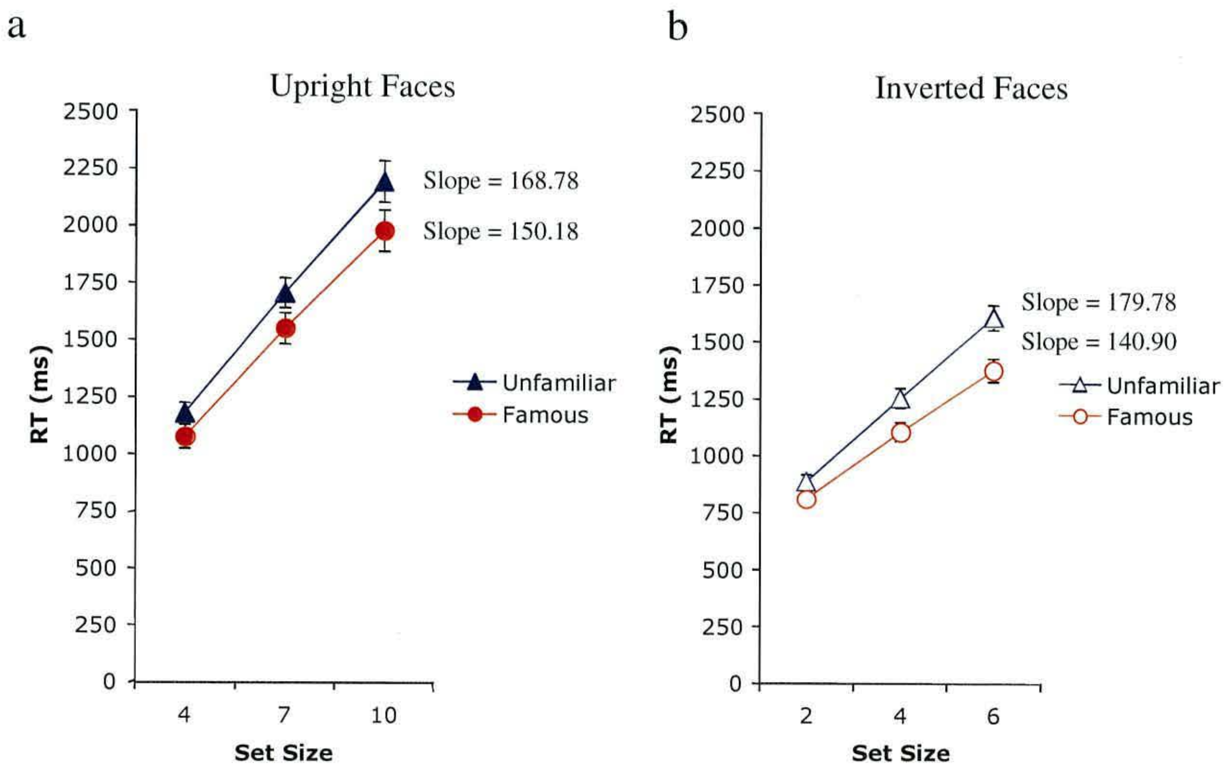


Figure 12. Mean reaction time (RT) to make a correct visual search response as a function of set size for (a) upright unfamiliar and famous faces (Experiment 4), and (b) inverted unfamiliar and famous faces (Experiment 5). Bars represent ± 1 standard error. Search slope values are also provided.

Experiment 5: Inverted Unfamiliar and Famous Faces in Visual Search

The aim of this experiment was to determine the relative contribution of featural and configural information to unfamiliar and familiar face processing. As outlined in Chapter 1, there is little evidence for differential inversion effects for familiar compared to unfamiliar faces, suggesting that featural and configural processes contribute relatively equally (e.g., Collishaw & Hole, 2000). Studies of object recognition, on the other hand, have shown that dog and Greeble experts demonstrated greater inversion effects than novices (Diamond & Carey, 1986; Gauthier & Tarr, 1997). Based on the assumption that inversion disrupts configural processing, it was concluded that expertise increases the reliance on configural information (i.e., configural precedence) used to recognise objects. Does familiarity similarly increase reliance on configural processing of faces? If reduced attentional requirement for familiar compared to unfamiliar faces reflects configural precedence, perhaps aided by compact visual codes, it is expected that disabling access to configural information by inverting the faces will remove the familiar face advantage found in Experiment 4, and equivalent search efficiency for famous and unfamiliar inverted faces is expected.

Methods

Participants.

Twenty-four participants (21 females, 3 males; mean age 21 years) volunteered to participate in exchange for course credits or monetary remuneration.

All reported normal or corrected to normal vision and were naïve to the purpose of their experiment. Informed consent was obtained prior to participation.

Stimuli, design, and procedure.

I used the same sets of 30 unfamiliar and 30 famous faces as used in Experiment 4 and inverted them by rotating them 180°. Set sizes 2, 4, and 6 were used. All other design and procedure parameters were employed as in Experiment 4.

Results and Discussion

Contrary to predictions derived from the object recognition literature, search was more efficient for famous than unfamiliar inverted faces and all analyses mirrored the familiarity advantage found for upright faces in Experiment 4. Figure 12b above shows RT data for unfamiliar and famous inverted faces as a function of set size. It is clear that search for famous faces was faster overall and did not slow as rapidly with increasing set size as search for unfamiliar faces. A repeated-measures ANOVA with familiarity (unfamiliar, famous) and set size (2, 4, and 6) as within factors revealed a significant main effect of familiarity, $F(1, 23) = 7.70, p < .05$ and a significant interaction between familiarity and set size, $F(2, 46) = 8.46, p < .01$. A comparison of the slopes for functions relating RT to set size revealed a significantly shallower search slope for famous than unfamiliar faces, $t(23) = 3.25, p < .01$. Search was also more accurate for famous ($M = 93\%$, $SE = 0.01$) than unfamiliar ($M = 90\%$, $SE = 0.01$) faces and a paired samples t-test confirmed that this difference was significant, $t(23) = 5.26, p < .001$.

The presence of a familiarity advantage in visual search when faces were inverted does not support the notion that configural information is more critical for familiar than unfamiliar face recognition: the disruption of configural information by inversion did not create a level playing field where encoding was equally efficient for famous and unfamiliar faces. Although inconsistent with the notion that expertise in object recognition promotes configural precedence, these results are consistent with previous research that found configural processes are no more utilised in familiar than unfamiliar face recognition (Collishaw & Hole, 2000).

The current results also support Tong and Nakayama's (1999) suggestion that familiar faces are more robust than unfamiliar faces. They demonstrated that search performance for inverted, three-quarter, and profile views was significantly more efficient for one's own face than an unfamiliar face, and concluded that familiar faces contain some view-invariant (or orientation-independent) information that is absent in unfamiliar faces.

Recall that object recognition studies suggested that featural processes are orientation-independent (unaffected by inversion) while configural processes are orientation-dependent (impaired by inversion). Is *featural* processing therefore *more* critical for familiar than unfamiliar upright face recognition, as opposed to configural processing? This would be a bold interpretation that contradicts all previous notions of the effect of familiarity on object and face recognition processes and, to my knowledge, such a suggestion has not been put forward to date.

A more sensible interpretation is that, when required (i.e., when faces are inverted), featural information might be utilised to greater advantage for familiar than unfamiliar faces in a manner than improves recognition. One possible mechanism for this is improved interhemispheric cooperation. Recall the bilateral field advantage found for the recognition of familiar but not unfamiliar faces (Mohr et al., 2002; Schweinberger et al., 2003). If the left hemisphere is involved in processing features and the right hemisphere is involved in processing configurations, better cooperation, or cross-talk, between these two processes might improve the transformation of featural information into configural information diagnostic for face recognition. To the best of my knowledge, the notion that *familiar* face recognition involves better cooperation between featural and configural processes than *unfamiliar* face recognition is unique.

In both upright and inverted experiments, it could be argued that search was more efficient for famous than unfamiliar faces due to the presence of stimulus artefacts. For example, what if face identification mechanisms were not engaged in the task and search performance was instead based on low-level featural differences between unfamiliar and famous face images such as luminance or contrast? To investigate this, search efficiency for upright versus inverted faces was compared. Based on the assumption that both unfamiliar and familiar upright face identification involves configural processing, inversion was expected to reduce search efficiency for both face types. Unfortunately the results are not clear-cut. A comparison of the slopes for functions relating RT to set size revealed non-significant inversion effects for unfamiliar and famous faces ($p > .3$ in both cases). Significant inversion-related decrements in search *accuracy* for both unfamiliar, $t(46) = 3.90$, $p < .001$, and famous, $t(46) = 3.19$, $p < .01$, faces was revealed, however. These results suggest that a speed-accuracy trade-off was present in the inverted faces data.[12]

Visual Search for Hanzi: Experiments 6a and 6b

The aim of the two experiments reported in this current section was to determine whether familiarity effects found for face stimuli could also apply to complex non-face stimuli (Hanzi). There are two reasons for investigating familiarity effects on Hanzi processing. First, the focus of this thesis is the effect of familiarity on attention and visual WM processes involved in stimulus recognition in general; I did not intend my research to address face recognition only. Second, due to suggestions that faces are processed differently from objects, I did not want to base any conclusions about familiarity solely on results from face experiments.

Before I report the experiments, I provide a rationale for using Hanzi as stimuli.

[12] In the visual WM experiments reported subsequently in Chapter 6, a sub-set of the same face images presented in the visual search experiments was used, and significant inversion impairments for both unfamiliar and famous faces were revealed. The presence of such inversion effects in the visual WM experiments suggests that the speed-accuracy trade-off found here is sufficient to account for the current lack of inversion effect. Nevertheless, it is difficult to rule out a low-level explanation for the familiarity effects found here. Further investigation that controlled for such artefacts might involve use of the same set of faces for which one participant group were familiar and another group were unfamiliar, in a similar vein to the British versus European groups in Experiment 2, Chapter 3. Alternatively, Fourier analysis of the currently used face images could determine whether there were any significant differences in spatial frequency (high versus low contrast) between unfamiliar and famous face images.

Why Use Hanzi?

Hanzi is the term for characters that form units of the Chinese logographic language system. An example is presented in Figure 13. Hanzi map onto *morphemes* (meaning) rather than *phonemes* (minimal sound units – as represented by English letters, for example).

Figure 13. Example of Hanzi. This one means *happy*.

The use of language-based, verbal stimuli to examine familiarity effects on *visual* processing might appear contradictory. However, Hanzi are particularly useful stimuli for the current purposes on three counts. First, Hanzi logographs undergo a high degree of visual processing. Each Hanzi is composed of a number of *strokes*, the smallest unit of Chinese language, that are packed into a square area according to learned stroke assembly rules. Visual and spatial properties of strokes are analysed to form a coherent percept. There is neurophysiological evidence that Hanzi recognition relies more heavily on visual processes than recognition of words derived from languages based on the Roman alphabet (e.g., English, French, Swedish, etc.). For example, Tan et al. (2001) found that English words were processed predominantly in left hemisphere verbal areas (i.e., Broca's area), while Hanzi were processed predominantly in right hemisphere regions such as the dorsolateral prefrontal gyrus, superior and inferior parietal lobules, and the occipital cortex – areas considered to serve visuospatial processes (e.g., Courtney, Petit, Maisog, Ungerleider, & Haxby, 1998; Schacter et al., 1995; Shen, Hu, Yacoub, & Ugurbil, 1999).

Second, the marked visual difference between faces and Hanzi allowed for clear investigation of whether the familiarity effects found for faces could apply to non-face objects. Previous studies that examined familiarity effects on object recognition used dogs, birds, and Greebles (as outlined in Chapter 1). While the results of such studies suggest that objects are processed in a similar way to faces if object expertise has been developed, it could be argued that dogs and birds are face-like because each has a face that is similarly configured to a human face (e.g., eyes, nose, mouth/beak). It has also been argued that Greebles are face-like (Kanwisher, 2000; although see Gauthier, Behrmann, & Tarr, 2004 for an alternative view). It is clear that Hanzi do not look like faces. Yet they are relatively complex stimuli and therefore serve as an ideal non-face comparison stimulus.

Third, the potential for stimulus artefacts to drive performance could be avoided by using the same set of stimuli with expert and novice participant groups. Physical stimulus attributes remained constant while only perceived familiarity differed between groups.

Despite evidence that Hanzi recognition involves a large degree of visual processing, I attempted to further minimise verbal processing of Hanzi within the expert group. A particular type of Hanzi was selected with which experts, although familiar with visual Hanzi form, had difficulty transforming into meaningful verbal information. These are called *traditional* Hanzi. Traditional Hanzi contain large numbers of strokes, are derived from ancient Chinese times, and are infrequently used in modern Chinese language. They contrast with *simple* Hanzi, derivatives of traditional Hanzi which contain fewer strokes. Simple Hanzi were developed in order to modernise and simplify the Chinese logographic language. Whereas the majority of young people in mainland China understand the meanings of, and are able to pronounce, simple Hanzi, they find meaning and pronunciation difficult to derive from traditional Hanzi.[13] To clarify this distinction, traditional Hanzi might be analogous to Middle English (circa 1150-1500) or Scots writing styles. Figure 14a shows an example of Middle English in the form of the first four lines from Geoffrey Chaucer's *The Canterbury Tales*, a collection of stories written between 1387-1400. Figure 14b shows an example of Scots in the form of the first verse from Robert Burns's poem *To A Mouse* (dated 1785). The modern English translation is provided on the right hand side of each figure.

[13] Thanks go to Xue Yingqi, a former Psychology Masters student at the University of Wales Bangor, for providing this useful information.

a

Middle English	Modern English translation
Whan that Aprill, with his shoures soote	When in April the sweet showers fall
The droghte of March hath perced to the roote	That pierce March's drought to the root and all
And bathed every veyne in swich licour,	And bathed every vein in liquor that has power
Of which vertu engendred is the flour;	To generate therein and sire the flower;

b

Scots	Modern English translation
Wee, sleekit, cowrin, tim'rous beastie,	Small, sleek, cowering, timorous beast,
O, what a panic's in thy breastie!	O, what a panic is in your breast!
Thou need na start awa sae hasty	You need not start away so hasty
Wi bickering brattle!	With hurrying scamper!
I wad be laith to rin an' chase thee,	I would be loath to run and chase you,
Wi' murdering pattle.	With murdering plough-staff.

Figure 14. As an analogy to how traditional Hanzi appear to Chinese participants familiar only with modern Hanzi, the two texts above illustrate how traditional English language forms such as Middle English and Scots can appear difficult to pronounce and interpret compared to modern English form. (a) An extract from Geoffrey Chaucer's Canterbury Tales (written 1387-1400). (b) An extract from Robert Burns's poem To A Mouse (written 1785).

As outlined below, Hanzi experts were recruited on the basis that they were from mainland China and were experienced with only simple, modern Hanzi. To recap, the key assumption was that expert participants were familiar with Hanzi visual form and structure in general, while novice participants had no experience in visually processing such stimuli.

Experiment 6a compared expert and novice visual search performance for upright Hanzi. Experiment 6b compared expert and novice visual search performance for inverted Hanzi. Similar methods were used in both experiments, therefore a General Methods section is provided below.

General Methods

Participants

Fourteen Hanzi experts (8 females, 6 males; mean age 28 years) and 14 Hanzi novices (7 females, 7 males; mean age 24 years) volunteered to participate in exchange for monetary remuneration. Hanzi experts were recruited on the basis that they were from mainland China and were familiar with only simple Chinese, not traditional Chinese language. Hanzi novices were required to have had no experience with Chinese language or any other similar logographic language such as Japanese. All reported normal or corrected to normal vision and were naïve to the purpose of their experiment. Informed consent was obtained prior to participation.

Apparatus

Stimuli were displayed on a 22-inch Mitsubishi DiamondPro 2060u monitor (32-bit true colour; resolution 1280 x 1024 pixels) and generated by E-Prime software (Version 1.0; Schneider et al., 2002). Viewing distance was 70 cm.

Stimuli

Traditional Hanzi were selected with the help of Xue Yingqi, a former Psychology Masters student at the University of Wales Bangor. Together we consulted an online Chinese dictionary (www.chinalanguage.com) and selected 30 Hanzi. As a young female from mainland China, Yingqi was able to select traditional Hanzi which she could not pronounce or interpret. The Chinese version of Powerpoint was used to produce each Hanzi and create high quality bitmap image files. Hanzi were comprised of between 13 and 27 strokes and were therefore highly physically complex items. All 30 Hanzi can be seen in Appendix C.

Design

Hanzi were black in colour, subtended approximately 2.5° by 2.5° of visual angle, and were randomly displayed within a white background region subtending approximately 17.0° x 14.4°, with the constraint that each was separated by at least 2.7° centre to centre. Twelve targets and 18 distractors were used. Set sizes 4, 6, and 8 were used for upright Hanzi in Experiment 6a; set sizes 2, 4, and 6 were used for inverted Hanzi in Experiment 6b. No Hanzi appeared more than once in any given display. There were 72 trials per set size (50% target present), yielding 216 trials in total per experiment. All participants, experts and novices, completed both upright and inverted experiments. Each experiment was completed on a different day and experiment order was counterbalanced.

Procedure

The visual search task for Hanzi was identical to that used for faces in Experiments 4 and 5 above, with the following exceptions (refer back to Figure 11 for an example of a typical trial). Hanzi stimuli were used instead of faces. Experiment instructions were provided in both English and (simple) Chinese to expert participants in order to remove any language barriers and ensure they fully understood the task. On completion of the subsequent visual WM experiment (reported in Chapter 6), expert participants were required to complete a questionnaire that probed three pieces of information. For all 30 Hanzi, they: (1) rated visual familiarity on a scale of 0-5 where 0 = not familiar at all and 5 = extremely familiar, (2) stated whether they understood the full meaning by responding "yes", "no", or "maybe". If they replied yes or maybe then they were invited to provide an interpretation, and (3) rated pronounceability on scale of 0-5 where 0 = unpronounceable and 5 = easily pronounceable. As expected, expert participants rated the traditional Hanzi low on visual familiarity ($M = 2.02$, $SE = 0.51$) and low on pronounceability ($M = 1.33$, $SE = 0.40$). Responses to the question on whether they understood the meaning of each Hanzi were numerically coded as follows: "No" = 1; "Maybe" = 2; "Yes" = 3. Group mean responses indicated that there was little meaningful interpretation of Hanzi ($M = 1.31$, $SE = 0.13$). Furthermore, only 47 interpretation responses were provided out of a possible 420 (14 participants x 30 Hanzi); only 9 of these interpretations were correct or nearly correct.

Data Analysis

RT data were analysed as in Experiments 4 and 5 above. The percentage of outliers that were removed from each set size and familiarity condition in each Hanzi experiment can be seen in Tables B3 and B4 in Appendix B. In the expert group, session order (upright-inverted; inverted-upright) did affect performance but only for inverted Hanzi. Using visual slope values, I found that visual search for inverted Hanzi was marginally significantly less efficient when experts performed this task in the second session, $t(12) = 1.85$, $p = .09$. As this is contrary to what might be expected by learning effects (i.e., learning should improve second session performance), this result is disregarded as an artefact.

Experiment 6a. Upright Hanzi in Visual Search: Experts Versus Novices

The aim of this experiment was to determine whether search for Hanzi was more efficient for experts than for novices. If familiarity with the visual Hanzi form increases encoding efficiency with little drain on attentional resource, search was expected to be faster and reveal a shallower search slope for experts than novices.

Results and Discussion

The results support the notion that familiar stimuli are encoded more efficiently than unfamiliar stimuli. Figure 15a shows RT data for experts and novices as a function of set size. It is clear that expert search for upright Hanzi was faster overall and did not slow as rapidly with increasing set size as novice search. A mixed design repeated-measures ANOVA with group (expert, novice) as a between factor and set size (4, 6, and 8) as a within factor revealed a significant main effect of group, $F(1, 26) = 14.74$, $p < .01$ and a significant interaction between group and set size, $F(2, 52) = 5.40$, $p < .05$. A comparison of the slopes for functions relating RT to set size revealed a significantly shallower search slope for experts than novices, $t(26) = 2.38$, $p < .05$. Unlike improved accuracy for famous compared to unfamiliar faces, there was a non-significant difference in Hanzi search accuracy between experts (M = 95%, SE = 1) and novices (M = 94%, SE = 0.5), $p > .3$. This suggests that novices were more careful in their responses and took longer in order to make a correct decision, indicative of laboured encoding.

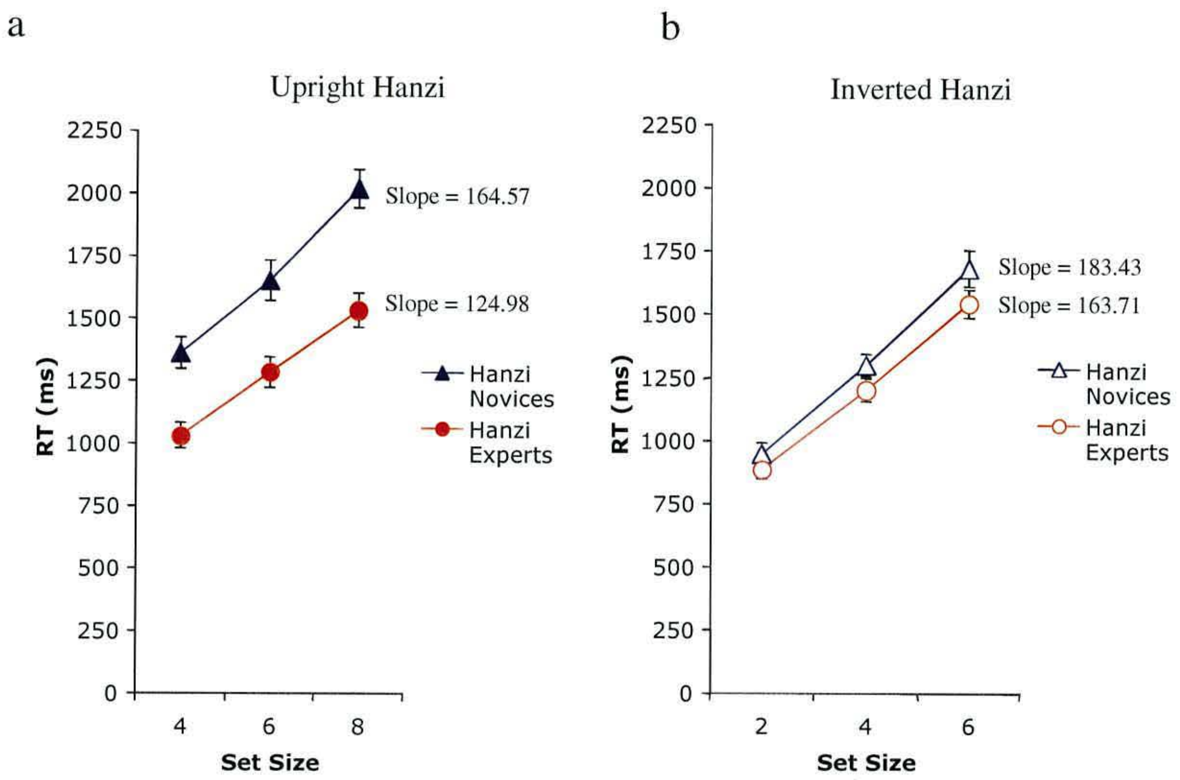


Figure 15. Mean reaction time (RT) to make a correct visual search response as a function of set size for (a) upright Hanzi, experts versus novices (Experiment 6a), and (b) inverted Hanzi, experts versus novices (Experiment 6b). Bars represent ± 1 standard error. Search slope values are also provided.

Overall, the results of this experiment confirm that the familiarity advantage in encoding upright faces found in Experiment 4 can be applied to the encoding of upright complex non-face stimuli.

Experiment 6b. Inverted Hanzi in Visual Search: Experts Versus Novices

The aim of this experiment was to determine the relative contribution of featural and configural information to expert and novice object processing. Recall from Experiment 5 that inverted famous faces were encoded more efficiently than inverted unfamiliar faces, a finding that might reflect the presence of an orientation-independent property of familiar face processing that is absent from unfamiliar face processing. I suggested that better cooperation between featural and configural processes might improve the transformation of featural information into configural information diagnostic for recognition. Does expert object processing share this orientation-independent property? If so, then search for inverted Hanzi is expected to be significantly more efficient among experts than novices.

Alternatively, how experts process Hanzi might not be equivalent to how we process familiar faces. As mentioned previously, object recognition studies suggested that novices process objects in a featural manner whereas experts process objects in a configural manner. If studies of expert object recognition are to be supported, disabling access to configural information by inverting the Hanzi will remove the expertise advantage found in Experiment 6a, and equivalent search efficiency for experts and novices is expected.

Methods

Stimuli, design, and procedure.

I used the same set of 30 Hanzi as used in Experiment 6a and inverted them by rotating them 180°. Set sizes 2, 4, and 6 were used. All other design and procedure parameters were employed as in Experiment 6a.

Results and Discussion

The results support previous accounts of expert object recognition: experts and novices showed equivalent search efficiency for inverted Hanzi. Figure 15b above shows RT data for experts and novices as a function of set size. It is clear that search speed and the rate of decline in RTs as set size increased were similar for experts and novices. A mixed design repeated-measures ANOVA with group (expert, novice) as a between factor and set size (2, 4, and 6) as a within factor revealed a non-significant main effect of group and a non-significant interaction between group and set size, $p > .1$ in both cases. A comparison of expert and novice search slopes for functions relating RT to set size was also non-significant, $p > .1$. Similar to upright Hanzi, there was a non-significant difference in inverted Hanzi search accuracy between experts ($M = 96\%$, $SE = 1$) and novices ($M = 95\%$, $SE = 1$), $p > .7$.

An investigation of inversion effects in each participant group revealed some illuminating results. A comparison of the slopes for functions relating RT to set size revealed a significantly steeper slope for inverted than upright Hanzi among experts $t(13) = 4.55$, $p < .01$, indicating that inversion decreased encoding efficiency. In contrast, the slopes for inverted and upright Hanzi were not significantly different among Hanzi novices, $p > .1$. Both groups showed a non-significant difference in accuracy for upright versus inverted Hanzi, indicating the absence of any speed-accuracy trade-off.

The results of Experiments 6a and 6b suggest that novices processed Hanzi in a featural, orientation-independent manner, whereas experts processed Hanzi in a configural, orientation-dependent manner. These findings are consistent with the object recognition literature.

Chapter Discussion

This chapter examined the effect of familiarity on visual search performance. In Experiments 4 and 5, visual search for unfamiliar versus famous faces was compared. When faces were upright (Experiment 4) search was significantly faster, showed a shallower decline with increasing set size, and was more accurate for famous faces. This suggests that less attentional resource is required for the perceptual encoding of familiar than unfamiliar faces and is consistent with Tong and Nakayama's (1999) search advantage for one's own face versus an unfamiliar face. When faces were inverted (Experiment 5), search was significantly faster, showed a shallower decline with increasing set size, and was more accurate for famous faces – a result that mirrored that found for upright faces. This suggests that processes used to identify inverted familiar faces might access an orientation-independent mechanism that processes used to identify inverted unfamiliar faces cannot. Such an orientation-independent mechanism might exhibit efficient cooperation between featural and configural processes; this cooperation might allow featural information in inverted faces to be translated into configural information that can then be used to recognise a face. If cooperation between featural and configural processes is less established in unfamiliar faces, the transfer of featural information into configural information might be ineffective, and might thus explain poorer encoding efficiency relative to familiar faces.

In Experiments 6a and 6b, visual search efficiency for complex non-face stimuli called Hanzi was compared between a group of Hanzi experts and a group of Hanzi novices. These experiments aimed to determine whether familiarity effects found for faces could also apply to non-face objects. When Hanzi were upright (Experiment 6a), search was significantly faster and showed a shallower decline with increasing set size for experts than novices. Accuracy did not differ between groups; this indicates that slower RTs in the novice group were likely to result from increased response cautiousness, probably due to greater encoding difficulty. These findings suggest that experts required less attentional resource than novices for the perceptual encoding of upright Hanzi and support the familiarity effect found for upright faces in Experiment 4.

When Hanzi were inverted (Experiment 6b) search results were notably different from those found in Experiment 5 using inverted faces. Both expert and novice groups showed equivalent search speed, rate of decline, and accuracy, indicating that inversion reduced expert search efficiency to the same level as novice search efficiency. In line with previous studies of expert object recognition, it appears that object expertise increases the reliance on configural information. In contrast to processes underpinning the identification of inverted familiar faces, processes involved in expert recognition of inverted Hanzi do not appear to have access to an orientation-independent mechanism.

There are two possible explanations for why expert recognition of inverted Hanzi and recognition of inverted familiar faces might have engaged different processes in the above studies. Perhaps face processing is special after all, at least in the sense that there is no other object category that is as frequently processed and important to recognise than faces. On the other hand, perhaps the use of traditional rather than simple Hanzi could account for the results. Traditional Hanzi were used in favour of simple, modern, and frequently used Hanzi in order to minimise verbal input on what could be considered a language-based task. Experts were familiar with traditional Hanzi in the sense that they were highly practiced at processing the Hanzi stimulus *category*. However, experts were not highly familiar with the *specific* Hanzi exemplars used: they were visually familiar with certain strokes and configurations of strokes, but were unable to interpret their meaning. This contrasts with the visual search task for famous faces in which participants were familiar with the specific exemplars used. Therefore, it is reasonable to assume that expert Hanzi recognition, in the current task at least, involved basic-level or subordinate recognition processes whereas familiar face recognition involved micro-subordinate recognition processes. Perhaps an orientation-independent mechanism involved in visual recognition processes is only available as a result of refined, micro-subordinate discrimination expertise for highly familiar stimuli.

This issue aside, in this chapter I have provided evidence that familiarity increases encoding efficiency for upright faces and complex non-face stimuli. This finding supports the AB findings in the previous chapter: less attention is required for processing familiar stimuli than unfamiliar stimuli.

While I interpret the familiarity advantage in visual search as suggestive of efficient, attention-friendly (serial) processing of each display item, there are two alternative explanations. First, it is possible that initial *parallel* processing of a whole familiar display is enhanced relative to a whole unfamiliar display, perhaps enabled by more distinct perceptual processing of robust item representations. Second, perhaps the familiarity advantage in visual search shown here was driven by enhanced encoding of the familiar compared to the unfamiliar target (cue) item presented at the start of each trial. Enhanced encoding of the pre-display target might have facilitated the orientation of attention to its location when presented in the search display and the rejection/inhibition of non-target locations when it was absent from the display.

Summary of Part 2: The Effect of Familiarity on Attention

The three attentional blink experiments reported in Chapter 3 revealed an AB for unfamiliar faces that was absent for famous faces. I concluded that familiar faces require less attentional resource for processing than unfamiliar faces. Reduced attentional requirement is interpreted to reflect facilitation of some stage of processing. Two possibilities were outlined: early perceptual encoding and postperceptual storage of information within visual WM.

Experiments 4 and 5, reported in this chapter, addressed the first possibility of an early perceptual familiarity advantage. I demonstrated that visual search for famous faces was significantly more efficient than that for unfamiliar faces, even when faces were inverted. Enhanced visual search for famous faces suggests that the lack of AB for famous faces could indeed partly be accounted for by efficient, early perceptual processes easing the burden on attentional resource. Experiment 6a revealed that encoding efficiency, and the related reduction in attentional demand, could be applied equally well to complex, non-face stimuli – Hanzi experts demonstrated significantly more efficient search than Hanzi novices. In contrast to the results obtained with faces, the familiarity advantage for Hanzi was removed when Hanzi were inverted.

