SPATIAL AND TEMPORAL FACTORS AFFECTING HUMAN VISUAL RECOGNITION MEMORY

Daniel James Robertson, BSc.

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy

Apr 2007

Abstract

The current thesis investigated the effects of a variety of spatial and temporal factors on visual recognition memory in human adults. Continuous recognition experiments investigated the effect of lag (the number of items intervening between study and test) on recognition of a variety of stimulus sets (common objects, face-like stimuli, fractals, trigrams), and determined that recognition of common objects was superior to that of other stimulus types. This advantage was largely eradicated when common objects of only one class (birds) were tested.

Continuous recognition confounds the number of intervening items with the time elapsed between study and test presentations of stimuli. These factors were separated in an experiment comparing recognition performance at different rates of presentation. D-prime scores were affected solely by the number of intervening items, suggesting an interference-based explanation for the effect of lag. The role of interference was investigated further in a subsequent experiment examining the effect of interitem similarity on recognition. A higher level of global similarity amongst stimuli was associated with a lower sensitivity of recognition.

Spatial separation between study and test was studied using same/different recognition of face-like stimuli, and spatial shifts between study and test locations. An initial study found a recognition advantage for stimuli that were studied and tested in the same peripheral location. However, the introduction of eye-tracking apparatus to verify fixation resulted in the eradication of this effect, suggesting that it was an artefact of uncontrolled fixation. Translation of both face-like and fractal stimuli between areas of different eccentricity, with different spatial acuities, did decrease recognition sensitivity, suggesting a partial positional specificity of visual memory. These phenomena were unaffected by 180° rotation. When interfering stimuli were introduced between study and test trials, translation invariance at a constant eccentricity broke down.

Acknowledgements

I would like to extend thanks to all of those who have supervised me during the course of my PhD, Tim Ledgeway, Amanda Parker, Alex Easton, and Helen Cassaday. Their advice and support has made it possible for me to complete the PhD, and has been greatly appreciated. In particular I would like to thank Tim whose approachable and knowledgeable presence has been invaluable.

I am also grateful to others within the department who have helped me out with various aspects of my work. Thanks to Prof. Andrew Derrington, whose face-like stimuli and advice I have made use of, Dr. Chris Vincent, for his assistance with Matlab programming and the art of surviving a PhD, and Gareth Williams, for research assistance and good humour.

I would like to express gratitude to my family. Thanks to Mum, Dad, Ellie, and my grandparents, whose encouragement has helped me greatly during the course of the PhD. Finally, a very big thank you to my partner Tina, for intelligent advice and loving support where they were needed.

Declaration

This thesis is the candidate’s own original work and has not, whether in the same or different form, been submitted to this or any other University for a degree. The candidate performed all experiments and data analysis.

Table of Contents

Chapter 1 General Introduction

1.1 Recognition memory

The word “memory” is used in psychology to describe “an enduring change in behaviour, or in the behavioural potential, that results from the individual’s behavioural experience” and also encompassing “the retention over time of learned information” (Dudai, 2002, p. 157, p.157). It is a broad term that has been proposed to encompass a wide range of different stores and processes, which may or may not be closely related to each other. Memory likely consists of a number of different entities relying on different brain systems. Generally memory may be divided into two subcategories, explicit and implicit memory. Explicit tasks require conscious awareness of the material being retrieved, i.e. recollection of facts and events, known as declarative memory, whilst implicit tasks involve a priming stimulus and no conscious recall of experience (Squire & Zola, 1996). Declarative memory can be defined as the acquisition, retention and retrieval of knowledge that has been consciously and intentionally remembered (N. J. Cohen & Squire, 1980).

In addition to this subcategorisation, memory has been divided according to the nature of the material memorised. Episodic memory is the name given to the storage and recall of information about specific experiences, or episodes, e.g. that you had a letter in the post this morning. It is distinguishable from semantic memory, which consists of generalisable facts (Tulving, 1983), e.g. that the post arrives between 9 and 10 o’clock, and procedural memory, the memory for specific motor plans and actions. Episodic memory is memory for experienced events and usually involves the retrieval of perceptual information in specific spatiotemporal settings. This type of memory is commonly assayed by one of three methods: free recall, cued recall and recognition. Free recall involves the learning and subsequent uncued retrieval of lists of stimuli. Cued recall is similar, however retrieval is primed with cues, e.g. the first three letters of a word. Recognition tests involve simply identifying whether a stimulus has been perceived before or not.

Recognition literally means to know again. To recognise is to perceive something as previously known (Mandler, 1980), although psychological usage of recognition is usually restricted to judgements about the prior occurrence of events and, therefore, episodic memory. Psychological research into recognition memory aims to ascertain how people come to make judgements that an item or event has previously been encountered. Rather than simply being an individual's ability to identify an item as belonging to a particular semantic category (e.g. 'that is a cat'), recognition memory refers to an individual's ability to identify previous experience of that specific item (e.g. 'that is my cat').

Recognition memory can be divided further into two components: familiarity and episodic recognition. To give an example of this separation, imagine that you pass someone on your way to work, that you recognise that person as someone you have previously met, but cannot recall who they are or where you met them. This is the perceptual identification component of recognition memory, a judgement based on a feeling of familiarity. You "know" you have seen the person before but cannot remember the specific details about the event. Alternatively, imagine that you recognised the person's identity and the event during which you met them previously. In addition to the familiarity component, this type of recognition involves recollective matching, or "remembering" (M. W. Brown & Aggleton, 2001; Tulving, 1983). This is an example of episodic recognition, because the specific episode of the first encounter is recalled.

The first functional theory of two such processes was proposed by Kintsch (1970) but this view has largely been revised since then (see Clarys, 2001 for a review). Mandler's (1980) theory of activation proposes that memory is made up of representational units and proposes that there are two methods of retrieval: feelings of familiarity and conscious memory search. Sensory integration of stimulus features, or intra-event integration, gives rise to feelings of familiarity, the automatic reactivation of the stimulus representation. This is contrasted with the elaboration upon the event within a context or within another event that initiates a conscious search. This is proposed to be a controlled process requiring cognitive effort. The two processes are theorised to work in tandem in the recognition of a stimulus, but familiarity is assumed to be faster.

Jacoby and Dallas (1981) proposed a similar theory, but suggested that previous exposure to a stimulus led to perceptual facilitation, enhancing subsequent recognition (familiarity). Memory is assumed to make use of a combination of attention-demanding conscious processes with automatic processes. The balance would depend on the nature of the task with explicit tasks demanding greater use of conscious processes and implicit tasks making greater use of automatic processes.

Tulving (1985) studied the relationship between specific states of consciousness and different types of memory. He suggested that semantic, procedural and episodic memories could be characterised with different states of consciousness. Amongst these are autonoetic (self-aware) consciousness, in which a 'mental journey' is made by a person until an event may be recalled in its context (Wheeler, Stuss, & Tulving, 1997). Noetic consciousness denotes a state in which a person is aware of information but not its origin. Explicit episodic recall may thus be associated with autonoetic consciousness, whilst the experience of feelings of familiarity can be thought of as a noetic state of consciousness

(Gardiner & Java, 1993; Gardiner, Ramponi, & Richardson-Klavehn, 1999; Gardiner & Richardson-Klavehn, 2000).

Recognition memory, including both the capacity for perceptual identification of events, and the judgement of their prior occurrence (Aggleton & Brown, 1999; Mandler, 1980), is an essential component in memory as a whole. However, whilst the proposed subdivisions of memory are theoretically useful, there is no doubt that the systems are semidiscrete, with frequent interactions. Semantic memory, for example, presumably develops from individual episodes from episodic memory, which become generalised. For example, your knowledge that postboxes are red might come from the combination of many episodes in which you saw red postboxes, or perhaps an episode in which you were told that postboxes are red. Also, whilst it is often useful to view familiarity and episodic recognition within recognition memory as separate, they have been described as shallower and deeper levels of recognition memory. Some authors, basing their work on signal detection theory, suggest that there is just one memory system, and that supposed differences between familiarity and recollection are quantitative (Donaldson, 1996; Hirshman & Master, 1997; Inoue & Bellezza, 1998). Decisions are made according to the strength of the memory trace, the putative neural change that represents a memory, and the criterion set by the participant. However, neuroimaging studies (e.g. Duzel, Yonelinas, Mangun, Heinze, & Tulving, 1997; Yonelinas, Otten, Shaw, & Rugg, 2005) suggest that familiarity and episodic recollection are served by discrete brain regions. Yonelinas (1994) studied the Receiver Operating Characteristic (ROC) curve, produced with hit and false alarm rates from signal detection theory analyses of memory experiments (see Green & Swets, 1966), and proposed that conscious recall is an ‘all or nothing’ process which could either successfully recall the context of encoding or not. The hit rate is the proportion of responses to previously seen stimuli that are correctly recognised, and the false alarm rate is the proportion of novel stimuli that are incorrectly assumed to be old. Responses based on familiarity were assumed to be more variable and to depend on the criterion for decision adopted by the participant.

To summarise, recognition is a component of episodic, declarative memory that allows the discrimination of familiar and novel experiences. Whilst, historically, there has been some debate over whether recognition memory may be further subdivided into separate familiarity-based and episodic recall components or not, recent evidence from neuroimaging supports the notion of two separate processes.

1.2 The continuous recognition paradigm

Human and animal experiments aiming to test recognition memory have used a large number of different tasks, with a corresponding variety of strengths and limitations. The most common basic design involves the presentation of a list of items for memorisation, and a subsequent test of the participant's ability to discriminate between items that have been seen before and those that are entirely novel. This may take the form of a single long study list, followed by a single long list from which items are recognised (e.g. Strong, 1912), or a series of short lists interspersed by repeated or novel probe stimuli requiring "old" (recognised) vs. "new" (not recognised) discriminations (e.g. Zhou, Kahana, & Sekuler, 2004). An inevitable factor to be taken account of when examining the results obtained from such memory tasks is serial position. An item's position within a list has a significant effect on subsequent recognition, in addition to any other manipulated factors, resulting in primacy (the tendency for items at the start of a list to be remembered better than others) and recency (the tendency for items studied most recently to be remembered better than others). These produce a characteristic U-shaped curve when accuracy is plotted against serial position, an effect that is especially pronounced in long lists, but is observed even when there are as few as 4 items in the list (Korsnes, Magnussen, & Reinvang, 1996; see Figure 1.1 for an example; Wright, Santiago, Sands, Kendrick, & Cook, 1985). Serial position must, therefore, be accounted for before the inference of further effects of manipulations of the stimuli.

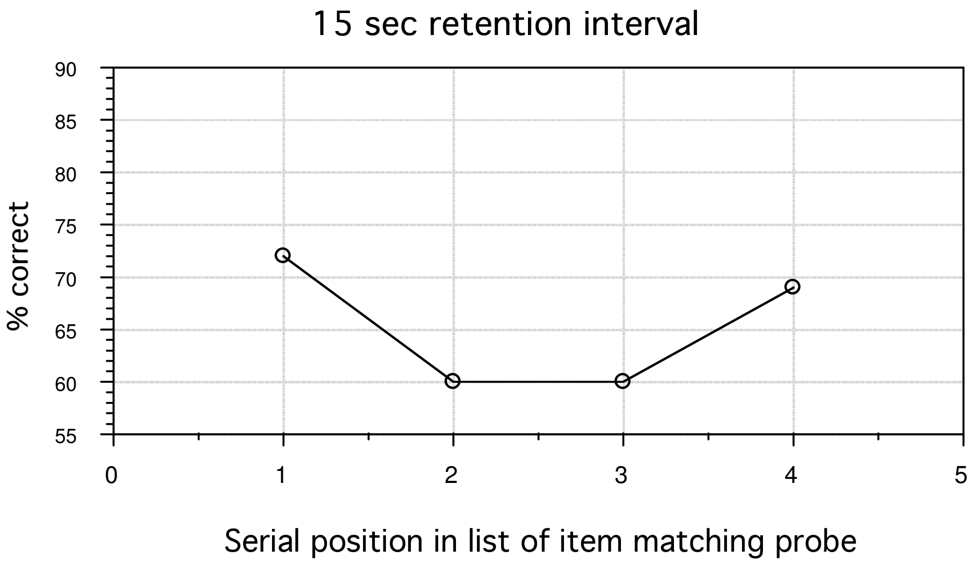

Figure 1.1: An example of a 'U-shaped' serial position curve. When a probe for recognition matched the first (1) or last (4) item in a list it was recognised more accurately than for recognition at either of the middle list positions. Figure adapted from Korsnes et al. (1996).

An alternative method of assessing recognition memory that avoids the serial position effect is the continuous recognition paradigm, introduced by Shepard and Teghtsoonian (1961). Their aim was to circumvent some of the limitations discussed above, and to provide a method for measuring recognition under “conditions approaching a steady state” (Shepard & Teghtsoonian, 1961). Whereas previous methods involved learning an isolated block of information, e.g. a word list, and subsequently retrieving as much of the information as possible, the continuous recognition task presented participants with a continuing sequence of information, and required their retention of that information throughout the course of the experiment by interleaving study and test trials. This procedure aimed to minimise the possibility of rehearsal of the material for retention, and maximise interference from previous material.

In Shepard and Teghtsoonian’s experiment, participants were required to turn over cards in a specially prepared pack one at a time. Each of the 200 cards had a three-digit number printed on it, and each number occurred twice in the pack. The participant was required to respond to each card as it was presented, indicating whether they thought the number had not been seen before (“new”), or that it had been seen previously (“old”). The lag, or number of intervening cards between repeats, was manipulated by the experimenter. The ability to manipulate the lag as an independent variable is an important aspect of the procedure, and allows the comparison of recognition after different study-test intervals and different levels of interference, without the confounding influence of the primacy and recency effects observed in list learning.

Shepard and Teghtsoonian made several important findings in their initial experimentation. In a finding that has been much replicated since, they discovered that probability of recognition of items was dependent on lag. Recognition accuracy dropped rapidly from almost 100% from lag 0 (no intervening stimuli) to lag 6, before assuming a much more gradual decay, regardless of position within the task. The experimenters also discovered that, in spite of their hopes of studying recognition under a steady state, this was never totally achieved during their experiments. The probability of participants making false alarms was still increasing gradually after 200 trials, although after a certain point the rate of increase was marginal (after 40 trials the increase in probability occurred at a rate of 0.0007 per card). The curves obtained by plotting recognition accuracy against lag suggested that memory traces decayed over time, and became decreasingly stimulus-specific.

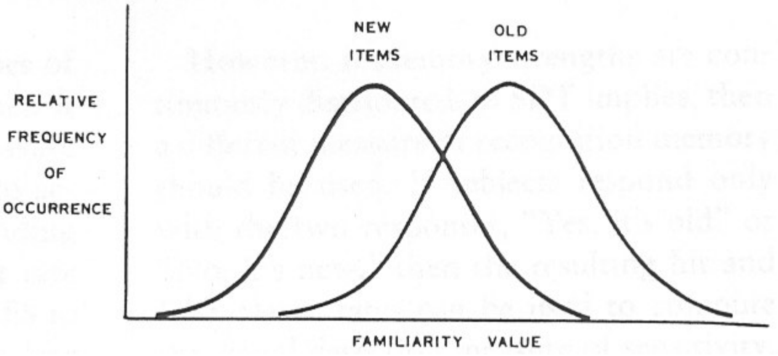

Figure 1.2: Theoretical distribution of familiarity values for old and new items predicted by signal detection theory. Figure reproduced from Swets (1964).

One of the advantages of the continuous recognition memory test is that the results are interpretable using signal detection theory, allowing *d'*, a measure of sensitivity, to be calculated in addition to the more traditional measure of proportion of correct responses. Noise, sensory input that is not part of the signal, is assumed to vary randomly over time and have a normal distribution. Signals (in the case of continuous recognition, old stimuli) add a fixed amount of sensory input to the noise present in a trial so signal plus noise is also normally distributed (Figure 1.2). Sometimes the signal is indistinguishable from noise because the two distributions overlap. In this case, the participant must adopt a criterion for making an old/new decision, based on the level of sensory input. This will inevitably result in errors in the form of false alarms (mistakenly registering a new stimulus as old) and misses (mistakenly registering an old stimulus as new). By calculating the Z-scores for hit and false alarm rates it is possible to obtain *d'* scores, a measure of discriminative ability independent of bias, for participants. The *d'* score is the difference between the noise mean and the signal + noise mean in standard deviation (Z-score) units.

The advantages of the continuous recognition paradigm have been put to extensive use since 1961, in all manner of memory studies. A number of continuous recognition-derived tests are used in neuropsychological assessment of human memory including the Continuous Recognition Memory Test (CRM; Hannay & Levin, 1988) and Continuous Visual Memory Test (CVMT; Trahan & Larrabee, 1988). These tests both employ a blocked design in which certain items from the first block recur in each of the remaining 6 blocks, and must be discriminated from categorically or perceptually similar items. Whilst both tests employ drawings as stimuli they differ in that the CRM uses drawings of living things whereas the CVMT uses abstract drawings. The CRM was first developed as a means to assess memory deficits in patients with closed-head injuries, as it is possible to administer to patients who are incapable of making written or multiword verbal responses. Performance is not related to age, sex or educational level (Hannay & Levin, 1989; Hannay, Levin, & Grossman, 1979), and the test has the advantage of being interpretable in terms of the theory of signal detection (Drake & Hannay, 1992). Factor analysis of the results of the CRM, administered as part of a battery of neuropsychological tests, found that correct recognition (hits) had a significant loading on a “learning and memory factor”, whereas CRM false alarms loaded on an “attention to visual detail” factor (Fuchs, Hannay, Huckeba, & Espy, 1999). A similar analysis of the CVMT found that *d’* scores were associated with “verbal“ and “visual/nonverbal” intellectual factors as well as attentional ability (Larrabee, Trahan, & Curtiss, 1992). Research has demonstrated the usefulness of continuous recognition memory tests in the study of closed-head injury (Brooks, 1972; , 1974a; , 1974b; , 1989; Hannay, Levin, & Grossman, 1979), dementia (Cutting, 1978; E. Miller & Lewis, 1977), alcoholic Korsakoff’s disease (Cutting, 1978; W. Riege, 1977), aphasia (W. H. Riege, Klane, Metter, & Hanson, 1982), and unilateral temporal-lobe lesions (Cutting, 1978; De Renzi, 1968; Kimura, 1963; Trahan & Larrabee, 1988).

Studies using continuous recognition in non-human primates have been informative in elucidating the neural circuits and brain regions underlying recognition memory (Gaffan, 1977), and have identified “lag-sensitive” neuronal populations (Brozinsky, Yonelinas, Kroll, & Ranganath, 2005; Xiang & Brown, 1998). The task is also being used to good effect in combination with neuroimaging to understand various neurological and psychiatric disorders, including obsessive-compulsive disorder (M. S. Kim et al., 2005), Alzheimer's disease (Joray, Herrmann, Mulligan, & Schnider, 2004), schizophrenia (M. S. Kim, Kwon, Kang, Youn, & Kang, 2004) and epilepsy (Guillem, N'Kaoua, Rougier, & Claverie, 1998). Most commonly, however, studies have investigated recognition in normal participants using standard English words as stimuli, and a wide range of factors and phenomena have been investigated. The influence of lags of various sizes, repetition (Hintzman, 1969; Hockley, 1982; Jessen et al., 2001; Okada, 1971; Van Strien, Hagenbeek, Stam, Rombouts, & Barkhof, 2005), feedback and base rates (Estes & Maddox, 1995a), and different types of stimuli (Doty & Savakis, 1997; Estes & Maddox, 1995b; Lehmann & Murray, 2005; Nickerson, 1965) have been investigated. Attempts have been made to investigate the neural mechanisms underlying recognition in combination with neuroimaging (Coney & MacDonald, 1988; Jessen et al., 2001; Van Strien, Hagenbeek, Stam, Rombouts, & Barkhof, 2005), in attempts to develop cognitive models of recognition separating familiarity and explicit recollection (Hockley, 1992; Jones & Atchley, 2002; B. B. J. Murdock & Hockley, 1989; Reder et al., 2000). Continuous recognition has also been used recently in order to attempt to very precisely model retention (Rubin, Hinton, & Wenzel, 1999). The findings of studies using continuous recognition with normal human participants are reviewed in the following sections.

1.2.1 Repetition

In the 20 or so years following the introduction of the continuous recognition paradigm, a number of researchers used continuous recognition to explore proposed theories and models of memory. By manipulating the lags used and modifying the paradigm to include stimuli occurring more than twice, resulting in more than one test trial, investigators hoped to describe the nature of recognition memory more precisely and thus determine whether previous models could be applied to this type of memory. In three highly similar studies, first Hintzman (1969), then Okada (1971), and then Hockley (1982), employed variations of the continuous recognition paradigm in which stimuli recurred two times (3 presentations in total) and, therefore, two lags could be manipulated. Naming the presentations P1, P2, and P3, one lag occurred between P1 and P2 (lag i), and one occurred between P2 and P3 (lag j). By including more than one repetition of items these studies revealed information about the effects of memory trace strength on recognition.