Rapid perceptual encoding might facilitate the transfer, or consolidation, of information into visual WM and ensure access is fully gained before other stimuli can interfere with processing. Once entry into visual WM has been successfully gained, information is presumed to require effective maintenance until such time as retrieval is necessary. Can familiarity enhance the *maintenance* of information within visual WM?

Part 3 of this thesis addresses this question. Chapter 5 provides a review of visual WM theory and assesses the evidence for familiarity effects in WM. Chapter 6 reports a series of visual WM experiments that measured visual WM capacity for faces and Hanzi. These experiments were based on the assumption that larger capacity reflects enhanced maintenance.

PART 3

THE EFFECT OF FAMILIARITY ON VISUAL WORKING MEMORY

CHAPTER 5

VISUAL WORKING MEMORY: THEORY, AND A ROLE FOR FAMILIARITY?

Visual WM is an active system that temporarily stores and manipulates visual information for a few seconds. This ability is particularly useful for maintaining a constant and coherent percept of the visual world in which eye gaze, head or body movement, and passing objects or people can interrupt the flow of visual information to the brain. Think back to Simon in the bakery talking to his neighbour. It is a busy lunchtime and many customers are jostling for room at the counter. The bakery assistant greets Simon and takes his order. She turns and moves away to retrieve his sandwich from the chill cabinet and place it in a bag. She returns to face the crowd of customers, places the sandwich on the counter, and addresses Simon for the money. Had she been unable to maintain and retrieve his face from visual WM, she might have requested the money from a customer who had not yet placed an order. Consider that the bakery assistant and Simon have known each other for years and are good friends. Would long-term visual familiarity that has been developed over the years affect her ability to maintain his face in visual WM in this situation?

In this chapter I review visual WM theory and ask whether familiarity can enhance visual WM. Unlike the attention literature, evidence for familiarity effects on WM is scant, and I am unaware of any studies that have investigated the influence of familiarity on *visual* WM specifically. Experiments 7-9, reported in Chapter 6, address this issue for the first time, and I draw hypotheses from closely related literature. Before I review the nature of visual WM, I describe the larger, multi-component general WM system of which it is part.

The Working Memory (WM) Model

Baddeley and Hitch's (1974) model is the most well-known and cited model of WM. Their model contains three major components: the *phonological loop*, the *visuospatial sketchpad* (labelled *slave systems*), and the *central executive*. The phonological loop serves to process verbal, speech-based information and is comprised of two subcomponents: a phonological store that holds the information, and an articulatory control process that can refresh the contents of the phonological store via vocal or subvocal repetition (*rehearsal*). The visuospatial sketchpad is used to temporarily store and manipulate visual and spatial information. The central executive is thought to act as an attentional control centre that supervises and coordinates operations in the phonological loop and visuospatial sketchpad (this component is discussed in more detail later). Figure 16a provides an illustration of this model.

The Baddeley and Hitch (1974) model was derived from an earlier concept of short-term memory (STM) developed by Atkinson and Shiffrin (1968). Atkinson and Shiffrin proposed that a STM mechanism stores information and controls certain processes related to information storage such as rehearsal, decision-making, and retrieval – an idea retained in Baddeley and Hitch's WM model. Probably one of the most critical findings of STM was its capacity limitation, i.e., the amount of information that can be retained in memory at any one time. Capacity limits were first discovered and quantified by Jacobs (1887, cited in Baddeley, 1990) who developed a technique that has become known as the *memory span* procedure. He presented participants with sequences of numbers that increased in length; participants were required to repeat each sequence back to him. He found that as the number of digits to be repeated increased, participants made more errors. The sequence length at which a participant was 50%correct was taken to indicate his or her memory span. Subsequently, Miller (1956) reported that STM span was limited to approximately 7 ± 2 items. I discuss WM capacity in more depth in a later section of this chapter.

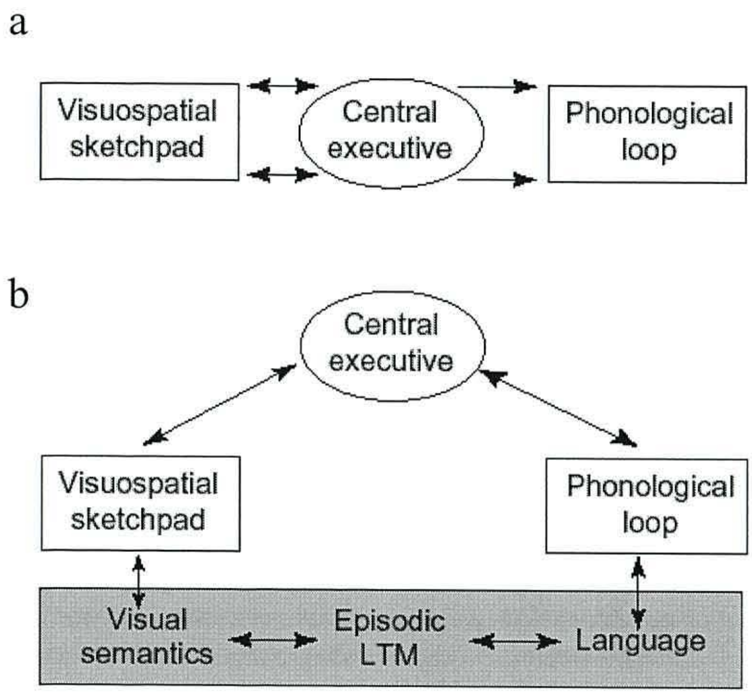


Figure 16. (a) Baddeley and Hitch's (1974) original model of working memory (WM). (b) Baddeley's (2000) updated WM model introduces the notion of an episodic buffer. Images were reproduced from Baddeley (2000).

The key difference between Atkinson and Shiffrin's (1968) STM model and Baddeley and Hitch's (1974) WM model is that the former considered STM to be unitary in that it operated on all kinds of tasks, whereas the latter proposed the fractionation of WM into the verbal and visuospatial systems outlined above.

The Fractionation of WM

The fractionation of WM was prompted by a series of dual-task experiments that yielded much less dramatic dual-task costs than predicted (Baddeley & Hitch, 1974). Participants repeated sequences of digits while concurrently performing a reasoning, learning, or comprehension task. Sequence length varied between 0-8 digits. The reasoning task, for example, involved true or false verification of sentence accuracy (e.g., "A follows B = BA?"). Based on a unitary account of WM, they predicted that as digit sequence length approached or exceeded capacity of approximately 7 items, the performance on the reasoning task would be dramatically impaired due to little, if any, residual WM resources. Contrary to this prediction, they found that participants were only 35% slower to complete the reasoning task when eight digits were repeated than when only one digit was repeated. Even more striking was that reasoning task error rates remained low at 5%, regardless of digit sequence length. Baddeley and Hitch surmised that digit span and reasoning task capacity limits originated from separate WM systems that control performance on these different types of task.

Both digit span and reasoning tasks described above appear to involve verbal, speech-based processes that require little, if any, storage or manipulation of visuospatial information in WM. So exactly how Baddeley and Hitch (1974) arrived at the notion of separate verbal and visuospatial WM stores from these tasks is unclear, and succeeding textbooks remain vague (e.g., Baddeley, 1986, 1990, 1997). Nevertheless, a variety of subsequent studies have provided support for the distinction between these two systems (e.g., Cocchini, Logie, Della Sala, McPherson, & Baddeley, 2002; Logie, Zucco, & Baddeley, 1990). It is beyond the focus of this thesis to provide a detailed review of evidence for separate verbal and visuospatial WM systems, but one study is described in full as an illustration.

Logie et al. (1990) required participants to concurrently conduct a primary and a secondary task that were either similar in the WM processes used (verbal-verbal; visuospatial-visuospatial) or different (verbal-visuospatial; visuospatial-verbal). The primary visuospatial task involved the observation of a matrix of black squares in which a selected proportion of squares turned white, one by one, at a rate of about 4 squares per second. After the full pattern of white squares had emerged, it remained on screen for 2 seconds, was removed from view for 2 seconds, and reappeared. On reappearance, one of the previously white squares had changed to black, and the participants were asked to point at the position in the matrix where they thought the change had occurred. Task complexity, and thus demand on capacity, was increased by means of increasing the number of squares in the matrix. The primary verbal task involved memorisation of a random sequence of consonants presented sequentially for 3 seconds each. Once the full sequence had been presented, a gap of 2 seconds occurred, after which the sequence of letters was repeated but with one of the letters replaced by another. Participants' task was to point to the screen whenever they detected a new (changed) letter. The visuospatial secondary task required participants to picture a 3 x 5 matrix of squares in their mind, and imagine squares being filled according to a sequence of instructions provided aurally by the experimenter. The instructions to fill or not fill each square started in the top left hand corner and proceeded along each row in turn. The resulting pattern of filled squares resembled a (digital style) number between 0-9; once instructions were complete, participants were required to report the number that had been created in their mind. The verbal secondary task involved mental arithmetic: participants were presented aurally with five single digits for mental addition and were required to respond vocally with the total. The primary task began first; the second (auditorily presented) task began approximately 2 seconds later. Logie et al. found that dual-task costs were significantly greater when the primary and secondary tasks shared similar WM processes than when they involved different processes. For example, performance on the verbal primary task was reduced from 80% when the secondary task was visuospatial, to 34% when the secondary task was verbal; performance on the visuospatial secondary task was reduced from 83% when the primary task was verbal, to 69% when the primary task was visuospatial.

Evidence for further fractionation of visuospatial WM into separate visual and spatial systems has also been provided (e.g., Della Sala, Gray, Baddeley, Allamano, & Wilson, 1999; Finke, Bublak, Neugebauer, & Zihl, 2005; Klauer & Zhao, 2004). For example, Klauer and Zhao found that a movement discrimination task (spatial; find the stationary item) significantly interfered with memory for dot locations (spatial; state the location of a previously presented dot), and a colour discrimination task (visual; state the dominant colour within an array of multiple colour patches) significantly interfered with memory for Chinese logographs (visual; state which of eight characters was previously presented). Interference between the two tasks was significantly lower when one task was visual and the other was spatial, compared to when both tasks shared similar processes. They concluded that a visual WM task was more disrupted by visual than spatial interference, and a spatial WM task was more disrupted by spatial than visual interference.

The most widely cited evidence for separate visual and spatial WM systems comes from the neuropsychological literature in which studies have found distinct involvement of different brain regions for visual and spatial WM tasks (known as the "*what* versus *where*" distinction). Visual WM tasks have been shown to activate ventral processing pathways (e.g., ventrolateral prefrontal cortex; VLPFC) whereas spatial WM tasks have been shown to activate dorsal processing pathways (e.g., dorsolateral prefrontal cortex; DLPFC) (see Haxby, Petit, Ungerleider, & Courtney, 2000 and Ungerleider, Courtney, & Haxby, 1998 for reviews).

The functional separation of verbal, spatial, and visual WM poses a problem however: how do these systems interact with each other? Many everyday tasks involve a combination of verbal, spatial, and visual processes. For example, the bakery assistant serving Simon has to retain verbal information about the change he is due ("forty-two pence"), while storing his face in visual WM, while keeping in mind where he is standing (if he is to receive his change swiftly). She also has to remember that it is Simon who is due forty-two pence change and not the man standing next to him. To account for the integration of information from various WM systems, Baddeley (2000) added a fourth component to the original Baddeley and Hitch (1974) WM model: the *episodic buffer* (see Figure 16b above). The episodic buffer is presented as "a limited capacity temporary storage system that is capable of integrating information from a variety of sources" (Baddeley, pp. 421). The episodic buffer is episodic in the sense that each piece of integrated information is linked to a particular, short-lived, time-locked episode or event. Baddeley also proposed this fourth component to involve the integration of information from LTM into WM. This suggestion was based on the finding that densely amnesic patient *PV*, who showed severe WM impairments (e.g., a word span of one), could accurately recall a sequence of up to five words if they combined to form a meaningful sentence (Vallar & Baddeley, 1984, cited in Baddeley). LTM preservation of grammatical knowledge, such as sentence structure, appears to have aided *PV*'s storage of information in verbal WM.

Another means by which these various WM components operate and interact is via the central executive (CE). The CE is commonly conceived of as the attentional command centre and is considered to be independent of task modality. The CE itself has been fractionated into a number of roles, summarised by Collette and Van der Linden (2002). These roles are: *updating* (the contents of WM storage systems are modified according to new incoming information), *inhibition* (irrelevant information is prevented access into WM), *shifting* (attention is directed towards and between task-relevant information, and away from task-irrelevant information), and *dual-task coordination* (considered the main function of the CE, attention is divided between two tasks to facilitate simultaneous encoding into, and storage within, WM). In terms of a neurological correlate, the CE does not appear to be *central* in that these functions do not derive from one specialised region. Rather, CE functions rely on a distributed cerebral network that involves prefrontal and posterior parietal regions (Collette & Van der Linden).

Having mentioned capacity limits briefly above, I now discuss this issue in more detail in the following section. Three WM processing stages have been defined: *encoding*, *maintenance*, and *retrieval*. Encoding involves the perception (e.g., visual or auditory) of information presented to our senses, and related processing such as stimulus recognition. Maintenance involves the storage of this information, considered to be located in the appropriate modality-specific WM system. Retrieval, in an empirical sense, typically involves the comparison of mentally stored to physically presented information. Typically, WM capacity limits are discussed in terms of the amount information that can be *maintained* at any one time.

WM Capacity Limits

Miller (1956) introduced the notion of a STM span of 7 ± 2 items. Subsequent evidence has suggested, however, that capacity is nearer 4 items (Henderson, 1972), an estimate that is still held today (see Cowan, 2001 for a review). The reduction in capacity estimate from 7 to 4 items appears to have stemmed from the identification of a process Miller termed *chunking*. He proposed that two or more individual items could be integrated into a single *chunk* of information, and that WM capacity is limited in the number of chunks, rather than individual items, that can be retained.

Based on ideas presented by Simon (1974), Cowan (2001) defined a chunk as "a collection of concepts that have strong associations to one another and much weaker associations to other chunks concurrently in use" (Cowan, pp. 89).

Recall amnesic patient *PV*, whose verbal WM span was only one word, but who was able to remember a sequence of five words when they combined to form a meaningful sentence. While it might appear that *PV*'s verbal WM capacity miraculously increased from one to five words, a more parsimonious explanation is that he had a constant capacity of one *chunk* of information; the ability to integrate multiple items into a single chunk, according to their degree of relatedness or association, enabled *PV* to store the five related words as one chunk – the sentence. Similarly, a sequence of four letters can be retained as one chunk if they spell a meaningful word (R-O-S-E), whereas four letters will more likely be retained in WM as four separate chunks if the letters are not meaningfully related (E-C-M-D). The integration of multiple items into one chunk does not have to rely on existing, long-term, meaningful associations, however. New associations can be formed between items. For example, if we are given the number 467312947565 to remember, we might parse it into three or four chunks – 4673 1294 7565 – to make it easier to retain and recall. And we might attach our own meaning to the letter sequence E-C-M-D (e.g., Eddie cooked Molly's dinner).

Baddeley (1990) summed up WM capacity as follows: "Memory span as measured in terms of items can be increased by increasing the number of items in each chunk" (Baddeley, pp. 42). To add to this definition, memory span as measured in terms of chunks appears to yield a relatively constant capacity of about four chunks.

The conceptual development of WM and chunking processes has been dominated by verbal WM studies. This is understandable because digit sequences, letters, words, sentences, and prose contain properties that can be easily separated and integrated; verbal WM tasks are relatively simple to design, implement, and measure. In contrast, visual images are not as obviously composed of discrete components; the extensive variety of image forms and complexities presents a vast heterogeneous set of stimuli with which to try and construct a theory of how information is organised in visual WM. Fortunately, significant empirical and theoretical advances in visual WM theory have been made within the last decade.

Visual WM

In a highly influential paper (extended in Vogel, Woodman, & Luck, 2001), Luck and Vogel (1997) used a change detection paradigm to measure visual WM capacity for simple visual objects. In the first set of experiments, a display of 1-12 randomly positioned coloured squares (the *sample*, or *memory* array) was presented for 100 ms, followed by a blank display interval of 900 ms (the *retention* interval). The squares display then reappeared (the *test* array) for 2000 ms, in which all the coloured squares were identical to the memory array or one square had changed colour. Participants were required to report, unspeeded, whether the memory and test arrays were the same or different. A typical trial is illustrated in Figure 17. They found nearly perfect performance at set sizes 1-3, but beyond this point performance declined systematically as set size increased.

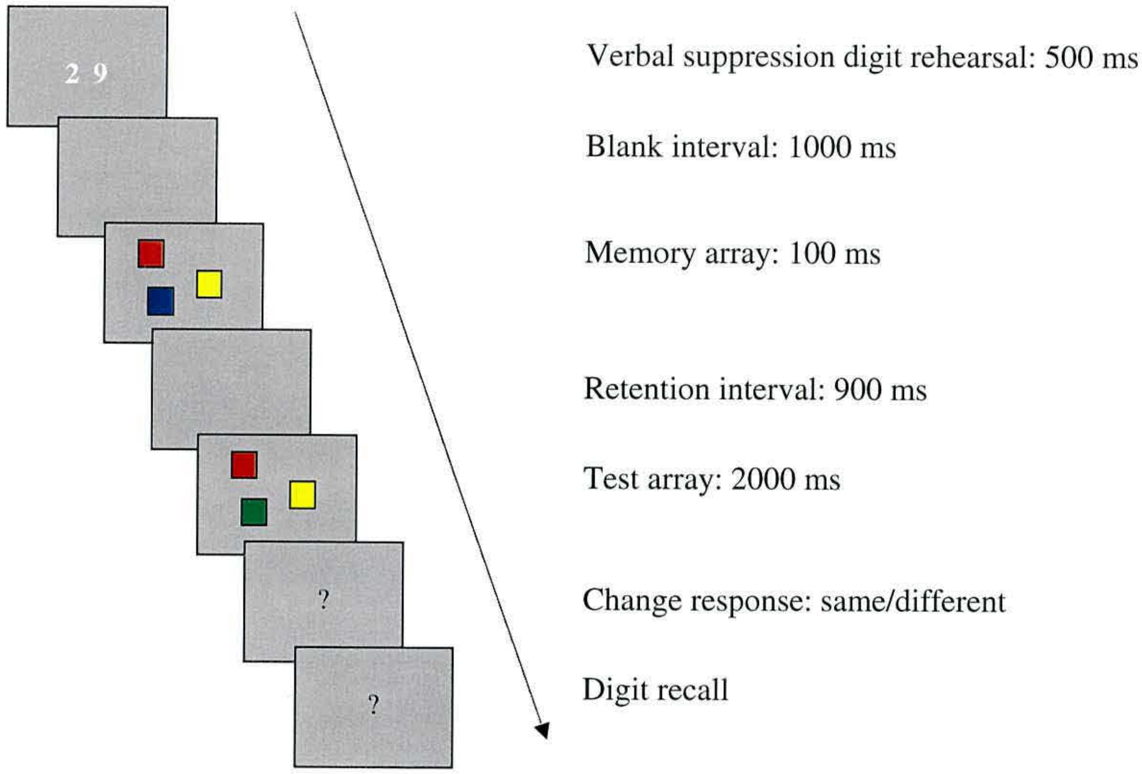


Figure 17. Example visual change detection trial procedure used by Luck and Vogel (1997) and Vogel et al. (2001) to measure visual WM capacity.

To demonstrate that these results reflected only visual WM capacity rather than the combined effort of verbal and visual WM, e.g., via colour naming, they added a concurrent verbal WM load to the change detection task. Participants were required to retain in WM two digits presented at the start of each visual change detection trial and report them at the end of the trial. They found that the addition of a verbal load made no difference to performance on the visual change detection task, and concluded that verbal WM processes did not influence their results.

In a subsequent set of experiments, Luck and Vogel (1997) also showed that increasing the duration of the sample array from 100 ms to 500 ms and using a bar orientation task rather than a colour task yielded similar results, suggesting that limitations in performance beyond set size 3 were not due to encoding speed restrictions, and that visual WM operates similarly for different kinds of objects and object tasks.

A decline in performance beyond set size 3 onwards suggests that visual WM capacity might have been about 3 items. In order to determine a more specific capacity estimate, Luck and Vogel (1997) applied a derivation of a mathematical formula developed by Pashler (1988), known as *Pashler's k*:

$$\boldsymbol{k} = [\mathbf{S} \times (\mathbf{H} - \mathbf{F})] / (\mathbf{1} - \mathbf{F})$$

where S = set size, H = hit rate, and F = false alarm rate (see Vogel et al., 2001). Vogel and colleagues explained this formula as follows: "...if a participant can hold *k* items in memory out of an array of *S* items, then the item that changed should be one of the items being held in memory on *k* / *S* trials, leading to correct performance on *k* / *S* of the trials on which an item has changed. To adjust for the effects of guessing, this approach also takes into account the false alarm rate" (Vogel et al., pp. 95).

Application of this formula to their results revealed a capacity of between 3 and 4 items, consistent with a chunk span of 4, and suggestive that each coloured square or orientated bar was stored as one chunk. This implies that a visual WM chunk contains simple information on one feature dimension: colour or orientation. Yet verbal WM studies have shown that relatively large amounts of information can be stored within one chunk (e.g., five words can be stored as one chunk if they form a meaningful sentence). Can each visual WM chunk contain more information if that information is associated in some way?

Luck and Vogel (1997) addressed the issue of information integration by measuring visual WM capacity for objects that contained two feature dimensions: colour and orientation (known as a *conjunction* task). Participants were required to look for a change in either colour or orientation and therefore had to store both dimensions of each object in visual WM in order to perform the task. A display of four coloured bars presented in four different orientations therefore yielded a display that contained eight feature dimensions. If each visual WM chunk is capable of storing only a single feature, performance was expected to dramatically decline when eight features were present in the display. Alternatively, if different features from the same object can be integrated into one chunk by benefit of association with the same discrete object, performance on the conjunction task was expected to remain similar to that on a single feature task. Their results favoured the integrated information hypothesis: despite the increased number of feature dimensions, they found that performance on the conjunction task was just as good as performance on single dimension tasks in which they were required to detect only a colour or orientation change. Luck and Vogel concluded that each visual WM chunk is able to store all the information contained within one object. This is analogous to the storage of one (meaningful) word as one chunk, rather than separate storage of each individual letter contained within that word.

Furthermore, objects defined by up to four conjunctions of colour, orientation, size, and shape (gap versus no gap in the centre of a square), and objects that contained more than one feature on the same feature dimension (i.e., colour-colour conjunction producing bi-coloured squares), were equally well retained in visual WM as single-featured objects, and yielded similar *k* capacity estimates. These findings concretised the concept of an *object-based* visual WM system in which features belonging to the same object are integrated into one visual WM chunk.

Raffone and Wolters (2001) offered support for Luck and Vogel's (1997) notion of integrated visual WM object representations. Via an electrophysiological simulation study, they present a model in which features within an object can be integrated into a single percept in visual WM, and in which non-associated features between objects are separated. They delivered simultaneous input to four neuronal assemblies, each with weak connections to one another and representing part of a multifeature object. They found that the four assemblies were bound into a single chunk of activity via synchronous firing. Raffone and Wolters termed this synchronous activity *within-chunk integration*. They also demonstrated that the onset of neural oscillation that coded for one object feature exerted strong, rapid, and transient inhibition on other neural assemblies that coded for features belonging to other objects. This inhibition, or desynchronisation, can be interpreted to reflect the suppression of irrelevant object features, ensuring the reduction of perceptual interference from other objects that could disrupt feature binding within a selected object. They termed this desynchronisation function *between-item segregation*. Finally, Raffone and Wolters combined within-chunk integration and between-item segregation functions. They provided input to 16 feature assemblies that were integrated into four chunk assemblies, each of which contained four feature assemblies. They demonstrated that activation of three out of the four chunk assemblies could be successfully maintained. Activity within the fourth chunk assembly was suppressed, suggesting that capacity might be nearer three than four items. This finding is consistent with previous *k* capacity estimates for simple stimuli, specifically 3.23 (± 0.83) and 3.40 (± 0.81) (Vogel et al., 2001).

A subsequent study by Wheeler and Treisman (2002), however, questioned the notion of a purely object-based visual WM system. Using the same change detection paradigm with largely similar design parameters, they failed to replicate Luck and Vogel's (1997) finding of equivalent change detection performance for single- versus bi-coloured squares. Instead, they showed that change detection accuracy for squares containing two colours was significantly reduced by about 20% overall relative to squares containing one colour. Similar performance impairments were reported for conjunctions of colour and location, and colour and shape. This has been labelled *feature-based* WM.

In contrast, when they changed the paradigm to involve comparison between a memory array and a single probe, as opposed to between two multi-item arrays, Wheeler and Treisman (2002) found that accuracy in the conjunction conditions was not significantly different to that in the single feature conditions. They suggested that conjunction impairments found when the whole display was presented at test were caused by perceptual interference between items in the shared display – attentional distraction might have impaired the integration, or *binding*, of object features. Wheeler and Treisman concluded that the maintenance of items in visual WM involves two separate mechanisms: (1) a basic capacity-limited storage facility that stores features from any dimension, and (2) a limited capacity attentional resource required for feature binding and the maintenance of bound features. They proposed that bound objects are vulnerable to deconstruction into their constituent parts when interference from neighbouring items reduces the amount of attention available for maintaining a bound representation.

The discrepancy between Luck and Vogel's (1997) object-based findings and Wheeler and Treisman's (2002) feature-based findings (for bi-coloured squares when a whole test display was used) remains unresolved. There are a handful of minor differences between their studies. In Luck and Vogel's study, different groups of participants performed in each single or bi-coloured condition, whereas in Wheeler and Treisman's study the same participants completed all conditions; Wheeler and Treisman provided feedback, whereas Luck and Vogel did not; articulatory suppression involved subvocal repetition of 2 digits with report at the trial's end (Luck and Vogel) or vocal repetition of the word *coca-cola* (Wheeler and Treisman); colours were brighter and presented on a darker background in Luck and Vogel's study; Luck and Vogel presented their sample array for 100 ms, whereas Wheeler and Treisman presented theirs for 150 ms. It is not obvious which of these differences might have caused the discrepancy in findings, if any. Wheeler and Treisman suggested that perhaps the bright colours used in Luck and Vogel's study resulted in perceptual colour blending, producing an impression of purple from a red and blue square, for example. Another possibility is that longer presentation of the sample array in Wheeler and Treisman's study might have allowed for bi-coloured squares to be perceived as comprised of two separable colours – that is, perhaps longer encoding time encouraged featural encoding. The availability of less time to process bi-coloured squares in Luck and Vogel's study might have encouraged global encoding.

Such an explanation would be consistent with the notion of global precedence as defined in the object recognition literature, which posits that processing of global form occurs more rapidly than processing of local features. Evidence of feature binding remains, however, in Wheeler and Treisman's single probe task using colour-location and colour-shape conjunctions, suggesting that time to encode might not have influenced the results in this way. They did not test a colour-colour conjunction using a single probe display, so interpretation of their results from whole displays of bi-coloured squares remains limited.

Alvarez and Cavanagh (2004) provide an alternative theory of how visual WM item capacity is influenced. They suggested that capacity is dependent on the amount of information contained within an object (I refer to this as the *complexity* account). They conducted a visual search task and a change detection task, using stimulus sets of single coloured squares, and multi-featured letters, object line drawings, Chinese characters, polygons, and shaded cubes. Each stimulus category was presented in separate blocks. The visual search task was used to provide a measure of stimulus complexity: Alvarez and Cavanagh proposed that the slower the search, the more complex the stimulus. A target item was presented at the start of each trial for 500 ms, followed by a 900 ms blank interval. An array of 4, 8, or 12 items was then presented, and participants indicated as quickly and accurately as possible whether the target item was present or absent in the array. The change detection task was used to provide a measure of visual WM capacity. They presented a sample array of between 1-15 items for 500 ms, followed by a 900 ms blank retention interval. A whole test array was then presented in which either all items were identical to the test array or one item had changed. Change detection hits and false alarms were used to calculate Pashler's *k*.

Alvarez and Cavanagh (2004) found a significant negative correlation between search rate and capacity: the slower the search, the lower the capacity. In other words, as stimulus complexity increased, visual WM capacity decreased. Capacity estimates were highest for single coloured squares (4.4), followed by letters (3.7), Chinese symbols (2.8), object line drawings (2.6), polygons (2.0), and lastly, 3-D shaded cubes (1.6). Their findings do not provide support for an object-based account of visual WM. If one chunk contains one object, capacity should have been approximately 4 for all stimulus categories tested. Instead, their results can be interpreted as an extension of the feature-based account, in as much as multiple features within each object did not appear to have been integrated into a coherent whole.

To summarise what is known about visual WM so far, upper capacity appears to be fixed at approximately 3-4 *chunks*: capacity for the simplest items, such as single coloured squares, does not appear to be able to exceed this upper limit. Visual WM capacity in terms of *items*, on the other hand, can vary significantly with stimulus complexity. It appears that, if separable features within an object are not integrated to form a whole percept, demand on capacity resources is increased; this increase in demand reduces the number of objects that can be stored. If separable features within an object are integrated within a single chunk, demand on capacity resources is decreased; this decrease in demand increases the number of objects that can be stored.

One particularly notable aspect of the findings described above is that visual WM chunks appear to be flexible resources in which a larger amount of information can be stored if the information is sufficiently integrated. This is an important consideration for my visual WM experiments reported in the next chapter. The notion of visual WM chunks as flexible resources is consistent with a study by Woodman, Vecera, and Luck (2003) in which perceptual grouping of items by proximity and connectivity was shown to influence which items were encoded into visual WM. They presented an attention-capturing cue at the location of one coloured square within an array of five other coloured squares. They found that squares that were close to or connected with the cued object by a boxed outline were more likely to be stored in WM than squares that were distant from or unconnected to the cued object. Integration of information in visual WM therefore appears to be able to occur both *within* objects in terms of feature binding and *between* objects if they are physically associated in some way.

Integration of information on a perceptual basis, such as feature associations within objects and spatial grouping associations between objects, is defined as integration that is driven by *bottom-up* processes. There is also evidence that information integration can be influenced by *top-down* mechanisms, such as LTM, that support strengthened conceptual, semantic, and perceptual associations via familiar representations.

In the final section of this chapter I review evidence for the influence of familiarity on WM. A handful of studies have shown a familiarity advantage in verbal WM, and there is some evidence for a familiarity advantage in visuospatial WM. No study has directly measured the influence of familiarity on visual WM specifically. I also discuss the potential role of LTM in WM processes in more detail.

The Effect of Familiarity on WM

In the verbal WM domain, a series of studies examining WM skills in bilingual participants found that serial recall performance for words was superior in the language of which individuals were most familiar (Chincotta & Hoosain, 1995; Chincotta & Underwood, 1996, 1997; Thorn & Gathercole, 1999, 2001). To account for this familiarity advantage, Thorn, Gathercole, and Frankish (2002) proposed that LTM aids in the reconstruction of decaying short-term memory traces. Bilingual participants were fluent in both primary and secondary languages tested, so LTM stores would unquestionably have existed for both. The superiority of primary language WM capacity therefore indicates a significant distinction between well-learned verbal material and exceptionally familiar verbal material. The nature of this distinction is unknown. Perhaps the familiarity advantage for one's first language is somewhat akin to the robustness of highly familiar face representations (i.e., one's own face) in visual search (Tong & Nakayama, 1999), or the representation for one's own name that appears relatively immune to interference from competing stimuli in the AB (Shapiro, Caldwell et al., 1997). This notion is also consistent with my finding from Experiment 2 that highly familiar faces did not show an AB effect whereas somewhat familiar faces did.

Chase, Simon, and Gobet's research into WM and the game of chess revealed that expert chess players can store more real chess patterns in WM than less skilled players (Chase & Simon, 1973; Gobet, 1998; Gobet & Simon, 1996). Chase and Simon proposed that chess experts are able to retain more chunks of information in WM, i.e., that their WM chunk capacity was expanded. They considered this capacity expansion to be supported by LTM, in which vast numbers of chess pattern representations are able to be stored free of capacity constraints. Alternatively, Gobet and Simon (1996) suggested that increased expert capacity did not reflect a larger number of chunks, but rather reflected larger amounts of information contained within each chunk. They proposed that chess expertise increases the number of conceptual associations between chess pieces and their locations, information about which is contained within each available chunk. Just as it was suggested that LTM aided amnesic patient *PV*'s verbal WM by facilitating chunking, it appears that LTM representations of chess configurations facilitated expert chess players' visual and spatial WM for chess patterns. When false chess patterns were used, expert chess players were no better at the WM task than novices.

To account for the facilitating effect of experience on WM capacity, Ericsson and Kintsch (1995) distinguished between *long-term working memory* (LT-WM) and *short-term working memory* (ST-WM). Based on a theoretical review of expert performance and memory literature, they proposed that LT-WM is a mechanism that enables skilled use of robust and stable representations stored in LTM to support more immediate tasks, information for which is temporarily stored in ST-WM. The basic tenet of their model is that large numbers of to-be-remembered items are rapidly encoded into LT-WM and stored in an associative manner by means of a hierarchical retrieval structure. This retrieval structure is linked to retrieval cues that are retained in ST-WM. As an example, Ericsson and Kintsch cited a study by Chase and Ericsson (1981) in which subject *SF* was able to memorise lists of 30 digits. In this study, it was identified that *SF* grouped, or chunked, the 30 digits into about three different hierarchies: the first level contained mnemonic encoding of digit groups (i.e., sets of about four digits were grouped according to athletic running times: 3596 = 3 minutes and 59.6 seconds); the second and third levels consisted of "supergroups" in which level 1 and level 2 information was associated by other numerical associations and spatial relations. Although somewhat unclear from their paper, Ericsson and Kintsch appear to conceive that information about how digits are organised (i.e., running times etc) is retained and recalled from ST-WM, and these retrieval cues enable sequential retrieval of the digits from LT-WM.

Ericsson and Kintsch's (1995) concept of LT-WM can be interpreted to involve rapid creation of *new* long-term information representations, based on experience in developing such hierarchical retrieval structures. This contrasts with the notion that *existing* LTM representations are used to support WM tasks.

Several researchers have argued that items held in WM are activated representations of existing items retrieved from a LTM store (Cowan, 1998, 2001; Crowder, 1993; Ruchkin et al., 2003). In this sense, WM and LTM representations are thought to be one and the same, but demonstrate different activity states. Ruchkin et al. reviewed electrophysiological and haemodynamic studies of the neural substrate of WM storage and reported findings of increased neural synchrony between prefrontal cortex (PFC) and posterior cortex. Based on evidence that LTM and WM systems are associated with the posterior cortex and PFC respectively, Ruchkin and colleagues proposed that the PFC directs attentional resource to the posterior cortex in order to maintain task-relevant representations. In addition, Ranganath, Johnson, and D'Esposito (2003) found that the same bilateral VLPFC and DLPFC regions were active during a WM task and a LTM task, and concluded that processes involved in both memory tasks shared similar neural pathways.

In summary, familiarity appears to facilitate storage of visuospatial and verbal information within WM, probably supported by LTM. LTM might support the maintenance of information in WM by reconstructing decaying information traces and/or enabling larger amounts of associated information to be integrated into each WM chunk.

In contrast to studies of visuospatial and verbal WM, there is little evidence of a familiarity advantage in *visual* WM specifically. Pashler (1988) examined familiarity effects on visual WM by presenting original (familiar) versus mirror reflected (less familiar) letters in a change detection paradigm. He found no difference in performance between original and reflected letters, suggesting that visual WM (for these verbal stimuli at least) is unaffected by familiarity. Similarly, Olson and Jiang (2004) found no difference in performance on a visual WM task for learned versus novel patterns and suggested that visual LTM does not facilitate visual WM. Why might familiarity effects be found in visuospatial and verbal WM studies but not visual WM studies? There are two possibilities. First, although Pashler showed that identification of reflected letters was more difficult compared to original letters, perhaps this impairment was not dramatic enough to reveal a related impairment in visual WM. Second, perhaps Olson and Jiang's visual pattern learning episode was insufficient to develop as strong and robust a familiar representation as those presumably formed for chess patterns by chess experts, and for first-language words through years of language development.

In the series of experiments reported in Chapter 6, I circumvented the potentially problematic issues present in the above studies and addressed the question of whether familiarity influences visual WM by (1) using stimuli that offer a clear unfamiliar-familiar distinction and (2) using familiar stimuli with which familiarity has been naturally developed over a long period of time. I conducted a series of visual WM experiments using unfamiliar and familiar faces (Experiments 7 and 8) and Hanzi (Experiments 9a and 9b), reported in Chapter 6. In addition to providing the first measure of familiarity effects in visual WM based on natural long-term learning, these experiments also provide the first systematic estimates of visual WM capacity for faces and highly complex non-face objects.

The use of faces stimuli also allowed me to address a more general question about the nature of information storage within visual WM, and extend existing empirical evidence that has been founded on research using relatively simple stimuli. Faces were particularly useful in addressing the debate on whether visual WM chunks contain whole objects or separate features on several counts: they contain multiple features, are highly physically complex, and are thought to involve holistic or configural processes. Based on the assumption that experts make use of more configural information to process Hanzi compared to novices, continued use of Hanzi stimuli is also valuable. Finally, the examination of face and Hanzi inversion effects (Experiments 8 and 9b) further elucidate the object-based versus feature-based debate. Specific hypotheses are outlined at the beginning of each experiment.

CHAPTER 6

THE EFFECT OF FAMILIARITY ON THE MAINTENANCE OF FACES AND COMPLEX NON-FACE OBJECTS IN VISUAL WORKING MEMORY: EXPERIMENTS 7-9

The final series of experiments reported in this chapter aimed to address whether familiarity can enhance the maintenance of information within visual WM. A change detection task identical to that used in Luck and Vogel (1997) was used to measure visual WM capacity. Interpretation of the results is based on the assumption that improved change detection performance reflects increased visual WM capacity, which in turn reflects enhanced maintenance.

To ensure that this task measured differences in visual WM *maintenance* rather than in encoding (or *consolidation*) speed between unfamiliar and familiar stimuli, sample displays that contained items to be memorised were presented for a sufficient length of time to allow effective perceptual encoding. The same participants were used in the visual WM experiments as were used in the corresponding visual search experiments. I purposely used the same groups so that I could set each individual's visual WM memory display duration according to RTs obtained in their related visual search experiment. In doing so, I was able to ensure not only that sufficient encoding time was provided but also that this was tailored to each participant, thus reducing any individual differences in encoding speed that might have intruded into the experiment had the same encoding time been given to all individuals. Each participant's mean RT to make a correct visual search response at the largest relevant set size was used to set the duration of the memory array in his/her succeeding WM task.

Visual WM for unfamiliar and famous upright faces was compared in Experiment 7; unfamiliar and famous inverted faces were compared in Experiment 8. Expert versus novice visual WM was compared using upright Hanzi (Experiment 9a) and inverted Hanzi (Experiment 9b). The examination of inverted stimuli allowed me to address the contribution of featural and configural processes to the way in which information is maintained within visual WM.

Face and Hanzi experiments are reported in two separate sections below, labelled accordingly.

Visual WM for Faces: Experiments 7 and 8[14]

Experiment 7 measured visual WM capacity for upright unfamiliar and famous faces; I also included a condition in which single coloured squares were used as stimuli in order to obtain a measure of capacity for single-featured items to which faces could be compared. Experiment 8 measured visual WM capacity for inverted unfamiliar and famous faces. A different group of participants completed each experiment. Specific hypotheses are presented at the start of each experiment section. Similar procedures were used in both experiments, therefore a General Methods section is provided below. (Note that specific methodological information about the squares condition is provided later in the section that reports Experiment 7.)

---

[14] Experiments 7 and 8 were presented as a poster at the Vision Sciences (VSS) Conference in May 2004, Florida. [Jackson, M. C. & Raymond, J. E. (2004). Visual Working Memory for Faces [Abstract], *Journal of Vision*, 4(8), 394a.]

General Methods

Apparatus

Stimuli were displayed on a 22-inch Mitsubishi DiamondPro 2060u monitor (32-bit true colour; resolution 1280 x 1024 pixels) and generated by E-Prime software (Version 1.0; Schneider et al., 2002). Viewing distance was 70 cm.

Stimuli

Subsets of 11 unfamiliar and 11 famous male faces were selected from the sets of 18 unfamiliar distractor faces and 18 famous distractor faces used in the visual search experiments (see Appendix A). In selecting the famous faces for this task I was careful to avoid possible pre-existing semantic connections, such as appearance in the same film or television programme, that could have led to semantic grouping effects. Faces subtended 2.8° x 3.3° and were displayed at random locations within a white region (approximately 17.0° x 14.4°), with the constraint that each was separated by at least 2.9° on the horizontal axis and 3.4° on the vertical axis, centre to centre.

Design

A standard visual change detection task (Luck & Vogel, 1997) with concurrent verbal memory load was used to measure visual WM capacity. Set sizes 1, 2, 3, 4, 5, 6, 8, and 10 were used for upright faces in Experiment 7; set sizes 2, 4, and 6 were used for inverted faces in Experiment 8. No face appeared more than once in any given display. There were 40 trials per set size (50% change), yielding 320 trials in Experiment 7 and 120 trials in Experiment 8. Set size and change/no change conditions were pseudo-randomised within each experiment. Participants in each experiment completed two sessions, one with unfamiliar faces and the other with famous faces. Each session was completed on a different day and session order was counterbalanced.

Procedure

Each WM trial was initiated by pressing the space bar and began with a 500 ms central presentation of two digits (selected at random from 0-9; font size 50). To suppress the use of verbal WM, participants were instructed to silently repeat these digits throughout the whole trial. After a blank interval of 1000 ms, the memory array of faces was presented. Memory array duration varied between experiments and across participants, but within a particular experiment and for a given participant the same array duration was used throughout. The range of durations used in each experiment is provided in the relevant procedure section of each experiment. The memory array was followed by a 900 ms blank retention interval. A test array of faces was then presented that was either identical to the memory array (no-change trial) or in which a single face had changed to a different face (change trial). Test array duration was one and a half times the memory array duration. No face changed location between memory and test arrays within a particular trial. After the test array disappeared, participants reported whether the memory and test arrays were the same or different (using key K labelled "same" and key S labelled "different"). They were then prompted to type the two digits that had been presented at the trial's start. All responses were unspeeded and no feedback was provided. A short practice session was provided before the main experiment began. Figure 18 illustrates a typical trial.

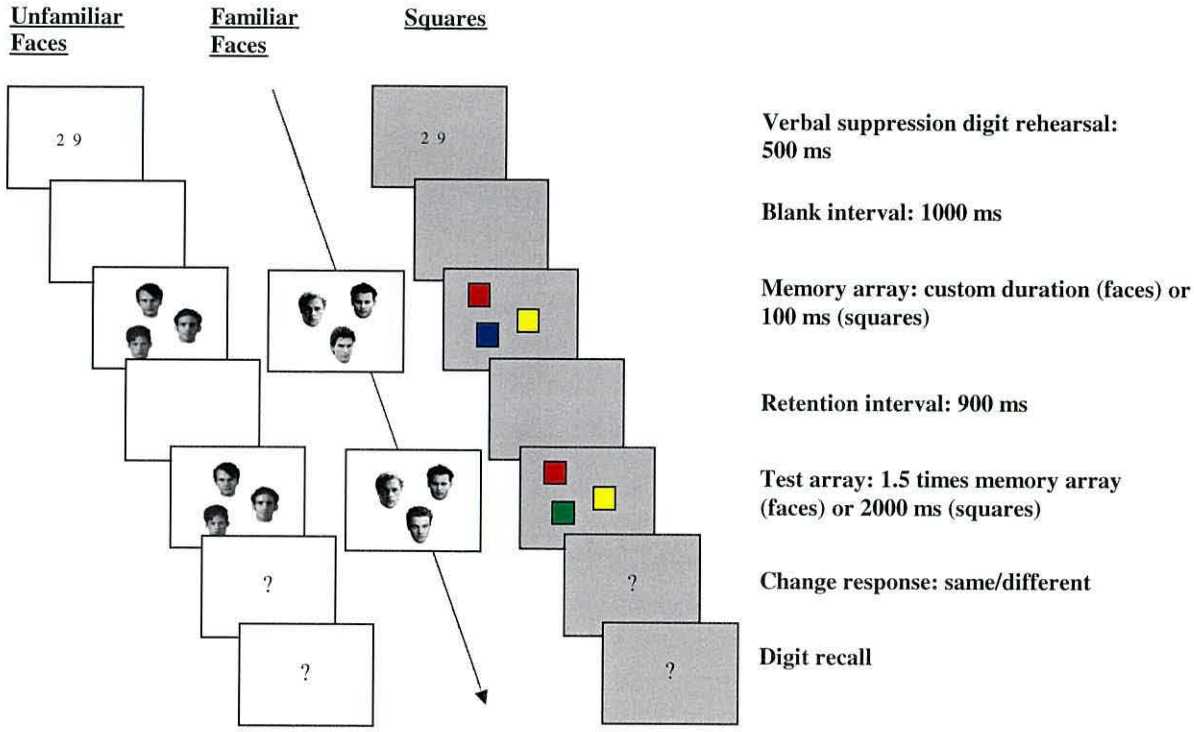


Figure 18. Example visual change detection trial procedure used to measure visual WM capacity. Experiment 7 used upright unfamiliar faces, upright famous faces, and single coloured squares (depicted here). Experiment 8 used inverted unfamiliar and inverted famous faces. Experiments 9a and 9b used upright and inverted Hanzi respectively and compared Hanzi experts to Hanzi novices. Face and Hanzi memory array durations were set according to each participant's mean RT to make a correct search response at the largest set size in the visual search task.

Data Analysis

Only trials on which participants correctly reported both digits in the verbal suppression task were included in the analyses. The percentage of incorrect digit response trials that were removed from each set size and familiarity condition in each experiment can be seen in Tables D1 and D2 in Appendix D.

False alarm (FA) rates for the WM change detection task varied significantly as a function of set size for all stimulus conditions across both experiments, indicating the need to convert proportion correct detection of change trials (hits) into *d'* (dprime) values. (Figures and statistics on the effect of set size on FA rate for all visual WM experiments in this chapter are provided in Appendix E.) Individual hit and FA rates for each set size and stimulus condition were used to calculate d' values. These data were then subjected to repeated-measures ANOVA, and post-hoc tests using Bonferroni corrections for multiple comparisons (where applicable).

In addition to reporting change detection performance, I also report *k* capacity estimates. The procedure for estimating *k* is presented in a separate section below. It warrants separate attention for two reasons: (1) I modified the application of previous *k* capacity formulae, and (2) I present a new, alternative formula to calculate capacity, which I call the *Bangor k*.[15]

Capacity estimate k.

As mentioned in Chapter 5, Luck and Vogel (1997) and Vogel et al. (2001) obtained visual WM capacity estimates by applying Pashler's *k* formula: **k = [S x (H–F)] / (1–F)**, where S = set size, H = hits (the proportion of change trials on which participants correctly stated that a change had occurred), and F = false alarm rate (the proportion of no change trials on which participants incorrectly stated that a change

[15] I developed the Bangor *k* formula in collaboration with my principal PhD supervisor Professor Jane Raymond.

had occurred). Cowan (2001) modified this formula to the following: **k = S (H–F)**. This has become known as Cowan's *k*. It is unclear which of these formulae is the more accurate measure of capacity.

Modified application of previous k formulae.

In calculating Pashler's *k* and Cowan's *k*, previous studies tended to use data from the largest set size that was presented. This seems reasonable if performance at large set sizes is sufficiently above chance. However, my change detection data for upright unfamiliar faces at the largest set size (i.e., set size 10) revealed near chance performance, and suggested that the use of such data points to calculate *k* might have resulted in under-estimation of capacity. Indeed, when I calculated *k* at each set size, within many of my visual WM experiments I found an inverted U-shaped function: *k* decreased as set size increased. I therefore used data from a lower set size – set size 6 – to calculate Pashler's *k* and Cowan's *k*. Although performance and *k* values were lower at set size 6 than at smaller set sizes, the benefit of using set size 6 is that this set size was common to all visual WM experiments conducted and I was therefore able to draw direct comparisons between different conditions.[16]

[16] In addition, I considered it useful to use set size 6 (as opposed to set size 4, for example) in order that Pashler's *k* and Cowan's *k* would be relatively comparable to the Bangor *k* formula that used a range of set sizes in which set size 6 was the maximum (see Appendix F).

New k formula: Bangor k.

The newly revised formula that I developed to measure visual WM is:

$$k = 10^{[(y-b)/m]}$$

where y = a threshold performance constant expressed in d' units (set arbitrarily here to a value of 3.45), and b and m are the intercept and slope respectively of a least squares line fit to d' values (from the change detection task) plotted as a function of logarithmic set size. Specific details of the procedures I used to develop Bangor k and apply it to my data are provided in Appendix F.

There are two key advantages of the Bangor k. First, it can be based on most or all of the data collected (see Appendix F for data selection criteria). Pashler's k and Cowan's k use just one set size to calculate capacity and thus fail to incorporate most of the data collected within the change detection task. Second, it allows for variable false alarm rates to be incorporated into capacity estimates. As mentioned above, I found that false alarm rates varied significantly as a function of set size (Appendix E). Pashler's k and Cowan's k use a false alarm rate from only one set size and therefore provide a crude measure of response bias.

For comparison and interest, in all subsequent visual WM studies I report Pashler's k, Cowan's k, and the new Bangor k formulae; I used paired- or independent-samples t-tests to compare k values between conditions of interest. I present only p values in the results sections – detailed statistics from comparisons involving each of the three capacity estimate measures in all experiments are provided in Appendix G.

Experiment 7. Visual WM for Upright Unfamiliar and Famous Faces

The primary aim of this experiment was to determine whether maintenance of familiar faces in visual WM is enhanced relative to unfamiliar faces, reflected by higher capacity. In one session, all faces were unfamiliar. In another session, all faces were famous. If familiarity increases capacity, significantly better change detection performance and higher capacity estimates for famous compared to unfamiliar faces are expected.

The secondary aim of this experiment was to examine the more general question of how information is stored within visual WM. Luck and Vogel (1997) proposed that visual WM stores whole objects. Wheeler and Treisman (2002) extended this suggestion; they proposed that only if all the information contained within an object is sufficiently bound could whole objects be stored in visual WM. Based on this idea I propose that capacity might reflect the extent of object binding, or what I call *within-item integration*. Note that this concept is similar to Raffone and Wolters's (2002) notion of within-chunk integration, a term they used to describe the integration of separate object features into one WM chunk. In addition to containing multiple features from one object, however, visual WM chunks appear to be able to contain more than one individual object. For example, objects that are in close proximity or connected to one another tend to be grouped into one chunk (Woodman et al., 2003). As such, within-chunk integration could be interpreted to reflect binding within or between objects. Therefore, I introduce the term within-item integration to specifically refer to binding of information within, not between, objects.

In addition, although the term within-item integration is essentially the same as the concept of binding, the term binding tends to be defined with regards to the connection between individual *features* and is commonly called *feature binding* (e.g., Treisman, 1992; Treisman & Gelade, 1980). The concept of *configural* relationships between features, that is their *spatial* relationships, appears to be conceived as separate to that of feature binding. Luck and Beach (1998) highlighted the limitation of feature binding theory in its lack of account for how objects that share the same features and differ only in their spatial configuration can be distinguished from one another; they termed this the *relationship problem*. I therefore use the term within-item integration to describe the way in which features *and* their configurations might be bound to create a coherent percept.

Alvarez and Cavanagh (2004) provided evidence of whole object-based storage for only the simplest of items. They suggested that visual WM capacity is dependent on the amount of information contained within an object, i.e., stimulus complexity. Perhaps full integration of multiple details within an object is not possible if the amount of detail is large.

How might facial information be organised in visual WM? Given that face identification is widely believed to recruit configural processes that span most, if not all, of the facial region (e.g., Tanaka & Farah, 1993; Bartlett et al., 2003), it is reasonable to imagine a high extent of integration among features and their configurations within each face. If integration is absolute, each face might be stored as a single chunk and capacity for around 3-4 faces is expected. Such a finding would support an object-based account of visual WM storage. Alternatively, if integration among face features and their configurations is not absolute, capacity for faces is expected to be less than 3-4. Based on Alvarez and Cavanagh's (2004) findings, a third alternative is proposed: if visual WM is dependent upon stimulus complexity, the vast amount of detail in a face might limit the extent of within-item integration and reveal a capacity *dramatically* lower than 3-4 items. Perhaps only one face can be maintained?

To introduce a simple stimulus measure with which faces could be directly compared, all participants completed a change detection experiment that used single coloured squares (outlined in the Methods section below). Each square was considered to represent a single chunk of information.

Methods

Participants.

The same 24 participants (16 females, 8 males; mean age 22 years) that completed Experiment 4 (visual search for upright unfamiliar and famous faces) completed this experiment. Informed consent was obtained prior to participation.

Stimuli.

All unfamiliar and famous faces were presented in their upright orientation. In addition to unfamiliar and famous faces, single coloured squares were used in a squares change detection session. Eleven filled squares (0.6° x 0.6°) each coloured with a single hue (white, black, red, green, blue, yellow, orange, brown, purple, turquoise, and pink) were displayed at random locations within a light grey region (approximately 9.8° x 7.3°), with the constraint that each square was separated by at least 2° centre to centre.

Design and procedure.

Each participant completed three sessions: squares, unfamiliar faces, and famous faces. Each session was completed on a different day and session order was fully counterbalanced across participants. The faces sessions proceeded as described in the General Methods section above. The squares session proceeded in an identical manner: no colour appeared more than once in any given display; set sizes 1, 2, 3, 4, 5, 6, 8, and 10 were used; there were 40 trials per set size (50% change), yielding 320 trials; set size and change/no change conditions were pseudo-randomised. Consistent with Luck and Vogel's (1997) study, each memory array of squares was presented for 100 ms, followed by a 900 ms blank retention interval, followed by a 2000 ms presentation of the squares test array. Across participants the memory array durations in the unfamiliar faces session ranged from 1307 ms to 3232 ms; the memory array durations in the famous faces session ranged from 1350 ms to 3031 ms.

In this and all subsequent visual WM experiments, there were non-significant effects of session order.

Results and Discussion

Figure 19a below shows the change detection results for squares, famous faces, and unfamiliar faces. Figure 19b shows Pashler's *k*, Cowan's *k*, and Bangor *k* values for each of the three stimulus types. Change detection performance and capacity estimates clearly varied among stimulus types and across set sizes. I report the results in two subsections: the first deals with the comparison between squares and faces, addressing the general issue of how information is organised within visual WM; the second deals with the comparison between unfamiliar and famous faces, addressing the primary question of whether familiarity can enhance maintenance in visual WM.

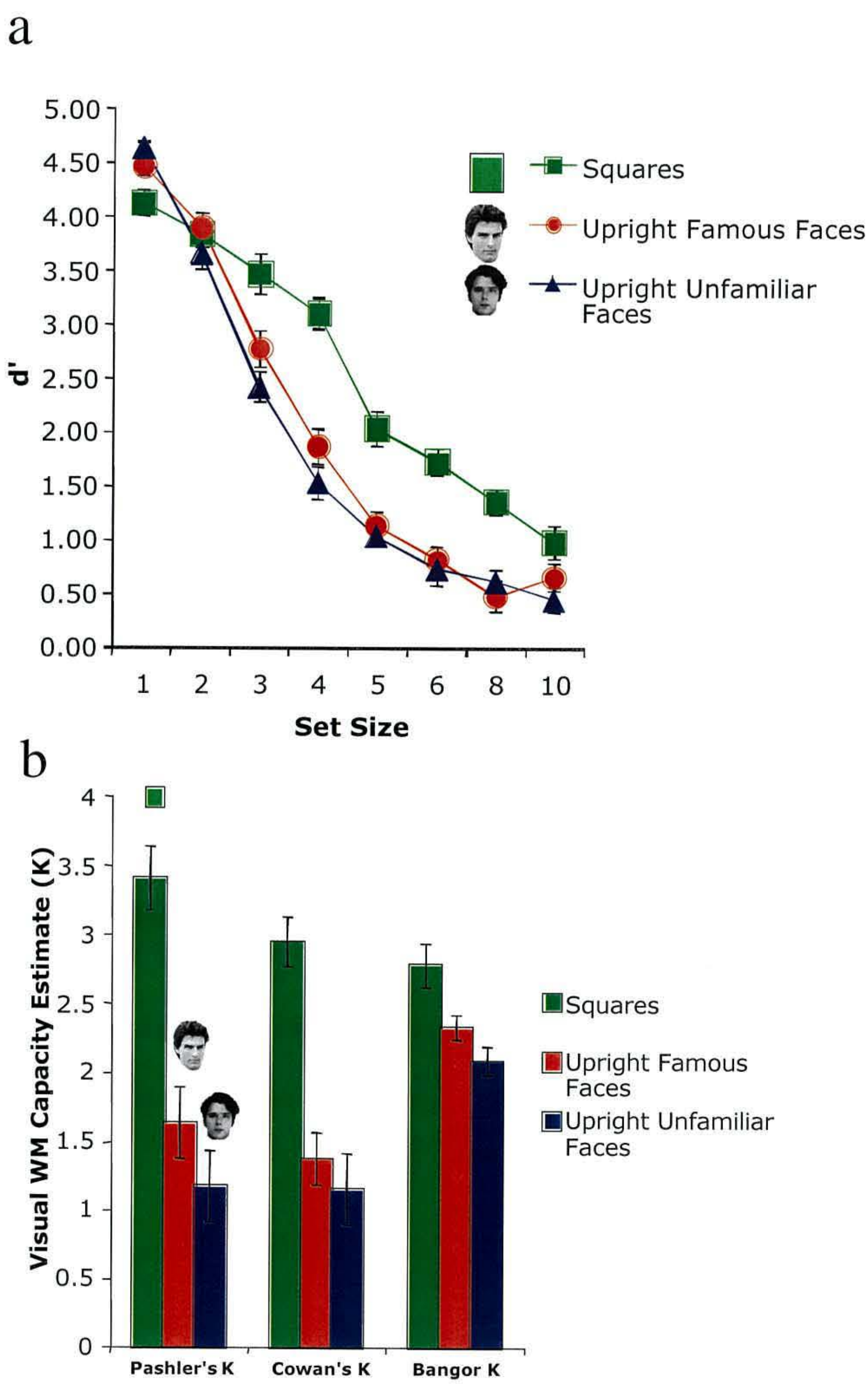


Figure 19. Results for Experiment 7: (a) Change detection performance (*d'*) as a function of set size for squares, upright famous faces, and upright unfamiliar faces. (b) Visual WM capacity estimates using Pashler's *k*, Cowan's *k*, and Bangor *k* for squares, upright famous faces, and upright unfamiliar faces.

Faces versus squares.

Figure 19a shows that the ability to detect a change between memory and test arrays was markedly better for squares than faces. This was supported by a significant stimulus by set size interaction, $F(14, 322) = 8.96, p < .01$, and post-hoc comparisons showed greater sensitivity for squares than unfamiliar or famous faces ($p < .01$ in both cases).

As can be seen in Figure 19b, capacity estimates ranged from 2.77 to 3.40 squares, consistent with previous visual WM capacity estimates of around 3-4 single coloured squares (Vogel et al., 2001). Capacity estimates for both unfamiliar and famous faces, however, were markedly and significantly lower than for squares for all capacity measures employed ($p < .05$ in all cases). On average, only about 2 faces could be maintained in visual WM. The finding that change detection performance and related capacity estimates were significantly lower for faces than for squares suggests that faces were not maintained as whole, maximally integrated items in visual WM. Thus, support for object-based storage of information (Luck & Vogel, 1997) is not provided using face stimuli. In addition, the results are inconsistent with Alvarez and Cavanagh's (2004) complexity account. Capacity for both unfamiliar and famous faces was greater than one. These face capacity estimates are similar to Alvarez and Cavanagh's estimate for polygons (2.0) and 3-D shaded cubes (1.6), stimuli that unarguably contain less visual detail than faces. I therefore cannot find support for the notion that visual WM storage limits are dramatically taxed by stimulus complexity. Instead, I interpret these results to suggest that information contained with each face appears to be integrated in visual WM to some extent, but not fully.

Unfamiliar versus famous upright faces.

It is clear from Figure 19a that the ability to detect a change between memory and test arrays was better for famous than unfamiliar faces. A repeated-measures ANOVA using familiarity (unfamiliar, famous) and set size (1, 2, 3, 4, 5, 6, 8, and 10) as within factors revealed a significant main effect of familiarity, $F(1, 23) = 7.55, p = .01$. Although the interaction between familiarity and set size was non-significant, $p > .2$, it is interesting to note that the most obvious performance advantage for famous over unfamiliar faces occurred between set sizes 2-4. From set size 5 onwards the familiarity advantage disappeared. Support for these observations was found when data from set sizes 2-4 and 5-10 were analysed separately. A significant main effect of familiarity between set sizes 2-4, $F(1, 23) = 10.32, p < .01$,[17] and a non-significant main effect of familiarity between set sizes 5-10, $F(1, 23) < 1$, was revealed. Interactions between familiarity and set size were non-significant in each comparison condition.

As can be seen in Figure 19b, all capacity measures revealed that more famous than unfamiliar faces could be maintained in visual WM. Pashler's *k* showed that capacity for famous faces (1.63) was 39% larger than that for unfamiliar faces (1.17) (*NS*).[18] Cowan's *k* showed that capacity for famous faces (1.37) was 19% larger than that for unfamiliar faces (1.15) (*NS*). Bangor *k* showed that capacity for

[17] An analysis of set sizes 1-4 also revealed a significant main effect of familiarity, $F(1, 23) = 6.60, p < .05$.

[18] *NS* denotes a non-significant result.

famous faces (2.32) was 12% larger than that for unfamiliar faces (2.08), a significant difference, $p < .05$.

Overall, these results are consistent with the notion that familiarity might enhance the maintenance of representations in visual WM. The lack of AB for famous faces found in Experiments 2 and 3 might be partly due to this advantage in visual WM. Support for this idea can be found in verbal WM studies that reported larger word span for first versus second language words among bilingual participants (e.g., Thorn et al., 2002), and in visuospatial WM studies that found chess experts could remember significantly more chess patterns than chess novices (Chase & Simon, 1973; Gobet, 1998; Gobet & Simon, 1996).