Hintzman (1969) employed high-frequency words (those with a frequency greater than 30 per million according to Thorndike and Lorge (1944) in a continuous recognition task where accuracy was near perfect (average 96%). Reaction time was measured as the dependent variable, and the effects of latency, frequency and spacing of repetitions on this variable were determined. All of the experimental words used in the experiment occurred three times with a fixed lag j of 16, and a variable lag i (1, 2, 4, 8, or 16). The major findings of this study were that error latencies were consistently longer than those for correct responses at each of the three presentations, that P2 recognition times increased with increasing length of lag, and that recognition times for the second repeat (P3) were shorter than those for the first (P2). P3 recognition times were also increased by increasing lag i. These findings made it clear that reaction times were dependent on lag even in the absence of significant variation in accuracy, and that repetitions of items facilitated the speed of recognition, as might be expected.

Okada (1971) was interested in what processes were involved in continuous recognition, and whether they were comparable to the serial comparison processes theorised to be involved in other memory tasks (Sternberg, 1966). Sternberg's theory explains recognition as a process of comparing the current stimulus with each of the previous stimuli one at a time. Okada's first experiment used a similar paradigm to Shepard and Teghtsoonian (1961) with only one repetition of stimuli, and lags of 0-7. Stimuli were common English words. Mean response latencies were found to increase exponentially with increasing lag. This appeared to suggest that continuous recognition relies on neither an exhaustive nor a self-terminating serial comparison process, as these would predict reaction times to be independent of lag or have a linear relationship with lag, respectively. Okada theorised that, instead of serial comparison, the trace strength of items was critical in determining the reaction time for recognition responses, since trace is hypothesised to decay exponentially with increasing lag. In a second experiment where words appeared once, twice or three times, reaction times were faster for P3 presentations than P2. Okada saw this as further support for the trace strength hypothesis as more repetitions would be expected to increase the strength of the memory trace.

Hockley's (1982) study utilised much longer lags than those investigated by either Hintzman or Okada, up to a maximum of 40 intervening items. Hockley used a noun/non-noun manipulation with the word stimuli employed, both between- and within-lists. Lists were made up of either nouns or non-nouns (between-list manipulation) or were made up of a mixture (within-list), to determine whether the two classes of word would affect reaction times. The effect of both manipulations was to produce a difference in the slope of the reaction time versus lag curve, without affecting the intercept. This suggested that retrieval time in this experiment was largely independent of the composition of word lists, as both pure- and mixed-lists gave rise to comparable measures of recognition. Hockley concluded that the effect of item repetition is more than a simple accumulation of memory strength, and suggested two models – multiple-observations (Pike, Dalgleish, & Wright, 1977) and diffusion process models (Ratcliff, 1978) – derived from signal detection theory, that were better able to account for the data.

Ratcliff's model proposes that each stimulus is simultaneously compared with all items in the memory search set, and a decision is made on the basis of 'relatedness' – the degree of similarity between an item in the memory and the probe item in the experiment. Relatedness is presumed to decrease over time, and correctly predicts the decreases in reaction time at longer lags, although the form of this relationship is not predicted. In the multiple-observations model, items are encoded as a set of features summed as a whole. The strength of the correlation between the probe item and the memory trace of previously observed items forms the basis of any recognition decision. The output of the correlation varies giving normal distributions for both members and non-members of the memory list. The difference between the means of the two distributions is assumed to be *d'*.

More recently, Jessen et al. (2001) used event-related functional magnetic resonance imaging (fMRI) to examine the encoding and retrieval processes involved in continuous recognition with two repetitions of stimuli. They found evidence of spatial differentiation of responses to encoding (study stimuli) and recognition (test stimuli) across a number of brain regions. In addition, frontal cortex activity was greater for the first repetition of a stimulus than for the second, and was ascribed to retrieval effort. This is evidence that the greater ease of recognising items a second time is related to altered neural processes in brain regions associated with memory.

A study measuring the event-related potentials (ERPs) of participants carrying out a continuous recognition task, involving 10 exposures to word stimuli, similarly found evidence of differences between activity for repeated and novel stimuli, and between stimuli repeated once and several times (Van Strien, Hagenbeek, Stam, Rombouts, & Barkhof, 2005). Increased positivity of the electroencephalogram (EEG) trace during the period 300-500ms after the stimulus was shown was found during correct recognition of an “old” word, in contrast to the trace for correct identification of a “new” word. In addition, correct recognition after increasing repetitions of stimuli correlated with increasing positivity in the 500-800ms period. This was accompanied by a linearly decreasing magnitude of induced delta power in the 375-750ms time window, correlated with greater speed and accuracy of behavioural responses. The authors cited these effects as evidence for a dual-processing model of recognition, suggesting that the “graded recollection state” could be dissociated from familiarity.

1.2.2 Stimulus attributes, base rates and feedback

Estes and Maddox (1995a) investigated the effects of several novel methodological adaptations of the continuous recognition procedure, giving some insight into how participants performed the task. They discovered that, when accurate feedback relating to the correctness of responses was given, the false alarm rate tended towards the old-new stimulus base rate of the experiment, suggesting that participants were manipulating the ratio of old:new responses according to the actual ratio of old:new trials. Without feedback false alarm rates remained independent of base rate. Estes and Maddox also used three different types of stimuli, with different levels of interitem similarity. Random digit triads (such as those used by Shepard & Teghtsoonian, 1961) were used for their high interitem similarity, random consonant trigrams had intermediate similarity, and common English words were used as examples of items with low similarity. Two experiments were carried out in which base rates were either 33% or 67% old stimuli, and informative feedback was either given or not given.

The most robust finding was that all of the response measures were dependent on the type of stimulus used, with the false alarm rate increasing from words to letters to digits, and therefore with increasing interitem similarity. An inverse effect was observed for hit rates, which decreased with increasing similarity. This would appear to demonstrate that increasing interitem similarity makes stimuli less memorable, and harder to discriminate from one another. Stimuli that are highly similar are less distinctive and there is higher global similarity between them. In addition, increased similarity of stimuli is likely to result in increased interference from stimuli intervening between study and test presentations. The false alarm rate for digit and letter stimuli, but not words, was dependent on the base rate when informative feedback was given. It could be suggested, therefore, that participants recognising more similar stimuli rely on feedback, at some level, to set the ratio of old:new responses that they make. This was not observed in the much easier word recognition task, perhaps as a result of the participants' greater confidence in their decisions.

These findings provide clear evidence that words are a very different class of stimuli to meaningless letter and number strings. Old-new decisions appear to be made with much greater accuracy when using word stimuli, although it is not clear from Estes and Maddox's study whether these findings are due to a lower interitem similarity, and subsequently greater discriminability, or whether participants' greater familiarity with these stimuli is responsible.

Interitem similarity was also studied by Raser (1972) who used words that were similar, either orthographically or acoustically, to study items as lures. The number of intervening stimuli between the study and the lure was manipulated in a manner similar to the lag in a classic continuous recognition experiment. The study determined that both types of similarity were responsible for increasing the false alarm rate, and that the effects were additive. The lag functions for the different conditions were similarly shaped, with false alarm rates rising rapidly at short lags and then reaching asymptote. A greater overall effect of orthographic similarity was noted than acoustic similarity, as many participants did not make any false alarms for acoustically similar stimulus pairs at all. For those who did, the effect was comparable to that of orthographic similarity.

1.2.3 Familiarity and explicit retrieval

An ongoing debate in recognition memory centres on the issue of whether there is more than one process involved. A large amount of evidence has accumulated in support of a dual process model of recognition (Jacoby & Dallas, 1981; Kintsch, 1970; Mandler, 1980; Tulving, 1985; Yonelinas, 1994), in which recognition may be based either on explicit recall of the encoding episode, or on feelings of familiarity.

One manner of conceptualising the dual process theory, is by assuming that familiarity can be equated with memory for items, whereas episodic recollection requires associative memory connecting events (Hockley, 1982). Mandler (1980) suggested that immediate recognition is based on item familiarity, and that as this information is lost through decay or interference, recognition becomes increasingly dependent on associative information. Murdock and Hockley (1989) found no diminution of forgetting rates for continuous recognition of associated pairs of items, even up to test lags of 26. This is surprising when compared with the results of studies examining forgetting for single items in a similar procedure (e.g. Hintzman, 1969; Hockley, 1982), which have determined a clear reduction in memory performance as lag increases. Hockley (1992) compared forgetting rates for item and associative information, by comparing continuous recognition for single words with that for associations between random word pairs (e.g. *forest-singing*), using both yes/no and 2-alternative forced choice procedures. The forgetting rates for word pairs were lower than those for single items, suggesting a distinction between these two types of memory.

Hockley explained the findings by assuming that the encoded information for associations between items is more distinctive than that for the items themselves. Mandler’s (1980) dual-process theory assumes that recognition of associative information involves a recall process, a controversial assumption with some evidential support (Humphreys, 1978; Humphreys & Bain, 1983; Mandler, 1980), but also with evidence against it (Gillund & Shiffrin, 1984; Gronlund & Ratcliff, 1989). The global matching models preferred by Hockley (Hintzman, 1984; B. B. Murdock, 1982; Pike, 1984; Raaijmakers & Shiffrin, 1981) do not incorporate this controversial recall process into recognition, preferring a matching of the test probe with memory representations and comparison of the resultant strengths of matches. The models can differentially emphasise item and associative information at encoding and recall, but the results of Hockley (1992) pose a problem for these models, as none has a mechanism that could differentially interfere with item and associative information resulting in the forgetting rates observed. However, item recognition involves more than just the discrimination of old events from new events as the items studied (words) have pre-experimental levels of familiarity. Their recognition, then, must be achieved in a context-dependent manner, excluding previous exposure to the words. Associations between words in word pairs are usually completely novel with no such pre-experimental familiarity. The MINERVA 2 global matching model (Hintzman, 1988) partitions item vectors into content and context elements. Item information thus consists of content and context features, whereas association information could consist primarily of content features. As the context changes between study and test one would expect a greater detriment to encoding based on both content and context compared with that based on content alone.

Reder et al. (2000) investigated familiarity and episodic recall in a continuous recognition-type procedure, in which participants had to classify whether their recognition judgements were based on specifically remembering the item ('remember') or whether they just knew that the item was familiar ('know'). In one experiment this discrimination was made instead of an 'old' response, and in another it was made following the 'old' response. Word stimuli occurred 1, 3, 5 or 10 times and were of either high frequency (normative mean frequency of 142) or low frequency (normative mean frequency of 1.6). Reder and colleagues were investigating the word frequency effect; the finding that high and low frequency word stimuli have different effects on both the hit rate (greater for high than low frequency words) and false alarm rate (also greater for high frequency words). The researchers found a significant main effect of word frequency (p<0.01), with low frequency words 'remembered' much more often than high frequency words. Fewer false 'remember' responses were encountered for low than high frequency words, but there were more 'know' responses for high frequency words.

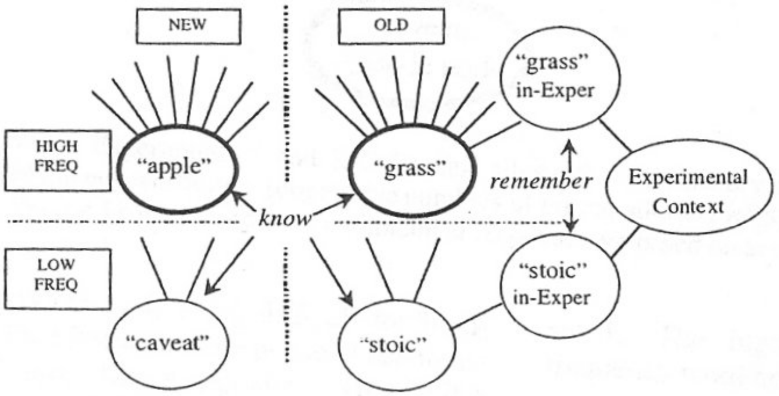

Figure 1.3: Schematic illustration of the differences between representations of high- and low-frequency words in recognition experiments. Figure reproduced from Reder et al. (2000).

Reder and colleagues explained these effects with their Source of Activational Confusion (SAC) model. This model assumes one memorial node to represent each word, and another to represent the encoded memory event for the word. Memory strength is represented by the level of activation of the word node, and the base level of activation is determined by previous experiences of the word, their frequency and recency. The base level of activation affects the general familiarity of a word, and explains why words of high frequency are more familiar to participants. The model also explains why the false alarm rate is higher for this class of words. For low frequency words there is a smaller ‘fan’ of contextual nodes associated to the word node making the experimental encoding event easier to retrieve, and contributing to the greater accuracy of ‘remember’ responses for these words (Figure 1.3). The explanation of the word frequency effect offered is certainly a satisfactory one, seeming to provide a mechanism for Mandler’s (1980) idea of a “counter” for familiarity, and a manner of combining item and context information in the manner suggested by Hockley (1992). Whether neural correlates of the proposed word and contextual nodes exist is an interesting question for future research.

Jones and Atchley’s (2002) examination of feature and conjunction error rates in continuous recognition did not find strong evidence for a dual-process account of recognition. The study used compound parent words (e.g. blackmail, jailbird) and conjunction lures (blackbird), presented at set lags after the parent, predicting that, according to the dual process model, familiarity and recollection should be placed in opposition to one another. Feelings of familiarity with components of the parent words would lead participants to make conjunction errors (false alarms for the novel conjunction lures) whereas explicit recollection of parent words would lead participants to correctly reject the novel conjunctions. The authors, drawing on the previous work of Hockley (1992) and Raser (1972), made the prediction that, under a dual process approach, recent information should be highly familiar but accessible by recollection. As such, they expected low conjunction error rates for short lags where conscious recollection was likely to succeed, but that the rate of conjunction errors would rise at longer lags. In fact, the opposite effect was observed in their first experiment, where conjunction error rate was highest at a lag of 1 and then decreased slightly as lag increased to 5. This pattern of results could be explained with reference to a simple one process familiarity-based model. Results from a second experiment, in which participants were instructed to use recollection to avoid making conjunction errors, gave some support to the dual-process theory. Error rate was lowest at a lag of zero, the shortest lag. Once again, however, the simple familiarity model fitted the pattern of results at greater lags. The poor recollective ability of participants was explained with reference to the divided attention thought to be produced by the continuous recognition task.

1.2.4 Modelling retention

While most of the studies detailed above have attempted to explain specific memory phenomena with manipulations of the continuous recognition paradigm, Rubin et al. (1999) were interested in mathematically modelling the retention process, without manipulation, as precisely as possible. The rationale behind elucidating a retention function was to enable practical estimation of the time period for retention of material, and also to inform psychologists about the nature of the retention process.

The experiments were carried out on a large number of participants (100 in each condition), over a wide range of lags (0, 1, 2, 4, 7, 12, 21, 35, 59 and 99), and with many repetitions of each lag (27 for recognition), giving a total of 600 trials per participant. The lags were chosen because they provided a wide range of points spaced evenly on a logarithmic scale. The large numbers of trials and participants ensured that the data collected were reliable, and that accurate curve-fitting could be carried out. There were two conditions for recognition, a standard continuous recognition paradigm and a remember-know adaptation, similar to that described for the Reder et al. (2000) study above. The stimuli used were digit-letter-digit trigrams. They could occur in any of the outer positions of a screen divided into a 3x3 matrix, with the middle position reserved for feedback. The data collected were then fit to a large number of functions, and from these a sum of exponentials was selected as the best fit to the data (Figure 1.4). The function, $y = a_1e^{-t/T1} + a_2e^{-t/T2} + a_3e^{-t/T3}$, described three components of the curve, defined by exponentials of the time constants T1, T2 and T3. T3 was infinite and T1 was set at 1.15, but the best fitting T2 value was different for old-new recognition (13.38) compared with 27.55 for remember and remember+know responses. With these parameters a curve with an r^2 value (goodness of fit) of .998 was obtained for remember judgements, and 0.996 for old-new recognition. The function fitted the data very well, although, given the 6 free parameters, this may not be too surprising.

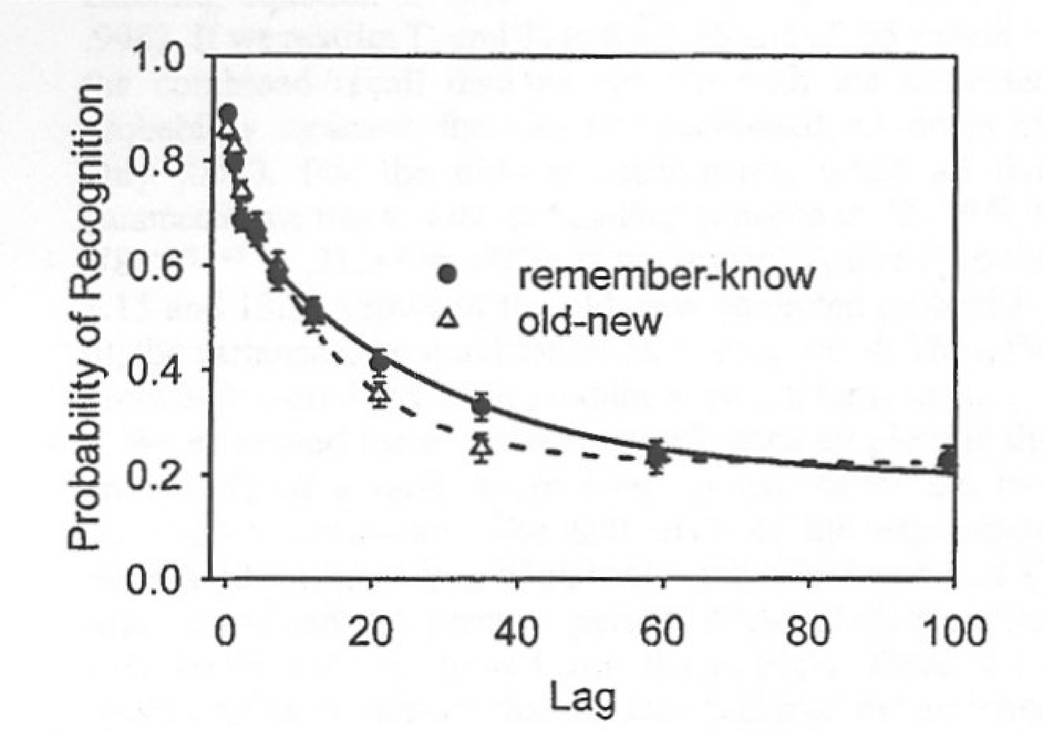

Figure 1.4: The probability of recognition for words as a function of lag, for both old/new, and remember/know/new judgements. Figure reproduced from Rubin et al. (1999). Probability of recognition = (hits - false alarms) / (1 - false alarms).

Rubin and colleagues argued that their results were the most precise data yet to be obtained for retention, and with such large numbers of participants and repetitions this seems a fair assumption. Certainly this precision allowed the authors to discriminate between functions to a greater degree than had been possible in their previous review of 100 years of forgetting data (Rubin & Wenzel, 1996).The function held true for several data sets and for both grouped and individual data. When tested using data from a study of implicit tasks (McBride & Dosher, 1997), such as cued-recall and stem-completion, the function fit the data with r^2 values in the range .86 to .99, suggesting that the retention function could be generalised to these types of task also.

Rubin et al. (1999) discussed some of the theoretical implications of the findings and suggested that the three components of the equation might correlate with different components of memory. The T1 component was proposed to describe working memory, consistent with a 2 sec phonological loop (A. D. Baddeley, 1997). In addition, the authors made the more controversial assumption that the other two components described either a divided long-term memory or intermediate- and long-term memory stores. They cite behavioural (Bahrick, 1965; Ericsson & Kintsch, 1995) and biological evidence (Ng et al., 1991; Rosenzweig, Bennett, Colombo, Lee, & Serrano, 1993) to support such a division, but caution must be made when inferring that the three-component descriptive function is sufficient evidence to warrant such a re-division of memory. A three-component curve does not necessitate a three-component memory. Nonetheless, the debate over the nature and time-course of consolidation of memory remains pertinent, and offers the prospect of interesting future research (see Dudai, 2004; Wixted, 2004 for reviews). The modelling has provided a deeper insight into the time course of retention, and introduced a very rigorous and precise manner of studying recognition.

1.2.5 Neural mechanisms

Whilst the above studies largely approached recognition from a cognitive viewpoint, the use of neuroimaging in combination with traditional tasks has allowed inferences to be made about the brain regions involved in different aspects of memory. Coney and Macdonald (1988) investigated hemispheric differences in recognition by presenting stimuli to either the left or right visual field, thereby confining presentations to visual processing by one cerebral hemisphere or the other. Subsequent test presentations were either crossed (in the opposite field to the study presentation) or uncrossed (in the same field). The aim of the experiment was to determine whether hemispheric differences, for example the superior ability of the left hemisphere to process verbal material, contributed to lateral asymmetries in memory performance, and over what time periods these asymmetries could be observed. Word stimuli were projected on a screen for 150ms, either to the left or right of a central fixation point.