Thorn et al. (2002) proposed that LTM might support the maintenance of information in WM by reconstructing decaying information traces. Consistent with this idea, studies have suggested WM and LTM might be connected in some way, either by shared neural pathways (Ranganath et al., 2003) or as different states of the same representation (e.g., Ruchkin et al., 2003). Alternatively, Gobet and Simon (1996) proposed that stable long-term information about visual and spatial associations between items, such as is required for the memory of chess patterns, could be used to increase the amount of information that is contained within one WM chunk. This could be interpreted to mean that familiarity acts to increase within-chunk integration *between* separate objects or information sources. Because I attempted to ensure that long-term associations between the famous faces used in the change detection arrays could not be made (i.e., I avoided using celebrities who commonly appeared in the same film or television programme), such between-item integration appears unlikely to account for enhanced maintenance in this study.

Instead, I suggest that if visual WM capacity is dependent on the extent of within-item integration, larger visual WM capacity for famous than unfamiliar faces accrues because information contained within familiar faces is more integrated than that contained within unfamiliar faces. I return to this suggestion in more detail in the Chapter Discussion.

Experiment 8. Visual WM for Inverted Unfamiliar and Famous Faces

The aim of this experiment was to use inverted faces to assess how featural and configural processes might contribute to the storage of familiar and unfamiliar information within visual WM. In one session, all faces were unfamiliar. In another session, all faces were famous. Recall that visual search was significantly more efficient for famous than unfamiliar faces, even when the faces were inverted (Experiment 5). I suggested that processes used to identify familiar faces might be able to recruit an orientation-independent mechanism; this mechanism might exhibit better cooperation between featural and configural processes in that one route can compensate visual processing when the other becomes unavailable. Based on this notion, significantly better change detection performance and higher capacity estimates were expected for inverted famous compared to inverted unfamiliar faces.

As mentioned in previous chapters, inversion is considered to disrupt configural processing and enforce greater use of featural information. In this sense, information contained within an inverted face might be less integrated than information contained within an upright face. Based on this assumption, I expected change detection performance to be poorer and visual WM capacity to be smaller for inverted compared to upright faces. Because both unfamiliar and familiar face identification is thought to involved configural processes, I expected this inversion effect to be observed for both unfamiliar and famous faces. Inversion effects were examined by comparing results from this experiment with those obtained from Experiment 7 (a between-group analysis).

Methods

Participants.

The same 24 participants (21 females, 3 males; mean age 21 years) that completed Experiment 5 (visual search for inverted unfamiliar and famous faces) completed this experiment. Informed consent was obtained prior to participation.

Stimuli, design, and procedure.

The stimuli were the same 11 faces used in Experiment 7, but they were rotated by 180°. Across participants the range of memory array durations for unfamiliar faces was 1188 ms to 2219 ms; the range of memory array durations for famous faces was 972 ms to 1972 ms. Set sizes 2, 4, and 6 were used; there were 40 trials per set size (50% change), yielding 120 trials; set size and change/no change conditions were pseudo-randomised. All other design and procedure parameters were as outlined in the General Methods.

Results and Discussion

I report the results in two subsections: the first deals with the comparison between unfamiliar and famous inverted faces in the current experiment; the second deals with inversion effects by reporting a between-experiment comparison of upright versus inverted unfamiliar and famous faces.

Unfamiliar versus famous inverted faces.

Figure 20a shows the change detection results for unfamiliar and famous inverted faces. A repeated-measures ANOVA using familiarity and set size as within factors revealed a significant main effect of familiarity, $F(1, 23) = 4.61, p < .05$. The interaction between face type and set size was non-significant, $p > .6$. The most obvious performance advantage for famous over unfamiliar faces occurred at set size 6. A paired samples *t*-test revealed this difference to be marginally significant, $t(23) = 1.99, p = .06$. This is somewhat surprising, given that for upright faces there was no performance difference for famous versus unfamiliar faces between set sizes 5-10. It is difficult to interpret why a familiarity advantage for inverted faces was observed at set size 6 only. If visual WM capacity for inverted famous faces were larger than for inverted unfamiliar faces, better change detection performance would be expected at smaller set sizes – contrary to this, an ANOVA using only set sizes 2 and 4 revealed a non-significant difference between unfamiliar and famous inverted faces, $F(1, 23) = 1.04, p > .3$. Perhaps the observed difference at set size 6 is simply an artefact.

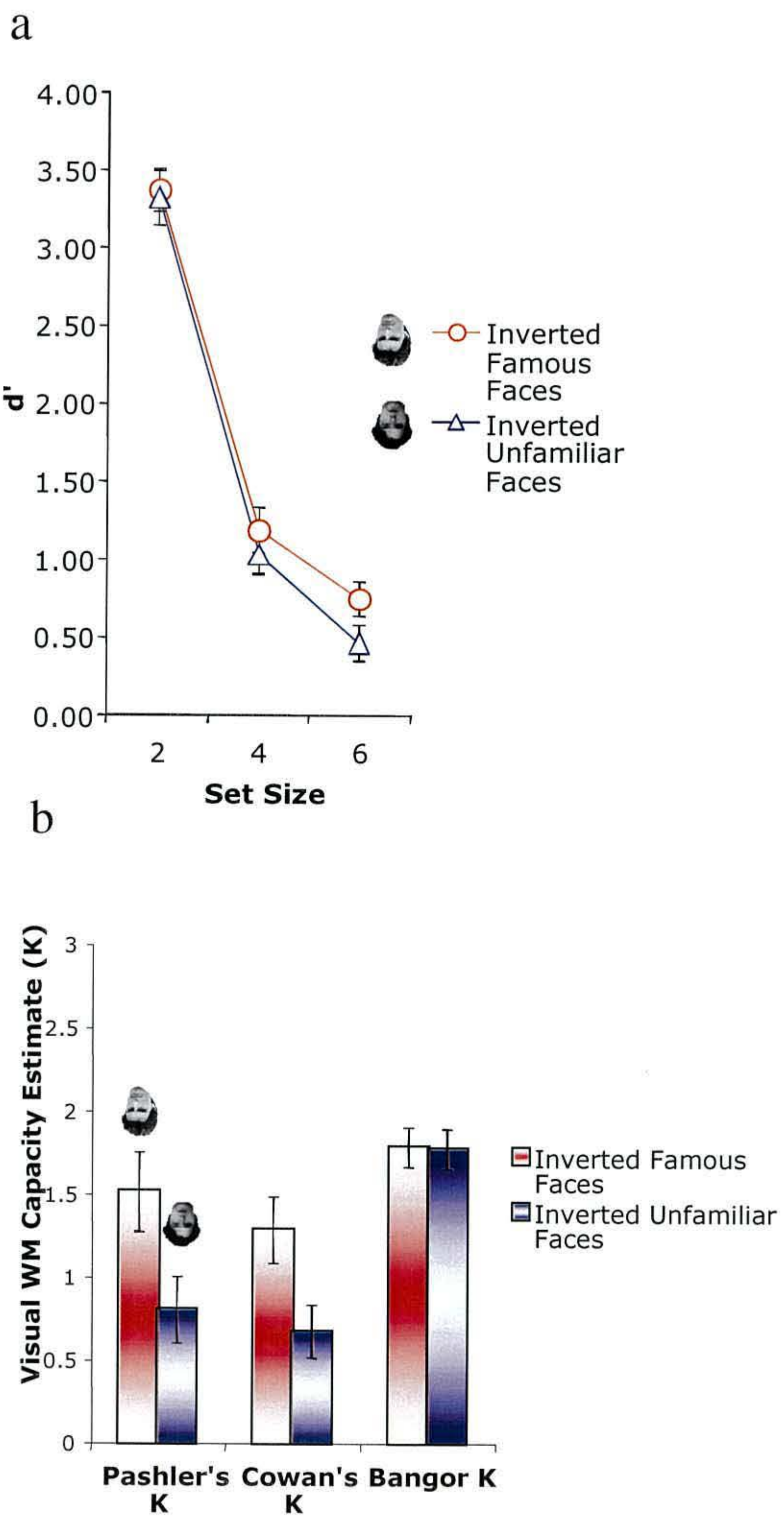


Figure 20. Results for Experiment 8: (a) Change detection performance (*d'*) as a function of set size for inverted famous faces and inverted unfamiliar faces. (b) Visual WM capacity estimates using Pashler's *k*, Cowan's *k*, and Bangor *k* for inverted famous faces and inverted unfamiliar faces.

As can be seen in Figure 20b, while Pashler's *k* and Cowan's *k* revealed higher capacity estimates for famous than unfamiliar inverted faces, Bangor *k* revealed equivalent capacity estimates for famous than unfamiliar inverted faces. Pashler's *k* showed that capacity for famous faces (1.51) was 89% larger than that for unfamiliar faces (0.80), $p = .05$. Cowan's *k* showed that capacity for famous faces (1.28) was 91% larger than that for unfamiliar faces (0.67), $p < .05$. Bangor *k*, however, showed that capacity for famous faces (1.78) was similar to that for unfamiliar faces (1.77) *(NS)*.

The disparity between Bangor *k* and the other two measures is likely to have been driven by the use of set size 6 to calculate Pashler's *k* and Cowan's *k* formulae, the only set size at which change detection performance differed for unfamiliar and famous faces. If this performance difference at set size 6 were an artefact, then Pashler's *k* and Cowan's *k* capacity estimates might be misleading. This highlights one of the key advantages of the Bangor *k* formula: it uses a wider range of set sizes to calculate a capacity estimate and thus minimises the potential for errors that might arise from employing a capacity measure based on only one particular portion of the data.

The most parsimonious interpretation of these results is that maintenance of inverted famous faces is not enhanced compared to inverted unfamiliar faces. This finding is inconsistent, however, with the visual search advantage found for famous over unfamiliar inverted faces in Experiment 5. The presence of a familiarity advantage for inverted faces in visual search and the absence of a familiarity advantage in visual WM suggests that, while familiar face identification processes might be able to make use of an orientation-independent mechanism during perceptual encoding of inverted faces (perhaps involving the transfer of featural into configural information), processes involved in the *maintenance* of inverted familiar faces in visual WM appear to lack access to this mechanism. Processes underpinning maintenance in visual WM therefore appear to be similar for familiar and unfamiliar inverted faces.

In summary, inversion appears to have removed the familiarity advantage revealed for upright faces in Experiment 7. I interpret the lack of a familiarity advantage for inverted faces to suggest that the extent of within-item integration is equivalent in unfamiliar and familiar inverted faces.

Inversion effects: Experiments 7 and 8 compared.

Figure 21 illustrates upright versus inverted change detection performance for unfamiliar (Figure 21a) and famous faces (Figure 21b). Mixed design repeated-measures ANOVA with orientation (upright, inverted) as a between factor and set size (2, 4, and 6) as a within factor revealed that change detection performance was significantly poorer for inverted than upright faces in both unfamiliar, $F(1, 46) = 6.75$, $p < .05$, and famous, $F(1, 46) = 11.23$, $p < .01$, face conditions. For unfamiliar faces the orientation by set size interaction was non-significant, $p > .6$. For famous faces the orientation by set size interaction was significant, $F(2, 92) = 3.63$, $p < .05$.

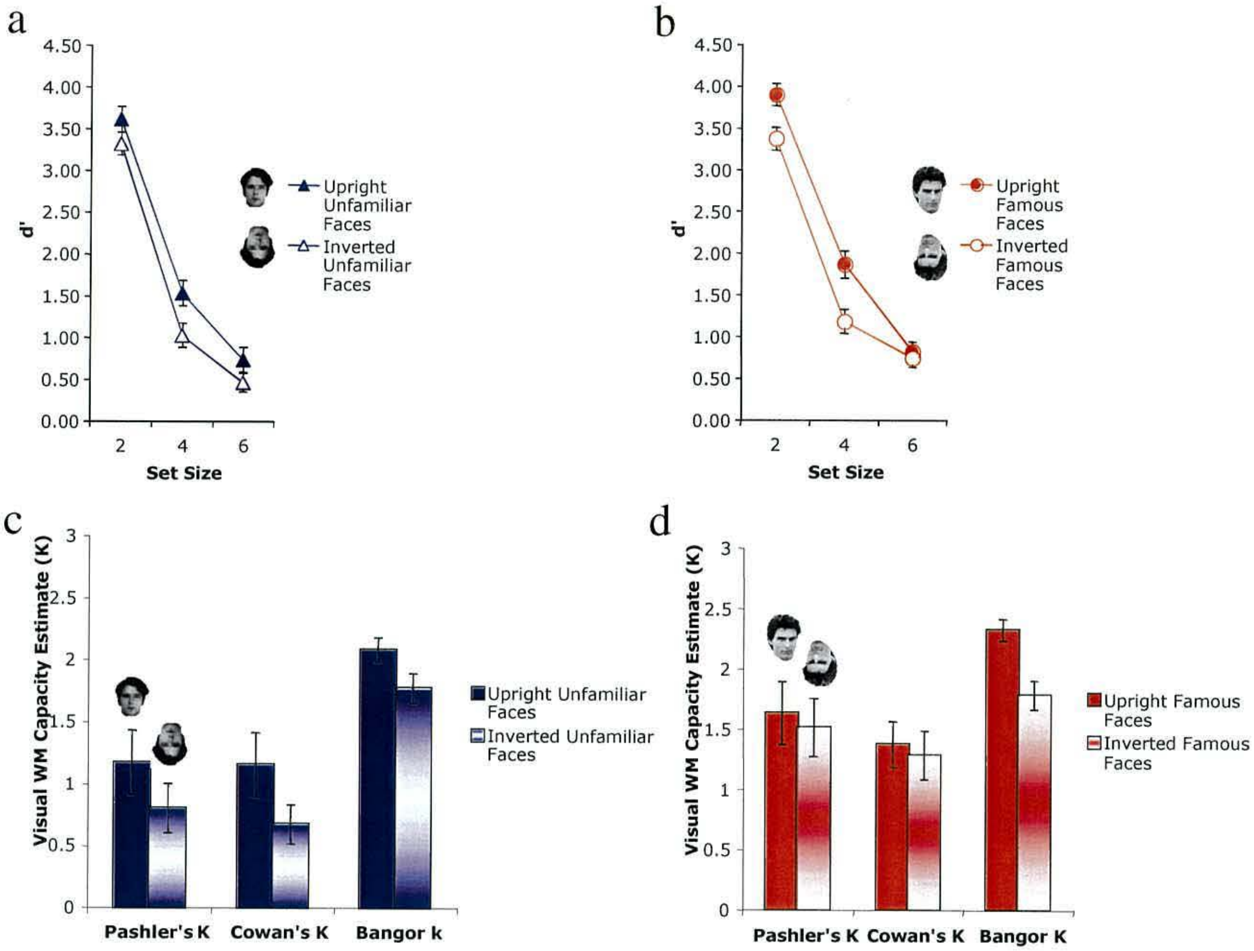


Figure 21. Face inversion effects (Experiments 7 and 8 compared): Change detection performance (*d'*) as a function of set size for (a) upright unfamiliar versus inverted unfamiliar faces, and (b) upright famous versus inverted famous faces. Visual WM capacity estimates using Pashler's *k*, Cowan's *k*, and Bangor *k* for (c) upright unfamiliar versus inverted unfamiliar faces, and (d) upright famous versus inverted famous faces.

Inversion reduced all capacity estimates for both unfamiliar (Figure 21c) and famous (Figure 21d) faces. Using Pashler's *k*, capacity was reduced by 32% from 1.17 upright unfamiliar faces to 0.80 inverted unfamiliar faces (*NS*); capacity was reduced by 7% from 1.63 upright famous faces to 1.51 inverted famous faces (*NS*). Using Cowan's *k*, capacity was reduced by 42% from 1.15 upright unfamiliar faces to

0.67 inverted unfamiliar faces (*NS*); capacity was reduced by 7% from 1.37 upright famous faces to 1.28 inverted famous faces (*NS*). Using the Bangor *k* formula, capacity was reduced by 15% from 2.08 upright unfamiliar faces to 1.77 inverted unfamiliar faces, a significant difference, $p < .05$; capacity was reduced by 23% from 2.32 upright famous faces to 1.78 inverted famous faces, a significant difference, $p < .01$.

One interpretation of impaired change detection performance and lowered capacity estimates for inverted versus upright faces might involve differences in the level of support provided from LTM. Famous faces in LTM might be more robustly represented in upright than inverted views because faces are more frequently seen in their upright, canonical orientation, than upside down. This view would imply that multiple representations of a face in all its encountered orientations are stored in LTM (i.e., template-matching theory, as discussed in the object recognition section in Chapter 1). But there is little evidence for this. Instead, I interpret the inversion effects to suggest that inversion resulted in the deconstruction of facial information from relatively well integrated to poorly integrated configurations of features, and that this placed greater demands on visual WM capacity resources.

A final point to note is that inversion effects were obtained using the same sets of faces in both upright and inverted face experiments, so it can also be concluded that face processing mechanisms were used to detect changes in the WM task, rather than low-level image artefacts such as luminance or contrast.

Visual WM for Hanzi: Experiments 9a and 9b[19]

In Chapter 4, I showed that the familiarity advantage in visual search for upright faces also applied to upright Hanzi (complex non-face objects): search was significantly more efficient for famous than unfamiliar faces, and among Hanzi experts than Hanzi novices. When stimuli were inverted, the familiarity search advantage remained for faces but was absent for Hanzi. I concluded that, unlike familiar face processing, processes involved in expert Hanzi recognition might not have access to an orientation-independent mechanism.

The aim of the two experiments reported in this current section was threefold. First, I wanted to determine whether the visual WM familiarity advantage found for upright faces could also apply to upright Hanzi. Second, I wanted to determine whether the absence of a familiarity advantage (and therefore the absence of an orientation-independent mechanism) found for inverted Hanzi in the visual search task could be evidenced in a visual WM task. Third, I wanted to compare the effect of inversion for Hanzi experts and Hanzi novices.

Experiment 9a compared expert and novice change detection performance for upright Hanzi. Experiment 9b compared expert and novice change detection performance for inverted Hanzi. Similar methods were used in both experiments, therefore a General Methods section is provided below. (Unless stated otherwise these General Methods are as outlined in the visual WM for faces section above.)

[19] Experiments 9a and 9b were presented as a poster at the Vision Sciences (VSS) conference in May 2005, Florida. [Jackson, M. C., & Raymond, J. E. (2005). Visual Working Memory: Capacity is Dependent on Perceived, not Physical, Stimulus Complexity [Abstract]. *Journal of Vision*, *5*(8), 621a.]

General Methods

Stimuli and Design

A subset of 9 Hanzi was selected from the set of 18 Hanzi distractors used in the visual search experiments (see Appendix C). Hanzi subtended approximately 2.5° by 2.5° of visual angle, and were randomly displayed within a white background region subtending approximately 17.0° x 14.4°, with the constraint that each was separated by at least 2.7° centre to centre. The change detection task was identical to that used for faces; concurrent verbal suppression was administered as before. Set sizes 1, 2, 3, 4, 5, 6, and 8 were used for upright Hanzi in Experiment 9a; set sizes 2, 4, and 6 were used for inverted Hanzi in Experiment 9b. No Hanzi appeared more than once in any given display. There were 40 trials per set size (50% change), yielding 280 trials in Experiment 9a and 120 trials in Experiment 9b. Set size and change/no change conditions were pseudo-randomised within each experiment.

Participants in each experiment completed two sessions, one with upright Hanzi and the other with inverted Hanzi. Each session was completed on a different day and session order was counterbalanced.

Procedure

The change detection task for Hanzi was identical to that used for faces in Experiments 7 and 8 above (refer back to Figure 18 for an example of a typical trial), with the following exceptions. Hanzi stimuli were used instead of faces; experiment instructions were provided in both English and (simple) Chinese to expert participants in order to remove any language barriers and ensure they fully understood the task.

As in the faces experiments, the duration of memory arrays varied between experiments and across participants, according to each participant's mean RT to make a correct response in the associated visual search experiment; the range of durations used is provided in the relevant procedure section of each experiment.

Data Analysis

The same data analysis procedure was applied to Hanzi data as it was to faces data: only trials on which participants correctly reported both digits in the verbal suppression task were included in the analyses. (The percentage of incorrect digit response trials that were removed from each set size, for each group, in each experiment can be seen in Tables D3 and D4 in Appendix D.) As in Experiments 7 and 8, FA rates for the WM change detection task varied significantly as a function of set size for all stimulus conditions across both experiments (see Figures E3 and E4 in Appendix E), indicating the need to convert proportion correct detection of change trials (hits) into *d'* values. As before, *d'* values were calculated using individual hit and FA rates for each set size and stimulus condition; these data were then subjected to repeated-measures ANOVA. Pashler's *k*, Cowan's *k*, and Bangor *k* capacity estimates are also reported.

Experiment 9a. Visual WM for Upright Hanzi: Experts versus Novices

The aim of this experiment was to determine whether experts were better able to maintain upright Hanzi in visual WM than novices, reflected by higher capacity. Based on the notion that familiarity increases capacity, significantly better change detection performance and higher capacity estimates for Hanzi experts compared to Hanzi novices were expected.

Methods

Participants.

The same 14 Hanzi experts (8 females, 6 males; mean age 28 years) and 14 Hanzi novices (7 females, 7 males; mean age 24 years) that completed Experiments 6a and 6b (visual search for upright and inverted Hanzi, respectively) completed this experiment. Informed consent was obtained prior to participation.

Stimuli, design, and procedure.

All Hanzi were presented in their upright orientation. As mentioned above, set sizes 1, 2, 3, 4, 5, 6, and 8 were used. The memory array durations for Hanzi novices ranged from 1423 ms to 2823 ms; the memory array durations for Hanzi experts ranged from 1155 ms to 2026 ms. All other design and procedure parameters were as outlined in the General Methods.

Results and Discussion

The results mirrored the familiarity advantage that was found for upright faces in Experiment 7. Figure 22a shows that the ability to detect a change between Hanzi memory and test arrays was markedly better for experts than novices. A mixed design repeated-measures ANOVA using group (novice, expert) as a between factor and set size (1, 2, 3, 4, 5, 6, and 8) as a within factor revealed a significant main effect of group, $F(1, 26) = 7.84$, $p < .05$. The interaction between group and set size was non-significant, $p > .3$.

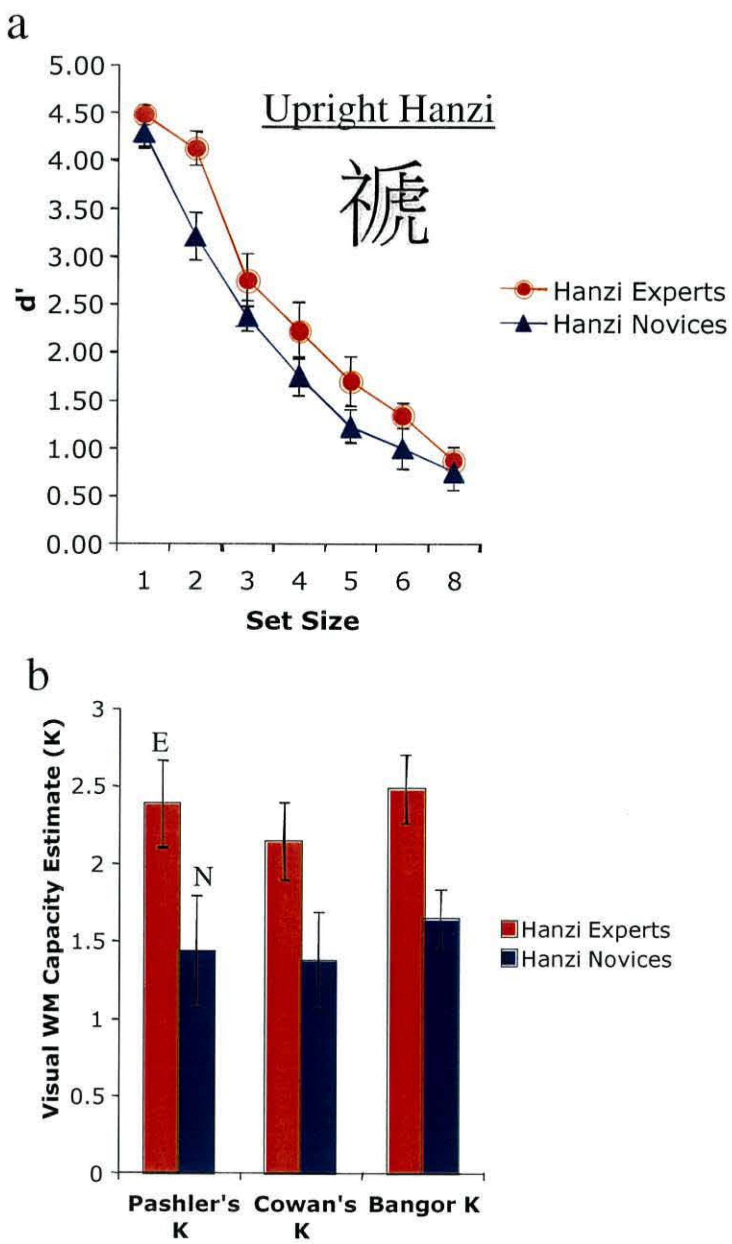


Figure 22. Results for Experiment 9a: (a) Change detection performance (*d'*) as a function of set size for upright Hanzi, comparing Hanzi experts and Hanzi novices. (b) Visual WM capacity estimates using Pashler's *k*, Cowan's *k*, and Bangor *k* for upright Hanzi, comparing Hanzi experts and Hanzi novices. (The letters E and N in figure b denote Experts and Novices respectively, and are included to ensure clarity if the graph is reproduced in black and white.)

As can be seen in Figure 22b, all measures of capacity revealed that experts were able to maintain more Hanzi in visual WM than novices. Pashler's k showed that expert capacity (2.38) was 40% larger than novice capacity (1.43), a significant difference, $p < .05$. Cowan's k showed that expert capacity (2.14) was 36% larger than novice capacity (1.37), a marginally significant difference, $p = .06$. Bangor k showed that expert capacity (2.48) was 34% larger than novice capacity (1.64), a significant difference, $p < .01$.

These findings concretise the suggestion that familiarity enhances the maintenance of representations within visual WM. If larger visual WM capacity is dependent on the extent of within-item integration, my results suggest that experts integrate information contained within Hanzi to a greater extent than novices, perhaps aided by LTM. This is discussed in more detail in the Chapter Discussion.

Experiment 9b. Visual WM for Inverted Hanzi: Experts versus Novices

The aim of this experiment was to assess how featural and configural processes might contribute to the storage of familiar and unfamiliar non-face information within visual WM. Recall that visual search efficiency for inverted Hanzi did not differ between experts and novices; while inversion significantly reduced expert Hanzi search efficiency, novice search was unaffected by inversion. Consistent with the object recognition literature, I concluded that, for experts, upright Hanzi were processed in a configural manner while inverted Hanzi were processed in a featural manner; for novices, both upright and inverted Hanzi were processed featurally.

Based on the above notions, I predicted that change detection accuracy and related visual WM capacity estimates would not significantly differ between experts and novices when Hanzi were inverted. Furthermore, I expected a significant inversion effect for Hanzi experts, and a non-significant inversion effect for Hanzi novices.

Methods

Stimuli, design, and procedure.

The stimuli were the same 9 Hanzi used in Experiment 9a, but they were rotated by 180°. As mentioned above, set sizes 2, 4, and 6 were used. The memory array durations for Hanzi novices ranged from 1391 ms to 2192 ms; the memory array durations for Hanzi experts ranged from 1169 ms to 1972 ms. All other design and procedure parameters were as outlined in the General Methods.

Results and Discussion

I report the results in two subsections: the first deals with the comparison between experts and novices in the current experiment; the second deals with inversion effects by reporting a between-experiment, within-group comparison of upright versus inverted Hanzi for both expert and novice groups.

Experts versus novices: inverted Hanzi.

Figure 23a shows the change detection results for experts and novices when inverted Hanzi were used. While the main effect of group was non-significant, F(1,

26) < 1, the interaction between group and set size was marginally significant, $F(2, 52) = 2.80$, $p = .07$, influenced by better performance for experts than novices at set size 2. An independent *t*-test between expert and novice performance at set size 2 revealed a significant difference, $t(26) = 2.10$, $p < .05$.

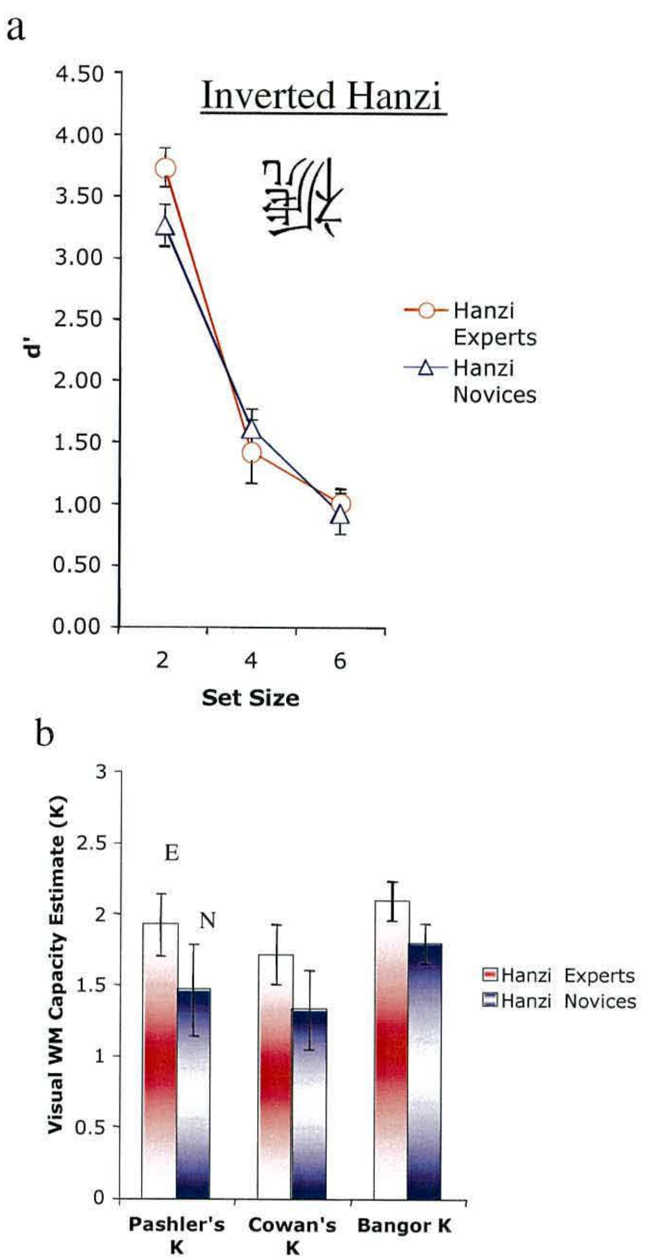


Figure 23. Results for Experiment 9a: (a) Change detection performance (*d'*) as a function of set size for inverted Hanzi, comparing Hanzi experts and Hanzi novices. (b) Visual WM capacity estimates using Pashler's *k*, Cowan's *k*, and Bangor *k* for inverted Hanzi, comparing Hanzi experts and Hanzi novices. (The letters E and N in figure b denote Experts and Novices respectively, and are included to ensure clarity if the graph is reproduced in black and white.)

As can be seen in Figure 23b, all measures of capacity revealed that experts were able to maintain more inverted Hanzi in visual WM than novices, although none of these differences reached significance. Pashler's *k* showed that expert capacity (1.92) was 24% larger than novice capacity (1.46) (*NS*). Cowan's *k* showed that expert capacity (1.71) was 23% larger than novice capacity (1.32) (*NS*). Bangor *k* showed that expert capacity (2.09) was 14% larger than novice capacity (1.79) (*NS*).

As were data from the inverted faces experiment, the results of this inverted Hanzi experiment are somewhat difficult to interpret. A statistically significant difference in change detection performance between experts and novices was evident at set size 2, yet the main effect of group was non-significant and the significance of the interaction between group and set size was marginal. And although *k* estimates for inverted Hanzi tended to be larger for experts than novices, none of the statistical comparisons between expert and novice *k* values reached significance.

If a notable visual WM advantage for experts over novices were interpreted from the inverted Hanzi data, an explanation for this advantage would be difficult to conceive. In the visual search task with inverted Hanzi (Experiment 6b), search efficiency was equivalent between experts and novices – why might familiarity aid the maintenance of inverted Hanzi in visual WM but not the perceptual encoding of inverted Hanzi in visual search? One possibility is that experts demonstrated rapid visual learning of Hanzi from the initial visual search session to the subsequent visual WM session. (Recall that a subset of Hanzi used as distractors in the visual search task was used in the WM task.) Such rapid learning might have resulted in the availability of an orientation-independent mechanism that served to enhance processing of inverted Hanzi in the WM task, and that was absent from the visual search task. This seems unlikely, however, given that Tong and Nakayama (1999) were unable to find evidence for rapid learning of faces over hundreds of presentations during their visual search task. Furthermore, a section analysis of the inverted Hanzi visual search data (the first versus the second half of trials) revealed a non-significant main effect of section half, $F(1, 13) = 2.89, p > .1$, and a non-significant interaction between section half and set size, $F(2, 26) = 1.98, p > .1$, suggesting that expert search efficiency did not improve over time and that rapid and substantial learning of inverted Hanzi did not occur.[20]

Given that there appeared to be little evidence of a familiarity advantage for inverted faces in visual WM (Experiment 8), and given that experts performed no better than novices when inverted Hanzi were presented in the visual search task (Experiment 6b), the most parsimonious interpretation of these findings is that maintenance of inverted Hanzi in visual WM is not particularly enhanced for experts compared to novices.

Inversion effects: Experiments 9a and 9b compared.

Figure 24 illustrates upright versus inverted change detection performance for novices (Figure 24a) and experts (Figure 24b). Repeated-measures ANOVA with orientation (upright, inverted) and set size (2, 4, and 6) as within factors, revealed that expert change detection performance was significantly poorer for inverted than upright Hanzi, $F(1, 13) = 14.71, p < .01$, whereas novice change detection performance was unaffected by inversion, $F(1, 13) < 1$. The interaction between orientation and set size was non-significant in both groups.

---

[20] Note that it is difficult to rule out completely the possibility that prior exposure effects from the initial visual search session influenced performance in the subsequent visual WM session.

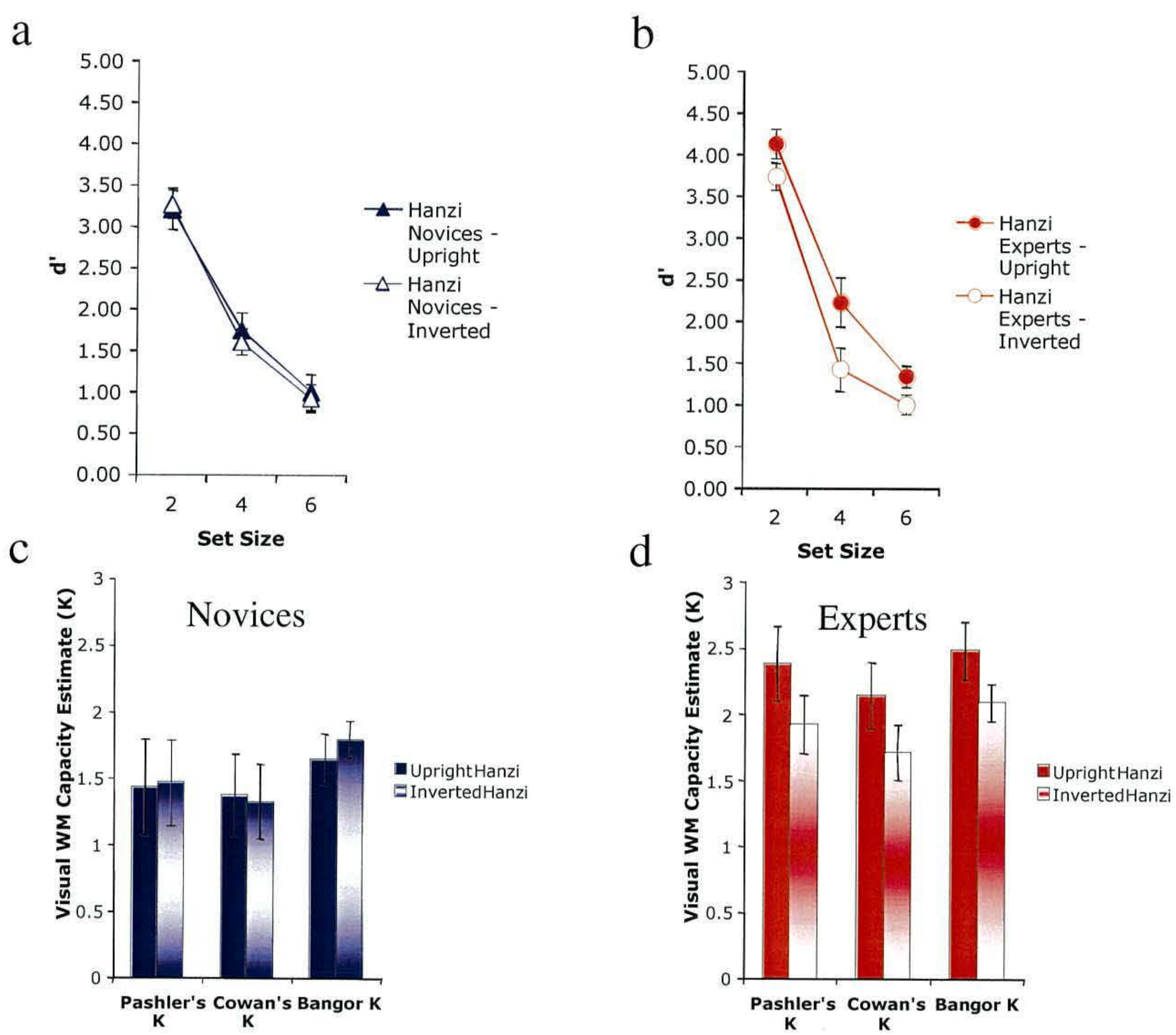


Figure 24. Hanzi inversion effects (Experiments 9a and 9b compared): Change detection performance (*d'*) as a function of set size for (a) upright versus inverted Hanzi for Hanzi novices, and (b) upright versus inverted Hanzi for Hanzi experts. Visual WM capacity estimates using Pashler's *k*, Cowan's *k*, and Bangor *k* for (c) upright versus inverted Hanzi for Hanzi novices, and (d) upright versus inverted Hanzi for Hanzi experts.

Figure 24c shows that inversion did not affect capacity estimates for novices. Using Pashler's *k*, novice capacity was increased by 2% from 1.43 upright Hanzi to 1.46 inverted Hanzi (*NS*). Using Cowan's *k*, novice capacity was reduced by 4% from 1.37 upright Hanzi to 1.32 inverted Hanzi (*NS*). Using Bangor *k*, novice capacity was increased by 9% from 1.64 upright Hanzi to 1.79 inverted Hanzi (*NS*). Note that none of these inversion-induced differences for novices were significant and minor increases in capacity estimates probably reflect noise in the data.

Figure 24d shows that, in contrast to novices, inversion reduced capacity estimates for experts. Using Pashler's k, expert capacity was reduced by 19% from 2.38 upright Hanzi to 1.92 inverted Hanzi (*NS*). Using Cowan's k, expert capacity was reduced by 20% from 2.14 upright Hanzi to 1.71 inverted Hanzi (*NS*). Using Bangor k, expert capacity was reduced by 16% from 2.48 upright Hanzi to 2.09 inverted Hanzi, a significant difference, $p < .05$.

The finding that inversion significantly impaired change detection performance and reduced capacity estimates for Hanzi experts supports the notion that configural information used in upright Hanzi processing was disrupted. This is consistent with the visual search results and with the nature of expert object recognition proposed in the literature. The finding that inversion did not significantly impair change detection performance and did not reduce capacity estimates for Hanzi novices supports the notion that novices used featural information to process both upright and inverted Hanzi, again consistent with the object recognition literature.

Chapter Discussion

This chapter examined the effect of familiarity on the maintenance of faces and non-face complex visual representations in visual WM. A change detection paradigm was employed and capacity estimates were derived. My interpretation of results is based on the assumption that better change detection performance and larger capacity estimates reflected enhanced maintenance. To ensure that the change detection results reflected maintenance, rather than perceptual encoding efficiency and hence perhaps the speed of transfer/consolidation of items into visual WM, memory displays were presented for a sufficient length of time to allow effective perceptual encoding. Memory display durations were derived from RT values produced at the largest set size in the equivalent visual search experiment. This was tailored to each participant and thus served to minimise individual differences in encoding speed that might have influenced the visual WM results.

In Experiment 7, visual WM for single coloured squares, unfamiliar faces, and famous faces was compared; all faces were upright. The ability to detect a change between memory and test arrays was significantly better for squares than for both unfamiliar and famous faces. Capacity estimates revealed that while approximately 3-4 squares could be maintained, only about 2 faces could be maintained. Based on the assumption that each square represented a single chunk of information within visual WM, I propose that each face was not stored as a single chunk. This finding does not support an object-based account of visual WM (Luck & Vogel, 1997), at least for face stimuli.

It is possible, however, that the use of a full test array of faces prevented support for object-based storage to be evidenced. Recall that Wheeler and Treisman (2002) found evidence of object-based visual WM when a single test item was presented, but not when a whole test array was presented. To account for the lack of object-based evidence, they proposed that neighbouring items in the whole test array interfered with feature binding. Perhaps my use of a whole test array masked the potential for object-based visual WM for faces to be revealed. I have evidence to suggest that this is unlikely, however. A Masters project in which I collaborated with a former Masters student, Bethany Wu, showed that even when a single probe design was used, as in Wheeler and Treisman's study, object-based storage of faces could not be found. In this experiment we presented between one and four (unfamiliar) faces in a 2 x 2 grid location matrix for 2000 ms. When fewer than four faces were presented, all other grid locations were occupied by a scrambled face. After a blank retention interval of 1000 ms we presented a single probe face in the centre of the screen, and participants reported whether this face had been present or absent in the previous display. Capacity estimates from the data revealed that participants could maintain about 1.5 faces in visual WM, similar to capacity estimates found in my current experiment for unfamiliar faces using a whole test array. The consistent presentation of four items (faces and scrambled faces) across all set size conditions in our experiment also precludes the argument that the load effects I found for faces here might have been caused by increased demand on perceptual resources rather than the demand on WM resources. For example, McConnell and Quinn (2000, 2004) proposed that the larger the number of items in a display, the greater the chance that some items will interfere with the perceptual encoding of other items and thus impair consolidation into WM.

These results do not support Alvarez and Cavanagh's (2004) proposal that complexity determines capacity. A capacity for about 2 faces is similar to Alvarez and Cavanagh's capacity estimate for simple polygons and 3-D shaded cubes; if their theoretical view is taken, this suggests that face representations might be *simplified* in some way and allow for more information to be stored within one chunk. One manner in which the vast amount of visual detail contained within a face might be simplified is via within-item integration. I extrapolate this idea from Wheeler and Treisman's (2002) paper, in which they suggested that only when features within an object are fully bound would object-based storage be evidenced. I use the term within-item integration to specifically refer to the process of binding both features *and* their configurations to form an individual percept.

I propose that the extent of within-item integration determines the maintenance effectiveness of representations within visual WM. The more integrated the information, the better it will be maintained. The ability to integrate facial information might be supported by significant expertise in processing faces. This could account for why capacity for faces was similar to that found for more simple objects for which expertise was presumably lacking (as in Alvarez & Cavanagh's 2004 study).

The primary aim of Experiment 7 was to determine whether familiarity could enhance the maintenance of faces in visual WM. Consistent with studies that showed a familiarity advantage in verbal and visuospatial WM, I found that change detection performance was significantly better for famous than unfamiliar upright faces. Capacity for famous faces was estimated to be up to 39% larger than capacity for unfamiliar faces. Experiment 9a showed that this familiarity advantage could also apply to complex non-face stimuli: for upright Hanzi, experts demonstrated significantly better change detection performance and larger capacity estimates than novices.

Previous interpretations of the benefit of familiarity in WM involved the influence of LTM. Thorn et al. (2002) suggested that representations stored in LTM aid the reconstruction of decaying representations maintained in WM. Consistent with this idea, neuropsychological research has shown that WM and LTM systems share similar neural pathways (Ranganath et al., 2003), and several researchers have proposed that WM and LTM representations are one and the same in that representations maintained in WM are active states of representations stored in LTM (Cowan, 1998, 2001; Crowder, 1993; Ruchkin et al., 2003). It is reasonable to consider that highly familiar faces, such as the ones used here, are strongly represented in LTM whereas unfamiliar faces are poorly represented in LTM. Thus, familiar faces might be better maintained in visual WM due to greater strength of signal from neurons coding these well-represented faces, activated on presentation of the memory array. In contrast, the signal elicited from neurons coding an unfamiliar face might be weaker, and thus a poorly represented face might decay more rapidly or suffer more interference from neural signals coding other faces.

Another way of interpreting this idea is to consider that maintenance is perhaps enhanced by improved *encoding quality* of familiar images. That is, stronger neural activation elicited on presentation of a familiar face might result in a more distinct and robust representation that is able to be clearly distinguished and protected from interference from other active face representations. Weaker signals elicited by presentation of an unfamiliar face might reflect a more *fuzzy* representation that is less easily distinguished from other faces in the display. One might expect that a stronger signal for a familiar face would elicit heightened levels of activity underlying face-related perceptual processes, for example, increased or speeded N170 amplitude or increased FFA activation. But, as discussed in Chapters 1 and 3, studies that examined these issues have found no evidence for such effects of face familiarity (Bentin & Deouell; Eimer, 2000a, 2000b; Gorno-Tempini & Price, 2001; Gorno-Tempini et al., 2000).

The results of Experiment 9a are inconsistent with the notion that WM and LTM representations are one and the same. Hanzi experts were not familiar with individual Hanzi exemplars used. Rather, they were familiar with the general shape and composition of strokes that can combine to form Hanzi. In this sense, it is conceivable that they would not have had previously stored LTM representations of each of the Hanzi used in this study. How then could LTM aid the maintenance of Hanzi in visual WM among expert participants? Similar to my proposal that the ability to integrate facial information might be supported by significant expertise in processing faces in general, perhaps knowledge, or the experience of, combining or integrating visual detail within each Hanzi enhanced visual WM. Hanzi novices, who presumably lacked this experience, might have been less able to integrate the visual detail contained within each Hanzi, and this inexperience placed a greater burden on visual WM resources for the storage of unconnected or weakly connected information. This notion returns to my proposal that familiarity might enhance the maintenance of representations by enabling a greater extent of within-item integration of features and their configurations.

How might LTM support such integration? Perhaps connections, or pathways, between neurons coding for featural and configural properties of a particular stimulus identity, or stimulus category, are *fortified* over time as the number of exposures to that stimulus or category increases. Or perhaps the firing of specific stimulus-related neurons becomes more *synchronous* with experience? These ideas are explored in more detail in the General Discussion (Chapter 7).

Recall in Chapter 4 that a visual search advantage for famous over unfamiliar faces was observed even when faces were inverted. I interpret this finding to suggest that familiar face processes might be able to activate an orientation-independent mechanism that unfamiliar face processes cannot. I propose that this mechanism might involve better cooperation between featural and configural processes; better cooperation might allow for featural information, predominant in inverted faces, to be translated into configural information that can then be used to more effectively identify a face.

To determine whether an orientation-independent mechanism could be evidenced for visual WM processes, in Experiment 8 I employed the change detection task with inverted unfamiliar and famous faces. In contrast to the familiarity advantage for encoding inverted faces in visual search, I found no clear evidence that famous faces were better maintained in visual WM than unfamiliar faces. This suggests that an orientation-independent mechanism that especially supports familiar face processing might only be employed for perceptual face identification processes and cannot support the maintenance of a familiar face representation in WM once that face has been identified.

I interpret the lack of a familiarity advantage in visual WM for inverted faces to reflect an equivalent extent of within-item integration for familiar and unfamiliar inverted faces. Inversion might have *unglued* integrated features and configurations in upright faces; famous faces appeared just as susceptible to such part deconstruction in visual WM as unfamiliar faces.

In Experiment 9b, Hanzi experts and novices completed the change detection task with inverted Hanzi. While there was some indication that maintenance of inverted Hanzi might have been more effective among experts than novices, the data were not clear-cut. Based on my finding of equivalent visual search efficiency for inverted Hanzi between experts and novices in Experiment 6b, and given that inverted famous faces were not better maintained in visual WM than inverted unfamiliar faces, I adopt the most parsimonious interpretation and suggest that there was no convincing evidence that experts were better able to maintain inverted Hanzi in visual WM than novices.

In summary, familiarity supports the maintenance of upright faces and complex non-face objects in visual WM. This familiarity advantage appears to have been removed when these stimuli were inverted.

A final point to note from these experiments is that novice capacity for upright Hanzi ($k = 1.37$-1.64) was similar to capacity for upright unfamiliar faces ($k = 1.15$-2.08). If we consider that novices had no prior experience with Hanzi, and that humans are "experts" at face processing, the similarity in capacity estimates might seem surprising. These findings suggest two things. First, perhaps faces are not a particularly special class of stimuli. Second, perhaps novices integrated information within Hanzi to some extent, despite the lack of pre-existing LTM representations. If integration of information within a *novel* class of stimuli can occur, this suggests that mechanisms other than LTM can support the maintenance of items within visual WM. Bottom-up processes, such as perceptual grouping by proximity and connectivity of features and configurations *within* items, might be naturally harnessed to promote within-item integration and ease the burden on visual WM resources.

PART 4

GENERAL DISCUSSION AND PRESENTATION OF A NEW MODEL, FLUENCY-BY-INTEGRATION

CHAPTER 7

GENERAL DISCUSSION

The objective of this thesis was to investigate how familiarity can ease visual processing and promote what Jacoby and Dallas (1981) called the *fluency heuristic*. I conducted a series of nine experiments that investigated the effect of familiarity on attention and visual WM processes involved in face and non-face complex object (Hanzi) processing. Three different paradigms were used: attentional blink, visual search, and change detection. I found that familiar (upright) stimuli required less attention for processing, were more efficiently encoded, and were more effectively maintained in visual WM than unfamiliar stimuli. I also examined the effect of inversion on encoding and maintenance of faces and Hanzi. While the familiarity advantage found for encoding upright faces remained for encoding inverted faces in visual search, there was no benefit of familiarity for encoding inverted Hanzi in visual search, or for the maintenance of inverted faces or Hanzi in visual WM.

The use of these paradigms, and the comparison of upright and inverted stimuli, allows me to make some inferences regarding how perceptual processes – such as the analysis of featural and configural information – might have differed between familiar and unfamiliar stimuli and eased the burden on attention and visual WM resources. In this final chapter, Part 4, I draw together the findings from all experiments and I present a model to illustrate how familiarity might promote fluent visual processing. I call this model *fluency-by-integration*.

In Part 1 (Chapter 1), I reviewed evidence for the effect of familiarity on the perceptual processes involved in object and face recognition. Studies of expert object recognition revealed that familiarity increased reliance on configural processing (what I call configural precedence), and increased activation in the FFA, a region of the brain that is especially active during face processing (e.g., Diamond & Carey, 1986; Gauthier et al., 1999; Gauthier & Tarr, 2002). Familiarity effects on face recognition are generally inconsistent with what might be predicted from the object recognition literature, however: there is little evidence to suggest greater reliance on configural processing (Collishaw & Hole, 2000) and little evidence for greater activity in the FFA (Gorno-Tempini & Price, 2001; Gorno-Tempini et al., 2000) for familiar compared to unfamiliar faces. Rather, studies have shown that compared to recognition of unfamiliar faces, familiar face recognition was less affected by changes in orientation (Pourtois et al., 2005; Schweinberger, Pickering, Jentzsch et al., 2002), activated a wider neural network (Leveroni et al., 2000), and showed a greater bilateral visual field advantage that was suggested to reflect greater hemispheric cooperation (Mohr et al., 2002; Schweinberger et al., 2003).

Whereas the effect of familiarity on perceptual processes involved in object and face recognition has been relatively well studied, the question of how familiarity influences higher-level cognitive processes such as attention and working memory has been less well addressed. In Part 2 of this thesis (Chapters 2-4), I investigated the effect of familiarity on attention. In Part 3 (Chapters 5-6), I investigated the effect of familiarity on visual WM. I summarise each part separately below. I also progress the general discussion to reflect upon how attention and visual WM systems might be related, and how they might combine to produce fluent processing.

Familiarity and Attention

A variety of studies have suggested that familiar stimuli require less attention for processing than unfamiliar stimuli. For example, inattentional blindness and attentional blink effects were reduced or eliminated when the target stimulus was one's own name compared to when it was a stranger's name (Mack & Rock, 1998; Shapiro, Caldwell et al., 1997). These findings also apply to more complex visual stimuli, such as faces. Tong and Nakayama (1999) showed that visual search was more efficient for one's own face versus a stranger's faces; Buttle and Raymond (2003) showed that the ability to detect a change between two faces was enhanced when one of the faces was famous. While authors of the above studies suggested that familiarity reduced the attentional requirement for stimulus processing, I proposed that some of these results could reflect attentional capture (pop-out) rather than reduced attentional demand, due to the singularity of a familiar item presented among unfamiliar items. In addition, I questioned whether the familiarity effect in Shapiro and colleagues' AB study and Tong and Nakayama's visual search study might particularly reflect an advantage in processing self-relevant stimuli rather than an advantage in processing familiar stimuli in general. I used AB and visual search paradigms to explore these issues.

The Effect of Familiarity on the AB

In Chapter 3, I reported a series of three experiments that used the AB paradigm to examine the effect of familiarity on attentional demand whilst manipulating the potential influence of attentional capture. I used unfamiliar faces and famous (i.e., non-self-relevant) faces. In Experiment 1, T2 and distractor faces were unfamiliar. In Experiment 2, distractor faces remained unfamiliar but T2 was a famous face. In Experiment 3, distractors were famous faces and T2 faces were either unfamiliar (potential pop-out, as in Experiment 2) or famous (no pop-out, as in Experiment 1). I demonstrated an AB effect for unfamiliar target faces and no AB effect for famous target faces, regardless of whether distractor faces were unfamiliar or famous. I suggest that familiar face processing requires less attentional resource than unfamiliar face processing, and that attentional capture cannot account for the results obtained. This is consistent with Shapiro, Caldwell et al.'s (1997) study in which one's own (familiar) name did not show an AB whereas other common names and nouns did. Furthermore, the lack of AB for famous faces indicates that familiarity, not just self-relevance, can reduce the demand on attention, and that this familiarity advantage can apply to highly complex visual stimuli.

The results of these AB experiments provide two additional theoretical advances with regards to the nature of face processing in general. First, an AB for unfamiliar faces suggests that face processing *per se* requires attention. Therefore, proposals that face processing is obligatory and attention-free (e.g., Farah, 1996; Farah, Wilson et al., 1995) are not supported. Second, the presence of an AB effect for unfamiliar faces when T1 was a featural task indicates that face processing does not require access to a specific configural attentional channel. This contrasts with the proposal that a deficit in face processing can only be observed if configural processes and related configural attentional channels are previously occupied (Awh et al., 2004; Palermo & Rhodes, 2002).

In summary, Experiments 1 and 3 provided clear evidence that an AB effect for unfamiliar faces could be found. Experiments 2 and 3 demonstrated that familiar faces were protected from the AB effect. Based on existing theories, it is possible that the lack of AB for famous faces reflected, at least in part, protection from interference from competing stimuli in the RSVP stream. Protection from interference might have arisen at an early perceptual encoding stage, or a later post-perceptual stage that involved visual WM. Experiments 4-6 (Chapter 4) involved a visual search paradigm and addressed whether familiarity can enhance perceptual encoding. Experiments 7-9 (Chapter 6) involved a change detection task and examined whether familiarity can enhance the maintenance of representations visual WM.

The Effect of Familiarity on Visual Search Efficiency

In Experiment 4, I presented unfamiliar and famous (upright) faces in a visual search paradigm. I found that search was significantly more efficient (i.e., search was speeded and more accurate) for famous than unfamiliar target faces. This suggests that perceptual encoding of familiar faces appears to be enhanced relative to unfamiliar faces. Furthermore, increased efficiency in visual search is considered to reflect reduced attentional requirements, and supports the lack of AB for famous faces found in Experiments 2 and 3. Although Tong and Nakayama (1999) had already demonstrated that visual search for a familiar (one's own) face was more efficient than search for an unfamiliar face, it was not clear whether their results reflected the effect of familiarity, or the effect of attentional capture – in their study, a familiar target face was the only familiar face in the search display and might have particularly stood out from its unfamiliar distractors. In Experiment 4, and all subsequent visual search (and visual WM) experiments, I controlled for attentional capture by presenting homogeneous displays in which all stimuli were either unfamiliar or familiar. Thus, I am confident that my results reflect the effect of familiarity rather than the effect of attentional capture.

In order to determine whether the familiarity advantage for faces in visual search could apply to complex non-face objects, in Experiment 6a I presented Hanzi in visual search and compared Hanzi experts with Hanzi novices. I found that search for Hanzi was significantly more efficient for Hanzi experts than Hanzi novices. This suggests that the benefit of familiarity for stimulus encoding efficiency, and the related ease on attentional resource, is not particularly unique to face stimuli.

In summary, reduced attentional demand for familiar compared to unfamiliar stimuli appears to be related, at least in part, to enhanced perceptual encoding.

How Might Familiarity Enhance Perceptual Encoding?

If the absence of an AB effect for famous faces and more efficient visual search for famous faces and for Hanzi among Hanzi experts reflects enhanced perceptual encoding, how might perceptual encoding be enhanced? One possibility is that familiar stimuli are processed using a compact visual code that contains only information diagnostic for identification, such as broad-scale configural information, and excludes finer featural details that might be unnecessary for recognition (Tong &

Nakayama, 1999). This is essentially the notion of configural precedence, a term I introduced to describe the dominance of configural processes over featural processes. To examine the influence of familiarity on featural and configural processes in more detail and to determine whether configural precedence could account for a reduced attentional requirement, I presented inverted faces and Hanzi in the visual search paradigm (Experiments 5 and 6b, respectively). To maintain identical low-level perceptual properties, the same faces and Hanzi were used in the inverted condition as in the upright condition. If familiarity increases configural precedence, I expected that by disabling access to configural information (via inversion) I would remove the familiarity advantage found for upright faces in Experiment 4 and upright Hanzi found in Experiment 6a, and yield equivalent search efficiency for familiar and unfamiliar inverted faces, and for Hanzi experts and novices.

Support for configural precedence was found for Hanzi – the familiarity advantage found for upright Hanzi was removed when Hanzi were inverted and experts and novices showed equivalent search efficiency. This is consistent with proposals from object recognition studies that experts process upright objects configurally and inverted objects featurally, whereas novices process both upright and inverted objects in a featural manner.

In contrast, visual search was more efficient for famous than unfamiliar inverted faces, a finding that is consistent with Tong and Nakayama's (1999) report of a persistent familiar face advantage for inverted and rotated three-quarter and profile views. A familiarity advantage for inverted faces is inconsistent with the notion of configural precedence and suggests instead that processes involved in familiar face identification might benefit from access to an orientation-independent mechanism that is unavailable to processes involved in unfamiliar face identification. Based on the finding that familiar but not unfamiliar face recognition was enhanced by bilateral field presentation compared to right or left visual field presentation alone (e.g., Mohr et al., 2002), and based on the notion from object recognition literature that the location of featural and configural processes is biased to the left and right hemispheres respectively (Heinze et al., 1998; Martinez et al., 1997; Yamaguchi et al., 2000; Delis et al., 1986), I propose that such an orientation-independent mechanism might involve superior cooperation between featural and configural processing. Improved cooperation between these processes might allow featural information in inverted faces to be efficiently translated into configural information that can then be used to recognise a face. The notion of improved cooperation between featural and configural processes for familiar face recognition is consistent with the dual-route hypothesis which states that face processing in general involves the analysis of both featural and configural information, and with Moscovitch and Moscovitch's (2000) more specific proposal that face recognition involves the *exchange* of featural and configural information (a process they called interactive activation).

The lack of evidence for configural precedence in my data argues against Tong and Nakayama's (1999) proposal of a compact visual code that contains only key configural information. That is not to say, however, that the notion of a compact visual code is not viable. I suggest that a compact code used for recognition might include key configural *and* featural information, rather than key configural information only. As such, when one type of information is unavailable for recognition, for example when an image is blurred (disrupting featural information) or inverted (disrupting configural information), key information from the remaining available processing route might be used to facilitate recognition. The ability to abstract necessary featural *or* configural information could result in flexibility of recognition processes under different viewing conditions.

The notion of abstraction is supported by a study of single neurons in the human brain. Quiroga, Reddy, Kreiman, Koch, and Fried (2005) showed that different pictures of a given famous person or landmark, or familiar animal or object, could preferentially activate a given neuron: that is, single neurons showed some degree of image invariance to the same highly familiar item. Using patients with pharmacologically intractable epilepsy who had been implanted with depth electrodes to localise the seizure onset site, Quiroga et al. analysed responses of neurons from the hippocampus, amygdala, entorhinal cortex, and parahippocampal gyrus to varied and repeated image presentations of the above item categories. (I will focus on their results from face stimuli.) They reported that a single neuron fired selectively on presentation of different pictures of a particular person (e.g., Halle Berry) whether these pictures were photographs or hand drawings and when they depicted that person from different viewpoints. Quiroga and colleagues suggested that an invariant, sparse, and explicit code might be used to transform complex visual percepts into long-term abstract memories. Unfortunately, Quiroga et al. did not conduct this experiment using unfamiliar faces. It would be of great interest to examine the activity of single neurons to the presentation of different pictures of unfamiliar (recently learned) faces and determine whether less invariant neural activity could be observed for unfamiliar than famous faces.

A final point to note from the inverted faces and Hanzi experiments is that although a familiarity effect was present for inverted faces but not for inverted Hanzi, I do not view this disparity as support for the notion that faces are unique. Although these findings could be interpreted to indicate that face identification processes *per se* can access a distinct orientation-independent mechanism that is unavailable to processes involved in non-face object recognition, they could also be interpreted to reflect differences in the level of familiarity for each type of stimulus used in my studies. In the faces tasks, participants were familiar with the famous faces on a micro-subordinate, individual level. In the Hanzi tasks, experts were familiar with Hanzi on a less refined, subordinate level at which they appeared able to effectively discriminate between different exemplars but not identify specific individual Hanzi. Recall that, in order to reduce intrusion from verbal identification processes (Hanzi are logographic units of Chinese language), I used traditional Hanzi with which experts were essentially unfamiliar and found difficult to pronounce and interpret. The key assumption was that expert participants were familiar with Hanzi visual form and structure in general, while novice participants had no experience in visually processing such stimuli. Such a difference between the faces and Hanzi used in my studies does not detract from the familiarity effect found for inverted faces. Rather, it enhances our understanding of how familiarity might affect visual stimulus recognition at different levels of processing. Orientation-independent properties of visual recognition might only emerge as a result of refined, micro-subordinate identification expertise for highly familiar stimuli. If modern Hanzi with which experts are familiar on a micro-subordinate level were used, perhaps a familiarity advantage for inverted Hanzi could be evidenced.