No main effect of the visual field to which the stimulus was initially presented was found in reaction time data. However, the effect of the visual field to which probe items were presented was significant, with faster reaction times when words were presented in the right visual field. There was also a significant interaction between the target and probe visual fields with congruent target and probe visual fields (i.e. presentation for target and probe in the same visual field) producing faster correct reaction times than incongruent fields (target and probe in different visual fields). Results for different lags suggested that hemispheric interaction varied significantly over time. At a retention interval of 3 sec (lag 1) no differences were observed between the 4 presentation conditions, suggesting that perceptual matching accounts for responses, as opposed to comparisons with memory. After 12 sec (lag 4) left hemispheric processing was clearly dominant, probably as a result of the left hemisphere’s superior verbal processing abilities. The effect of crossed presentations only diverged significantly from uncrossed presentations at intervals of 32 sec (lag 8) and 96 sec (lag 32).

The results implied that memory traces were initially generated in both brain hemispheres in response to stimulus presentation, as no difference in recognition between crossed and uncrossed conditions was found, until a latency of 32 sec. These representations are likely to differ, however, in terms of their overall level of activation, with activation likely to be weaker in the indirectly-activated hemisphere than that directly activated (contralateral to the visual field of presentation). Coney and Macdonald suggested two possible explanations for later hemispheric asymmetries. First, they suggested that decay of the two representations may have occurred to such an extent that the difference between direct and indirect traces affected retrieval time. Alternatively, once a trace had decayed to a critical level it may no longer have been effectively retrieved at all.

The fMRI study of Jessen and colleagues (Jessen et al., 2001) mentioned previously, found evidence of different activation patterns in participants when comparing encoding and recognition. Recognition of a test stimulus was associated with stronger activation of left parahippocampal and inferior frontal gyri than the initial study presentation of the same stimulus. These findings are consistent with the literature on amnesia and animal models of amnesia, which have found evidence of parahippocampal (Aggleton & Brown, 1999; Buffalo et al., 1999; Meunier, Bachevalier, Mishkin, & Murray, 1993) and frontal lobe (Bachevalier & Mishkin, 1986; Janowsky, Shimamura, Kritchevsky, & Squire, 1989; Owen, Sahakian, Semple, Polkey, & Robbins, 1995) involvement in recognition, and functional interaction between the two (Parker & Gaffan, 1998a). Furthermore, when comparing the first and second repetitions of ‘old’ items, bilateral activation of frontal areas was stronger on the first repetition. This decreased frontal activity during the second repetition was thought to be an indication of the correspondingly reduced retrieval effort associated with recognition of items repeated a second time.

Whilst fMRI can give a good indication of the spatial profile of the anatomical substrate underlying the recognition process, better temporal resolution can be achieved with EEG measurements. Van Strien et al. (2005) examined changes in brain electrical activity during continuous recognition by EEG, and studied the resultant ERPs and induced band power, once again comparing novel items with their repetition, and repeated items at different levels of exposure. Recognition is associated with ‘old/new’ effects in the ERP, consisting of altered responses to repeated items compared to novel ones. Van Strien et al. found such an effect between 300 and 500 ms after exposure to a stimulus, when potentials for old items were associated with significantly greater positivity than those for new items. This effect was most pronounced around the midline parietal electrode position. In the time period 500-800 ms after presentation, multiple repetitions of an item were associated with linearly increasing positivity, most pronounced at the midline central and fronto-central electrodes. Stronger memory for an item can therefore be related to increased positivity in this period. Similarly, induced bandpower (IBP) data showed evidence of higher bandpower in lower-2 alpha, theta, and delta bands for old items compared with new, and the induced delta activity was lessened with increasing repetitions in the period 375-750 ms after presentation. These effects constituted evidence, the authors suggested, for a dual-process interpretation of recognition, as familiarity was discernable from a graded recollection state dependent on repetitions of the stimulus.

1.2.6 Different classes of stimuli

Estes and Maddox's (1995b) research had a bearing on an aspect of recognition memory research hitherto largely neglected. The amount of material memorised, and the process of retrieval, appear to be dependent on the type of stimulus used by the researcher. Most research using the continuous recognition paradigm until this point had been carried out using common English words as stimuli. However, other stimuli have been used in the paradigm, not just digit trigrams (Estes & Maddox, 1995a; Shepard & Teghtsoonian, 1961) but also pictorial and abstract visual stimuli (Doty & Savakis, 1997; Nickerson, 1965), and multisensory stimuli (Lehmann & Murray, 2005).

Nickerson's (1965) research, utilised a continuous recognition-type experiment with a series of black and white photographs, selected from photography periodicals. The paradigm was not strictly continuous recognition as the first 200 presentations were passively viewed new stimuli, followed by 400 presentations which contained 200 new and 200 repeated photographs. Very long lags (40, 80, 120, 160 and 200) were employed. The aim of the experiment was to investigate the long-term memory capacity for complex meaningful stimulus configurations. Nickerson's results suggested that these stimuli were recognised significantly better than those that had been previously investigated (i.e. words, numbers and nonsense syllables), with participants performing at an average of 95% accuracy. This experiment served to demonstrate that, given the appropriate materials, the capacity of human recognition memory is very high indeed. The fact that these images were meaningful and perceptually rich undoubtedly contributed to the very high recognition rates obtained. Meaningfulness of stimuli has also been demonstrated to be critical in the recognition of word vs. letter stimuli in investigations such as Estes and Maddox's (1995a), where strings of random letters were not memorised as effectively as meaningful words.

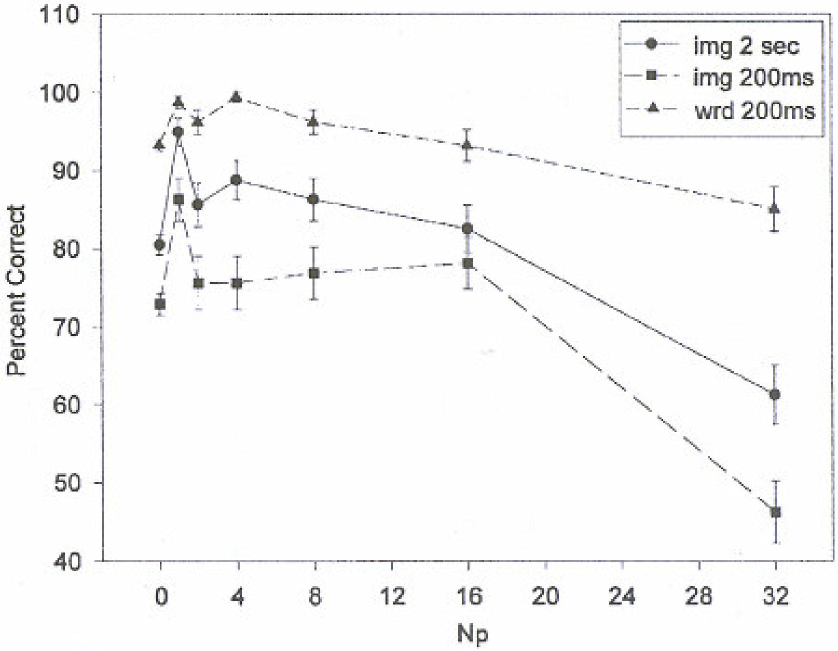

Figure 1.5: Percent correct responses as a function of lag (Np) for images (2s presentation and 200ms presentation) vs. words (200ms presentation). Figure reproduced from Doty and Savakis (1997).

In contrast, abstract and meaningless pictorial stimuli are remembered with much lower accuracy. Doty and Savakis (1997) assayed visual memory with unique, non-objective images, and compared performance to that for common 4-letter words. This was done in an attempt to determine whether recognition for the two different stimulus types was carried out via the same or different processes. Visual and verbal processes are known to involve, at least to some degree, separate neuronal systems in different brain regions, and with focuses in different hemispheres. Visual memory is largely independent of learning and language, and is thought not to differ greatly between humans and macaques. Doty and Savakis attempted to generate stimuli that did not suffer from the drawbacks of those used in previous experiments (e.g. Lewine, 1989; Shepard, 1967; Standing, 1973) in that they were full colour and not readily nameable items. The images proved to be extremely difficult to retain, but nonetheless their recognition was very similar to that of words in terms of both accuracy and patterns of reaction times at the lags used (0, 1, 3, 7, 15 and 31) (see Figure 1.5). Despite the differences in the y-intercepts of the curves plotted, the shape of the curves were very similar. Doty and Savakis concluded that these findings implied a basic commonality of the mnemonic neuronal processes involved in the recognition of word and picture stimuli.

In addition to studies employing single modality stimuli, Lehmann and Murray (2005) have used multisensory stimuli in conjunction with continuous recognition. They found that recognition of repeated images was significantly impaired if they were initially presented paired with an auditory tone, when compared with recognition of items studied without a tone. However, recognition was significantly improved when images were paired with a congruent sound (e.g. a picture of a gun with the sound of a gunshot) at the first presentation. The authors proposed that object-based multisensory interactions are particularly sensitive to the identity and semantic attributes of stimuli, and reveal the opposing effects of semantic and episodic contexts in auditory-visual multisensory events. The pairing of an image with an unrelated sound appears to impair the episodic encoding of that image, whereas the pairing of the image with a semantically-related sound enhances its recognition.

1.2.7 Conclusions

Research conducted so far with the continuous recognition paradigm has provided many insights into the workings of recognition memory under conditions approaching steady state. Whilst the research summarised above has provided many explanations of the processes of recognition, it appears that research into this field has thus far been limited by the choice of study materials. Experiments have been carried out predominantly with word stimuli which, as Estes and Maddox (1995a) have demonstrated, are over-learned as a stimulus set, and appear to be memorised in a manner different to stimuli charged with fewer semantic associations. In Estes and Maddox's experiment, trigrams and nonsense words were used, but these too rely on 'reading' (e.g. phonological recoding) in order to be effectively memorised. Word frequency effects, such as those investigated by Reder et al. (2000), suggest that a participant's general familiarity with a word is likely to affect the accuracy of its recognition. The advantages of using non-meaningful, parametrically-defined visual stimuli to probe short-term memory are discussed in Section 1.4.

1.3 Studying memory with word stimuli

Historically, the majority of studies of human memory have been carried out using word stimuli. Whilst there are advantages of practicality involved with such study materials, it should be clear from some of the evidence discussed above that words are not ideal stimuli for a number of reasons. The most important of these is that words are familiar stimuli, and that this familiarity varies according to a person's previous experience of these words. There are ways of attempting to control the level of prior familiarity, for example, by the use of word frequency measures which give an approximation of the mean occurrence of words in writing, but every individual's level of exposure to, and associations with a word are different. As well as individual differences there are systematic differences according to word frequency. High frequency words are associated with a higher base level of familiarity than low frequency words and this affects their recognition profile (e.g. Glanzer and Adams, 1990).

There is also the problem of interpreting the level at which words are processed and memorised. If they are presented as written words, they must be recognised by the visual system, but the encoding process usually involves reading, e.g. phonological recoding. This is also true of letters and numbers. In addition, with the use of word stimuli there is the added complication of semantic associations with the stimuli. As well as visual and phonological associations, words have meanings associated with them. These different modalities of representation and their semantic associations constitute complex entities that may rely on interactions between semantic and episodic memory. More simple, single modality stimuli may be more straightforward to study, and may yield results that are easier to interpret. The study of recognition using simple visual stimuli that are not amenable to naming, and vary only in their physical parameters, has been suggested as an appropriate avenue for future research (Kahana & Sekuler, 2002).

1.3.1 Word frequency effects

One of the most commonly reported phenomena that demonstrates the variability of recognition of word stimuli, is the word frequency (or mirror) effect (e.g. Glanzer & Adams, 1985; Glanzer & Adams, 1990; Glanzer, Adams, Iverson, & Kim, 1993). This is the much-replicated finding of very different response profiles for recognition of common (high normative frequency) and rare (low normative frequency) words. If the study items in a recognition task are common words, participants are more likely to respond ‘old’ than they are for less common words, giving rise to more false alarms with these words. Despite this, the hit rate (i.e. correct recognition) is actually higher for low frequency words.

These effects are assumed to be due to participants making decisions based on feelings of familiarity with these common words, as opposed to explicit recognition. Rarer words are less likely to be familiar but are more distinctive, and subsequently, are better recognised. The more frequently a word has been encountered, the greater the number of associations that have been made with that word. Reder et al. (2000) posit the theory that the greater the number of episodes associated with a word, the less distinctive the specific encoding episode during an experiment is, and, as a result, recollection of the experience becomes more difficult. This may explain the lower hit rate encountered for high frequency words. The high base rate of familiarity for common words makes them harder to identify as novel within the experimental context, compared with relatively rare words, resulting in a higher rate of false alarms.

The word frequency effect is an example of a mirror effect. This is the consistent finding that conditions that give rise to better recognition of old items as old also give rise to better recognition of new items as new, and can be understood as a series of distributions according to signal detection theory. Old responses for condition A (improved recognition) will be distributed further towards the old end of the decision axis than for condition B, whilst new responses for the two conditions will have the opposite arrangement (closer to the new end of the decision axis for condition A). This mirror effect holds true across a wide range of variables (Glanzer & Adams, 1985), although it is not predicted by many of the strength theories of memory (Atkinson & Juola, 1974; Bower, 1972; Gillund & Shiffrin, 1984; Glanzer & Bowles, 1976; Hintzman, 1988; Hockley & Murdock, 1987; Juola, Fischler, Wood, & Atkinson, 1971; Kintsch, 1967; Mandler, 1980; B. B. J. Murdock & Dufty, 1972; Parks, 1966). These theories propose that memory strength (or familiarity, amount of marking, number of representations, or amount of cuing) defines the decision axis. Glanzer et al. (1993) suggest an attention/likelihood model to account for the observed effects, focusing on the role of attention in the participant's learning and the role of likelihood ratios, derived from information about old and new items, in the recognition decision. In addition to the noting of memory strength the participant evaluates the likelihood of the stimulus being new or old, and this evaluation determines the decision. Because information about both old and new distributions is incorporated into the likelihood ratio, and therefore the decision about each old or new item, anything that affects either old or new likelihood distribution affects the other as well. This theory appears to account for the mirror effect better than older strength-based models.

The word frequency effect clearly indicates that memory for words can be affected by semantic influences, and that recognition of stimuli that occur frequently in everyday life, are harder for participants to recognise in a specific experimental episode.

1.3.2 Levels of processing

Experiments manipulating the level at which information is processed demonstrate how important the type of processing a stimulus undergoes is to its memorability. Studies using semantic and graphemic orienting tasks to manipulate the level of processing word stimuli undergo (McBride & Dosher, 1997; Roediger, Weldon, Stadler, & Riegler, 1992) have elicited differences in participants' performance at explicit memory tasks. Studying items at a semantic level (deep level of processing) resulted in significantly better recognition for those items than that for items studied at the level of physical appearance (shallow level of processing). However, no significant difference between performance at the two levels of processing was observed for implicit tasks, providing further evidence for a dual process approach to memory. Some experimenters, however, have been able to elicit level of processing differences in implicit tasks through changes in experimental design (e.g. the use of between-participants design) (Challis & Brodbeck, 1992; Thapar & Greene, 1994). Indeed, Brown and Mitchell's (1994) review of 38 studies manipulating level of processing in both implicit and explicit tasks suggests that 79% of the studies showed greater performance for semantic study than graphemic study in implicit tasks.

As a result of the dual processing of words, both visually and recoded phonologically, short-term memory (STM) for words appears to be affected by manipulations of either orthographic or phonological length (Coltheart, Mondy, Dux, & Stephenson, 2004). Recall and recognition of words presented in lists at either STM rate (1 per second) or rapid serial visual presentation (RSVP) rate (8 per second) were affected by orthographic length (measured by the number of letters), but only presentation at the STM rate was affected by phonological length (measured by the number of syllables). This was assumed to be because phonological recoding was impossible at the fast presentation times in RSVP. Recognition of words appears to be affected by manipulations at visual, phonological and semantic levels. This is another factor that makes the study of memory with words problematic.

1.4 The use of visual stimuli in recognition experiments

In addition to the classic literature on recognition derived from word experiments, an increasing body of literature is examining recognition of other visually presented material. As with words, the study of visual objects is problematic due to the human propensity for naming images where possible. However, a new approach is to take well-characterised stimuli from vision research (e.g. sinusoidal grating patterns) and use them in memory experiments (e.g. Kahana & Sekuler, 2002). This combination of visual psychophysics and memory appears to be a fruitful avenue for future research.

1.4.1 Categories and naming

Studies examining recognition memory for both words and pictures in the same paradigm have sometimes suggested that memory for pictures is superior to that for words (e.g. Standing, 1973). However, as Goldstein and Chance (1970) pointed out, the recognition of heterogeneous sets of pictures may be dependent on verbal labelling of the stimuli. In order to prevent phonological recoding and semantic classification there is a need to avoid this. Certainly there is a need to ensure that, when comparing pictorial memory to memory for words, that the test items must be discriminated as being unique stimuli, rather than being recognised on the basis of class membership. Making the stimulus set relatively homogenous minimises the opportunity for verbal labelling and ensures that a task is oriented towards visual discrimination rather than naming.

Picture naming, is an example of referential processing from nonverbal to verbal information (Paivio, Clark, Digdon, & Bons, 1989). The naming process is well-characterised, as are the features of a visual object that affect the process (see Johnson, Paivio, & Clark, 1996 for a review). Naming can occur at several levels of generality, the subordinate (e.g. a Granny Smith), the intermediate or basic (e.g. an apple) and the superordinate (e.g. a piece of fruit), and this affects naming performance. Names at the basic level are given faster than names at either the superordinate or subordinate conditions (Jolicoeur, Gluck, & Kosslyn, 1984), and these form the basis of common categories. A selection of stimuli from different basic categories can be readily named. However, stimuli from within a common category are much harder to give unique labels. If the stimuli are unfamiliar it is likely that participants will have no existing names for them, and will only be able to distinguish between them verbally with complex labels. In this scenario it is predicted that processes of purely visual discrimination would allow a much more effective way of discriminating between stimuli than by the process of naming.

Serial order in visual memory for pictures has been shown to be supported by phonological codes (G. Cohen, 1972; Manning & Schreier, 1988). Nelson et al. (1976) found that the phonological similarity between labels impaired serial order recall for pictures, and that this strategy was spontaneously adopted by participants. Visual similarity also reduced memory performance, suggesting that both modalities were involved.

Naming pictures appears to increase recognition performance, as shown by Wright et al. (1990) in a study using abstract kaleidoscope pictures. Learning names for these images resulted in a recognition profile similar to that encountered with travel slides, pictures of nameable objects, people and scenes, in that increasing interstimulus interval (ISI) was associated with increasing memory performance. Recognition of the kaleidoscope patterns without naming was not related to ISI. Participants were interviewed about the strategies they used to remember stimuli and were subsequently divided into those who attended only to sensory aspects of the stimuli, those who repeated a verbal label for the current stimulus, and those who ‘chained’ verbal labels together into a list. Those who carried out the task with sensory features alone showed a flat ISI function, as opposed to those using verbal labels for whom ISI had a significant effect on their recognition accuracy. The chaining strategy resulted in a greater effect of ISI, and was the most effective strategy for accurate recognition. It was effective even at a presentation rate of a 6-item list in 0.88 sec, too fast to actually carry out verbal rehearsal of the items. These results appear to demonstrate that the development of naming and rehearsal strategies qualitatively alters participants’ recognition of pictures.

In a similar study comparing rapid and slowly presented lists of pictures, Coltheart (1999) found evidence of phonological similarity effects at the slow rate but not the rapid. Items were presented at RSVP rate (8 per second) or STM rate (1 per second), and phonological similarity of the picture names decreased memory performance at STM rate but not RSVP rate, suggesting that pictures shown this rapidly are not phonologically recoded. However, when the written names of the pictures were presented, a phonological similarity effect did occur at this rate. This gives some indication of the time course of automatic naming. Potter and Levy (1969) have previously found evidence of picture comprehension, even at presentation rates of 10-12 per second. Presumably, at these very rapid rates of presentation, naming was not possible and a purely visual representation of stimuli was memorised.

The findings of the studies described above appear to suggest that visual sensory and verbal information are memorised in different ways, but a recent study suggests functional equivalence for these modalities (Ward, Avons, & Melling, 2005). Both recognition and serial order memory were tested for unfamiliar faces and heard nonwords, and similar memory profiles were observed across modalities. Recognition of both types of stimuli in a 2-alternative forced choice (2AFC) test found limited recency and no primacy, whereas serial reconstruction was associated with U-shaped serial position curves. The authors suggested that these findings are evidence that the type of information memorised (i.e. information about items vs. information about their order) has more effect on memory performance than the modality of the stimuli. Performance in the reconstruction of order task is dependent on prior responses, and is therefore serially dependent, whereas performance in the 2AFC task is not. The authors also inferred that the differences found in previous studies may have resulted from comparisons of novel pictures with familiar words.

A comparison of memory for sequences of colours or tones found that, although memory for visual sequential memory was poorer it produced a similar U-shaped curve to that for the auditory sequence (McFarland & Cacace, 1995). This would appear to suggest that the serial position curve is not dependent on modality, but is a general property of memory for sequences.

To return to the investigation of memory for the items studied, rather than the sequence in which they are presented, previous investigations of visual memory for novel patterns, rather than nameable pictures, are not associated with a U-shaped curve. The typical serial position function of recognition of such stimuli has no primacy effect and a recency effect limited to the most recent item only (e.g. Broadbent & Broadbent, 1981; Hines, 1975; W. A. Phillips & D. F. Christie, 1977; W. A. Phillips & D. F. M. Christie, 1977). Phillips and Christie (1977; , 1977) theorised that the recency effect was the product of a short-term visualisation process, and that the earlier part of the curve could be attributed to long-term components, in their duplex interpretation of visual STM. This one-item recency has also been observed in memory for scenes (Weaver & Stanny, 1978), and memory for spatiotemporal sequences (P. Walker, Hitch, & Duroe, 1993). If the visualisation process proposed by the duplex theory encodes spatial location as well as configuration of the item, it can also explain the findings of Walker and colleagues.

In order to examine memory for novel visual stimuli in the absence of serial position effects one must turn to continuous recognition studies. One such study using visual stimuli selected for their ‘unnameability’ (Doty & Savakis, 1997) also appears to demonstrate that recognition of visual stimuli is comparable in form to recognition of words, although it is slower and less accurate. A major problem with Doty and Savakis’ stimuli, however, was that they varied randomly. There was no systematic configuration or framework underlying their construction, and as such, it is hard to determine whether recognition in their experiments was true recognition of individual items within a class. It would be very difficult with such stimuli to control for distinctive items. Ideally stimuli should have the same basic configuration, so that they may be recognised as members of a class, but they should have features that vary within well-defined limits, giving rise to individual items within a population.

1.4.2 Combining vision research and memory

An avenue in visual recognition research that has potential for future exploration is the integration of findings from visual research, into discrimination and categorisation, into memory research. Whilst memory research represents items and uses decision processes similar to those involved in visual discrimination and classification models (e.g. Hockley & Murdock, 1987), the area has so far failed to ground abstract stimuli in perceptually-defined structures, with the resulting problems mentioned above. Likewise, vision research has only recently begun to acknowledge the importance of memory in simple perceptual tasks (e.g. Blake, Cepeda, & Hiris, 1997; Kahana & Bennett, 1994; Magnussen, 2000). Understanding vision and memory depends crucially on how they interact. The vast majority of sensory information available to individuals to form memories (more than 90%) is visual, and its storage (visual memory) can only be understood through its encoding (the results of visual processing). Conversely, lower functions like visual processing may be affected by higher functions through perceptual learning (e.g. Ahissar & Hochstein, 2002). A recent paper by Kahana and Sekuler (2002) has attempted to reconcile the two worlds by using elemental visual stimuli (2D textures composed of summed sinusoidal gratings) in a recognition memory experiment (Figure 1.6).

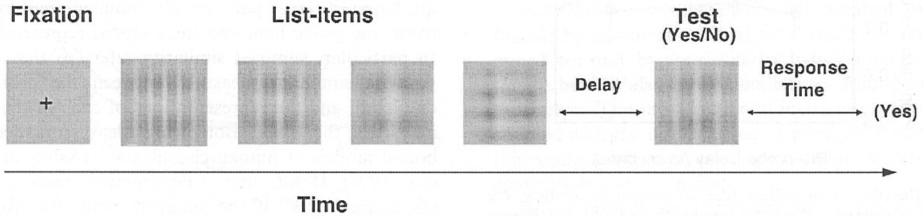

Figure 1.6: Schematic illustration of the experiments of Kahana and Sekuler, including examples of their stimuli. Figure reproduced from Kahana and Sekuler (2002).

The advantage of using such stimuli is that the neuronal processes associated with their perception are well-characterised (De Valois & De Valois, 1988), and it has been suggested that these dimensions may form the basis of organisation in short-term visual memory (Kosslyn et al., 1999; Magnussen, 2000). As these gratings have a number of well-defined properties such as spatial frequency, interitem similarity may be readily quantified and manipulated with such stimuli. The stimuli are also very resistant to symbolic encoding, the problem that has dogged the use of words and pictures.

Kahana and Sekuler's (2002) study has given an insight and an introduction into a more specific and systematic approach to the problems of visual memory research, avoiding the problems of symbolic recoding of visual stimuli, and providing a way in which interitem similarity may be quantifiably tested. Interitem similarity within short lists of stimuli, was used by participants in their recognition judgements, in addition to probe-item similarity. Certainly such systematic techniques point to a possible future in memory research, in which the stimuli employed will be parametrically defined and mathematically quantifiable. This will enable the study of memory from the bottom up, identifying the basic components and processes accurately rather than tackling the more diffuse semantic components of memory.

1.5 Visual processing in the inferotemporal cortex

In order to properly understand visual memory it is necessary to have an understanding of both higher level vision and memory, and how the two nominally distinct systems interact with one another. Indeed higher level visual object processing and memory storage are intimately connected in a way that makes them hard to separate. Described and discussed in this section are the neural mechanisms that contribute to the perception of objects from individual visual features.

Visual processing in primates, beyond the primary visual cortex, divides into two major pathways, or ‘streams’, devoted to two separate tasks (Desimone & Ungerleider, 1989; Ungerleider & Mishkin, 1982). These have been defined as the ‘what?’ and ‘where?’ streams, as their purported purposes are to identify what an object is, and to represent the space in which objects are detected and manipulated, respectively. The identification of objects relies on the successful separation of individual objects from a scene. The signals that represent the component parts of an object must be combined into a coherent representation of the object that can be remembered, and used in interactions with that object. This processing is carried out in the ventral stream, a pathway roughly corresponding to the inferior longitudinal fasciculus, and passing ventrally in extrastriate cortex to reach the inferotemporal (IT) cortex. This stream comprises areas V2, V4, and areas TEO, TE, and the perirhinal cortex, within IT cortex (see Figure 1.7). The second pathway, or dorsal stream, roughly corresponds to the superior longitudinal fasciculus, and passes dorsally in extrastriate cortex to end in the posterior parietal lobule and frontal lobe. It consists of areas V2, V3, MT, MST, and PO, and is responsible for the encoding of spatial parameters including the representation of an animal’s environment, the objects within it, and the space in which movement and manipulation of objects occurs, although it is a distinct possibility that information from ventral stream processing of objects is necessary for the construction of a scene.

The representation of visual objects in memory appears to rely heavily on the output of ventral visual processing, and of particular interest are the later stages in the chain of processing, especially those mediated by IT cortex.

1.5.1 Anatomy

IT cortex is a large region extending from just anterior to the inferior occipital sulcus to an area a few millimetres posterior to the temporal pole, extending from the fundus of the superior temporal sulcus to the fundus of the occipito-temporal sulcus (Logothetis & Sheinberg, 1996). This area roughly corresponds to Brodmann areas 20 and 21, or can be subdivided into areas TEO (posteriorly), TE (anteriorly) (Iwai & Mishkin, 1969; Von Bonin & Bailey, 1950), and perirhinal cortex (Brodmann areas 35 and 36).

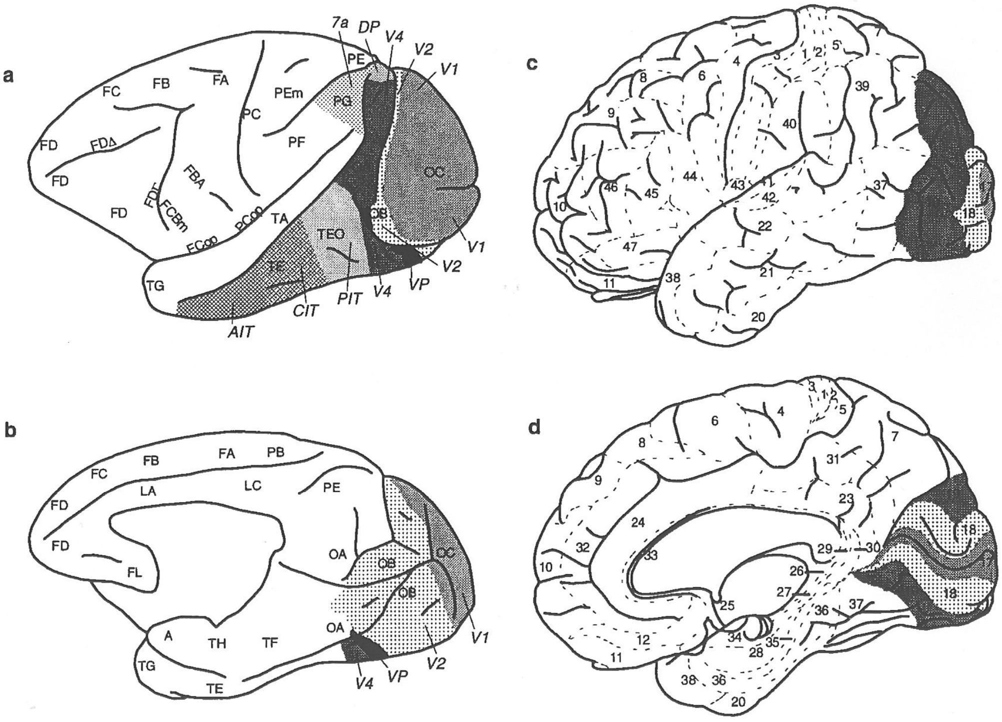

Figure 1.7: Diagrams of the visual areas of the brains of Macaque monkeys (*Macaca mulatta*) (a and b), and humans (c and d). Shown are Von Bonin and Bailey's (1947)

maps of the lateral (a) and medial (b) surfaces of the macaque brain with major visual areas superimposed by Felleman and Van Essen (1991). Labelled are visual areas 1, 2 and 4 (V1, V2, V4), ventral posterior (VP), posterior, central and anterior inferotemporal (PIT, CIT, AIT), and dorsal parietal (DP) cortical regions. The human brain's lateral (c) and medial (d) surfaces are also shown with Brodmann's areas numbered. Reproduced from Logothetis and Sheinberg (1996).

Area TEO is a strip from the edge of the superior temporal sulcus, to a few millimetres medial to the occipto-temporal sulcus. Information, largely from V4 but also from V2 and V3, is received by TEO, which contains a map of the contralateral visual field. However, later areas in the stream are not visuotopically organised in this manner (Desimone, Fleming, & Gross, 1980). Interhemispheric connections are mediated via the corpus callosum. All of these areas also receive feedback connections from TEO (Distler, Boussaoud, Desimone, & Ungerleider, 1993; Rockland, Saleem, & Tanaka, 1994). TEO receives feedback from the parahippocampal area TH.

The average receptive field (RF) size of neurones in TEO, that is the area of the visual field in which an effective stimulus may be detected by neurones, is larger than that of neurones in V4, and this trend continues, into TE. RF sizes have been estimated as 8° of visual angle in V4, 20° in TEO, and 50° in TE, and increase by a factor of approximately 2.5 for each level of processing (Boussaoud, Desimone, & Ungerleider, 1991).

Area TE extends further anteriorly from TEO, to about the sphenoid (Logothetis & Sheinberg, 1996). The RFs of TE almost always include the centre of gaze, sometimes encompass the whole contralateral visual field, and often extend into areas of the ipsilateral visual field (Desimone & Gross, 1979; Gross, Rocha-Miranda, & Bender, 1972). TE has direct connections to the amygdaloid complex (Amaral & Price, 1984; Iwai & Yukie, 1987) and to the hippocampus (Yukie & Iwai, 1988), as well as an indirect connection to the hippocampus via the parahippocampal gyrus (Van Hoesen, 1982). Both TEO and TE receive inputs from several nuclei of the thalamus, the hypothalamus, locus coeruleus, reticular formation, and the dorsal and median raphe nuclei. The region is critical for object identification and recognition (Dean, 1976), and appears to be perceptual rather than memory-related (Gaffan, Harrison, & Gaffan, 1986). However, the diverse subcortical connections of many areas posterior to IT, indicate that object-related information does not have to pass through this area to reach motor and memory systems.

Perirhinal cortex receives the majority of its input from TE and projects to the hippocampus via the entorhinal cortex. In addition to these connections there are direct pathways from V2 to TEO, and V4 to TE, and each area sends feedback signals to the area before it in the stream (W. Suzuki, Saleem, & Tanaka, 2000).

1.5.2 Lesion studies

Total removal of IT cortex in monkeys results in severe impairments of visual discrimination learning and retention, in the absence of changes to sensory thresholds such as acuity (see Dean, 1976, for reviews; Mishkin, 1966). Performance on tasks that require judgements of recency (working memory), or stimulus familiarity (recognition memory), is impaired, as the memory of a visual stimulus decays significantly over a minute or two, and may occur even more rapidly when stimuli intervene between study and test (Mishkin, 1982).

Bilateral lesions of TEO in the monkey result in deficits of colour, texture, and shape perception (Iwai & Mishkin, 1969), although ablation of both TEO and TE leads to much more significant impairments (Gaffan, Harrison, & Gaffan, 1986). TEO lesions do not appear to effect performance on tasks involving shape distortion, colour and oddity discrimination, or perceptual grouping (Huxlin, Saunders, Marchionini, Pham, & Merigan, 2000).

TE, however, is critical for object identification and recognition (Dean, 1976), and deficits appear to be perceptual rather than memory-related (Gaffan, Harrison, & Gaffan, 1986). TE lesions result in a permanent deficit in colour threshold. Animals appear to be able to perform normally in tasks requiring the memorisation of one object pair, but when larger numbers of objects must be remembered impairment is evident (Buffalo et al., 1999). These authors also found that this impairment was present for visual object recognition whilst tactual recognition memory was spared, suggesting that primarily visual recognition is affected. Altogether this evidence points to a major involvement of IT in both higher-level visual perception and visual recognition.

Perirhinal cortex is considered to be part of a circuit of structures involved in memory, the Delay-Brion system, on the basis of lesion studies (Buffalo, Stefanacci, Squire, & Zola, 1998; Squire & Zola-Morgan, 1991), but there is growing evidence that the region has a perceptual function as well (see Buckley, 2005). Murray and Bussey (1999) propose that perirhinal cortex represents objects whose visual features are represented in area TE, as monkeys with perirhinal ablations are impaired at object memory tasks when the perceptual difficulty of the task is increased. For example, monkeys with bilateral perirhinal ablations are unimpaired in their learning of 20 two-choice object-reward associations, but increasing the number of distracter objects, or increasing the number of problems to be learnt, impaired the lesioned animals' performance (Buckley & Gaffan, 1997). Similarly, showing the objects from different viewpoints on every trial (Buckley & Gaffan, 1998a) or presenting real objects in naturalistic scenes (Buckley & Gaffan, 1998b), which also increase the perceptual difficulty of the tasks, resulted in impairments specific to the operated animals. More recently, Bussey et al. (2003) demonstrated impairment in the discrimination of very similar morphed images of objects in the macaque, and this has also been observed in humans with perirhinal damage (A. C. Lee, Barense, & Graham, 2005). The discrimination of stimuli with high levels of feature overlap is also impaired in monkeys (Bussey, Saksida, & Murray, 2002, 2005). These findings have led to the proposal that the perirhinal cortex has both mnemonic and perceptual roles (Buckley, 2005; Buckley & Gaffan, 1998c; E. A. Murray & Bussey, 1999). Buckley (2005) concludes that the perirhinal cortex is specialised for the processing of stimuli at the object level and the binding of object features. This function is not only essential for the perception of objects but has a mnemonic role in the maintenance of associative linkages between constituent object features necessary for configural and paired-associate learning, as well as recognition.

1.5.3 Neurophysiology

Neurophysiological measurements of the firing rates of neurones in TEO in macaque monkeys have demonstrated that firing is driven by more complex stimuli than those effective for V4 neurones (Kobatake & Tanaka, 1994). The neurones are selective for object features including shape, colour, and texture (Komatsu & Ideura, 1993), suggesting a role for TEO in object perception and recognition, although such a role has not been confirmed in lesion studies.

TE neurones are similar in that they respond to complex visual stimuli, and there is evidence that cells responding to similar object features cluster together in columns, in a similar manner to cells in earlier visual areas (Fujita, Tanaka, Ito, & Cheng, 1992). As with other visual areas, most neurones in IT respond to many different visual stimuli and, therefore, cannot be considered as narrowly tuned ‘detectors’ of particular objects. The output of any one individual cell is inherently ambiguous with regards to the identity of the stimulus that gave rise to its change in firing rate. However, those cells that are stimulus-selective are usually selective along dimensions of shape, colour, or texture of a stimulus

(Desimone, Albright, Gross, & Bruce, 1984), although the level of complexity necessary to drive cells appears to be much greater than that for cells in TEO.

It has been suggested that the function of the large RFs found in IT cortex is the mediation of perceptual equivalence of objects across retinal translation, as the same stimuli can be detected across a large area of the visual field (Gross & Mishkin, 1977; Seacord, Gross, & Mishkin, 1979). In addition, some cells respond to the same stimuli at different sizes within their RFs (Desimone, Albright, Gross, & Bruce, 1984; Schwartz, Desimone, Albright, & Gross, 1983). The relative preferences to particular stimuli remain, although with both position and size changes the firing rates may alter.

1.5.3.1 Stimulus selectivity

Ito et al. (1995) have found that IT cells are selective, across their large receptive fields, for the shape of critical features, as determined by the reduction method. This method attempts to reduce an effective stimulus to the minimum complexity required to drive the cell. Preferences for faces, hands, and complex geometrical shapes have been found in cells of TE (Tanaka, 1996), and their responses are dependent on the configuration of specific features (Desimone, Albright, Gross, & Bruce, 1984). These face-selective neurones respond 2-20 times more to faces than to a wide range of gratings, simple geometrical stimuli, or complex 3-D objects (see Rolls, 1984, 1992), and reflect more information about face stimuli (average of 0.4 bits) than about non-face stimuli (0.07 bits) in a mixed set (Tovee & Rolls, 1995).

Tanaka (1996) presents data demonstrating the large effects of stimulus rotation on the responses of TE cells to effective stimuli. A rotation of 90° decreased neuronal responses by more than half of the cells they recorded, although the remaining cells were tuned more broadly. For some of these cells even rotations of 180° did not result in a change in firing response comparable to that measured for the first group. The responses to changes of size were more noticeable, with 21% cells responding to change of size less than 4 octaves of the critical features, and 43% cells responding to changes of less than 2 octaves. TE cells seem to maintain a response-selectivity to shapes across differences of luminosity, and coarseness of texture (Sary, Vogels, & Orban, 1993).

Logothetis and his colleagues (Logothetis & Pauls, 1995; Logothetis, Pauls, Buelthoff, & Poggio, 1994; Logothetis, Pauls, & Poggio, 1995) have used combined psychophysical üand neurophysiological experiments in macaques in an attempt to determine whether the configural selectivity observed for hands and faces is generated for other, novel object classes. Extensive training of the monkeys using novel computer-generated wire and spheroidal objects, with no biological relevance for the animals, resulted in the animals learning to discriminate the objects from highly similar distracters (Logothetis, Pauls, Buelthoff, & Poggio, 1994). The objects were all composed of highly similar parts, so discrimination was only possible on the basis of subtle shape variation, and the authors assert that this is similar to the categorisation of, for example, birds, or identification of specific faces. The suggestion that IT is involved in categorisation is backed up by a more recent experiment by Vogels (1999), who demonstrated interaction between cells coding for specific members of a category and those responding to all members of a category, but not to cells responding to other categories of objects.

Recordings from neurones in IT, near the anterior medial temporal sulcus, have revealed a subpopulation of neurones that respond to views of the unfamiliar objects used in the Logothetis et al. (1994) study (Logothetis & Pauls, 1995; Logothetis, Pauls, & Poggio, 1995). Certain views were optimal for activation of the cells, and as the object was rotated in 3-D the firing rate dropped off. The gradual decline in the responses of these cells to rotated optimal stimuli means that they act like “blurred templates” (Logothetis & Sheinberg, 1996).

Whilst the representations of objects encoded in these subpopulations do not appear to generalise to all viewpoints, they are also not specific to a single viewpoint, so some generalisation is possible.

Kobatake et al. (1993) have induced stimulus selectivity in cells of IT cortex in monkeys, through extensive training at a discrimination task with 28 simple shapes composed of geometric primitives. IT cells were subsequently tested using a battery of stimuli, and a much higher proportion of cells recorded responded to the test stimuli than other objects. This suggests that the monkeys' experiences modified the response properties of cells in IT cortex, resulting in changes to their tuning characteristics.

Altogether, this evidence demonstrates that IT cortex neural representations may contribute to the recognition of objects at subordinate level, i.e. of individual items from the same basic category, and that this appears to be dependent on experience. High stimulus selectivity does not appear to be limited to faces and other biological forms, but may involve similar mechanisms to those that produce face-selective cells.

1.5.3.2 Columnar organisation

IT is proposed to have a columnar organisation, similar in some respects to that observed in primary visual cortex. This organisation was inferred from simultaneous recordings of cells located close together in the cortex (Fujita, Tanaka, Ito, & Cheng, 1992). The second cell often responded to the optimal and suboptimal stimuli of the first cell, but with slightly differing selectivities. Cells recorded along penetrations made tangenitally to the surface of the cortex usually responded to the same critical features of the first cell, through all cortical layers. Those recorded from penetrations made oblique to the cortical surface showed similarities to other cells in an area of approximately 400 μm. Cells outside of this area did not appear to have similar response profiles to stimuli effective for cells within the 400 μm area. The TE region of IT, therefore, appears to be composed of columnar “modules” of cells responding to similar visual features. By dividing the area of TEd (dorsal TE) into 500 x 500 μm squares, an estimate of 1300 such modules was obtained.