In summary, it appears that efficient perceptual encoding, perhaps enabled by better cooperation between featural and configural processes, could account for the reduced attentional requirement for micro-subordinate familiar face identification and partly explain the lack of AB for famous faces. For complex non-face objects processed at a subordinate level, it appears that a familiarity advantage might stem from greater use of configural than featural information.

Familiarity and Visual WM

Many accounts of the cause of the AB effect propose that WM plays a significant part in whether T2 is successfully reported or not. It is thought that accurate report of T2 depends on successful consolidation into and effective maintenance within a WM storage buffer, which protects from interference from competing items (e.g., Chun & Potter, 1995; Shapiro, Arnell et al., 1997; Shapiro et al., 1994). While some evidence has been provided that familiarity can enhance verbal and visuospatial WM processes, evidence for the effect of familiarity on *visual* WM specifically was lacking. To address this issue, I used faces and Hanzi to examine the effect of familiarity on visual WM capacity, with the assumption that increased capacity reflects enhanced maintenance. To ensure that the change detection task measured differences in visual WM maintenance rather than in encoding (or consolidation) speed between unfamiliar and familiar stimuli, memory arrays were presented for sufficient length of time (pre-determined by the associated visual search task) to allow for effective perceptual encoding.

In Chapter 5, I reviewed literature on the effect of familiarity on WM capacity. A series of studies that examined verbal WM skills in bilingual participants found that serial recall for words was superior in the language of which individuals were most familiar (Chincotta & Hoosain, 1995; Chincotta & Underwood, 1996, 1997; Thorn & Gathercole, 1999, 2001). Research into the game of chess revealed that expert chess players could store more real chess patterns in visuospatial WM than less skilled players (Chase & Simon, 1973; Gobet, 1998; Gobet & Simon, 1996). Both sets of studies proposed that familiarity could increase capacity due to support from LTM. Thorn et al. (2002) proposed that LTM supported the reconstruction of decaying representations in WM. Gobet and Simon proposed that long-term knowledge of visuospatial and conceptual associations between chess pieces and their locations aided chunking of this information. Chunking is a process that is considered to involve the integration of associated information within one WM chunk. While the maximum number of WM chunks is thought to remain at approximately four, chunking appears to be able to increase item capacity (for example chess patterns) if items are closely associated with one another.

To the best of my knowledge, there has been no examination of the effect of familiarity on visual WM capacity specifically; Experiments 7-9, reported in Chapter 6, were the first to address this issue. I used a change detection task, identical to that used in Luck and Vogel's (1997) study, to measure participants' ability to maintain information in visual WM, and derived capacity estimates using Pashler's *k*, Cowan's *k*, and a new Bangor *k* formula that I developed. A concurrent verbal suppression task was administered to reduce input from verbal WM. Interpretation of the results was based on the assumption that improved change detection performance and larger visual WM capacity estimates reflects enhanced maintenance.

In Experiment 7, I presented displays of between 1-10 single coloured squares, unfamiliar faces, and famous faces (all faces were upright). Change detection performance was significantly better for squares than faces; capacity estimates revealed that approximately 3-4 squares and 2 faces could be maintained in visual WM. Lower capacity for faces than squares is interpreted to reflect that faces were not maintained as wholly integrated items, inconsistent with an object-based account of visual WM (Luck & Vogel, 1997). Alvarez and Cavanagh's (2004) proposal that stimulus complexity determined capacity was also unsupported because capacity for faces found in my study was similar to their estimate of capacity for polygons and 3-D shaded cubes, stimuli that are unarguably less complex than faces.

Rather, I propose that detailed information contained within a face is bound, or integrated, to some degree, but not fully. I introduced the term within-item integration to describe the binding of both features and their configurations; I use this term in favour of the term binding in order to distinguish the process of integrating features and their configurations from the more common definition of binding as involving only the integration of features. Based on Wheeler and Treisman's (2002) suggestion that object-based visual WM (i.e., capacity of 3-4 items) can be evidenced if objects are stored as wholly bound items, I propose that within-item integration probably varies by extent, and that the extent of within-item integration determines visual WM capacity.

The key finding from Experiment 7 was that change detection was significantly more accurate, and capacity estimates were higher, for famous than unfamiliar faces. This familiarity advantage was also found in Experiment 9a using complex non-face objects: Hanzi experts showed significantly better change detection performance and higher capacity estimates than Hanzi novices. These experiments provide a systematic measurement of WM capacity for faces and highly complex non-face objects. Furthermore, they reveal that familiarity can enhance maintenance in visual WM, reflected in better change detection performance and larger capacity estimates for familiar compared to unfamiliar stimuli.

How Might Familiarity Enhance Maintenance in Visual WM?

If capacity is dependent on the extent of within-item integration, perhaps a larger capacity for familiar than unfamiliar stimuli reflects better integration of information contained within each familiar item. There are two ways in which better integration might be achieved: (1) greater use of configural information (configural precedence), i.e., spatial relational information spans a wider region of the face, or (2) better integration of features and their configurations. As I mentioned earlier, previous research has found no support for configural precedence in familiar face processing. Collishaw and Hole (2000) compared the effect of inversion and scrambling (considered to disrupt configural processing) on familiar and unfamiliar face recognition. They found that recognition of familiar and unfamiliar faces was impaired to a similar degree when inverted or scrambled, and interpreted this result to reflect equivalent use of configural information for familiar and unfamiliar face recognition processes. As discussed above, the results of my visual search experiments with upright and inverted faces similarly found no evidence for greater configural processing for famous compared to unfamiliar faces.

Therefore, rather than consider enhanced within-item integration to involve greater use of configural processes, I propose that it might involve better integration of features and their configurations, supported by existing representations stored in LTM. This is consistent with my proposal that familiar stimulus recognition might involve improved cooperation between featural and configural processes compared to unfamiliar stimulus recognition.

Whereas Thorn and colleagues (2002) described LTM support in terms of reconstruction, I suggest that of the role of LTM in the maintenance of items in visual WM might be to prevent deconstruction of just previously activated visual representations. It is thought that when a face or object is presented in the visual field, visual detail contained within must be bound, or integrated, for its representation to be perceived as a whole percept rather than a collection of separate features (e.g., Treisman, 1992). When it is removed from view and its representation is required to be maintained in visual WM, the integrated nature of information contained within must be maintained as such. In other words, the information contained within each item representation must remain glued together. If integration cannot be maintained effectively then an item's representation might suffer deconstruction (i.e., become unglued). If visual WM capacity depends on the degree of within-item integration, a deconstructed representation can be conceived of as a poorly integrated representation, and capacity will be reduced. LTM might serve to more effectively sustain a familiar bound representation during maintenance in visual WM than an unfamiliar item that is less well represented in LTM.

In Experiment 8, I examined the effect of inversion on visual WM for unfamiliar and famous faces. A comparison with the results from Experiment 7 revealed that inversion significantly impaired change detection performance and lowered capacity estimates for both unfamiliar and famous faces; this suggests that both face types suffered part deconstruction when they were upside down. In addition, I found that inversion removed the familiarity advantage revealed for upright faces. This suggests that while the extent of within-item integration might have been greater for famous than unfamiliar upright faces maintained in visual WM, the extent of within-item integration appeared equivalent in unfamiliar and famous inverted faces. Whereas LTM appeared to support the perceptual encoding of inverted famous faces in the visual search task and yield a familiarity advantage, LTM seemed unable to effectively support the maintenance of inverted faces in visual WM. It is difficult to interpret the disparity between identification of inverted faces in visual search and maintenance of inverted faces in visual WM. Perhaps inversion disrupts face processing to such a degree that a physically present visual percept (as in visual search) is necessary to support identification of inverted famous faces, possibly via translation of featural into configural information; when this percept is removed from view (as in visual WM), the translated information might be rapidly lost resulting in deconstruction of the representation that renders the extent of within-item integration equivalent to that of inverted unfamiliar faces.

In Experiment 9b, I examined the effect of inversion on visual WM for Hanzi, comparing Hanzi experts with Hanzi novices. A comparison with the results from Experiment 9a (upright Hanzi) revealed that while inversion significantly reduced visual WM maintenance effectiveness among Hanzi experts, Hanzi novice performance was unaffected. This suggests that expert Hanzi processing suffered part deconstruction when Hanzi were upside down, but novice Hanzi processing did not. Perhaps novices demonstrated poor within-item integration for upright Hanzi that could not be made any worse by inverting the Hanzi. Similar to the lack of familiarity effect for inverted faces in visual WM, there was no convincing evidence that maintenance of inverted Hanzi was enhanced for experts compared to novices.

In summary, the results using upright faces and Hanzi in visual WM are clear: familiarity appears to increase the effectiveness with which these items can be maintained, resulting in better change detection performance and larger item capacity estimates. This familiarity advantage in visual WM is consistent with the effect of familiarity on attention: attentional blink and visual search experiments with upright stimuli showed that familiar faces and Hanzi required less attention for processing and were more efficiently encoded. The degree of within-item integration might therefore not only have affected the maintenance of faces and Hanzi in visual WM, but also the demand on attentional resource.

The Link Between Attention and Visual WM

Over the course of this thesis, I have already given brief mention to several ways in which attention and WM systems might interact. For example, attention is thought to be required for the transfer, or consolidation, of information into WM, as demonstrated in studies of the AB effect (e.g., Luck et al., 1996; Vogel & Luck, 2002). Attention is also thought to be required for information binding (e.g., Luck & Beach, 1998; Treisman, 1992) and possibly for maintaining bound representations within visual WM (Wheeler & Treisman, 2002). In Baddeley and Hitch's (1974) model of WM, the central executive component is considered an attentional centre that selects relevant information for storage in WM, updates the contents of WM as new and relevant information arrives, prevents irrelevant information from entering WM, and coordinates the simultaneous encoding and storage of information from more than one task or item. Attention appears therefore to select information for processing, bind (integrate) it into a coherent percept, transfer it into WM, and possibly maintain it in WM as an integrated percept until the information is required again at a later point in time. If sufficient attention is available for these tasks, the selected items are thought to enjoy enhanced protection from decay or interference from other items competing for such encoding and maintenance resources, and the item is retrieved relatively unscathed. When insufficient attention is available, any one or more of these stages might be impaired, the representation suffers over time, and retrieval is poor. In terms of attention and WM tasks, within-item integration could therefore be considered necessary from the first perceptual input until the final response output. In this sense, if familiarity improves within-item integration, most, if not all, attention and WM processes outlined above might be facilitated. I expand on this idea in the fluency-by-integration model below.

The above theoretical account of the relationship between attention and WM is based on the premise that a unidirectional relationship exists between the two systems: that is, attentional resource dictates what and how information gets into WM, and what and how information remains within it. Could this relationship work the other way? Could the contents of WM impact attentional resource? Suppose attention were required to consolidate and maintain items in WM. If this were true, then when a large number of items are maintained in WM and capacity limits are approached, there would be less attention available for other items competing for access and storage. This would decrease the efficiency with which these other items could be consolidated and maintained. In this sense, attention and WM processes might interact with one another in a bi-directional manner.

A study by de Fockert, Rees, Frith, and Lavie (2001) provides some support for the suggestion that the contents of WM can influence attention. In their study, participants performed a selective attention task that involved the classification of famous written names as pop stars or politicians while ignoring distractor faces presented in the background (i.e., a name was superimposed on a face). The identity of the distractor faces either matched the written names (congruent trials) or did not (incongruent trials) and distractor face processing was assessed by comparing classification performance on congruent versus incongruent trials. This selective attention task was flanked by a working memory task that varied in terms of the demand on WM resources. In the WM task, a sequence of digits was presented at the start of a trial which participants were required to maintain during the selective attention task. After the attention task response, a single probe digit was presented and participants were required to state the digit that was presented after that probe in the previous sequence. In the low WM load condition, the digit sequence was numerically ordered, e.g., 0, 1, 2, 3, 4; in the high load condition, the digit sequence did not follow numerical order, e.g., 0, 3, 1, 2, 4. De Fockert et al. showed that when WM resources were heavily taxed by the high load digit task, distractor faces interfered significantly with the name classification task; when the WM task did not substantially tax capacity limits, distractor interference was reduced. De Fockert and colleagues proposed that a high WM load engaged a large proportion of attention and left insufficient attentional resource for inhibiting the intrusion of the distractor faces on the name classification task.

In the final section I present a new model, fluency-by-integration, to illustrate how familiarity-enhanced within-item integration might promote fluent visual processing and reduce the burden on both attention and visual WM resources.

A Proposed Model of Fluency-by-Integration

I define fluency as efficient perceptual and cognitive processes that combine to yield improved task performance. The fluency-by-integration model is illustrated in Figure 25 below. In this model, I propose the following: (1) Attention is required for binding, or integrating, features and their configurations – a process I term within-item integration. (2) Attention might also be required for maintaining integrated representations in visual WM and preventing deconstruction. (3) Familiarity with a stimulus increases the extent of within-item integration via some neural means, possibly enhanced neural synchrony. (4) Familiarity might enhance these neural processes as a result of repeated and possibly varied processing experiences that might strengthen the connection between neurons coding for a particular stimulus and create a well-rehearsed, efficient neural response pattern. (5) A familiar representation that already enjoys significant within-item integration will require less attention for processing when present in the visual field compared to an unfamiliar representation in which features and configurations are less well integrated. (6) If less attention is needed for processing a familiar, highly integrated stimulus, more attention will be available for maintaining that and other well-integrated representations in visual WM. (7) Items in visual WM that are well served by sufficient attentional resource will be effectively maintained, resulting in accurate retrieval and enhanced item capacity. Thus, I propose that familiarity promotes fluency by virtue of reducing the burden on attention and visual WM resources.

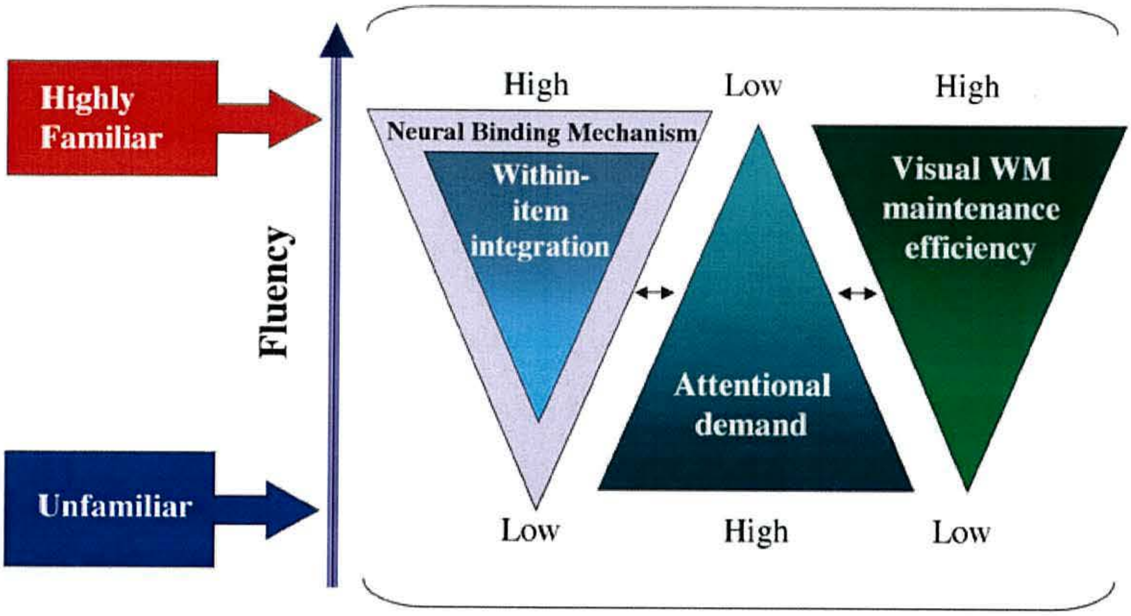


Figure 25. Illustration of the proposed fluency-by-integration model, developed to conceptualise how familiarity might ease visual processing. I consider that information within a highly familiar stimulus representation is highly integrated, or glued together, perhaps enabled by enhanced neural synchrony. Neural synchrony might be enhanced by activation of a robust representation in LTM, created by strong and well-used neural connections that have developed as a result of numerous rich and possibly varied exposures to a particular stimulus. If the degree of within-item integration is high, less attention is required for perceptual binding and maintaining a bound representation in visual WM. As a result, familiar items are more effectively maintained in visual WM and reveal a larger WM capacity. Unfamiliar stimuli, within which information is less well integrated, require more attention for creating and maintaining a bound representation than familiar stimuli. This reduces the effectiveness with which they are maintained in visual WM and results in lower WM capacity.

An additional property could be included in the fluency-by-integration model – that of between-item segregation. Raffone and Wolters (2001) introduced this term to describe the strong, rapid, and transient inhibition that was exerted on neural assemblies coding for a different object to the one that was selected for within-item integration. This inhibition, or desynchronisation, was suggested to reflect the suppression of irrelevant object features and reduce the potential perceptual interference from other objects that could disrupt feature binding within a selected object. In other words, desynchronisation might reduce between-item binding errors that might arise from the intrusion of one object's features into another object. Based on their suggestion that better binding (i.e., within-item integration) results in enhanced between-item segregation, I propose that both these processes might combine to produce the familiarity advantage found for familiar over unfamiliar stimuli in my series of experiments.

How might the combination of enhanced within-item integration and between-item segregation protect against an AB effect, improve search efficiency, and enhance maintenance within visual WM? As outlined in Chapter 2, Gross et al. (2004) demonstrated that a strong synchronised neural network response to T2 was related to successful report of T2 (an AB effect was absent), whereas weaker synchronisation to T2 was related to impaired T2 report (an AB effect was present). Based on their findings, I propose that in an RSVP stream it is possible that the presentation of a highly integrated familiar T2 item elicits a highly synchronised response that exerts sufficient inhibition on competing items; inhibition might protect T2 from interference that could cause an AB, and might occur during perceptual encoding, consolidation into WM, or maintenance within visual WM. Recall that I found no AB for famous T2 faces when distractors were unfamiliar (Experiment 2) or when distractors were famous (Experiment 3). It is possible therefore that the level of inhibition exerted by a highly synchronised neural response to a famous T2 face might have been sufficient for desynchronising highly competitive neural responses to famous distractors.

In the visual search paradigm, perhaps enhanced search efficiency for a familiar target among familiar distractors compared to an unfamiliar target among unfamiliar distractors reflects rapid and effective inhibition of familiar distractors, following a highly synchronised episode in which they were identified as a non-target match.

Finally, enhanced within-item integration coupled with enhanced between-item segregation might serve to better maintain familiar items in visual WM in much the same way as familiar targets are protected from the AB effect – perhaps the ability to maintain distinct and separable representations in WM reduces the potential for integration errors that might occur when properties from two different (poorly integrated) representations might intrude on one another and cause confusion.

While there is evidence to support some properties of the model, other properties are based on indirect extrapolation of ideas from various studies. There is evidence that binding (i.e., within-item integration) might be achieved by synchronised firing of neurons that code elements belonging to the same object – known as *binding-by-synchrony* (e.g., Milner, 1974; von der Malsburg, 1995). Oscillations in the range of 30-70 Hz (gamma-band range) are thought to aid synchronisation (Singer et al., 1997; Singer & Gray, 1995), with induced activity at 40 Hz considered to reflect the construction of a coherent object percept (Tallon-Baudry & Bertrand, 1999). Bottom-up processes related to Gestalt grouping properties (e.g., similarity, proximity, continuation, closure, and figure-ground segregation) appear to modulate gamma activity levels. For example, Tallon-Baudry, Bertrand, Delpuech, and Pernier (1996) showed that an illusory triangle (called a

Kanizsa triangle) and a real triangle induced similar levels of gamma activity, whereas gamma activity was much weaker in response to a "non-triangle" (see Figure 26a). Their findings can be interpreted to suggest that higher levels of gamma activity reflect enhanced within-item integration, resulting in a coherent object percept.

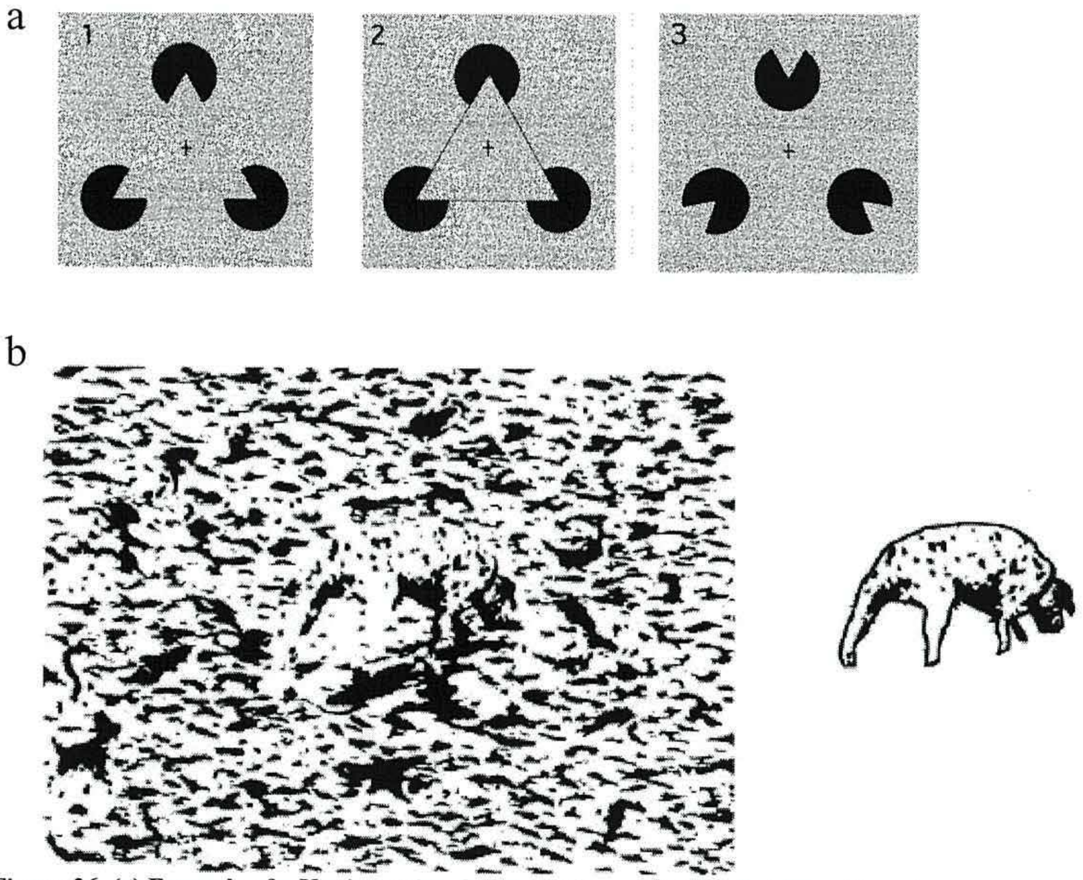


Figure 26. (a) Example of a Kanisza triangle (image 1), a real triangle (image 2), and a non-triangle (image 3) as used in (and reproduced from) Tallon-Baudry et al. (1996). (b) An illustration of the hidden Dalmation dog image (on the left) and the actual Dalmation that is hidden (on the right) as used in (and reproduced from) Tallon-Baudry et al. (1997).

There is also evidence that induced gamma activity levels might reflect the extent of within-item integration in more complex stimuli such as faces. Existing studies of gamma activity and face processing have predominantly focused on the effects of autism and Williams Syndrome (WS) on visual processing. People with autism and WS are commonly described as having difficulty integrating perceptual features, interpreted as a dominance of featural over global or configural processing and sometimes known as weak central coherence (Frith, 1989, cited in Grice et al., 2001). They show impaired face processing, typically reflected in poor face discrimination and recognition performance and delayed N170 activity (Dawson, Webb, & McPartland, 2005; McPartland, Dawson, Webb, Panagiotides, & Carver, 2004). Grice et al. demonstrated that impaired face processing was correlated with reduced gamma activity. They used EEG to measure gamma activity while participants with autism and WS, and control participants, passively viewed (unfamiliar) face images. Participants with autism and WS demonstrated lower gamma activity than controls. While Grice and colleagues interpret their results to suggest that reduced gamma activity reflects an impaired ability to bind featural information within a face, I propose that it might reflect impaired binding of features *and* their configurations (i.e., impaired within-item integration).

Grice and colleagues (2001) also examined the effect of face inversion on gamma activity. They found that gamma activity in participants with autism and WS was unaffected by inversion. In contrast, and perhaps more intriguing, gamma activity was reduced for inverted relative to upright faces among normal controls. Reduced gamma activity for inverted faces might reflect weaker within-item integration, a notion that I presented earlier to interpret lower visual WM capacity for inverted compared to upright faces (both unfamiliar and famous) and Hanzi (for Hanzi experts).

In addition to a bottom-up influence on gamma activity, there is evidence that top-down processes, such as those involved in LTM, can modulate gamma activity. Tallon-Baudry, Bertrand, Delpuech, and Pernier (1997) presented participants with a picture in which an image of a Dalmation dog was hidden (see Figure 26b). When participants were naïve to the presence of the dog, they were unable to perceive a coherent image and gamma activity was weak. In contrast, after participants were trained to perceive and detect the hidden dog, a much larger induced gamma response was recorded, irrespective of whether the picture contained the Dalmation. They suggested that activation of a mental representation from LTM, whether externally or internally driven, involves oscillatory neural synchronisation. Their finding that active LTM representations can elicit gamma activity introduces the possibility that perhaps the strength or quality of representations stored in LTM might modulate gamma activity and reflect different degrees of within-item integration. To my knowledge, the effect of familiarity on gamma activity has not yet been directly investigated. An EEG study to compare gamma activity between unfamiliar and familiar faces might elucidate the question of whether familiarity increases within-item integration. If familiarity increases within-item integration, higher gamma activity might be expected in response to presentation of familiar compared to unfamiliar faces.

A final property of gamma-band activity reported by Tallon-Baudry and colleagues is that it can be sustained during short-term maintenance of internal representations (Tallon-Baudry, Bertrand, Peronnet, & Pernier, 1998). This suggests that a combined EEG and visual WM study using familiar and unfamiliar stimuli might shed greater light on my proposal that the extent of within-item integration modulates the effectiveness with which visual representations are maintained in WM. If highly integrated familiar representations are more effectively maintained in visual WM than poorly integrated unfamiliar representations, one might expect a correlation between heightened gamma activity levels during the retention interval and enhanced change detection performance.

In summary, the experiments reported in this thesis have highlighted several visual processing advantages of familiar over unfamiliar faces and complex non-face objects: they require less attentional resource, are more rapidly and efficiently encoded, appear less susceptible to interference from competing stimuli, and are more effectively maintained in visual WM. Familiar representations might benefit from a high degree of within-item integration (binding between features and their configurations) supported by LTM and perhaps enabled by enhanced neural synchronisation. I introduced a model of fluency-by-integration to explain how familiarity might promote fluent visual processing and reduce the burden on attention and visual WM resources.

The benefit of robust, integrated representations are clear – we are able to interact with familiar people and objects in our environment easily and efficiently, and suffer little confusion with competing stimuli.

There might also be benefits for holding less well-integrated, unfamiliar representations. Rather than think of information within an unfamiliar representation to be weakly integrated, consider that it might be *flexibly* integrated. Perhaps flexibility in the relationship between features and their configurations is essential for visual learning to progress – room is allowed for adaptation to the image as it changes over time and under different viewing conditions. Perhaps flexible representations enable “fluent learning”, a process which might involve labile neural networks in which, once created, connections can be modified. It is possible that some neural connections might be strengthened and some made redundant as we learn which visual details are most diagnostic for stimulus recognition, the result of which might be a compact and highly integrated visual code that promotes fluent visual processing. It would be interesting to explore whether there are any tasks that could show an “unfamiliarity benefit” that might be reflective of such fluent learning.

REFERENCES

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15*(2), 106-111.

Anderson, A. K., Christoff, K., Panitz, D., De Rosa, E., & Gabrieli, J. D. E. (2003). Neural correlates of the automatic processing of threat facial signals. *Journal of Neuroscience, 23*(13), 5627-5633.

Arnell, K. M., & Jolicoeur, P. (1999). The attentional blink across stimulus modalities: Evidence for central processing limitations. *Journal of Experimental Psychology: Human Perception & Performance, 25*(3), 630-648.

Arnell, K. M., Shapiro, K. L., & Sorensen, R. E. (1999). Reduced repetition blindness for one's own name. *Visual Cognition, 6*, 609-635.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *The Psychology of Learning & Motivation, 2*, 89-195.

Awh, E., Serences, J., Laurey, P., Dhaliwal, H., van der Jagt, T., & Dassonville, P. (2004). Evidence against a central bottleneck during the attentional blink: Multiple channels for configural and featural processing. *Cognitive Psychology, 48*, 95-126.

Baddeley, A. D. (1986). *Working Memory*. Oxford: Oxford University Press.

Baddeley, A. D. (1990). *Human Memory: Theory and Practice*. Boston: Allyn and Bacon.

Baddeley, A. D. (1997). *Human Memory: Theory and Practice (Revised Edition)*. Hove: Psychology Press Ltd.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417-423.

Baddeley, A. D., & Hitch, G. (1974). Working Memory. *The Psychology of Learning & Motivation, 8*, 47-89.

Baker, C. I., Behrmann, M., & Olson, C. R. (2003). Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. *Nature Neuroscience, 5*(11), 1210-1216.

Bar, M., & Biederman, I. (1998). Subliminal visual priming. *Psychological Science, 9*(6), 464-469.

Barnard, P. J., Scott, S., Taylor, J., May, J., & Knightley, W. (2004). Paying attention to meaning. *Psychological Science, 15*(3), 179-186.

Bartlett, J. C., Searcy, J. H., & Abdi, H. (2003). What are the routes to face recognition? In M. A. Peterson & G. Rhodes (Eds.), *Perception of Faces,* *Objects, and Scenes: Analytic and Holistic Processes.* New York: Oxford University Press.

Baylis, G. C., & Rolls, E. T. (1987). Responses of neurons in the inferior temporal cortex in short-term and serial recognition memory tasks. *Experimental Brain Research, 65*(3), 614-622.

Begleiter, H., Porjesz, B., & Wang, W. (1995). Event-related brain potentials differentiate priming and recognition to familiar and unfamiliar faces. *Electroencephalography & Clinical Neurophysiology, 94*, 41-49.

Behrmann, M., Moscovitch, M., & Winocur, G. (1994). Intact visual imagery and impaired visual perception in a patient with visual agnosia. *Journal of Experimental Psychology: Human Perception & Performance, 20*(5), 1068-1087.

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience, 8*(6), 551-565.

Bentin, S., & Deouell, L. Y. (2000). Structural encoding and identification in face processing: ERP evidence for separate mechanisms. *Cognitive Neuropsychology, 17*(1/2/3), 35-54.

Besner, D., Stoltz, J., & Boutilier, C. (1997). The Stroop effect and the myth of automaticity. *Psychological Bulletin & Review, 4*, 221-225.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*(2), 115-147.

Biederman, I., & Cooper, E. E. (1991). Evidence for complete translational and reflectional invariance in visual object priming. *Perception, 20*, 585-593.