Further study of the putative columnar organisation using optical imaging (Wang, Tanaka, & Tanifuji, 1996) has uncovered more about this arrangement. The optical imaging technique utilises light with a wavelength of 605 nm, which is absorbed more by deoxygenated than oxygenated haemoglobin. When the light is shone on exposed cortex, active areas are darker than the surrounding illuminated cortex, because the active cortex uses oxygen from the blood leaving the surrounding capillaries with a greater concentration of deoxygenated haemoglobin. This technique was combined with neurophysiological recordings from single cells.

First, critical features for the cells were determined by the reduction method and neuronal recording. The 605 nm light revealed 2-10 dark spots within the imaged region in response to previously determined critical features, and each time one of the spots covered the position of the initial electrode penetration. Spots were approximately 490 μm in diameter, and are further evidence that cells responding to similar complex features are clustered.

Interestingly, some partial overlapping of different critical features was observed in one optical imaging session. These three features were similar in that they included two combinations of colours of different luminosities and a gradation of colour, and all evoked dark spots around the original electrode penetration. For each critical feature, a spot of 500 μm diameter was observed, along a region about 1100 μm long. A similar overlap was found for face stimuli at different viewing angles. Five different viewing angles of a face were used and all evoked spots around the original penetration. As the face was turned from the left profile to the right profile the position of the dark spot moved in one direction along the surface of the cortex, suggesting a corresponding axis of representation in TE, 800 µm long. Together these findings suggest that related stimulus features may be represented in adjacent columns forming larger scale units. The findings from face rotation suggest that certain complex features may be continuously mapped within these larger units. Tanaka (1996) suggests that this is only the case for faces, as these are the only complex stimuli that are critical features for TE cells. As such, non-face objects must be represented across multiple cortical sites and different locations are activated by different views of these objects.

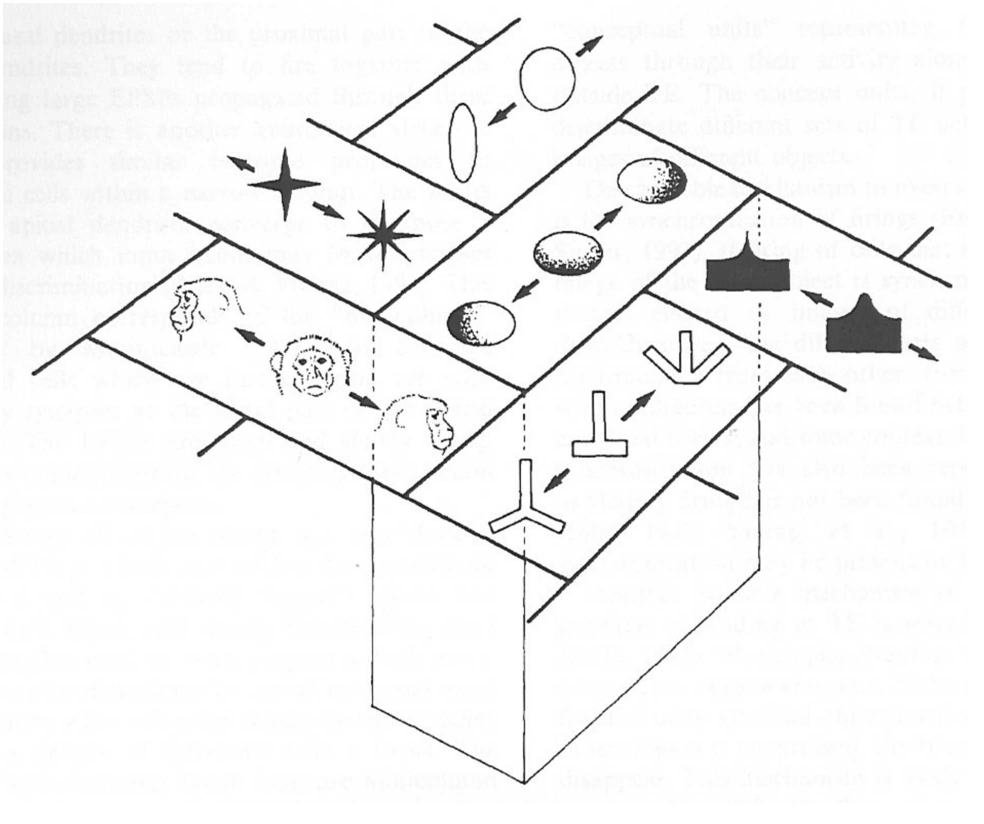

Figure 1.8: Schematic diagram of columnar organisation in TE, detailing the cortical surface's subdivision into areas responsive to similar stimuli. Figure reproduced from Tanaka (1996).

The columnar organisation suggested by the studies detailed above suggest that, in TE, object features are represented by the activity of many cells in columnar “modules” (Figure 1.8). This organisation may allow robustness to subtle changes in sensory input (i.e. a certain amount of transformation invariance), combined with preciseness of representation. It has been suggested that the clustering of cells with overlapping but slightly different selectivities may act as a buffer for slight changes to the input image, allowing some degree of transformation invariance (Tanaka, 1996). Whilst individual cells in this region are usually selective for size, orientation, and contrast polarity, the columnar modules may contain cells with selectivity for different sizes, orientations and contrast polarities for the same object feature. A more precise representation may also be achieved by many cells with overlapping selectivities, than summations of representations from individual cells. Tanaka (1996) suggests a mechanism similar to that proposed for hyperacuity (R. P. Erickson, 1968), whereby the difference between activities for nearby cells is thought to be used to obtain a greater resolution than that possible from the simple summation of outputs. Whilst activation within a column might represent a particular feature, subtle changes to that feature may be precisely represented by differences in the activities of cells with different selectivities.

1.5.3.3 Feature binding

Binding individual visual features to form coherent objects is an essential process in perception. Individual cells in IT only respond to moderately complex object features, so the information provided by individual columns is only ever partial. To represent an object in its entirety information from many columns must be combined. This is problematic where more than one object must be represented simultaneously, as features from different nearby objects must be discriminated from one another. The receptive fields of neurones in IT are too large to discriminate objects according to their retinal locations. Several mechanisms have been proposed to deal with this, including synchronisation of firing (Engel, Konig, Kreiter, Schillen, & Singer, 1992; Singer, 1993), attentional selection (Crick, 1984), and the formation of loops of activity with earlier areas in the visual pathway (Kawato, Hayakawa, & Inui, 1993).

Synchronisation of firing evoked by one object and desynchronisation of firing evoked by other objects allows the two sets of responses to be differentiated. Synchronisation of firing with oscillations has been observed in cells in cat visual cortex, and is proposed to be context dependent. Oscillatory firing has not been found in TE, but nonperiodic synchronisation may be present.

Another possible solution to the feature binding problem is selective attention. Only one, or possibly a few, objects can be attended to at a time, and if the features of an attended object are enhanced relative to other objects, it may be differentiated from them. Strong effects of attention on the responses of TE cells have been reported (Chelazzi, Miller, Duncan, & Desimone, 1993; Moran & Desimone, 1985; Richmond & Sato, 1987; Spitzer, Desimone, & Moran, 1988), suggesting that this is a distinct possibility.

Finally, there exists the possibility that representations of features in IT are combined with retinotopically organised areas at earlier stages in the ventral pathway by the formation of loops of activity. Feedback projections between TE and TEO, V4, V2 and V1 exist (Rockland, Saleem, & Tanaka, 1994; Rockland & Van Hoesen, 1994) and there are also feedback connections between each stage and areas immediately posterior in the pathway. Whilst these three solutions have been suggested independently there is a distinct possibility that more than one mechanism might work in combination.

1.5.3.4 Organisation through experience

Whilst much is now known about the columnar organisation of IT, less is known about how this arrangement may arise. Erickson et al. (2000) suggest that the perirhinal cortex has a dynamic functional architecture, and that it is moulded by experience. From studying the activity of pairs of simultaneously recorded neurones whilst monkeys viewed novel and familiar objects, they found a difference between the patterns of neuronal responses for the different classes of objects. When familiar objects (seen a few dozen times on the previous day) were viewed, neurones within about 100 µm of one another frequently responded similarly. This trend was not observed for responses to entirely novel objects. The similarity was such that, if a cell responded to e.g. 5 objects from a set of 16 familiar objects, nearby neurones would have a tendency to respond to the same 5 items. This is evidence that perirhinal neurones undergo rapid experience-related development to form functional groups, like the columns described by Fujita et al. (1992), requiring only a few dozen prior experiences. This ongoing plasticity appears to be essential for the function of this cortical area, and suggests its involvement in memory. This plasticity is certainly not unique within the cortex; Zohary et al. (1994) have described experience-dependent sensitivity to motion in cells of visual area MT, and neurones in other regions of IT that are responsive to temporal association have been described by Miyashita (1988).

Erickson et al.'s (2000) results suggest that the perirhinal cortex may represent not simply an object's features, as is the case in earlier areas of the visual system, but also an animal's experience with the object. They add further weight to the hypothesis that this brain region is involved in object categorisation, essential for higher cognitive functions. Objects associated with one another by experience within the experimental context might form a category, represented by neurones grouped together in the same region of cortex.

Miller (2000) speculates that these results mean that the organisation of perirhinal cortex is not only local, but may have some global pattern, drawing from the organisation of primary visual cortex (V1), which has both local, columnar organisation based on the similarity of visual features, and a global organisation that maps the retina on the surface of the cortex. Such a retinotopic mapping has not been found in IT but there may be a larger scale organisational pattern to the cortex that is still undiscovered. The mechanism by which the local organisation occurs so rapidly is unknown, although it may occur through the strengthening of lateral connections between neurones.

1.5.3.5 Learning in IT neurones

Miller et al. (1991) proposed a neural mechanism for working memory and recognition memory in IT cortex, from recordings made in monkeys required to retain items held in memory whilst viewing a series of stimuli. Study stimuli were presented at fixation and followed by one to five test stimuli. The animals were required to release a bar when a stimulus matching the sample was presented. For most cells that showed selectivity for matching stimuli, a response was apparent even when the maximum four items intervened, and it appeared that this was caused by an active matching process that was ‘reset’ between trials. The responses of these cells to matching and nonmatching items tended to become more difficult to distinguish from one another as the number of intervening items increased. However, the results of a separate experiment showed that a difference was still detectable when up to 6 items intervened between study and test. In fact, the authors suggested that there may be no limit on the ‘memory span’ of these neurones.

Logistic regression was used to estimate the functions relating cells’ response magnitude and the probability that a stimulus was matching. Whilst these individual cell functions were not good predictors of whether a stimulus was matching or nonmatching, in principle the success rate could be improved to behavioural performance levels (90%) by averaging over populations of neurones.

Additionally, Miller et al. (1991) used repetitions of initially novel sample stimuli throughout a 200-400 trial session, to examine the effects of increasing familiarity with these stimuli. For many neurones (over one third of a sample of 72), a systematic decline was observed as familiarity with these stimuli increased across the course of the experiment. Again, the magnitude in the decrease of response was dependent on the number of intervening trials between presentations of the same sample, which is not predicted by simple fatigue of the neurones. Familiarity with stimuli could still, in principle, be coded by a response decrement even after 140 intervening stimuli. Miller and colleagues conclude that "IT neurons may be acting as adaptive mnemonic filters that seek to preferentially pass information about new, unexpected, or not recently seen stimuli," (E. K. Miller, Li, & Desimone, 1991, p. 1379, p.1379). These functions have undoubted importance for the formation of memories about new objects and the selection of behaviourally relevant information for further processing.

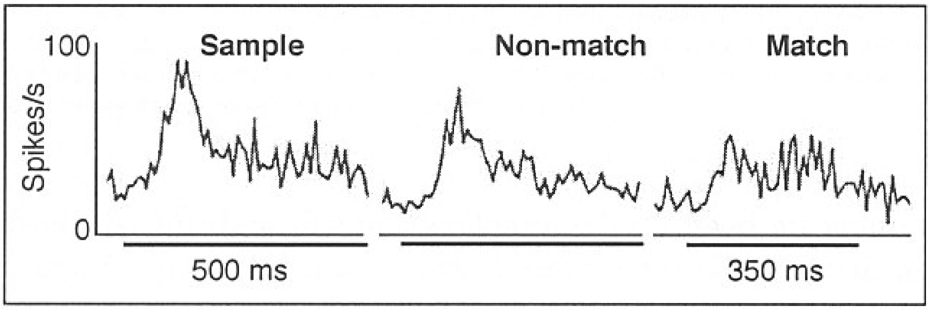

Figure 1.9: An example of repetition suppression in macaque IT cortex. Averaged responses to stimuli appearing as samples (first presentation), non-matches (first presentation of a new stimulus after a sample) or matches (repeated presentation of the sample stimulus). The bar beneath each graph shows stimulus presentation. Figure reproduced from Li et al. (1993).

The properties of IT neurones observed by Miller et al. (1991) have been described as stimulus specific adaptation (SSA): large changes in cell firing rates between initial and repeated presentations of effective stimuli. Such learning in both TE and perirhinal cortex is thought to occur rapidly, and repetition effects have been reported both in individual neurones (e.g. Li, Miller, & Desimone, 1993; E. K. Miller & Desimone, 1994) and also fMRI measurements in humans representing the activity of millions of neurones (e.g. Buckner et al., 1995). The changes in response to repetition may code for the familiarity of objects. There is some variety in the persistence of this response, and its latency, but in many cells there is a clear change in firing rate between the initial and second presentation of a stimulus. The commonest form of SSA is repetition-suppression (RS) whereby cells' firing rates are suppressed on repetition of a previously seen stimulus (Grill-Spector, Henson, & Martin, 2006; see Figure 1.9). RS effects have been observed in awake behaving monkeys performing match-to-sample tasks (E. K. Miller, Li, & Desimone, 1993) and recognition memory tasks (M. W. Brown & Xiang, 1998; Sobotka & Ringo, 1993), as well as in anaesthetised animals (E. K. Miller & Desimone, 1993). Because RFs in this region are very large the nature of the information processed is hard to identify, but the RS can be considered to be stimulus-specific in that it does not appear to reflect global changes in the firing rate of neurones to subsequent stimuli. However, neural RS has been demonstrated to be invariant with regards to certain changes in stimulus dimensions, including size and position of an object within the RF (Lueschow, Miller, & Desimone, 1994). RS persists even when many stimuli intervene between the initial and subsequent presentations of an item (E. K. Miller, Li, & Desimone, 1993; Xiang & Brown, 1998) and increases with further repetitions (Li, Miller, & Desimone, 1993). The mean latency of RS in IT has been estimated at 150ms and is thought to occur for about 50-67% of neurones responsive to visual stimulation (E. K. Miller, Li, & Desimone, 1993; Ringo, 1996).

In cells of the perirhinal cortex certain cells can be classified as familiarity and recency neurones according to the pattern of response observed. Response latencies in this region can be very rapid: as fast as 70-80 ms in some neurones (Xiang & Brown, 1998). Familiarity cells are thought to code whether a stimulus has ever been seen before (absolute novelty), whereas recency cells code whether an object has been seen in the recent past. In most of the perirhinal cells exhibiting SSA, responses are highest for new stimuli and decline with repeated presentations (Xiang & Brown, 1998), and this novelty response is thought to be involved in establishing the representation of a novel stimulus. These firing properties combined with extensive feedback connections to TE suggest that the perirhinal cortex may integrate simpler object features from earlier areas in the ventral stream, into object representations.

In a detailed study of these different types of cell, Fahy et al. (1993) carried out recordings from cells of the entorhinal, perirhinal and IT cortex, during monkeys' performance on a serial recognition task. The stimuli used were complex pictures of abstract and naturalistic scenes and objects, and, of the 2705 neurones that were visually responsive, only 120 (9.7%) showed significant RS. Of these, most were found in perirhinal cortex as well as areas TE1 and TE2, whilst they were not as common in TEO. The responses of 14.4% of neurones exhibiting RS showed significantly greater responses to unfamiliar compared with highly familiar stimuli and these cells were found in perirhinal and lateral entorhinal cortices, as well as areas TE1, TE2 and TE3. Decrements were observed even with long study-test lags and even after 24hrs in 6 of the neurones tested (Figure 1.10). The authors classified 7 neurones as familiarity neurones and 58 as recency neurones, based on their response profiles, and suggested that such neurones capable of signalling information useful for recognition were found in cortex close to the rhinal sulcus.

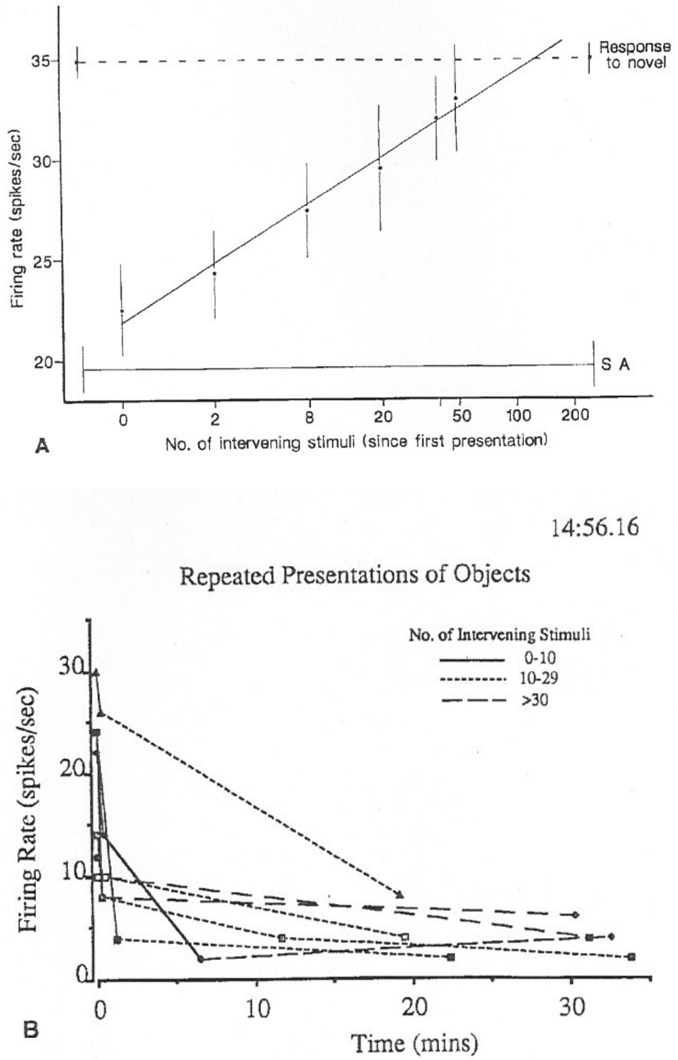

Figure 1.10: The effects of number of intervening items (lag) (A) and elapsed time (B) between study and test, on the responses of visual neurones. A) Responses of a single TE neurone to old items during continuous recognition. The broken line indicates the response to new items. The intercept of the regression line indicates a memory span of up to 120 intervening items. SA = spontaneous activity. B) Responses of a single perirhinal neurone to successive presentations of unfamiliar objects. The response never recovered to the value for an objects' first presentation even after more than 30 min. Figure reproduced from Fahy et al. (1993).

Judgements of whether a stimulus is entirely novel or has been seen before (is familiar) can be used to determine the recurrence of objects that are entirely novel when first presented. For items that are already familiar, judgements of recency (whether an item has been seen recently) are required to make a decision about whether or not the item has been seen during the current experimental context. The information coded by the familiarity and recency neurones studied by Fahy and colleagues is therefore essential for accurate performance on recognition, and working memory tasks. In addition, they are thought to be useful for priming memory, as initial large responses facilitate future performance on tasks.

1.6 Neural mechanisms of memory

Visual memory, the ability to store and retrieve aspects of visual experience, relies heavily on the same areas involved in higher visual processing, as described in the section above. Whilst this is especially true for memory for visual objects and their properties, memory also necessitates the integration of visual information into specific episodes, involving entire visual scenes. Described below are some of the brain regions and processes that are thought to achieve the different forms of memory in humans.

The study of the neural basis of memory has historically been centred on the medial temporal lobe (MTL), as damage to this region in human patients has been associated with anterograde amnesia, and recognition memory impairment (e.g. Scoville & Milner, 1957). Anterograde amnesia is the inability to form memories of new episodes experienced by the patient. For many years the hippocampus was viewed as the key structure for the encoding and storage of memory, however, more recent studies in animals have suggested that the perirhinal cortex may be more important for recognition. Groups studying the effects of hippocampal and perirhinal lesions in monkeys and rats tend to agree that perirhinal lesions are more detrimental to recognition than hippocampal damage (Aggleton & Brown, 1999; Gaffan & Murray, 1992; Meunier, Bachevalier, Mishkin, & Murray, 1993; Meunier, Hadfield, Bachevalier, & Murray, 1996; E. A. Murray & Mishkin, 1998; W. A. Suzuki, Zola-Morgan, Squire, & Amaral, 1993; Zola et al., 2000). Consequently, a new model of separate memory systems is emerging, in which areas such as IT and perirhinal cortex have major roles.

The IT and perirhinal cortices have important roles in visual memory, and interact with specialised memory structures of the Delay-Brion circuit. The type of memory encoded is related to the sensory information relevant to the task performed, and three major systems are thought to be present in this area. Object memory is thought to be located in the ventral stream itself (i.e. in IT), with memory about the properties of objects (semantic memory in humans) also thought to reside in the cortex of the temporal lobe. Episodic memory, the memory for personally experienced events and their context, is considered to involve the areas involved in object memory, as all events contain objects. In addition, other cortical areas that process the spatial layout of the environment are required, as is the hippocampus and the Delay-Brion system of cortical and subcortical structures.

The mechanisms within IT thought to be responsible for object memory have already been discussed in greater detail above. Primates have an ability to rapidly form detailed memories of novel objects, and these representations are thought to be encoded and stored within the ventral visual stream. The precise mechanisms via which this occurs are poorly understood, however, the repetition suppression observed for many neurones in IT, when objects are seen more than once, is a mechanism that may be of importance.

Knowledge about object properties enables the development of categorical knowledge about different types of objects. In humans this is linked to language and is known as semantic memory. The study of human semantic memory is complicated by the fact that humans can learn about objects without directly perceiving them. However, the development of a semantic memory category is likely to involve repeated exposures to individual members of a class of objects. This is thought to result in a distributed representation of a category composed of specific feature representations (e.g. orangeness as a feature of carrots). Miyashita (2000) has demonstrated IT neuronal responses to temporally contiguous pairs of visual stimuli regularly occurring together, even though the two items are separated by a delay of 1-3 sec. These ‘pair-coding’ neurones are thought to be involved in making connections between specific object features, as part of a category within semantic memory. In addition, Erickson et al. (2000) have demonstrated that neurones in perirhinal cortex adapt with increased familiarity with objects, such that cells responding to familiar objects are likely to be physically proximal to one another in the cortex. When the same objects were first observed, neurones responsive to the objects were more widely distributed. This suggests a reorganisation of the cortex, perhaps in order to categorise the objects observed as a result of experience.

Human patients with damage to anterior temporal cortex show a memory deficit without episodic memory impairments but are deficient at general knowledge, or semantic memory. Whilst they are able to remember specific events such as the visit of a family member they may be deficient at identifying and describing that family member (Hodges, Patterson, Oxbury, & Funnell, 1992). The deficits are thought to be caused by damage to the perirhinal cortex, as the impairments are similar to those caused by lesions of the perirhinal cortex in monkeys. Monkeys with such lesions appear to have a disorder of knowledge about objects, as revealed by impaired ability to discriminate between objects for reward (Gaffan, 1994a), and impaired ability to match target stimuli to a sample (Gaffan & Murray, 1992). These deficits are only present when animals have to deal with more than one pair of objects (Eacott, Gaffan, & Murray, 1994). Deficits in delayed non-matching-to-sample (DNMS) have been observed in perirhinal cortex lesioned animals for both simple objects (Meunier, Bachevalier, Mishkin, &

Murray, 1993) and discrimination learning with scenes (Gaffan, 1994a). Together, these data suggest a crucial role for the perirhinal cortex in learning about objects.

Episodic memory is more complicated in that it appears to rely on both object memory and knowledge about the objects, as well as information about the environmental context in which these objects are experienced. In humans episodic memory is memory for personally experienced events involving the retrieval of perceptual information in spatiotemporal settings, and the re-experiencing of these events is a defining feature. The integration of objects into scenes is of particular importance. Research with animals suggests that other species can remember events in their contexts (Clayton & Dickinson, 1998), although whether this resembles human episodic memory is impossible to say. Work is underway to elucidate episodic-like memory in animals in order to model the human system.

The different functional memory systems can be adapted to different tasks, allowing for some redundancy and the compensation of damage to one system. Each independent system can have independent access to behavioural output as well as combined output through episodic memory.

1.6.1 The medial temporal lobe and the Delay-Brion circuit

The medial temporal lobe (MTL) system, consisting of the hippocampus, entorhinal, perirhinal, and parahippocampal cortices (Figure 1.11), has been implicated in episodic memory ever since studies of amnesic patients suggested a link between their pathology and sustained damage to this area of the brain. Human patients with extensive damage to the MTL are profoundly amnesic, whereas those with less extensive damage centred on the hippocampus are less so. The structures that form the MTL system are part of the ‘Papez’ circuit of cortical and subcortical structures, first suggested as involved in episodic memory by Benedek and Juba (1940). Delay and Brion (1969) proposed that diencephalic and temporal lobe amnesia are the result of interruptions to the circuit at different points, and Kopelman (1995), in reviewing the evidence accrued since then, suggests that this circuit should include the hippocampus, entorhinal and perirhinal cortex, the mamillary bodies, mamillo-thalamic tract, and the anterior nucleus of the thalamus. The lack of patients with damage limited to individual structures, in whom pre-morbid memory ability is unknown, make studies of patients with damage to the Delay-Brion circuit difficult to interpret. As a result, much effort has been devoted to the development of animal models of amnesia.

Studies investigating macaque monkeys with ablations to regions of the Delay-Brion circuit have revealed significant impairments to their memory for complex spatially organised scenes. For example, impairments in spatial memory in mazes (E. A. Murray, Davidson, Gaffan, Olton, & Suomi, 1989), memory for the location of hidden food rewards (Gaffan & Harrison, 1989), memory for complex naturalistic scenes (Gaffan, 1992), and memory for artificial computer-generated scenes (object-in-place task, Gaffan, 1994b), have been observed. However, object memory, independent of the context in which objects are placed, is spared. The Delay-Brion system, then, appears to be specialised for remembering specific events, which likely correspond to human episodic memory.

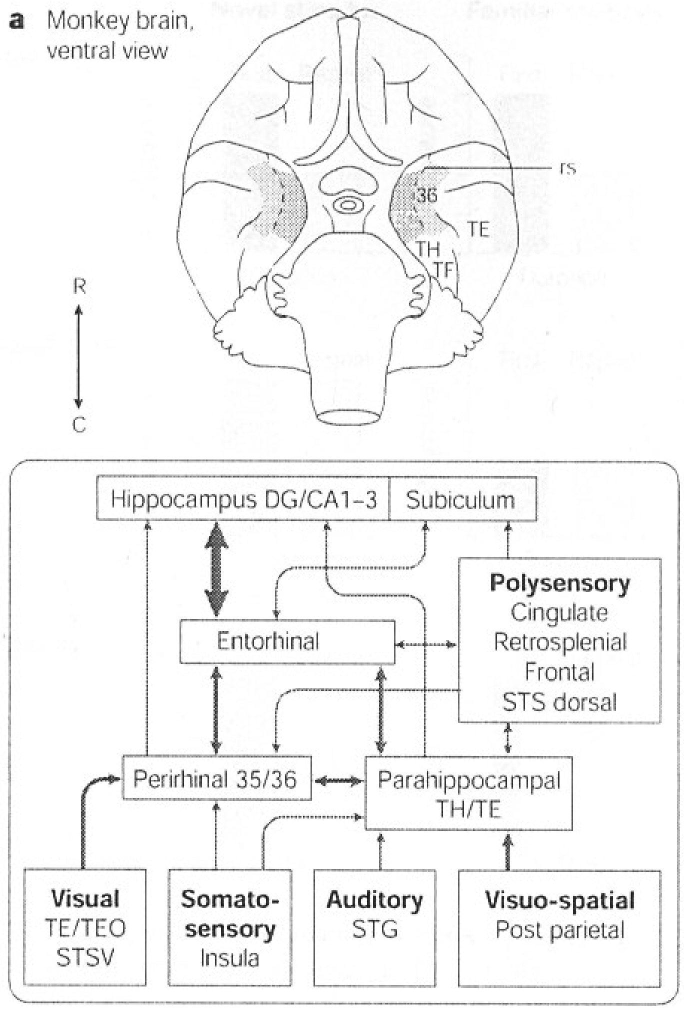

Figure 1.11: Diagram of medial temporal lobe areas involved in memory in the monkey (macaque) brain. The hippocampus refers to the dentate gyrus (DG), subfields CA 1-3 of the hippocampus, and the subicular complex. The thickness of arrows indicates the size of the projection. EC = entorhinal cortex, rs = rhinal sulcus, STG = superior temporal gyrus, STS = superior temporal sulcus. Reproduced from Witter et al. (1989).

Parker and Gaffan (1998b) presented the results of a series of experiments in which the memory performance of monkeys with lesions of specific structures of the Delay-Brion circuit were compared. The object-in-place task required the monkey to select a particular object from a pair of objects for a reward. The object always occupied a particular position in a background randomly composed of shapes and colours, and the monkey was required to learn lists of these scenes. Impairments in monkeys' performance of the task were observed following fornix transection (Gaffan, 1994b), lesion of the mamillary-body (Parker & Gaffan, 1997b), and lesion of the anterior thalamus (Parker & Gaffan, 1997a). Anterior thalamus lesions have also been shown to result in anterograde amnesia in humans (Daum & Ackermann, 1994; Hankey & Stewart-Wynne, 1988; M. H. Kim, Hong, & Roh, 1994), in the absence of object recognition memory impairment. Lesions of the cingulate cortex, a region included by Papez in the originally proposed circuit (Papez, 1995), did not impair monkeys' performance on the object-in-place task. Whilst cingulate cortex damage has been implicated in the development of amnesia in some human patients, e.g. the patient studied by Valenstein et al. (1987), it is possible that this patient also sustained damage to the fornix. However, the retrosplenial cortex has reciprocal connections with anterior nuclei of the thalamus, and the subiculum and presubiculum of the hippocampus, suggesting some role in memory (Devinsky & Luciano, 1993). Comparison of mean increases in error rates for macaques in the Gaffan and Parker studies suggests that the cingulate gyrus is not critical to the circuit. Whilst the increases in error score for fornix (13.0%), mamillary-body (18.3%), and anterior thalamus (12.2%) lesions were similar, that for lesion of the cingulate (3.6%) was significantly lower.

The hippocampus appears to be critical for episodic memory encoding and retrieval but is not essential for visual object memory. The hippocampus receives input from the entorhinal cortex, the presumed endpoint of the ventral visual stream. In addition, a pathway passes through the parahippocampal cortex, which receives inputs from V4, TEO, and TE of the ventral visual stream, and parietal areas 7a and LIP, to the hippocampus. This pathway is less well characterised than the entorhinal pathway, but due to the combination of information from visual processing and parietal cortex, it is likely that it may be involved in the integration of visual scenes in memory. A study of humans using fMRI has revealed activation of this region when participants are studying natural scenes (Menon, White, Eliez, Glover, & Reiss, 2000) supporting this putative role.

The parahippocampal gyrus and hippocampus have recently been found to contain neurones exhibiting repetition suppression similar to that previously discovered in perirhinal and entorhinal cortex in an fMRI study (Brozinsky, Yonelinas, Kroll, & Ranganath, 2005). These responses were also sensitive to lag in a continuous recognition task, only occurring when the repetition interval was relatively short. The relationship of these responses to memory performance could not be determined as there were insufficient trials for analysis, but they suggest a role for these regions in recognition.

The Delay-Brion system's function in episodic-like memory can be dissociated from the semantic-like memory of the perirhinal cortex through selective lesions. Interruption of the Delay-Brion system, e.g. by fornix transection, results in a severe impairment to memory for the spatial organisation of scenes with only a mild impairment in matching-to-sample with objects, whereas ablation of the perirhinal cortex results in the opposite pattern of effects (Gaffan, 1994a). The different computational tasks performed by the two systems relate back to the two types of information processed by the two visual streams, discussed previously. One system is required to solve the problem of storing information required for object perception and recognition (ventral visual stream and perirhinal cortex), and one system for the perception and memorisation of the animal's position in space (dorsal visual stream and Delay-Brion system). However, there are certain tasks that require the flow of information between the two systems, e.g. the object-in-place task, which requires the animal to remember spatial arrangements of multiple objects. The perirhinal cortex is reciprocally connected to the subiculum of the hippocampus (Amaral & Insausti, 1990) and can also exchange information via the entorhinal cortex (Insausti, Amaral, & Cowan, 1987; Witter & Amaral, 1991). When surgical disconnection of the two regions is carried out, normal performance on the object-in-place task is severely disrupted (Gaffan & Parker, 1996).

1.6.2 The encoding circuit

In order for memory encoding of objects or episodes to occur in the areas discussed above, modulation of neurones that are active during perception must occur. This modulation is thought to be achieved by a circuit of structures that connect representations of goals (e.g. food rewards for animal subjects) in the frontal cortex (E. K. Miller, 2000), with the representations of objects and scenes in the temporal lobe. The interaction and communication of goals with structures involved in encoding is achieved through subcortical connections, and current evidence implicates the basal forebrain. This region contains cholinergic neurones that project to IT and MTL and disconnection of this region from the temporal cortex results in dense amnesia (Gaffan, Parker, & Easton, 2001). There are three main routes from the basal forebrain to the temporal lobe: via the temporal stem, the amygdala, and the fornix, and disruption of any of these pathways results in severe anterograde amnesia (Gaffan, Parker, & Easton, 2001).

1.6.3 The neural bases of recognition memory

As has been described previously, recognition memory requires both the ability to identify objects and events, and also the judgement of their prior occurrence (Mandler, 1980). Whilst some accounts view recognition as a unitary process, in which recognition memory is an integral part of the memory lost in amnesia (Donaldson, 1996; Haist, Shimamura, & Squire, 1992; Hirshman & Master, 1997), an alternative theory posits two component processes (Gardiner & Parkin, 1990; Jacoby & Dallas, 1981; Mandler, 1980). Only one of these processes should be lost in anterograde amnesia (Aggleton & Brown, 1999). The suggestion from recent studies with animals is that episodic recollection is served by the Delay-Brion system, whilst familiarity may be coded in the perirhinal and IT cortex.

The repetition suppression (RS) effect in responses of neurones of the temporal lobe observed in electrophysiological studies of these neurones (M. W. Brown, Wilson, & Riches, 1987; M. W. Brown & Xiang, 1998; Fahy, Riches, & Brown, 1993; Li, Miller, & Desimone, 1993; E. K. Miller, Li, & Desimone, 1993; Sobotka & Ringo, 1993; Xiang & Brown, 1998) carries information useful for judgements of the prior occurrence of stimuli. These responses occur most frequently in anterior IT, especially the perirhinal cortex, and are much less common in the hippocampus (Brown et al., 1987; Rolls et al., 1989; Riches et al., 1991; Miller et al., 1993; Sobotka and Ringo, 1993; Xiang and Brown, 1998; Brown and Xiang, 1998). Indeed, cells showing these responses in the hippocampus were found at no more than chance levels in two studies (Riches, Wilson, & Brown, 1991; Xiang & Brown, 1998). A recent fMRI study has discovered lag-sensitive cells exhibiting RS in the hippocampus (Brozinsky, Yonelinas, Kroll, & Ranganath, 2005) although their significance is yet to be understood, and may relate to other aspects of the continuous recognition task. On the basis of the evidence accumulated thus far, anterior IT, especially the perirhinal region, appears to contain the majority of neurones with familiarity signalling properties.

Within the perirhinal cortex, neurones that exhibit RS show very rapid (~75 ms) familiarity and recency discrimination of individual stimuli, and are capable of single-trial learning, with a relatively long-term (>24 hr), and high capacity of storage (E. K. Miller, Li, & Desimone, 1993; Xiang & Brown, 1998). Thus RS is thought to be a mechanism of long-term memory storage, and this is supported by the performance of perirhinal lesioned monkeys at two variants of a delayed matching-to-sample (DMS) task (Eacott, Gaffan, & Murray, 1994). In the trial unique stimulus variation, where long-term memory is required, animals are significantly impaired, whereas performance of the variant in which stimuli repeat frequently and working memory is taxed, performance is unimpaired. Together this is strong evidence for the involvement of perirhinal cortex and area TE in the discrimination of the recency and familiarity of visual stimuli, independently of the hippocampus and other Delay-Brion structures. The latency for discrimination of prior occurrence is as fast as the latency for identification within monkey perirhinal cortex and TE (E. K. Miller, Li, & Desimone, 1993; Xiang & Brown, 1998), which excludes the possibility of top-down input from either prefrontal areas or the hippocampus. These findings tally with the evidence available from perirhinal lesion studies, which have noted impaired recognition memory for individual objects (Meunier, Bachevalier, Mishkin, & Murray, 1993; Meunier, Hadfield, Bachevalier, & Murray, 1996; W. A. Suzuki, Zola-Morgan, Squire, & Amaral, 1993).

Electrophysiological recordings from the hippocampus have suggested its role in transmitting information about the spatial environment of the animal. Some authors have found neurones signalling the familiarity of a visual stimulus occurring in a particular spatial position (Eichenbaum, 2000; Rolls et al., 1989). These neurones, then, may perform context-dependent recognition; the object-in-place memory described by Gaffan and Parker (1996). The entorhinal cortex, the anatomical region between the hippocampus and perirhinal cortex, may contain neurones encoding information about both stimulus familiarity and spatial information (M. W. Brown & Xiang, 1998; W. A. Suzuki, Miller, & Desimone, 1997; Xiang & Brown, 1998), and thus, may act as a junctional region between the two.

In hippocampal lesioned animals, standard object recognition memory task performance is only mildly impaired compared with perirhinal lesioned subjects (Aggleton & Brown, 1999). Murray and Mishkin (1998) found no change in performance at all following hippocampal lesion, whilst other studies report varying deficits most apparent at long retention delays (Alvarez, Zola-Morgan, & Squire, 1995; Beason-Held, Rosene, Killiany, & Moss, 1999; Zola-Morgan, Squire, Rempel, Clower, & Amaral, 1992; Zola et al., 2000). More reliable effects of hippocampal damage have been observed in Gaffan and Parker's (1996) object-in-place task. The hippocampus appears to have a role in recognition memory when familiarity judgements depend on associations between items, which are often spatial.

1.6.3.1 Evidence from clinical studies

The dissociable effects of perirhinal and hippocampal lesions in animals suggest that there may be similar dissociations in the abilities of amnesic patients with damage to these different structures. There is some evidence of amnesics with spared recognition memory in the literature (Aggleton & Shaw, 1996; Hirst et al., 1986; McMackin, Cockburn, Anslow, & Gaffan, 1995), although few studies have comprehensively examined the role of task difficulty. Some case studies have demonstrated a sparing of recognition memory irrespective of severity of amnesia or task difficulty (Hanley & Davies, 1997; Parkin, Dunn, Lee, O'Hara, & Nussbaum, 1993; Parkin, Rees, Hunkin, & Rose, 1994). In one case (Parkin, Rees, Hunkin, & Rose, 1994) recognition appeared to rely on judgements of stimulus familiarity, although the amnesia did not appear to be caused by damage to the hippocampus.

Studies of patients with pathology confined to the hippocampus or fornix have, again, demonstrated single instances where recognition based on familiarity is spared (Aggleton et al., 2000; Mayes, Van Eijk, Gooding, Isaac, & Holdstock, 1999; McMackin, Cockburn, Anslow, & Gaffan, 1995; Vargha-Khadem et al., 1997). In one particular case (Holdstock et al., 2000; Mayes, Van Eijk, Gooding, Isaac, & Holdstock, 1999), a patient with bilateral atrophy of the hippocampus and apparent sparing of adjacent regions, recognition memory appears to be preserved but is accompanied by persistent episodic amnesia. Recognition memory deficits are observed only where associative memory is required (e.g. memory for specific object pairings), and loss of this kind of memory has been associated with early hippocampal damage (Vargha-Khadem et al., 1997).

Use of Yonelinas' (1994) method of separating familiarity and recollection in recognition memory, the dissociation of processes procedure, has found loss of both processes in amnesics with extensive pathology (Yonelinas, Kroll, Dobbins, Lazzara, & Knight, 1998). Patients with pathology centred in the hippocampus have also been shown to be impaired at both components (Knowlton & Squire, 1995; Squire & Zola, 1998), supporting a single-process model of recognition.

1.6.3.2 Evidence from human imaging studies

Important findings on the nature of recognition are starting to emerge from ERP and fMRI studies. Rugg et al. (1998) manipulated the level of processing at which stimulus words were encoded, and then examined ERPs during subsequent recognition. The recorded activity suggested that three functionally dissociable populations of neurones responded to recognised stimuli. One was insensitive to both the accuracy of recognition and depth of processing and was thought to reflect priming. Another population, recorded above the left parietal cortex, was sensitive to level of processing giving rise to ERPs occurring 500 ms after stimulus onset, and thought to reflect explicit recollection. The third group, recorded over the frontal scalp, and present 300-500 ms after onset, was insensitive to level of processing, but was sensitive to whether items had been seen before or not, only being present for old items. This third response was thought to reflect familiarity. Similar findings of a later (400-800 ms) parietal recollective ERP and an earlier (300-500 ms) frontal familiarity ERP have been made in a subsequent study (Curran, 1999). When ERP differences have been studied using the ‘remember’/’know’ paradigm, enhanced ERPs were recorded over left parietotemporal and bilateral frontal sites for ‘remember’ (R) compared to ‘know’ (K) responses (Duzel, Yonelinas, Mangun, Heinze, & Tulving, 1997). These studies suggest that different neuronal populations signal different aspects of recognition, although this technique does not have the spatial resolution to locate these responses with any high degree of specificity.