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

Boutet, I., Gentes-Hawn, A., & Chaudhuri, A. (2002). The influence of attention on holistic face encoding. *Cognition, 84*, 321-341.

Broadbent, D. E. (1958). *Perception and Communication*. London: Pergamon Press.

Broadbent, D. E. (1982). Task combination and the selective intake of information. *Acta Psychologia, 50*, 253-290.

Brown, A. S., Jones, T. C., & Mitchell, D. B. (1996). Single and multiple test repetition priming in implicit memory. *Memory, 4,* 159-173.

Brown, V., Huey, D., & Findlay, J. M. (1997). Face detection in peripheral vision: Do faces pop out? *Perception, 26*(12), 1555-1570.

Bruce, V. (1989). *Recognising Faces.* Hove: Lawrence Erlbaum Associates.

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

Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video: Evidence from security surveillance. *Psychological Science, 10*(3), 243-248.

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

Carey, S., & Diamond, R. (1994). Are faces perceived as configurations more by adults than by children? *Visual Cognition, 1*(2-3), 253-274.

Cave, C. B. (1997). Very long-lasting priming in picture naming. *Psychological Science, 8*, 322-325.

Chase, W. G., & Ericsson, K. A. (1981). Skilled memory. In J. R. Anderson (Ed.), *Cognitive Skills and their Acquisition* (pp. 141-189). Hillsdale NJ: Erlbaum.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual Information Processing* (pp. 215-281). New York: Academic Press.

Chincotta, D., & Hoosain, R. (1995). Reading rate, articulatory suppression, and bilingual digit span. *European Journal of Cognitive Psychology, 7*, 201-211.

Chincotta, D., & Underwood, G. (1996). Mother tongue, language of schooling and bilingual digit span. *British Journal of Psychology, 87*, 193-208.

Chincotta, D., & Underwood, G. (1997). Bilingual memory span advantage for Arabic numerals over digit words. *British Journal of Psychology, 88*, 295-310.

Chun, M. M. (1997). Types and tokens in visual processing: A double dissociation between the attentional blink and repetition blindness. *Journal of Experimental Psychology: Human Perception & Performance, 23*(3), 738-755.

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception & Performance, 21*, 109-127.

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

Cocchini, G., Logie, R. H., Della Sala, S., McPherson, S. E., & Baddeley, A. D. (2002). Concurrent performance of two memory tasks: Evidence for domain-specific working memory systems. *Memory & Cognition, 30*(7), 1086-1095.

Collette, F., & Van der Linden, M. (2002). Brain imaging of the central executive component of working memory. *Neuroscience & Biobehavioral Reviews, 26*(2), 105-125.

Collin, C. A., & McMullen, P. A. (2005). Subordinate-level categorization relies on high spatial frequencies to a greater degree than basic-level categorization. *Perception & Psychophysics, 67*(2), 354-364.

Collishaw, S. M., & Hole, G. J. (2000). Featural and configurational processes in the recognition of faces of different familiarity. *Perception, 29*, 893-909.

Courtney, S. M., Petit, L., Maisog, J. M., Ungerleider, L. G., & Haxby, J. V. (1998). An area specialized for spatial working memory in human frontal cortex. *Science, 279*(5355), 1347-1351.

Cowan, N. (1998). Visual and auditory working memory capacity. Commentary. *Trends in Cognitive Sciences, 2*(3), 77-78.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioural & Brain Sciences, 24*, 87-185.

Crowder, R. G. (1993). Short-term memory: Where do we stand? *Memory & Cognition, 21*, 142-145.

Dawson, G., Webb, S. J., & McPartland, J. (2005). Understanding the nature of face processing impairment in autism: Insights from behavioural and electrophysiological studies. *Developmental Neuropsychology, 27*(3), 403-424.

De Fockert, J. W., Rees, G., Frith, C. D., & Lavie, N. (2001). The role of working memory in visual selective attention. *Science, 291*, 1803-1806.

De Renzi, E. (1989). Prosopagnosia: A multi-stage, specific disorder? In H. D. Ellis (Ed.), *Handbook of Research on Face Processing*. Amsterdam: North-Holland.

Delis, D. C., Robertson, L. C., & Efron, R. (1986). Hemispheric specialization of memory for visual hierarchical stimuli. *Neuropsychologia, 24*(2), 205-214.

Della Sala, S., Gray, C. M., Baddeley, A. D., Allamano, N., & Wilson, L. (1999). Pattern span: A tool for unwelding visuo-spatial memory. *Neuropsychologia, 37*, 1189-1199.

Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences of the United States of America, 93*, 13494-13499.

Desimone, R., & Ungerleider, L. G. (1989). Neural mechanisms of visual processing in monkeys. In J. Grafman (Ed.), *Handbook of Neuropsychology* (Vol. 2, pp. 267-299). New York: Elsevier.

Devlin, J. T., Russell, R. P., Davis, M. H., Price, C. J., Moss, H. E., Fadili, M. J., et al. (2002). Is there an anatomical basis for category-specificity? Semantic memory studies in PET and fMRI. *Neuropsychologia, 40*(1), 54-75.

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.

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science, 11*(1), 86-89.

Donchin, E. (1981). Surprise! ... Surprise? *Psychophysiology, 18*(5), 493-513.

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral & Brain Sciences, 11*(3), 357-374.

Downing, P. E., Liu, J., & Kanwisher, N. (2001). Testing cognitive models of visual attention with fMRI and MEG. *Neuropsychologia, 39*, 1329-1342.

Driver, J. & Tipper, S. P. (1989). On the nonselectivity of "selective" seeing: Contrasts between interference and priming in selective attention. *Journal of Experimental Psychology: Human Perception & Performance, 15*(2), 304-314.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*(3), 433-458.

Eastwood, J. D., Smilek, D., & Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Perception & Psychophysics, 63*(6), 1004-1013.

Edelman, S. (1998). Representation is representation of similarities. *Behavioral & Brain Sciences, 21*(4), 449-498.

Eimer, M. (2000a). Effects of face inversion on the structural encoding and recognition of faces: Evidence from event-related brain potentials. *Cognitive Brain Research, 10*, 145-158.

Eimer, M. (2000b). Event-related potentials distinguish processing stages involved in face perception and recognition. *Clinical Neurophysiology, 111*, 694-705.

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.

Ellis, H. D., & Young, A. W. (1989). Are faces special? In H. D. Ellis (Ed.), *Handbook of Research on Face Processing.* Amsterdam: Elsevier.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*(2), 211.

Farah, M. J. (1990). *Visual Agnosia: Disorders of Object Recognition and What They Tell Us About Normal Vision.* Cambridge, MA: MIT Press.

Farah, M. J. (1996). Is face recognition 'special''? Evidence from neuropsychology. *Behavioural Brain Research, 76*(1-2), 181-189.

Farah, M. J., Tanaka, J. N., & Drain, M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception & Performance, 21*(3), 628-634.

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

Fink, G., Halligan, P., Marshall, J., Frith, C., Frackowiak, R., & Dolan, R. (1997). Neural mechanisms involved in the processing of global and local aspects of hierarchically organized visual stimuli. *Brain, 120*(10), 1779-1791.

Finke, K., Bublak, P., Neugebauer, U., & Zihl, J. (2005). Combined processing of what and where information within the visuospatial scratchpad. *European Journal of Cognitive Psychology, 17*(1), 1-22.

Frith, U. (1989). *Autism: Explaining the Enigma*. Oxford: Blackwell.

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

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

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

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

Gauthier, I., & Tarr, M. J. (2002). Unravelling mechanisms for expert object recognition: Bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception & Performance, 28*, 431-446.

Giesbrecht, B., & Di Lollo, V. (1998). Beyond the attentional blink: Visual masking by object substitution. *Journal of Experimental Psychology: Human Perception & Performance, 24*(5), 1454-1466.

Gobet, F. (1998). Expert memory: A comparison of four theories. *Cognition, 66*(2), 115-152.

Gobet, F., & Simon, H. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology, 31*(1), 1-40.

Gorno-Tempini, M., & Price, C. (2001). Identification of famous faces and buildings: A functional neuroimaging study of semantically unique items. *Brain, 124*(10), 2087-2097.

Gorno-Tempini, M., Price, C., Josephs, O., Vandenberghe, R., Cappa, S., Kapur, N., et al. (2000). The neural systems sustaining face and proper-name processing. *Brain, 123*(2), 419-419.

Grice, S. J., Spratling, M. W., Karmiloff-Smith, A., Halit, H., Csibra, G., de Haan, M., et al. (2001). Disordered visual processing and oscillatory brain activity in autism and Williams Syndrome. *Neuroreport, 12*(12), 2697-2700.

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience, 7*(5), 555-562.

Gross, J., Schmitz, F., Schnitzler, I., Kessler, K., Shapiro, K., Hommel, B., et al. (2004). Modulation of long-range neural synchrony reflects temporal limitations of visual attention in humans. *Proceedings of the National Academy of Sciences of the United States of America, 101*(35), 13050-13055.

Gruber, T., & Muller, M. (2005). Oscillatory brain activity dissociates between associative stimulus content in a repetition priming task in the human EEG. *Cerebral Cortex, 15*(1), 109-116.

Hamm, J. P., & McMullen, P. A. (1998). Effects of orientation on the identification of rotated objects depend on the level of identity. *Journal of Experimental Psychology: Human Perception & Performance, 24*(2), 413-426.

Haxby, J., Petit, L., Ungerleider, L., & Courtney, S. (2000). Distinguishing the functional roles of multiple regions in distributed neural systems for visual working memory. *Neuroimage, 11*(5), 380-391.

Haxby, J., Ungerleider, L., Horwitz, B., Maisog, J., Rapoport, S., & Grady, C. (1996). Face encoding and recognition in the human brain. *Proceedings of the National Academy of Sciences of the United States of America, 93*(2), 922-927.

Heinze, H. J., Hinrichs, H., Scholz, M., Burchert, W., & Mangun, G. R. (1998). Neural mechanisms of global and local processing: A combined PET and ERP study. *Journal of Cognitive Neuroscience, 10*(4), 485-498.

Henderson, L. (1972). Spatial and verbal codes and the capacity of STM. *Quarterly Journal of Experimental Psychology, 24*, 485-495.

Henson, R. N. A., Goshen-Gottstein, Y., Ganel, T., Otten, L., Quayle, A., & Rugg, M. (2003). Electrophysiological and haemodynamic correlates of face perception, recognition and priming. *Cerebral Cortex, 13*(7), 793-805.

Henson, R. N. A., Shallice, T., & Dolan, R. J. (2000). Neuroimaging evidence for dissociable forms of repetition priming. *Science, 287*(5456), 1269-1272.

Henson, R. N. A., Shallice, T., Gorno-Tempini, M., & Dolan, R. (2002). Face repetition effects in implicit and explicit memory tests as measured by fMRI. *Cerebral Cortex, 12*(2), 178-186.

Hillger, L. A., & Koenig, O. (1991). Separable mechanisms in face processing: Evidence from hemispheric specialization. *Journal of Cognitive Neuroscience, 3*(1), 42-58.

Hillyard, S. A., & Picton, T. W. (1987). Electrophysiology of Cognition. In F. Plum (Ed.), *Handbook of Physiology: Section 1. The Nervous System.* (Vol. 5, pp. 519-584): Waverly Press.

Hole, G. J. (1994). Configurational factors in the perception of unfamiliar faces. *Perception, 23*(1), 65-74.

Holmes, A., Vuilleumier, P., & Eimer, M. (2003). The processing of emotional facial expression is gated by spatial attention: Evidence from event-related brain potentials. *Cognitive Brain Research, 16*, 174-184.

Holscher, C., Rolls, E., & Xiang, J. (2003). Perirhinal cortex neuronal activity related to long-term familiarity memory in the macaque. *European Journal of Neuroscience, 18*(7), 2037-2046.

Humphreys, G. W., & Muller, H. J. (1993). Search via Recursive Rejection (ERR): A connectionist model of visual search. *Cognitive Psychology, 25*, 43-110.

Jacobs, J. (1887). Experiments on 'prehension'. *Mind, 12*, 75-79.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110*(3), 306-340.

Jeffreys, D. (1996). Evoked potential studies of face and object processing. *Visual Cognition, 3*(1), 1-38.

Jenkins, R., Lavie, N., & Driver, J. (2003). Ignoring famous faces: Category-specific dilution of distractor interference. *Perception & Psychophysics, 65*(2), 298-309.

Johnston, W. A., Hawley, K. J., Plewe, S. H., Elliot, J. M. G., & DeWitt, M. J. (1990). Attention capture by novel stimuli. *Journal of Experimental Psychology: General, 119*(4), 397-411.

Jolicoeur, P. (1999). Restricted attentional capacity between sensory modalities. *Psychonomic Bulletin & Review, 6*(1), 87-92.

Kahneman, D. (1973). *Attention and Effort*. New Jersey: Prentice-Hall.

Kampf, M., Babkoff, H., & Nachson, I. (2005). Laterality in familiar face recognition: A meta-analysis. *International Journal of Neuroscience, 115*(1), 23-33.

Kanwisher, N. (1987). Repetition blindness: Type recognition without token individuation. *Cognition, 27*, 117-143.

Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience, 3*, 759-763.

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

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

Kanwisher, N., & Potter, M. C. (1990). Repetition blindness: Levels of processing. *Journal of Experimental Psychology: Human Perception & Performance, 16*(1), 30-47.

Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Annual Review of Neuroscience, 23*, 315-341.

Katanoda, K., Yoshikawa, K., & Sugishita, M. (2000). Neural substrates for the recognition of newly learned faces: A functional MRI study. *Neuropsychologia, 38*(12), 1616-1625.

Kim, J. J., Andreasen, N. C., O'Leary, D. S., Wiser, A. K., Ponto, L. L. B., Watkins, G. L., et al. (1999). Direct comparison of the neural substrates of recognition memory for words and faces. *Brain, 122*, 1069-1083.

Kimchi, R. (1992). Primacy of wholistic processing and global/local paradigm: A critical review. *Psychological Bulletin, 112*(1), 24-38.

Klauer, K. C., & Zhao, Z. M. (2004). Double dissociations in visual and spatial short-term memory. *Journal of Experimental Psychology: General, 133*(3), 355-381.

Kobatake, E., Wang, G., & Tanaka, K. (1998). Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. *Journal of Neurophysiology, 80*(1), 324-330.

Kuehn, S. M., & Jolicoeur, P. (1994). Impact of quality of the image, orientation, and similarity of the stimuli on visual-search for faces. *Perception, 23*(1), 95-122.

Lamme, V. A. F. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences, 7*, 12-18.

Lavie, N., Ro, T., & Russell, C. (2003). The role of perceptual load in processing distractor faces. *Psychological Science, 14*(5), 510-515.

Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 53*(2), 513-536.

Leveroni, C., Seidenberg, M., Mayer, A., Mead, L., Binder, J., & Rao, S. (2000). Neural systems underlying the recognition of familiar and newly learned faces. *Journal of Neuroscience, 20*(2), 878-886.

Li, L., Miller, E., & Desimone, R. (1993). The representation of stimulus familiarity in anterior inferior temporal cortex. *Journal of Neurophysiology, 69*(6), 1918-1929.

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: An MEG study. *Nature Neuroscience, 5*(9), 910-916.

Liu, J., Higuchi, M., Marantz, A., & Kanwisher, N. (2000). The selectivity of the occipitotemporal M170 for faces. *Neuroreport, 11*(2), 337-341.

Logie, R. H., Zucco, G. M., & Baddeley, A. D. (1990). Interference with visual short-term-memory. *Acta Psychologica, 75*(1), 55-74.

Love, B. C., Rouder, J. N., & Wisniewski, E. J. (1999). A structural account of global and local processing. *Cognitive Psychology, 38*(2), 291-316.

Luck, S. J., & Beach, N. J. (1998). Visual attention and the binding problem: A neurophysiological perspective. In R. D. Wright (Ed.), *Visual Attention* (pp. 455-478). New York: Oxford University Press.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279-281.

Luck, S. J., Vogel, E. K., & Shapiro, K. L. (1996). Word meanings can be accessed but not reported during the attentional blink. *Nature, 383*, 616-618.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences, 4*(11), 432-440.

Macho, S., & Leder, H. (1998). Your eyes only? A test of interactive influence in the processing of facial features. *Journal of Experimental Psychology: Human Perception & Performance, 24*(5), 1486-1500.

Mack, A., & Rock, I. (1998). *Inattentional Blindness*. Cambridge Mass: MIT Press.

Marcel, A. J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15*(2), 197-237.

Marcel, A. J. (1983). Conscious and unconscious perception: An approach to the relations between phenomenal experience and perceptual processes. *Cognitive Psychology, 15*(2), 238-300.

Mari-Beffa, P., Estevez, A. F., & Danziger, S. (2000). Stroop interference and negative priming: Problems with inferences from null results. *Psychonomic Bulletin & Review, 7*(3), 499-503.

Marois, R., Yi, D. J., & Chun, M. M. (2004). The neural fate of consciously perceived and missed events in the aftentional blink. *Neuron, 41*(3), 465-472.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* San Francisco: W.H. Freeman.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of spatial organization of 3-dimensional shapes. *Proceedings of the Royal Society of London Series B: Biological Sciences, 200*(1140), 269-294.

Marsolek, C. J. (1999). Dissociable neural subsystems underlie abstract and specific object recognition. *Psychological Science, 10*(2), 111-118.

Martens, S., Wolters, G., & van Raamsdonk, M. (2002). Blinks of the mind: Memory effects of attentional processes. *Journal of Experimental Psychology: Human Perception & Performance, 28*(6), 1275-1287.

Martinez, A., Moses, P., Frank, L., Buxton, R., Wong, E., & Stiles, J. (1997). Hemispheric asymmetries in global and local processing: Evidence from fMRI. *Neuroreport, 8*(7), 1685-1689.

McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience, 9*(5), 605-610.

McConnell, J., & Quinn, J. (2000). Interference in visual working memory. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 53*(1), 53-67.

McConnell, J., & Quinn, J. (2004). Complexity factors in visuo-spatial working memory. *Memory, 12*(3), 338-350.

McDermott, K. B., Buckner, R. L., Petersen, S. E., Kelley, W. M., & Sanders, A. L. (1999). Set- and code-specific activation in the frontal cortex: An fMRI study of encoding and retrieval of faces and words. *Journal of Cognitive Neuroscience, 11*(6), 631-640.

McMullen, P. A., & Jolicoeur, P. (1992). Reference frame and effects of orientation on finding the tops of rotated objects. *Journal of Experimental Psychology: Human Perception & Performance, 18*(3), 807-820.

McPartland, J., Dawson, G., Webb, S., Panagiotides, H., & Carver, L. (2004). Event-related brain potentials reveal anomalies in temporal processing of faces in autism spectrum disorder. *Journal of Child Psychology & Psychiatry, 45*(7), 1235-1245.

Medin, D. L., Ross, B. H., & Markman, A. B. (2005). *Cognitive Psychology* (4th ed.). Hoboken, NJ: John Wiley & Sons.

Miller, E. K., Li, L., & Desimone, R. (1991). A neural mechanism for working and recognition memory in inferior temporal cortex. *Science, 254*(5036), 1377-1379.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

Milner, P. (1974). A model for visual shape recognition. *Psychological Review, 81*, 521-535.

Miyashita, Y., Date, A., & Okuno, H. (1993). Configurational encoding of complex visual forms by single neurons of monkey temporal cortex. *Neuropsychologia, 31*(10), 1119-1131.

Mohr, B., Landgrebe, A., & Schweinberger, S. (2002). Interhemispheric cooperation for familiar but not unfamiliar face processing. *Neuropsychologia, 40*(11), 1841-1848.

Moore, C. M., & Wolfe, J. M. (2001). Getting beyond the serial/parallel debate in visual search: A hybrid approach. In K. L. Shapiro (Ed.), *The Limits of*

Attention: Temporal Constraints in Human Information Processing. (pp. 178-198). New York: Oxford University Press.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology, 11*, 56-60.

Moscovitch, M., & Moscovitch, D. (2000). Super face-inversion effects for isolated internal or external features, and for fractured faces. *Cognitive Neuropsychology, 17*(1-3), 201-219.

Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Neuroscience, 9*(5), 555-604.

Musen, G., & Treisman, A. (1990). Implicit and explicit memory for visual patterns. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 16*, 127-137.

Navon, D. (1977). Forest before trees: Precedence of global features in visual perception. *Cognitive Psychology, 9*(3), 353-383.

Nothdurft, H. C. (1993). Faces and facial expressions do not pop out. *Perception, 22*(11), 1287-1298.

Olson, I. R. & Jiang, Y. (2004). Visual short-term memory is not improved by training. *Memory & Cognition, 32*, 1326-1332.

Palermo, R., & Rhodes, G. (2002). The influence of divided attention on holistic face perception. *Cognition, 82*, 225-257.

Pashler, H. E. (1988). Familiarity and visual change detection. *Perception & Psychophysics, 44*(4), 369-378.

Pashler, H. E. (1998). *The Psychology of Attention*. Cambridge Massachusets: MIT Press.

Pessoa, L., Kastner, S., & Ungerleider, L. G. (2002). Attentional control of the processing of neutral and emotional stimuli. *Cognitive Brain Research, 15*, 31-45.

Pessoa, L., McKenna, M., Gutierrez, E., & Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *Proceedings of the National Academy of Sciences of the United States of America, 99*(17), 11458-11463.

Peterson, M. A., & Rhodes, G. (Eds.). (2003). *Perception of Faces, Objects, and Scenes: Analytic and Holistic Processes.* New York: Oxford University Press.

Pfutze, E., Sommer, W., & Schweinberger, S. (2002). Age-related slowing in face and name recognition: Evidence from event-related brain potentials. *Psychology & Ageing, 17*(1), 140-160.

Pinker, S. (1984). Visual cognition: An introduction. *Cognition, 18*, 1-63.

Poggio, T., & Edelman, S. (1990). A network that learns to recognize 3-dimensional objects. *Nature, 343*(6255), 263-266.

Poldrack, R. A., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1998). The neural basis of visual skill learning: An fMRI study of mirror reading. *Cerebral Cortex, 8*(1), 1-10.

Poldrack, R. A., Selco, S., Field, J., & Cohen, N. J. (1999). The relationship between skill learning and repetition priming: Experimental and computational analyses. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 25*, 208-235.

Pourtois, G., Schwartz, S., Seghier, M., Lazeyras, F., & Vuilleumier, P. (2005). View-independent coding of face identity in frontal and temporal cortices is modulated by familiarity: An event-related fMI study. *Neuroimage, 24*(4), 1214-1224.

Quiroga, R., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature, 435*(7045), 1102-1107.

Raffone, A., & Wolters, G. (2001). A cortical mechanism for binding in visual working memory. *Journal of Cognitive Neuroscience, 13*(6), 766-785.

Rakover, S. S., & Teucher, B. (1997). Facial inversion effects: Parts and whole relationship. *Perception & Psychophysics, 59*(5), 752-761.

Ranganath, C., Johnson, M. K., & D'Esposito, M. (2003). Prefrontal activity associated with working memory and episodic long-term memory. *Neuropsychologia, 41*(3), 378-389.

Raymond, J. E. (2003). New objects, not new features, trigger the attentional blink. *Psychological Science, 14*(1), 54-59.

Raymond, J. E., O'Donnell, H. L., & Tipper, S. P. (1998). Priming reveals attentional modulation of human motion sensitivity. *Vision Research, 38*(19), 2863-2867.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: an attentional blink? *Journal of Experimental Psychology: Human Perception & Performance, 18*, 849-860.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1995). Similarity determines the attentional blink. *Journal of Experimental Psychology: Human Perception & Performance, 21*(3), 653-662.

Reber, P. J., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (2005). Priming effects in the fusiform gyrus: Changes in neural activity beyond the second presentation. *Cerebral Cortex, 15*(6), 787-795.

Reed, C. L., Stone, V. E., Bozova, S., & Tanaka, J. (2003). The body-inversion effect. *Psychological Science, 14*(4), 302-308.

Reeves, A., & Sperling, G. (1986). Attention gating in short-term visual memory. *Psychological Review, 93*(2), 180-206.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8*, 368-373.

Rhodes, G. (1993). Configural coding, expertise, and the right-hemisphere advantage for face recognition. *Brain & Cognition, 22*(1), 19-41.

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

Riesenhuber, M., & Poggio, T. (1999). Are cortical models really bound by the "binding problem"? *Neuron, 24*(1), 87-93.

Rolke, B., Heil, M., Streb, J., & Hennighausen, E. (2001). Missed prime words within the attentional blink evoke an N400 semantic priming effect. *Psychophysiology, 38*, 165-174.

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

Rossion, B., Delvenne, J. F., Debatisse, D., Goffaux, V., Bruyer, R., Crommelinck, M., et al. (1999). Spatio-temporal localization of the face inversion effect: An event-related potentials study. *Biological Psychology, 50*(3), 173-189.

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

Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., et al. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: An electrophysiological account of face-specific processes in the human brain. *Neuroreport, 11*(1), 69-74.

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

Rossion, B., Schiltz, C., & Crommelinck, M. (2003). The functionally defined right occipital and fusiform "face areas" discriminate novel from visually familiar faces. *Neuroimage, 19*(3), 877-883.

Ruchkin, D. S., Grafman, J., Cameron, K., & Berndt, R. S. (2003). Working memory retention systems: A state of activated long- term memory. *Behavioral & Brain Sciences, 26*(6), 709-777.

Sagiv, N., & Bentin, S. (2001). Structural encoding of human and schematic faces: Holistic and part-based processes. *Journal of Cognitive Neuroscience, 13*(7), 937-951.

Sakai, K., & Miyashita, Y. (1994). Neuronal tuning to learned complex forms in vision. *Neuroreport, 5*(7), 829-832.

Schacter, D. L., & Buckner, R. L. (1998). Priming and the brain. *Neuron, 20*(2), 185-195.

Schacter, D. L., Reiman, E., Uecker, A., Polster, M. R., Yun, L. S., & Cooper, L. A. (1995). Brain regions associated with retrieval of structurally coherent visual information. *Nature, 376*(6541), 587-590.

Schneider, W. X., Eschman, A., & Zuccolotto, A. (2002). *E-Prime User's Guide.* Pittsburgh: Psychology Software Tools Inc.

Schneider, W. X., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, & attention. *Psychological Review, 84*, 1-66.

Schweinberger, S. R., Baird, L. M., Blumler, M., Kaufmann, J. M., & Mohr, B. (2003). Interhemispheric cooperation for face recognition but not for affective facial expressions. *Neuropsychologia, 41*(4), 407-414.

Schweinberger, S. R., Pfutze, E. M., & Sommer, W. (1995). Repetition priming and associative priming of face recognition: Evidence from event-related potentials. *Journal of Experimental Psychology: Learning Memory, & Cognition, 21*(3), 722-736.

Schweinberger, S. R., Pickering, E. C., Burton, A. M., & Kaufmann, J. M. (2002). Human brain potential correlates of repetition priming in face and name recognition. *Neuropsychologia, 40*, 2057-2073.

Schweinberger, S. R., Pickering, E. C., Jentzsch, I., Burton, A. M., & Kaufmann, J. M. (2002). Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Cognitive Brain Research, 14*, 398-409.

Schyns, P. G., & Gosselin, F. (2003). Diagnsotic use of scale information for componential and holistic recognition. In M. A. Peterson & G. Rhodes (Eds.), *Perception of Faces, Objects, and Scenes: Analytic and Holistic Processes.* New York: Oxford University Press.

Sergent, J. (1986). Microgenesis of face perception. In A. Young (Ed.), *Aspects of Face Processing*. Dordrecht: Martinus Nijhoff.

Sergent, J., & Signoret, J. L. (1992). Varieties of functional deficits in prosopagnosia. *Cerebral Cortex, 2*(5), 375-388.

Shaffer, L. (1975). Multiple attention in continuous verbal tasks. In S. Dornic (Ed.), *Attention and Performance* (pp. 157-167). New York: Academic Press.

Shah, N. J., Marshall, J. C., Zafiris, O., Schwab, A., Zilles, K., Markowitsch, H. J., et al. (2001). The neural correlates of person familiarity: A functional magentic resonance imaging study with clinical implications. *Brain & Cognition, 124*, 804-815.

Shapiro, K. L., Arnell, K. M., & Raymond, J. E. (1997). The attentional blink. *Trends in Cognitive Sciences, 1*, 291-296.

Shapiro, K. L., Caldwell, J. I., & Sorensen, R. E. (1997). Personal names and the attentional blink: The 'cocktail party' effect revisited. *Journal of Experimental Psychology: Human Perception & Performance, 23*, 504-514.

Shapiro, K. L., Driver, J., Ward, R., & Sorensen, R. E. (1997). Priming from the attentional blink: A failure to extract visual tokens but not visual types. *Psychological Science, 8*(2), 95-100.

Shapiro, K. L., Raymond, J. E., & Arnell, K. M. (1994). Attention to visual pattern information produces the attentional blink in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception & Performance, 20*, 357-371.

Shen, L. M., Hu, X. P., Yacoub, E., & Ugurbil, K. (1999). Neural correlates of visual form and visual spatial processing. *Human Brain Mapping, 8*(1), 60-71.

Simon, H. A. (1974). How big is a chunk? *Science, 183*, 482-488.

Singer, W. (1999). Neuronal synchrony: A versatile code for the definition of relations? *Neuron, 24*(1), 49-65.

Singer, W., Engel, A. K., Kreiter, A., Munk, M. H. J., Neunschwander, S., & Roelfsema, P. R. (1997). Neuronal assemblies: Necessity, signature, and detectability. *Trends in Cognitive Sciences, 1*, 252-261.

Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience, 18*, 555-586.

Soto-Faraco, S., Spence, C., Fairbank, K., Kingstone, A., Hillstrom, A. P., & Shapiro, K. (2002). A crossmodal attentional blink between vision and touch. *Psychonomic Bulletin & Review, 9*(4), 731-738.

Sternberg, S. (1966). High speed scanning in human memory. *Science, 153*, 652-654.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-661.

Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in Cognitive Sciences, 3*(4), 151-162.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *The Journal of Neuroscience, 16*(13), 4240-4249.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1997). Oscillatory gamma-band (30-70 Hz) activity induced by a visual search task in humans. *The Journal of Neuroscience, 17*(2), 722-734.

Tallon-Baudry, C., Bertrand, O., Peronnet, F., & Pernier, J. (1998). Induced gamma-band activity during the delay of a visual short- term memory task in humans. *Journal of Neuroscience, 18*(11), 4244-4254.

Tan, L. H., Liu, H. L., Perfetti, C. A., Spinks, J. A., Fox, P. T., & Gao, J. H. (2001). The neural system underlying Chinese logograph reading. *Neuroimage, 13*(5), 836-846.

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

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

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

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

Tarr, M. J. (1995). Rotating objects to recognize them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin & Review, 2*, 55-82.

Tarr, M. J., & Bulthoff, H. H. (1995). Is human object recognition better described by geon structural descriptions or by multiple views: Comment on Biederman and Gerhardstein (1993). *Journal of Experimental Psychology: Human Perception & Performance, 21*(6), 1494-1505.

Tarr, M. J., & Bulthoff, H. H. (1998). Image-based object recognition in man, monkey and machine. *Cognition, 67*(1-2), 1-20.

Tarr, M. J., & Gauthier, I. (2000). FFA: A flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience, 3*(8), 764-769.

Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology, 21*, 233-282.