Recognition has also been studied using fMRI, in order to achieve a better understanding of where functionally different neuronal populations might exist. Brewer et al. (1998) observed different bilateral medial temporal lobe activations during encoding of visual scenes, which could be used to predict whether the stimuli would be successfully recognised, and whether they would be remembered (R) or described as feeling familiar (K). Examination of retrieval found evidence of a dissociation amongst recognition memory components for words (Eldridge, Knowlton, Furmanski, Bookheimer, & Engel, 2000; Henson, Rugg, Shallice, Josephs, & Dolan, 1999). Henson et al. (Henson, Rugg, Shallice, Josephs, & Dolan, 1999)(1999) found evidence of different activity in the frontal cortex for R and K responses, and an increased response in the left posterior hippocampus for ‘remember’ vs. ‘not remember’ responses. Eldridge et al. (2000) found that increased activity in the hippocampus was only found during conscious recollection (R). In other studies, familiarity with scenes was found to be associated with decreased activity in the parahippocampal region (Gabrieli, Brewer, Desmond, & Glover, 1997). Reduced activity associated with familiarity with items has also been reported in IT (Jiang, Haxby, Martin, Ungerleider, & Parasuraman, 2000; Stern et al., 1996; Tulving, Markowitsch, Kapur, Habib, & Houle, 1994; Vandenberghe, Dupont, Bormans, Mortelmans, & Orban, 1995).

These responses contrast with those of the hippocampus, which are increased when recognising objects, studied as pictures, from test words naming the objects (Gabrieli, Brewer, Desmond, & Glover, 1997; Stark & Squire, 2000). This is a task involving associative memory of the type previously demonstrated to involve the hippocampus, although Stark and Squire (2000) show that this increased activity is not limited to conditions requiring association between items. It is not yet known whether such activity is present when conscious recollection is precluded.

Familiarity (K) decisions are made faster than recollect (R) decisions (Hintzman, Caulton, & Levitin, 1998; McElree, Dolan, & Jacoby, 1999; Seeck et al., 1997), and this parallels the finding of faster ERPs associated with familiarity than recall (Curran, 2000). The familiarity system might provide rapid and accurate detection of novelty but is unable to provide associative recollection. A second associative system centred on the hippocampus is required to remember associations with a stimulus, or the formation of new categories of stimuli.

Differences of neuronal responses to novel and familiar stimuli are common in the perirhinal cortex, whilst they are rarely found in the hippocampus, and do not persist over long intervals. Hippocampal neurones carry spatial or associational information, whereas perirhinal neurones do not seem as important for these processes. Neuroimaging has established qualitative differences between brain potentials and regions involved in the signalling of different aspects of recognition.

1.7 Translation invariance in memory and perception

Having explored current knowledge about the neural mechanisms involved in higher level visual processing and memory, it is important to examine one of the key problems of object perception and how it may be solved. Translation invariance in recognition is the ability to recognise the same object at different retinal locations, sizes, viewing angles, illumination conditions, etc. Whilst under different conditions objects may cause entirely different patterns of activation of the retina, higher level processing can usually identify that the object remains the same. Answering the question of how translation invariance is achieved, and under what conditions it breaks down, are likely to reveal much about how object identity is represented by the brain. Reviewed below are some of the key findings from studies of invariance in retinal location, one of the most widely studied phenomena in this field.

Electrophysiological recordings and lesion studies in monkeys have found evidence of both translation invariance and positional specificity of recognition-related responses. In monkeys where the optic chiasm has been sectioned, inter-ocular transfer of discriminations learnt in one visual hemifield is critically dependent on IT cortex (Seacord, Gross, & Mishkin, 1979), suggesting that this brain area is essential in the attainment of translation invariance. Studies of the electrophysiological properties of IT neurones can be divided into those showing some degree of translation invariance (E. K. Miller, Li, & Desimone, 1991; Tovee, Rolls, & Azzopardi, 1994) and findings of positional specificity (Chelazzi, Duncan, Miller, & Desimone, 1998; DiCarlo & Maunsell, 2003; Lueschow, Miller, & Desimone, 1994).

Desimone et al. (1984) found that most IT neurones respond to a variety of visual stimuli although their responses do appear to be selective along dimensions of shape, colour or texture. This selectivity was maintained throughout the neurones' receptive fields. As receptive fields of IT neurones are often large (median size 26° x 26°), and usually extend into both visual hemifields (Desimone & Gross, 1979; Gross, Rocha-Miranda, & Bender, 1972), these can be considered to be relatively translation invariant responses to specific properties of objects. This conclusion is supported by the findings of Tovee et al. (1994) that firing rates of temporal visual neurones in response to an effective image were not altered when the edge of the object was shifted up to 4° eccentrically from fixation. Even at this eccentricity there were only small decreases in responses.

However, not all IT neuronal responses are so invariant for changes in position. In a study of object recognition in which monkeys were required to discriminate between target images and distracters, changes in object location of 1.5° from fixation had minimal effects on behavioural accuracy and speed of recognition (DiCarlo & Maunsell, 2003). However, the anterior IT neurone responses were demonstrated to have a much greater sensitivity to positional change, showing a mean 60% decrease in response between locations. In a similar finding in a DMS task, Lueschow and colleagues found that, whilst the order of neurones' stimulus preferences did not change, 69% of cells recorded preferred a given retinal location (Lueschow, Miller, & Desimone, 1994). The change in location was much larger in this study (5°). Of the cells exhibiting repetition-sensitive responses, those with putative mnemonic capacity, only 7% of cells were not invariant for location. The authors suggested that retinal location is treated like a feature of the object by some IT neurones.

Psychophysical studies of memory for stimuli occurring at different positions in the visual field also provide evidence that is apparently contradictory. Some studies have suggested that the mnemonic representation is translation invariant – that once an item is encoded it will be recognised equally at any positional location (e.g. Biederman & Cooper, 1991; Bricolo & Bulthoff, 1993), whilst others suggest that a change of position between study and test incurs cognitive costs. Biederman and Cooper (1991) carried out an experiment in which pictures of readily nameable objects were presented to participants twice in two separate blocks for identification by naming. The interval between study and test presentations was approximately 7 min. Priming of the first presentation on an object's subsequent naming, measured as both a faster reaction time and a reduced error rate, was found to be independent of whether the initial presentation was in the same or opposite left/right or upper/lower visual hemifield. Performance for items with the same basic-level name that were different exemplars of the category (e.g. a blackbird and a sparrow) was poorer than for identical objects, suggesting that at least some of the priming was visual. The authors concluded that the activation of an object's basic-level concept and name were mediated by a position invariant representation. Bricolo and Bülthoff's (1993) study also suggested that object recognition is independent of test position. In a single interval forced-choice design participants were trained with wire-line objects at one position, 2.5° left or right, or up or down from fixation. Training position had no effect on later object recognition, although with a mean recognition rate of 90% any effect of translation may have been obscured by a ceiling effect. The translation invariant recognition that these two studies appear to support may well be dependent on the types of neuronal responses recorded by Desimone et al. (1984) and Tovee et al. (1994).

Evidence against complete translation invariance in humans comes largely from studies of pattern recognition. Kahn and Foster's (1981) study of participants' ability to discriminate sequentially presented dot patterns found that the distance between the two patterns was most important in determining the accuracy of responses. Stimuli were presented either at fixation or 0.5° to the left or right. Subsequently there were three possible separation distances: same position, 0.5° (centre-left/right), and 1° (left-right). D-prime was highest for identical study and test positions, and then decreased as a function of increasing distance. It should be noted, however, that better discrimination was measured when both study and test were at fixation, than when they occurred at an identical, but peripheral, location to either the left or right of fixation. There was also no attempt to separate trials on the basis of shift, taking into account the study and test positions, although the order of the two may be of importance to discriminability.

Nazir and O'Regan (1990) trained participants to discriminate dot patterns or columns of grey squares from two non-target distracters, at a fixed location to the left or right of fixation. Subsequently, discrimination was tested at three test locations – to the left or right of fixation, or at fixation (see Figure 1.12). Recognition rates dropped when stimuli were presented at novel positions, regardless of whether the eccentricity was 2.4°, 0.86°, or 0.49°. Later experiments demonstrated a similar pattern for stimuli presented above and below fixation. Error rates were greatest for the largest distance between stimuli (at 2.4° eccentricity). The results suggest that a change in location causes an increase in error rate, presumably due to a decreased ability to recognise the target. However, the error rates were similar whether the target item occurred at fixation or in the opposite visual hemifield, suggesting that it was the change in position that caused a decrement rather than its magnitude.

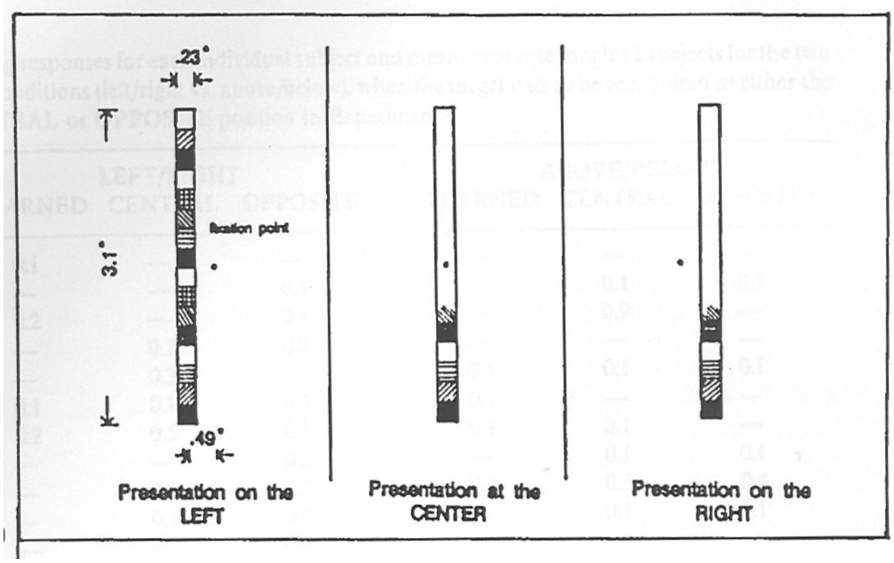

Figure 1.12: An example of the presentation positions of stimuli in a classic positional translation experiment. Reproduced from Nazir and O'Regan (1990).

In a similar experiment, Dill and Fahle (1997) confirmed the finding of Nazir and O'Regan, that improvement in discrimination performance at one retinal location does not transfer to new locations. This positional specificity seems to be found for novel patterns and demanding discriminations, suggesting that the achievement of positional invariance is memory-intensive. However, as accuracy following transfer from the learned to the unlearned position was significantly above chance in this study, some translation invariant recognition ability appears to be present.

One problem with all of the three preceding studies (Dill & Fahle, 1997; Kahn & Foster, 1981; Nazir & O'Regan, 1990) is that it is impossible to discern whether their results reflect a global positional specificity or whether this phenomenon is linked to the presentation of items at particular locations in the visual field (i.e. at fixation vs. outside fixation, left vs. right hemifield). Several studies have examined comparable shifts outside the region of fixation and including within-hemifield shifts (Dill & Fahle, 1998; Gratton, Corballis, & Jain, 1997; Hornak, Duncan, & Gaffan, 2002). The Gratton study compared the effects of both horizontal and vertical positional shifts of the same magnitude, so that stimuli appeared at different locations around fixation. A significant recognition advantage was found for same/different recognition of line pattern stimuli presented in the same visual hemifield, regardless of the distance between locations. In a further experiment systematically more negative recognition-related event-related potentials (ERPs) were found over the contralateral cortical hemisphere to the hemifield in which the stimulus was presented at study. The authors suggested that visual memories are contralaterally organised.

Hornak and colleagues (2002) used similar shifts in their design where stimuli could appear at any of the four corners of an imaginary square around the fixation point. Each point was equidistant from fixation, allowing the comparison of vertical and horizontal shifts at a constant retinal eccentricity. During learning, stimuli (pictures of nameable objects, containing different exemplars of the same categories) were presented in pairs in diametrically opposite positions. In a later test phase, pairs contained either two novel pictures or one learned and one novel picture, and participants were required to discriminate between the two (Figure 1.13). Old stimuli appeared either in an identical position, or shifted horizontally or vertically. The results showed a significant decrement in recognition of horizontally shifted stimuli when presentation times were 100ms (too short for the initiation of a saccade), compared to both the no change condition and the vertical shift. This appears to provide further evidence for a hemispheric organisation of memory, as shifts between left and right visual hemifields cause a greater recognition decrement than shifts between upper and lower visual field.

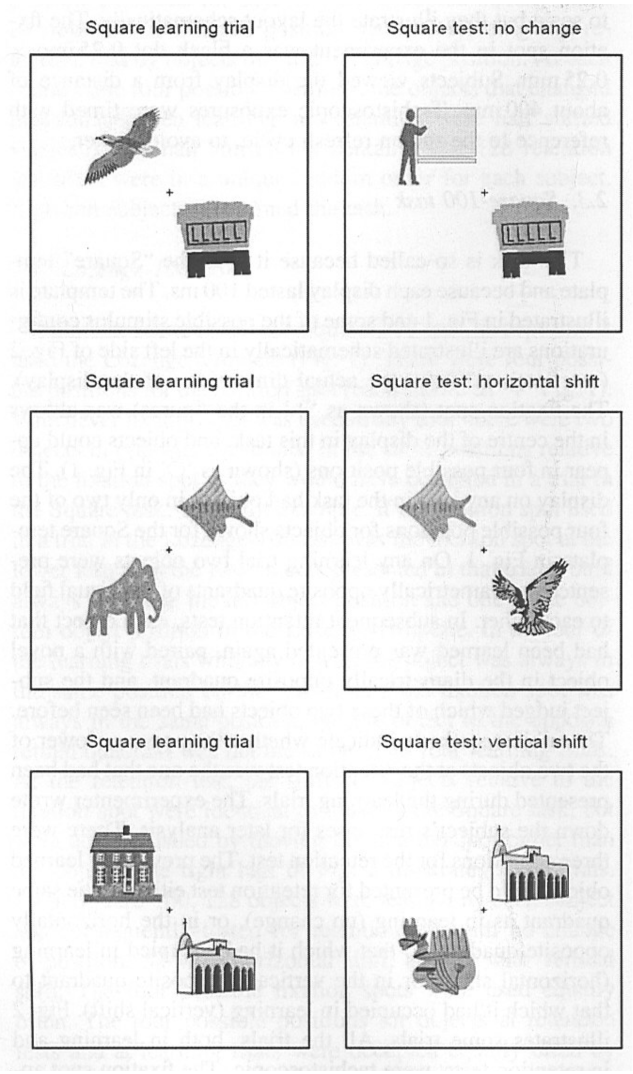

Figure 1.13: Possible configurations of learning and test trials from Hornak et al.'s (2002) 'square' experiment. Reproduced from Hornak et al. (2002).

The findings of Dill and Fahle (1998), however, seem to contradict the idea that shifts between hemifields produce a decrement greater than comparable vertical shifts. They found that horizontal and vertical shifts were equally effective in decreasing performance in a same/different task. Vertical and horizontal displacements of 0°, 0.5°, 1°, 1.5°, and 2° from an initial position 1° in the parafovea were examined with different types of stimuli (dot clouds and checkerboards) and different difficulty levels. Increasing decrements in performance were observed with increasing displacement, and this effect was independent of the similarity of the patterns. Larsen and Bundesen (1998) carried out a same/different experiment in which the patterns to be discriminated were presented simultaneously with varying spatial separation. In an experiment in which objects could be transformed by rotation as well as position, but the rotational component of the target was 0 (i.e. it was not rotated), *d’* was a monotonically decreasing function of spatial separation. This is a similar effect to the effect of displacement described by Dill and Fahle (1998) although the finding that it occurred without any temporal separation between the two items suggests that there is a perceptual component to the effect.

Despite evidence for translation invariant properties of IT neurones, and the recognition of pictures and 3-D objects (relatively naturalistic stimuli) occurring seemingly without regard for the position in which they were learnt, it would appear that certain discriminations utilise representations that are position specific. From the human experimental data we can ascertain that these are discriminations that are more demanding due to the abstract and/or highly similar nature of the stimuli employed (e.g. dot clouds, checkerboards). These patterns are unlikely to have been seen by participants before engaging in the experiment, and they are unlikely to have experience with similar objects. This suggests that some degree of perceptual learning of similar objects before an experiment may facilitate the formation of translation invariant representations. With utterly novel stimuli no such framework exists, and the perceptual apparatus must relearn the stimuli at each new location. This view appears to be consistent with the finding of experience-dependent organisation in the perirhinal cortex (C. A. Erickson, Jagadeesh, & Desimone, 2000). Perhaps translation invariance is a consequence of higher visual neurones adapting to many experiences with objects or categories of object. If this is the case, then discovering how much, and what kind of experience is required to achieve translation invariance is an important avenue for future research.

One study that combined naturalistic stimuli with a more demanding task found evidence of both translational invariance and positional effects depending on the stimuli employed (Dill & Edelman, 2001). In same/different experiments there was no effect of translation for ‘animal-like’ stimuli, regardless of their interstimulus similarity. The animal-like stimuli were composed of a standard number of computer-generated features such as legs and heads, and different ‘animals’ were defined both by the identity of their features and the features’ positional configurations. Translation invariance was observed when stimuli were ‘scrambled’ by randomising the identity of the component features whilst maintaining the global configuration of an animal, but not when the stimuli were made to differ in the locations but not the shapes of corresponding parts (Figure 1.14). This suggests that the representation of the identity of features may be position invariant, although the representation of their configuration may be specific to learning position.

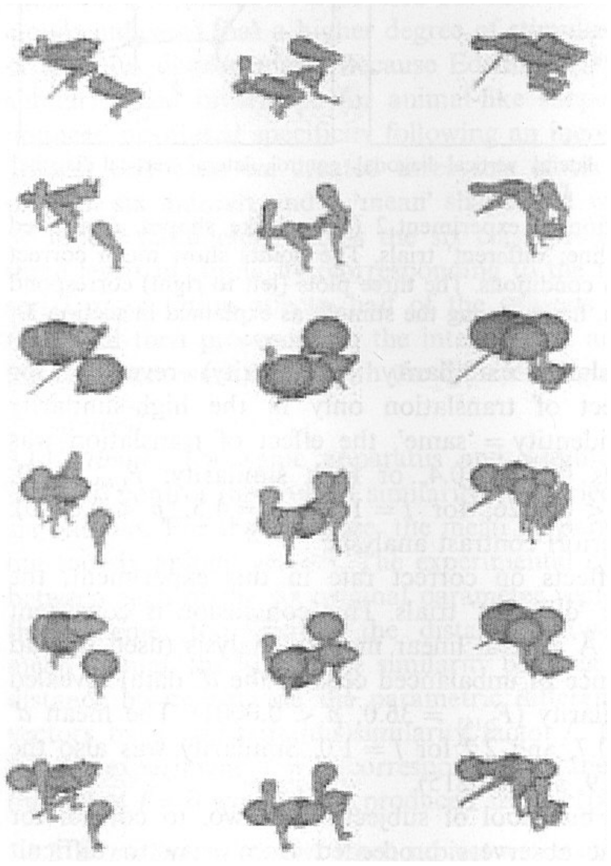

Figure 1.14: Example of 'scrambled' animal-like stimuli. Each column contains the same spatial configuration but the component features are different in each example. Each row shows examples of stimuli containing the same features, but in different spatial configurations. Figure reproduced from Dill and Edelman (2001).

Logothetis and Sheinberg (1996) conclude that object recognition occurs differently at the level of basic object categories and at the subordinate level. Whilst recognition at the basic level appears largely invariant to image transformations, and this may be carried out in the columnar modules of IT, recognition at the subordinate level may involve different types of representation using different neural mechanisms. Recognition at this level is initially highly dependent on specific views of objects, with generalisation thought to occur through perceptual learning. Dissociation between these two types of recognition is evident in agnosic patients. Whilst they are occasionally able to recognise objects with distinct shapes belonging to different categories, the identification of specific members of classes is typically impossible. The putative generalisation that results in translation invariance may occur through experience of objects changing position in the visual field, enabling feature analysers at different stages of cortical processing to become associated with one another through the strengthening of synaptic connections (Wallis & Rolls, 1997). As is the case with other kinds of reorganisation of visual areas, the questions that remain to be answered are how much and what kind of experience are required in order to achieve the changes.