Thorn, A. S. C., & Gathercole, S. E. (1999). Language-specific knowledge and short-term memory in bilingual and non-bilingual children. *Quarterly Journal of Experimental Psychology, 52A*, 303-324.

Thorn, A. S. C., & Gathercole, S. E. (2001). Language differences in verbal short-term memory do not exclusively originate in the process of subvocal rehearsal. *Psychonomic Bulletin & Review, 8*, 357-364.

Thorn, A. S. C., Gathercole, S. E., & Frankish, C. R. (2002). Language familiarity effects on short-term memory: The role of output delay and long-term knowledge. *The Quarterly Journal of Experimental Psychology, 55A*(4), 1363-1383.

Tipper, S. P., & Cranston, M. (1985). Selective attention and priming: Inhibitory and facilitatory effects of ignored primes. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 37*(4), 591-611.

Tong, F., & Nakayama, K. (1999). Robust representations for faces: Evidence from visual search. *Journal of Experimental Psychology: Human Perception & Performance, 25*, 1016-1035.

Townsend, J. T. (1990). Serial vs parallel processing: Sometimes they look like tweedledum and tweedledee but they can (and should) be distinguished. *Psychological Science, 1*(1), 46-54.

Townsend, J. T., & Fific, M. (2004). Parallel versus serial processing and individual differences in high-speed search in human memory. *Perception & Psychophysics, 66*(6), 953-962.

Treisman, A. (1992). Perceiving and reperceiving objects. *American Psychologist, 47*(7), 862-875.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12*, 97-136.

Tsukiura, T., Fujii, T., Fukatsu, R., Otsuki, T., Okuda, J., Umetsu, A., et al. (2002). Neural basis of the retrieval of people's names: Evidence from brain-damaged patients and fMRI. *Journal of Cognitive Neuroscience, 14*(6), 922-937.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247*, 301-306.

Ullman, S. (1989). Aligning pictorial descriptions: An approach to object recognition. *Cognition, 32*, 193-254.

Ungerleider, L. G., Courtney, S. M., & Haxby, J. V. (1998). A neural system for human visual working memory. *Proceedings of the National Academy of Sciences of the United States of America, 95*, 883-890.

Vallar, G., & Baddeley, A. D. (1984). Fractionation of working memory: Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning & Verbal Behaviour, 23*, 151-161.

Visser, T. A. W., Bischof, W. F., & Di Lollo, V. (1999). Attentional switching in spatial and nonspatial domains: Evidence from the attentional blink. *Psychological Bulletin, 125*, 458-469.

Vogel, E. K., & Luck, S. J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology, 37*(2), 190-203.

Vogel, E. K., & Luck, S. J. (2002). Delayed working memory consolidation during the attentional blink. *Psychonomic Bulletin & Review, 9*(4), 739-743.

Vogel, E. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception & Performance, 24*(6), 1656-1674.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception & Performance, 27*, 92-114.

von der Malsburg, C. (1995). Binding in models of perception and brain function. *Current Opinion in Neurobiology, 5*(4), 520-526.

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: An event-related fMRI study. *Neuron, 30*(3), 829-841.

Vuilleumier, P., Henson, R. N., Driver, J., & Dolan, R. J. (2002). Multiple levels of visual object constancy revealed by event- related fMRI of repetition priming. *Nature Neuroscience, 5*(5), 491-499.

Welford, A. (1980). The single-channel hypothesis. In A. Welford (Ed.), *Reaction Time* (pp. 215-252). New York: Academic Press.

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience, 18*(1), 411-418.

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General, 131*(1), 48-64.

Wiggs, C. L., Martin, A., & Sunderland, T. (1997). Monitoring frequency of occurrence without awareness: Evidence from patients with Alzheimer's disease. *Journal of Clinical & Experimental Neuropsychology, 19*(2), 235-244.

Wojciulik, E., Kanwisher, N., & Driver, J. (1998). Covert visual attention modulates face-specific activity in the human fusiform gyrus: fMRI study. *Journal of Neurophysiology, 79*(3), 1574-1578.

Wood, N. & Cowan, N. (1995). The cocktail party phenomenon revisited: How frequent are attention shifts to one's own name in an irrelevant auditory channel? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*, 255-260.

Woodman, G. F., & Luck, S. J. (2003). Serial deployment of attention during visual search. *Journal of Experimental Psychology: Human Perception & Performance, 29*(1), 121-138.

Woodman, G. F., Vecera, S. P., & Luck, S. J. (2003). Perceptual organisation influences visual working memory. *Psychonomic Bulletin & Review, 10*(1), 80-87.

Xiang, J. Z., & Brown, M. W. (1998). Differential neuronal encoding of novelty, familiarity and recency in regions of the anterior temporal lobe. *Neuropharmacology, 37*(4-5), 657-676.

Yamaguchi, S., Yamagata, S., & Kobayashi, S. (2000). Cerebral asymmetry of the "Top-Down" allocation of attention to global and local features. *Journal of Neuroscience, 20*(9), 1-5.

Young, A. W., Ellis, A. W., Flude, B. M., McWeeny, K. H., & Hay, D. C. (1986). Face name interference. *Journal of Experimental Psychology: Human Perception & Performance, 12*(4), 466-475.

Young, A. W., Hay, D. C., & McWeeny, K. H. (1985). Right cerebral hemisphere superiority for constructing facial representations. *Neuropsychologia, 23*(2), 195-202.

Young, A. W., 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*, 737-746.

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

APPENDICES

APPENDIX A

UNFAMILIAR AND FAMOUS FACE STIMULI

Unfamiliar Faces Used in Experiments 1 and 2

Some faces were used for more than one role across experiments and conditions. Below each face image I denote which experiment it was used in, which role(s) it played, and which T2 face it related to (where appropriate) according to the following key:

1 = Experiment 1	3 = Experiment 3	D = Distractor	TB = Tony Blair
2 = Experiment 2	M = T2 Mask	F = T2 Filler	PC = Prince Charles

No face fulfilled more than one role in any one experimental condition.

1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D

1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D

1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D

1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D

1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D

1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D

1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D

1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D 1,2 D

1 2 D 1,3 M 1,3 M 1,3 M 1,3 M 1,3 M* 1,3 M* 1,3 M 1,3 M 1,3 M

1,3 F 1,3 F 1,3 F 1,3 F 1,3 F 1,3 F 1,3 F 1,3 F 2,3 F, TB, PC 2,3 F, TB, PC

* These two faces were used as both T2 and as T2 masks in two different participant groups in Experiment 1. When one of the faces was T2, the other was a mask, and vice versa.


2,3 F, TB, PC


2,3 F, TB, PC


2,3 F, TB, PC


2,3 F, TB, PC


2,3 F, TB, PC


2,3 F, TB, PC


1,2,3 D M PC

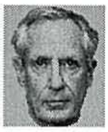

1,2,3 D M PC


1,2,3 D M PC

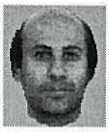

1,2,3 D M PC


1,2,3 D M PC


1,2,3 D M PC


1,2,3 D M PC


M PC


1,2,3 D M TB


1,2,3 D M TB

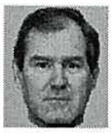

1,2,3 D M TB


1,2,3 D M TB


1,2,3 D M TB


1,2,3 D M TB


1,2,3 D M TB


M TB

Famous Faces Used in Experiment 3

Some faces were used for more than one role across experiments and conditions. Below each face image I denote which experiment it was used in, which role(s) it played, and which T2 face it related to according to the following key:

D = Distractor

UM = Unfamiliar T2 Mask

TBM = Tony Blair T2 Mask

PCM = Prince Charles T2 Mask

UF = Unfamiliar T2 Filler

TBF = Tony Blair T2 Filler

PCF = Prince Charles T2 Filler

No face fulfilled more than one role in any one experimental condition.

D	D	D	D	D	D	D	D	D	D

D	D	D	D	D	D	D	D	D	D

D	D, PCF	D	D	D	D	D	D	D	D

D | D | UM, D | UM, D | UM, D | UM, D | UM, D | UM, D | UM, D | UM, D

TBM, D, PCF | TBM, D PCF | TBM, D, PCF | TBM, D, PCF | TBM, D | TBM, D | TBM, D, PCF | TBM, D, PCF | PCM, TBF | PCM, D

PCM, D, TBF | PCM, TBF | PCM, D, TBF | PCM, D, TBF | PCM, TBF | PCM, D TBF | UF, D | UF, D | UF | UF, D

UF, D | UF, D | UF, D | UF, D | D, TBF, PCF

Unfamiliar Faces Used in Experiments 4, 5, 7, and 8

Some faces were used for more than one role across experiments 4 and 5 (visual search) and 7 and 8 (visual working memory). Below each face image I denote which role(s) it played:

VST = Visual Search Target

WM = Working Memory face

VSD = Visual Search Distractor

No face fulfilled more than one role in any one experimental condition.

WM, VSD	WM, VSD	WM, VSD	WM, VSD	WM, VSD	WM, VSD	WM, VSD	WM, VSD	WM, VSD	WM, VSD

WM, VSD	VSD	VSD	VSD	VSD	VSD	VSD	VSD	VST	VST

VST	VST	VST	VST	VST	VST	VST	VST	VST	VST

Famous Faces Used in Experiments 4, 5, 7, and 8

Some faces were used for more than one role across experiments 4 and 5 (visual search) and 7 and 8 (visual working memory). Below each face image I denote which role(s) it played:

VST = Visual Search Target

WM = Working Memory face

VSD = Visual Search Distractor

No face fulfilled more than one role in any one experimental condition.

WM, VSD	WM, VSD	WM, VSD	WM, VSD	WM, VSD	WM, VSD	WM, VSD	WM, VSD	WM, VSD	WM, VSD

WM, VSD	VST	VST	VST	VST	VST	VST	VST	VST	VST

VST	VST	VST	VSD	VSD	VSD	VSD	VSD	VSD	VSD

APPENDIX B

VISUAL SEARCH OUTLIER EXCLUSIONS

Table B1. Percentage of visual search outlier trials (± 2 SD) removed from Experiment 4.

	Upright unfamiliar faces	*Upright famous faces*
Set Size 4 target present	4.05% (35/864)	3.47% (30/864)
Set Size 4 target absent	3.70% (32/864)	4.05% (35/864)
Set Size 7 target present	3.70% (32/864)	4.05% (35/864)
Set Size 7 target absent	4.98% (43/864)	3.82% (33/864)
Set Size 10 target present	3.01% (26/864)	3.59% (31/864)
Set Size 10 target absent	4.75% (41/864)	4.17% (36/864)

Table B2. Percentage of visual search outlier trials (± 2 SD) removed from Experiment 5.

	Inverted unfamiliar faces	*Inverted famous faces*
Set Size 2 target present	4.17% (37/864)	4.40% (38/864)
Set Size 2 target absent	4.28 % (37/864)	4.75% (41/864)
Set Size 4 target present	3.70% (32/864)	3.36% (29/864)
Set Size 4 target absent	3.59% (31/864)	3.24% (28/864)
Set Size 6 target present	2.89% (25/864)	4.05% (35/864)
Set Size 6 target absent	4.05% (35/864)	4.17% (36/864)

Table B3. Percentage of visual search outlier trials (± 2 SD) removed from Experiment 6a.

	Upright Hanzi: Experts	*Upright Hanzi: Novices*
Set Size 4 target present	2.78% (24/864)	2.20% (19/864)
Set Size 4 target absent	2.43% (21/864)	2.20% (19/864)
Set Size 6 target present	1.97% (17/864)	2.08% (18/864)
Set Size 6 target absent	2.78% (24/864)	2.08% (18/864)
Set Size 8 target present	2.20% (19/864)	1.39% (12/864)
Set Size 8 target absent	1.74% (15/864)	2.08% (18/864)

Table B4. Percentage of visual search outlier trials (± 2 SD) removed from Experiment 6b.

	Inverted Hanzi: Experts	*Inverted Hanzi: Novices*
Set Size 2 target present	2.55% (22/864)	2.66% (23/864)
Set Size 2 target absent	2.78% (24/864)	2.08% (18/864)
Set Size 4 target present	2.31% (20/864)	2.89% (25/864)
Set Size 4 target absent	2.55% (22/864)	2.31% (20/864)
Set Size 6 target present	2.20% (19/864)	2.43% (21/864)
Set Size 6 target absent	1.97% (17/864)	2.08% (18/864)

APPENDIX C

HANZI STIMULI

Some Hanzi were used for more than one role across experiments 6a and 6b (visual search) and 9a and 9b (visual working memory). Below each Hanzi image I denote which role(s) it played:

VST = Visual Search Target WM = Working Memory Hanzi

VSD = Visual Search Distractor

No Hanzi fulfilled more than one role in any one experimental condition. Dictionary definitions and hanyu pinyin (roman script phonetic notation and transliteration) are also provided (www.chinalanguage.com).

Hanzi	**Dictionary Definition**	**Hanyu Pinyin**	**Experimental Role**
菙	Thorn, bramble	chui2	WM, VSD
竄	Escape, run away, flee / change, alter, revise, edit / banish, execute, expel	cuan4	WM, VSD
璗	Pure (gold, gems)	dang4	WM, VSD
魬	Flounder, sole (type of fish)	fan3	WM, VSD
龜	Turtle, tortoise	gui1	WM, VSD

Hanzi	Dictionary Definition	Hanyu Pinyin	Experimental Role
躒	Sore, ulcer / pestilence	li4	WM, VSD
廩	Granary, stockpile, supply (foodstuffs)	lin3	WM, VSD
蒱	Gambling game with dice	pu2	WM, VSD
濳	Conceal, hide / hidden, secret / latent, dive	qian2	WM, VSD
灝	Bean soup	hao4	VSD
護	Music	hu4	VSD
嘂	Scream, cry loudly, yell	jiao4	VSD
髡	Shear trees, shave hair	kun1	VSD
禠	Happy	si4	VSD
磑	Stone mill / grind / break apart	wei4	VSD

Hanzi	Dictionary Definition	Hanyu Pinyin	Experimental Role
髹	A kind of dark red paint, laquer	xiu1	VSD
猷	Plan, scheme, plot / way, path / draw, paint / like, similar	you2	VSD
鑿	Chisel. Dig, bore, pierce through / scuttle	zao4	VST
盦	Buddhist cloister	an3	VST
籌	Tally / chip / plan, prepare / raise money / assess, estimate	chou2	VST
鶨	A bird known as the stupid bird	dun4	VST
躉	A whole batch or amount / sell or buy wholesale	dun3	VST
薧	Dried food	kao4	VST
躒	Pebble, gravel, shingle	li4	VST
櫺	Carved or patterned window railings, sills	ling2	VST

Hanzi	Dictionary Definition	Hanyu Pinyin	Experimental Role
慔	Exert oneself	mo4	VST
繶	The leather binding of a shoe	yi4	VST
彘	Pig	zhi4	VST
籀	Deduce / a style of calligraphy	zhou4	VST
葄	Mat	zuo4	VST

APPENDIX D

INCORRECT VERBAL SUPPRESSION TRIAL EXCLUSIONS ON THE VISUAL WM CHANGE DETECTION TASKS

Table D1. Percentage of incorrect verbal suppression trials excluded from the visual WM change detection analysis: Experiment 7.

	Squares	*Upright unfamiliar faces*	*Upright famous faces*
Set Size 1	3.85% (37/960)	3.54% (34/960)	4.06% (39/960)
Set Size 2	4.38% (42/960)	4.48% (43/960)	5.94% (57/960)
Set Size 3	8.33% (80/960)	6.46% (62/960)	5.10% (49/960)
Set Size 4	3.85% (37/960)	5.52% (53/960)	7.92% (76/960)
Set Size 5	5.31% (51/960)	6.67% (64/960)	6.88% (66/960)
Set Size 6	5.42% (52/960)	6.04% (58/960)	8.13% (78/960)
Set Size 8	3.75% (36/960)	6.77% (65/960)	7.19% (69/960)
Set Size 10	4.17% (40/960)	6.46% (62/960)	5.73% (55/960)

Table D2. Percentage of incorrect verbal suppression trials excluded from the visual WM change detection analysis: Experiment 8.

	Inverted unfamiliar faces	*Inverted famous faces*
Set Size 2	8.96% (86/960)	8.65% (83/960)
Set Size 4	11.04% (106/960)	7.50% (72/960)
Set Size 6	11.98% (115/960)	10.83% (104/960)

Table D3. Percentage of incorrect verbal suppression trials excluded from the visual WM change detection analysis: Experiment 9a.

	Upright Hanzi: Experts	*Upright Hanzi: Novices*
Set Size 1	2.32% (13/560)	5.18% (29/560)
Set Size 2	1.79% (10/560)	4.11% (23/560)
Set Size 3	4.11% (23/560)	5.71% (32/560)
Set Size 4	7.50% (42/560)	5.00% (28/560)
Set Size 5	8.39% (47/560)	4.82% (27/560)
Set Size 6	8.75% (49/560)	4.82% (27/560)
Set Size 8	6.96% (39/560)	5.18% (29/560)

Table D4. Percentage of incorrect verbal suppression trials excluded from the visual WM change detection analysis: Experiment 9b.

	Inverted Hanzi: Experts	*Inverted Hanzi: Novices*
Set Size 2	4.64% (26/560)	3.04% (17/560)
Set Size 4	7.50% (42/560)	4.29% (24/560)
Set Size 6	7.32% (41/560)	1.79% (10/560)

APPENDIX E

FALSE ALARM RATES AS A FUNCTION OF SET SIZE - VISUAL WM CHANGE DETECTION TASKS

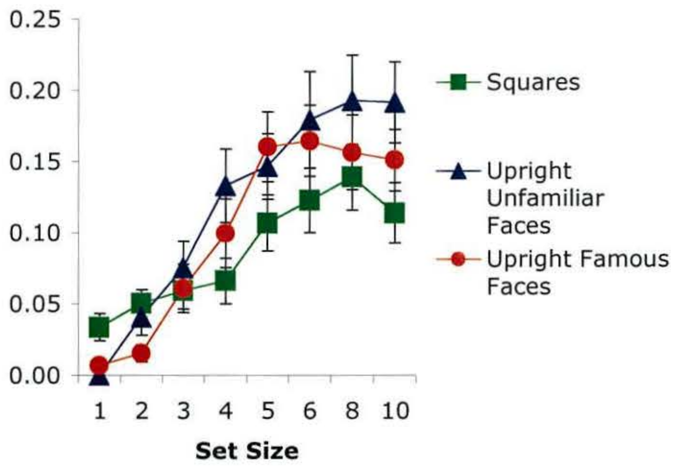


Figure E1. False alarm rate as a function of set size for squares, upright unfamiliar faces, and upright famous faces in Experiment 7. For each stimulus type, a repeated-measures ANOVA with set size (1, 2, 3, 4, 5, 6, 8, 10) as a within-factor revealed a significant main effect of set size: Squares, $F(7, 161) = 6.53, p < .001$; Upright unfamiliar faces, $F(7, 161) = 14.66, p < .001$; Upright famous faces, $F(7, 161) = 16.44, p < .001$.

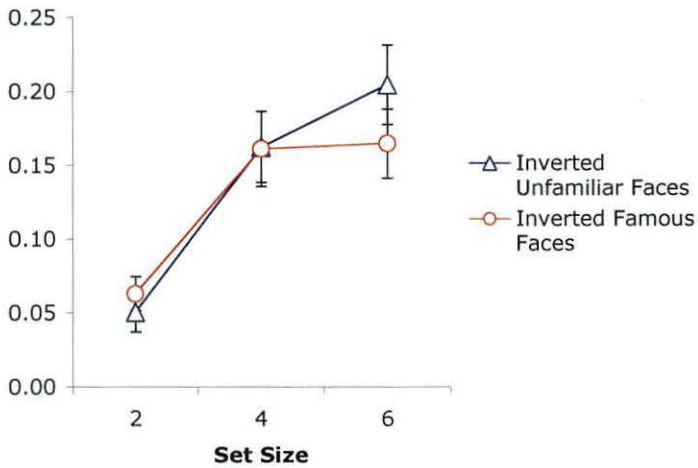


Figure E2. False alarm rate as a function of set size for inverted unfamiliar faces and inverted famous faces in Experiment 8. For each stimulus type, a repeated-measures ANOVA with set size (2, 4, 6) as a within-factor revealed a significant main effect of set size: Inverted unfamiliar faces, $F(2, 46) = 20.48, p < .001$; Inverted famous faces, $F(2, 46) = 15.23, p < .001$.

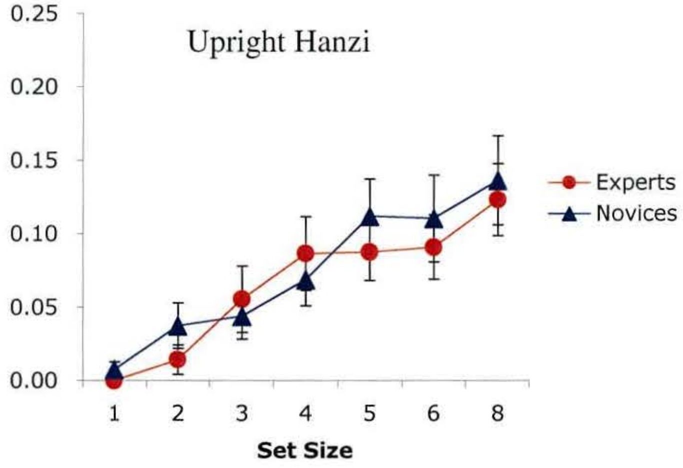


Figure E3. False alarm rate as a function of set size for upright Hanzi – experts versus novices – in Experiment 9a. For each group, a repeated-measures ANOVA with set size (1, 2, 3, 4, 5, 6, 8) as a within-factor revealed a significant main effect of set size: Experts, $F(6, 78) = 6.80, p < .001$; Novices, $F(6, 78) = 6.22, p < .001$.

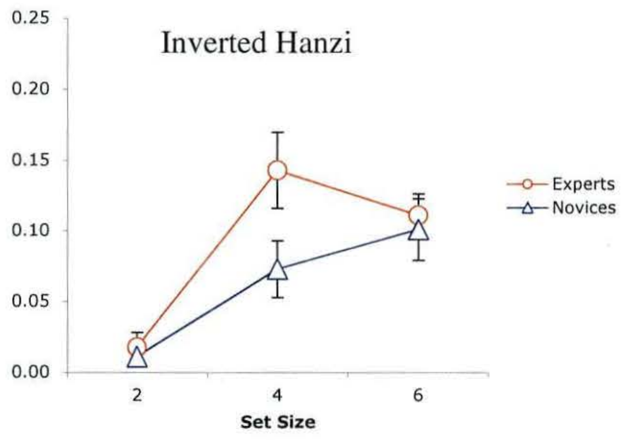


Figure E4. False alarm rate as a function of set size for inverted Hanzi – experts versus novices – in Experiment 9b. For each group, a repeated-measures ANOVA with set size (2, 4, 6) as a within-factor revealed a significant main effect of set size: Experts, $F(2, 26) = 19.91, p < .001$; Novices, $F(2, 26) = 11.28, p < .001$.

APPENDIX F

BANGOR *K*: A NEW FORMULA TO ESTIMATE VISUAL WM CAPACITY

Bangor *k*

The newly revised formula that I developed to estimate visual WM capacity is:

$$k = 10\text{^} \, [(y-b)/m]$$

where y = a threshold performance constant expressed in *d'* units (set arbitrarily here to a value of 3.45), and *b* and *m* are the intercept and slope respectively of a least squares line fit to *d'* values (from the change detection task) plotted as a function of logarithmic set size. An illustration of how the Bangor *k* formula performs is provided in Figure F1 below.

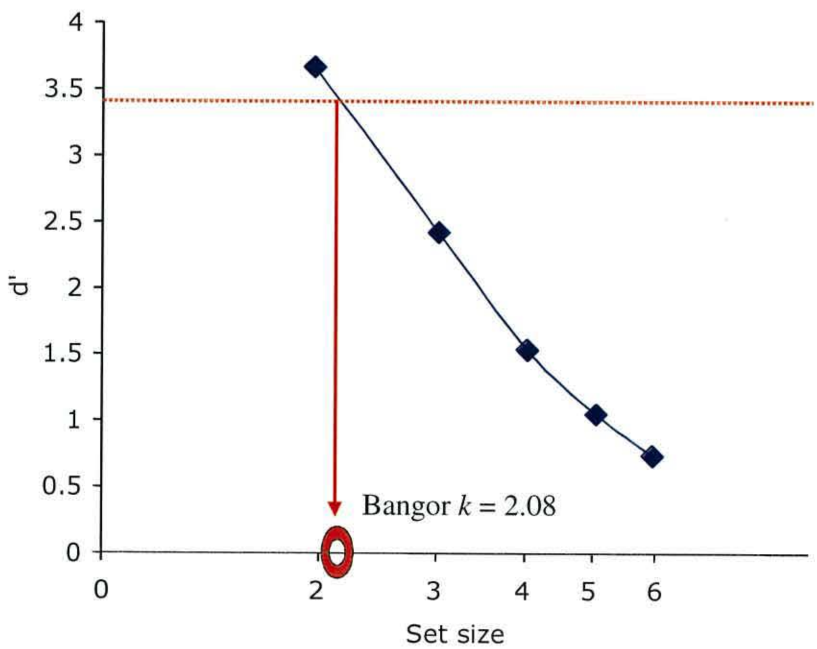


Figure F1. An illustration of the Bangor *k* formula – actual data from the unfamiliar upright faces condition in Experiment 7 are used here. *d'* values were plotted as a function of log transformed set size. The formula 10^ [(y-b)/m] calculates the value on the x-axis (set size) at which performance fall below a predefined threshold value of *d'* = 3.45. In this illustration, the capacity estimate was for 2.08 unfamiliar upright faces.

To calculate Bangor *k*, I used *d'* scores between set sizes 2-6. Although half of the visual WM experiments (reported in Chapter 6) employed a larger range of set sizes between 1-10, I observed that *d'* scores tended to asymptote from set size 6 onwards in most conditions where set sizes larger than 6 were used. To validate the observed asymptote I conducted paired-sampled *t*-tests on *d'* scores for each neighbouring set size. That is, I compared performance at set sizes 1 versus 2, 2 versus 3, 3 versus 4, and so on. For upright unfamiliar faces (set sizes 1-8 and 10, Experiment 7), I found significant differences between each set size comparison up to set size 6. The differences between set size 6 and 8, and between 8 and 10 were non-significant ($p > .1$). Thus I concluded that asymptote was reached at set size 6. For upright famous faces (set sizes 1-10, Experiment 7) asymptote was reached slightly later at set size 8. For squares (set sizes 1-10, Experiment 7), asymptote was not reached, that is performance at set size 10 was significantly lower than at set size 8. For upright Hanzi (set sizes 1-6 and 8, Experiment 9a), Hanzi experts did not reach asymptote; Hanzi novices reached asymptote at set size 6. Despite differences in asymptote between conditions, I decided to use set sizes no larger than 6 in order to ensure that Bangor *k* capacity estimates did not incorporate any flattening of *d'* performance at larger set sizes in any condition. I also chose to exclude set size 1 performance from the calculations as performance tended to reach ceiling across all conditions and might have skewed the line fit.

An additional advantage of selecting set sizes 2-6 was that Bangor *k* estimates for upright faces and Hanzi could be compared with inverted faces and Hanzi that were only presented in displays of sizes 2, 4, and 6 items. Thus, the range of set sizes used to calculate Bangor *k* was constant and comparable across all conditions.

After detailed analysis, I determined that least squares line fits to *d'* scores from set size 2 through 6 were better (that is, r^2 values were higher) when set size was log transformed than when linear values were used. This was true in all conditions but the squares condition. Table F1 shows the log and linear r^2 values for each visual WM experimental condition.

Table F1. r^2 values for linear and logarithmic line fits to *d'* scores between set sizes 2-6.

	Linear trend	*Log trend*
Squares (Expt. 7)	0.9569	0.8620
Upright Unfamiliar Faces (Expt. 7)	0.9336	0.9773
Upright Famous Faces (Expt. 7)	0.9631	0.9927
Inverted Unfamiliar Faces (Expt. 8)	0.8912	0.9658
Inverted Famous Faces (Expt. 8)	0.8708	0.9534
Upright Hanzi – Experts (Expt. 9a)	0.9252	0.9952
Upright Hanzi – Novices (Expt. 9a)	0.9503	0.9946
Inverted Hanzi – Experts (Expt. 9b)	0.8339	0.9290
Inverted Hanzi – Novices (Expt. 9b)	0.9316	0.9869

The y-constant *d'* value of 3.45 was chosen to reflect an overall performance accuracy threshold of 90%, the average of *d'* values obtained using hit rate probabilities between 0.90 – 0.99 and false alarm probabilities between 0.00-0.09 (see Table F2). The threshold can of course be modified, with a higher constant value producing more conservative estimates of capacity, and a lower value producing more lenient capacity estimates. The resultant value obtained from (y-b)/m was then translated back into a linear value by an inverse logarithmic function (10^).

Table F2. Illustration of the method for calculating a *d'* performance threshold value, used as the y constant in the Bangor *k* formula.

Hit rate	*False Alarm*	*d' equivalent*
0.90	0.00	3.61
0.91	0.01	3.67
0.92	0.02	3.46
0.93	0.03	3.36
0.94	0.04	3.30
0.95	0.05	3.28
0.96	0.06	3.30
0.97	0.07	3.36
0.98	0.08	3.46
0.99	0.09	3.67
		***M* = 3.45**

A final caveat to the application of Bangor *k* is that set sizes that exceed upper capacity limits must be included. If capacity limits are not exceeded, *d'* performance values are likely to remain above the specified performance threshold, and the formula would yield a capacity estimate of zero.

APPENDIX G

STATISTICAL ANALYSES USING VISUAL WM CAPACITY ESTIMATE *K*

Table G1. Statistical analyses output from paired and independent *t*-tests for each condition comparison using *k* values.

	Pashler's k	*Cowan's k*	*Bangor k*
Squares vs. upright unfamiliar faces	$t(23) = 6.10, p < .01$	$t(23) = 7.94, p < .01$	$t(23) = 3.39, p < .01$
Squares vs. upright famous faces	$t(23) = 5.44, p < .01$	$t(23) = 6.55, p < .01$	$t(23) = 2.33, p < .05$
Unfamiliar vs. famous upright faces	$t(23) = 1.20, p = .24$	$t(23) = .8, p = .43$	$t(23) = 2.38, p < .05$
Unfamiliar vs. famous inverted faces	$t(23) = 2.05, p = .05$	$t(23) = 2.54, p < .05$	$t(23) = 1.00, p = .92$
Upright vs. inverted unfamiliar faces	$t(46) = .86, p = .40$	$t(46) = 1.59, p = .12$	$t(46) = 2.03, p < .05$
Upright vs. inverted famous faces	$t(46) = .39, p = .70$	$t(46) = .34, p = .74$	$t(46) = 3.65, p < .01$
Upright Hanzi: Experts vs. novices	$t(26) = 2.15, p < .05$	$t(26) = 2.00, p = .06$	$t(26) = 3.01, p < .01$
Inverted Hanzi: Experts vs. novices	$t(26) = 1.24, p = .23$	$t(26) = 1.17, p = .25$	$t(26) = 1.62, p = .12$
Upright vs. inverted Hanzi: Experts	$t(13) = 1.63, p = .13$	$t(13) = 1.61, p = .13$	$t(13) = 2.53, p < .05$
Upright vs. inverted Hanzi: Novices	$t(13) = .04, p = .97$	$t(13) = .11, p = .92$	$t(13) = .66, p = .52$