1.8 Summary

This literature review has examined recognition memory, a form of memory requiring the identification of objects and events, and knowledge of their prior occurrence. The continuous recognition paradigm has been focused on as a powerful tool for the examination of recognition due to its avoidance of serial position effects, and the possibility of calculating signal detection measurements from data obtained with this procedure. However, a detailed examination of the mechanisms underlying forgetting in this type of task is yet to be undertaken. The difficulties of comparing data from studies with word stimuli and visual objects have been examined, and differences between visual and verbal memories considered. A new approach to visual memory combining techniques from vision research and memory psychophysics has been outlined as a promising avenue in future research. The brain regions and neural mechanisms involved in both higher level visual processing and the different categories of memory have been discussed, and their functions in producing translation invariant object recognition considered. In particular, the invariance of recognition to changes in object position has been examined. The question of under what conditions this invariance breaks down is of importance in determining how different representational systems operate under different task conditions.

Chapter 2 Retention of information during continuous recognition of a range of visual stimuli

2.1 Experiment 1: Retention of information during continuous recognition of pictures, faces, fractals and trigrams

2.1.1 Introduction

The goal of mathematically modelling the time course of memory retention is over 100 years old (e.g. Ebbinghaus, 1964; see Rubin and Wenzel, 1996). The purposes of obtaining a function, or collection of functions, to describe the process whereby memory performance declines from almost 100% accuracy to chance performance, are both practical and theoretical. The ability to accurately predict individuals' retention of information is of practical importance, whilst the elucidation of the function(s) would reveal important information regarding the components contributing to memory output.

As has been detailed in Chapter 1, attempts to evaluate retention of information using list-based memory tasks have been complicated by the presence of serial position effects. In a typical list-based task, the list of items is presented for memorisation, followed by a recognition or recall task. Plotting memory performance against serial position of study reveals that items towards the start and the end of lists are remembered better than those in the middle (primacy and recency effects), typically resulting in a U-shaped curve. This is observed even for lists as short as 4 items (e.g. Korsnes, Magnussen, & Reinvang, 1996; Wright, Santiago, Sands, Kendrick, & Cook, 1985) but is more pronounced for lists long enough to enable the plotting of the time course of retention.

An alternative method of assessing memory that avoids this complication is the continuous recognition paradigm, first used by Shepard and Teghtsoonian (1961). By intermixing study and test trials in a continuous stream of information, the authors were able to examine recognition at a relatively ‘steady state’. Each stimulus occurs twice, once as a novel stimulus and then as an ‘old’ stimulus, and participants are required to distinguish between the two. An initial unscored buffer of trials serves to prevent primacy effects, and after this period a steady state is assumed to have been reached. Whilst this is not always strictly the case, as demonstrated by Shepard and Teghtsoonian’s (1961) own discovery that false alarm rates very gradually increased throughout the experiment, the effects of serial position are minimal compared with those observed in list-based tasks. The separation of study and test trials is normally controlled, and the number of trials intervening between the two is known as the ‘lag’. Because the stimuli intervening between study and test are randomly selected, study-test pairs of the same lag throughout the experiment are assumed to be equivalent, and hit rates can be obtained for each lag. By the inclusion of a wide range of lags in an experiment it is possible to plot a retention curve of performance against lag.

This experimental paradigm has been used to good effect by Rubin et al. (1999) in the search for precise functions for the retention of information, tested by both recall and recognition. By using a very wide range of lags (0, 1, 2, 4, 7, 12, 21, 35, 59 and 99), many repetitions of each lag (27), and using a large number of participants (100 per condition), the authors achieved very precise retention curves for both recall and recognition of trigrams. By fitting the data obtained to a wide range of functions, informed by the authors’ previous fitting of data from 100 years of previous memory experiments (Rubin & Wenzel, 1996), a series of exponentials was selected as the best fitting function. This function, y =

$a_1e^{-t/T1} + a_2e^{-t/T2} + a_3e^{-t/T3}$, contains three time constants (T1, T2, and T3). Of these T1 was set at 1.15, and T3 was infinite. T2 varied according to the memory measure employed, being 27.55 for cued-recall and remember-know recognition, whereas 13.38 was better for old-new recognition, reflecting the different shaped curves plotted. Coupling this difference with the apparent difference between functions obtained for most data sets and those from studies of autobiographical memory in Rubin and Wenzel's (1996) study, it may be inferred that different memory processes produce retention curves that differ qualitatively as well as quantitatively.

Adopting the methodology of Rubin et al. (1999) seems to offer a precise and powerful method of comparing retention of information for different classes of stimuli. In the current experiment, a range of visual stimuli was tested in order to systematically compare and contrast the retention curves produced. In addition to the trigrams used by Rubin and colleagues, cartoon pictures, algorithm-generated fractals, and parametric face-like stimuli were also tested. The "clipart" cartoon images were chosen because they represented common objects, and are a stimulus type employed frequently in studies of visual memory (Barbarotto, Laiacona, Macchi, & Capitani, 2002; Biederman & Cooper, 1991; Hornak, Duncan, & Gaffan, 2002; Proverbio, Burco, del Zotto, & Zani, 2004; Snodgrass & Vanderwart, 1980; van Turennout, Bielamowicz, & Martin, 2003; Wan, Aggleton, & Brown, 1999). They were selected from several different basic level categories, with many exemplars from each (e.g. 10 oranges, 10 umbrellas, etc.). The fractals were chosen as an example of abstract stimuli that would be resistant to naming. As they are relatively complex, but do not resemble commonly encountered objects, recognition of these stimuli must largely rely on visual discrimination. Unlike the picture stimulus set, the fractal stimulus set can be considered to be composed of stimuli of a single category, and what is tested, therefore, is true recognition of individuals from a homogenous group, rather than discrimination between categories (Goldstein & Chance, 1970). Fractals similar to those used here have been used by Miyashita et al. (1993) in the study of stimulus-selectivity for complex visual forms of IT neurones. The face-like stimuli were also generated in such a way that they could be considered a homogenous category. These stimuli had a similar configuration to human faces, but were generated from a series of manipulated ellipses, whose parameters could be precisely controlled (Prof Andrew Derrington, personal communication). Whilst, as with the fractals, recognition of these stimuli might be expected to be based on visual discrimination as individual stimuli were relatively homogenous, participants' familiarity with the configuration of the features might be expected to result in a different recognition profile.

In addition to the expected differences between recognition of the different sets of visual stimuli, a difference between the visual stimuli and the verbal trigram stimuli was also expected. Whilst the findings of Ward et al. (2005) suggest that the form of memory for verbal and visual memory is similar, a previous continuous recognition task comparing the two modalities suggests that verbally encoded stimuli are recognised faster and more accurately than abstract visual stimuli (Doty & Savakis, 1997). Whether this finding, obtained with common 4-letter words, could be generalised to the more complex digit-letter-digit trigrams, was a matter of some considerable interest.

2.1.2 Methods

2.1.2.1 Participants

All of the participants were students at the University of Nottingham. There were 50 participants in each of the four stimulus conditions (faces, fractals, pictures, and trigrams), and the mean age of participants was 21-years-old in each group. In order to exclude data from participants who ‘gave up’ on the tasks, during the relatively long 30 min period, criteria were set for inclusion of the data in analyses. These were adapted from those used by Rubin et al. (1999). Participants were required to make fewer than 25 ‘no response’ trials (those where no response was measured during the allotted time), to prevent the inclusion of data from participants who had stopped responding. In addition they had to surpass the criterion of achieving recognition measures of at least 0.5 for lag 1 and lag 0 combined, and have a false alarm rate lower than 0.8, to eliminate participants who always responded with ‘old’. The recognition measure referred to here is the same measure used by Rubin and colleagues, [(hits-false alarms)/(1-false alarms)]. Mean numbers of ‘no responses’ for the remaining participants were 4.06 (faces), 4.14 (fractals), 3.04 (pictures), and 0.82 (trigrams). The mean probabilities of recognition for lags 0 and 1 combined for the remaining participants were 0.83 (SD=0.14) (faces), 0.81 (0.12) (fractals), 0.93 (0.07) (pictures), and 0.82 (0.13) (trigrams). Mean false alarm rates were 0.27 (SD=0.10) (faces), 0.23 (0.13) (fractals), 0.10 (0.06) (pictures), and 0.38 (0.09) (trigrams). To obtain 50 participants with data meeting the inclusion requirements in each group (200 in total), 54 were tested for the faces, 53 for fractals, 51 for bitmaps, and 54 for trigrams. Those excluded in each group were all removed for having failed to achieve mean recognition scores of 0.5 across lags 0 and 1.

2.1.2.2 Stimuli

Four different types of stimuli were generated: faces, fractals, pictures, and trigrams.

2.1.2.2.1 Faces

The faces were a set of 200 computer-generated ‘cartoon’ faces, generated by programs written by Andrew Derrington (see Figure 2.1). A seed for the random number generator was selected, based on the computer’s clock. The faces were generated in sequence by defining, and superimposing, 16 ‘egg-shaped’ ellipses, configured to represent the outline of a face, eyes, nose, hair, mouth, cheeks and eyebrows. Each ellipse was defined by random determination of properties such as length, position, height, curve, angle and RGB values (to define the ellipse’s colour), within pre-defined limits to ensure that the resulting structure resembled a face. The resultant ‘egg-shaped’ ellipses differed from the standard ellipse by having a 2nd harmonic component on the long axis, meaning that the ellipse could be fatter at one end than the other. In addition ellipses could be made asymmetrical along the long axis, by giving the two halves of the short diameter different lengths, which could be negative giving rise to a crescent-shaped ellipse. Details of each face’s parameters and the seed for the random number generator were stored in a data file.

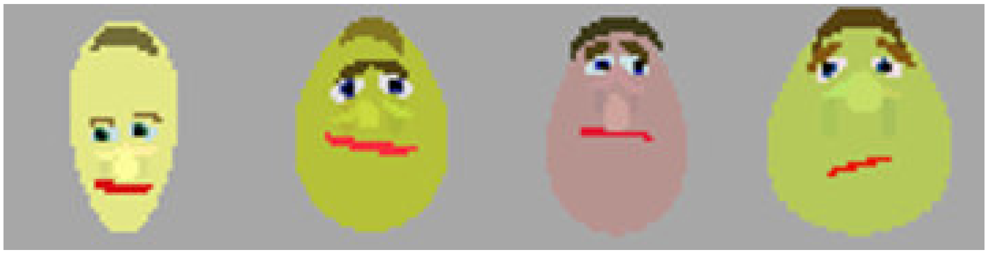

Figure 2.1: Example face stimuli.

2.1.2.2.2 Fractals

Fractals were generated by a program adapted from an algorithm detailed by Miyashita et al. (1991) on a Viglen PC running Matlab v6.1, equipped with the image processing toolbox (see Figure 2.2). Briefly, a seed for a random number generator was entered and set, in order that the fractals generated could be regenerated at any time, provided the seed and other variables entered were the same. A series of minimum and maximum levels for various properties of the fractals were entered enabling adjustment of the program to produce fractals of sufficiently different appearance. For each fractal the recursion limits and number of superpositions were randomly set between the minimum and maximum values specified. For each superposition random red, green and blue values were generated to define its colour, number of edges was randomly set between the minimum and maximum values, and then a regular or irregular deflection subroutine was called, according to a ratio of regular to irregular set at the start of the program. Both subroutines calculated the co-ordinates of a regular polygon and then carried out deflections on its sides in accordance with the transformations detailed in the appendix of the Miyashita paper. The only difference between the two subroutines was that the regular subroutine always carried out uniform deflections for each edge of the polygon, whereas the irregular routine did not have this constraint. Finally, superpositions were normalised so that each was centred on the same point. Each fractal was composed of a number of such superpositions, which became progressively smaller, in order that early superpositions were not obscured by later ones. Fractals were plotted in a Matlab figure window, then saved as bitmap images.

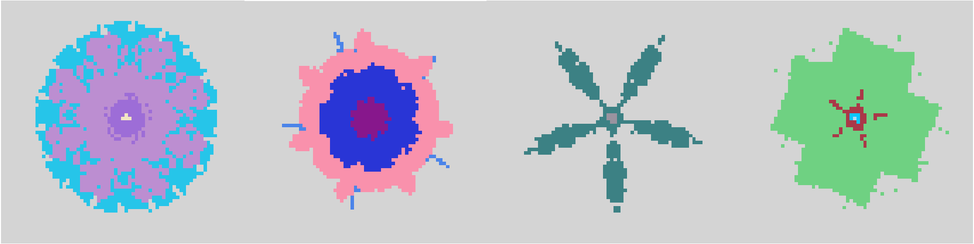

Figure 2.2: Example fractal stimuli.

2.1.2.2.3 Pictures

The pictures consisted of 200 unique clipart images of readily nameable objects. The 200 images were made up from 20 categories (see Figure 2.3): animals, apples, briefcases, bananas, glasses of beer, birds, burgers, butterflies, cats, clowns, coffee pots, dogs, fish, grapes, keys, oranges, shoes, strawberries, suns, and umbrellas. Ten images of each class of object were included (see Figure 2.4). The first 25 participants were tested with full colour bitmaps, but the experimenters noted that almost perfect recognition was obtained for this condition. In an attempt to increase task difficulty and minimise potential ceiling effects, the latter 25 participants were tested with greyscale bitmaps; exactly the same images but with colour information removed.

Figure 2.3: Examples of each of the 20 categories of picture stimuli.

Figure 2.4: Examples of stimuli within a category. The figure shows the 10 ‘apple’ stimuli employed in the experiment.

2.1.2.2.4 Trigrams

The trigrams employed were a randomly generated list of 200 ‘legal’ trigrams according to the rules of Rubin et al. (1999). That is that they were digit-letter-digit, with the digit zero excluded, and only letters K, V, W, Y and Z allowed. The trigrams were presented in white letters on a black background, such that the width of the trigram was approximately 2°.

2.1.2.3 Presentation

All stimuli were presented using an *Apple Macintosh G3* computer (300 Mhz, 384 Mb RAM) with a *ATI Radeon 7000* (32 Mb) graphics card, on a 21" *Mitsubishi* colour display monitor (size: 1024 x 768 pixels, resolution: 72 x 72 dpi, refresh rate: 75 Hz). Stimuli were presented in the centre of the screen using *Matlab* v5.2.1 (Mathworks UK), running the psychophysics toolbox (Brainard, 1997). All bitmap images (bitmaps, faces and fractals) were converted to a standard size of 59 x 59 pixels, covering an area of 2 cm x 2 cm on the screen when displayed at the resolution described above. Participants were seated at a distance of 57.5 cm from the screen so stimuli subtended an area of approximately 2° x 2° of visual angle.

2.1.2.4 Session design

A pseudorandomly determined frame of 120 trials was generated for each participant, providing 9 learning and 9 test trials at each of 6 lags (see Figure 2.5). These lags were chosen to be at regular intervals on a logarithmic scale being 0, 1, 2, 4, 6 and 9. Twelve filler trials made up the spaces in the frame. The filler trials were unscored study-test pairs, of variable lag. These were arranged so that the first unfilled space in the frame was a ‘study’ filler trial, the second was its ‘test’, the third was the second ‘study’ filler, and so on. The lag denotes the number of trials intervening between the learning and test presentations. Filler trials consisted of unscored learning and test presentation pairs. This frame was repeated three times, with novel stimuli each time, yielding 27 scored recognition tests for each of the lags. In addition, the experiment began with a buffer of 40 unscored filler trial pairs to prevent the occurrence of primacy effects. The effects encountered here, therefore, reflect memory performance where interference is high. This yielded a total of 400 trials. Once the order of trials for the entirety of a session had been generated, the 200 stimuli were randomly assigned to the 200 pairs of trials, so that each participant experienced the stimuli in a different order.

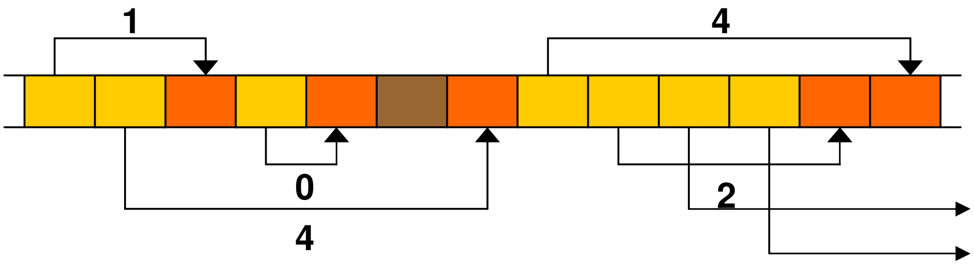

Figure 2.5: Schematic of an example region of the pseudorandom order frame. Assuming time proceeds from left to right, with each square representing a distinct trial. Yellow squares represent study trials connected by arrows to their corresponding (red) test trial trials. The number with each arrow is the value of the lag separating the pair. The brown square represents a filler trial.

2.1.2.5 Procedure

For each trial stimuli appeared on the screen for 2 sec (see Figure 2.6**Error! Reference source not found.**). During this period participants were instructed to respond either ‘old’ or ‘new’ with the left or right mouse button respectively, ‘old’ to previously seen stimuli and ‘new’ to novel items. In order that participants did not forget which button corresponded to which response during the experiment, a clear notice of which was which was placed next to the mouse. After 2 sec the stimulus was replaced with a blank screen for 1 sec, and then feedback for 0.5 sec. Feedback consisted of either the word ‘Right’ in green letters if the response was correct, or ‘Wrong’ in red letters if it was incorrect, or no response was registered. Finally, a further 1 sec blank screen separated the feedback from the next trial, bringing each trial to a total of 4.5 sec. For 400 trials, the experiment therefore lasted 30 min.

2.1.2.6 Scoring

As mentioned above, participants' responses were only collected during the 2 sec presentation period of each stimulus. Once a response had been made it was final, and no opportunity for the correction of responses was allowed. Hit rates and false alarm rates were calculated from test and study trials respectively. From these scores, *d'* was calculated. In addition, reaction time was measured as the latency from the start of the trial until the detection of the response.

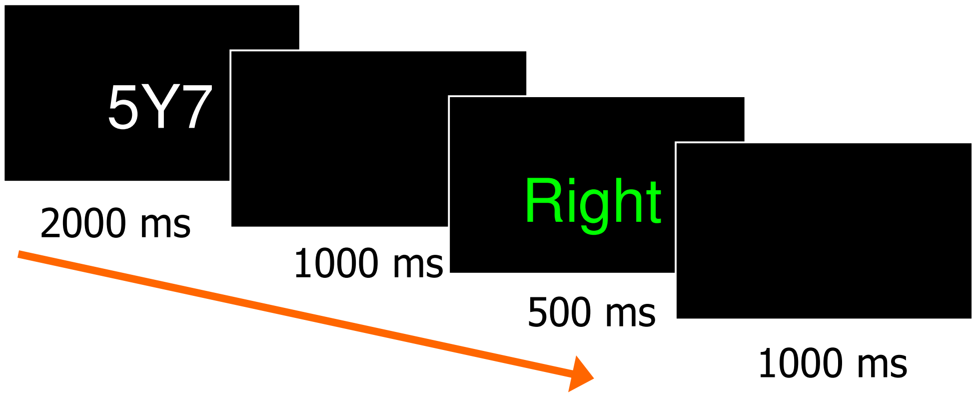

Figure 2.6: Schematic of the procedure for each trial. Sequence from top left to bottom right. Initially the stimulus was presented in the centre of the screen for 2000 ms during which time a response was required. This was followed by a blank screen for 1000 ms, appropriate feedback for 500 ms, and a final blank screen for 1000 ms before the start of the next trial.

2.1.3 Results

The number of correct responses to test presentations for each lag was divided by 27 (the number of test trials), to obtain the hit rate, or proportion of correct responses. False alarm rates for each participant were calculated by dividing the number of false alarms (‘old’ responses to novel stimuli) by the number of scored study trials. From these measures *d’*, the signal detection measure of sensitivity, could be calculated for recognition at each lag. The *d’* score is independent of the participant’s bias for answering ‘old’.

Recognition scores for the colour and greyscale pictures were compared using a 2 (colour vs. greyscale) x 6 (lag) mixed design analysis of variance (ANOVA). No significant difference between the two conditions was observed, and the data for both greyscale and colour pictures were included in further analyses as a single ‘pictures’ condition.

2.1.3.1 D-prime scores

The results for probability of recognition are shown in Figure 2.7. A 4 (stimulus type) x 6 (lag) mixed ANOVA was performed on *d’* data. Mauchly’s test for sphericity was significant for lag, and Greenhouse-Geisser epsilon corrected ANOVA results are reported for this factor. The ANOVA revealed a significant main effect of stimulus type ($F(3,20)=118$, $MSe=2.12$, $p<0.001$). As expected, participants found some classes of stimulus easier to recognise than others. Tukey HSD post-hoc tests revealed significant differences between pictures and all other stimulus types (all $p<0.001$), and between trigrams and all other stimulus types (all $p<0.001$) but not between faces and fractals. Picture stimuli were more accurately recognised than the three other classes of stimuli, and recognition of trigrams was significantly poorer.

A significant main effect of lag was also revealed (F(4.12, 161)=99.2, MSe=0.288, p<0.001). Tukey's post-hoc tests were carried out to reveal that recognition at lag 0 was significantly better than at all greater lags (all p<0.001), recognition at lag 1 was better than at lags 4, 6 and 9 (all p<0.001), recognition at lag 2 was better than at lags 6 (p<0.01) and 9 (p<0.001), and that recognition at lag 4 was better than at lag 9 (p<0.01).

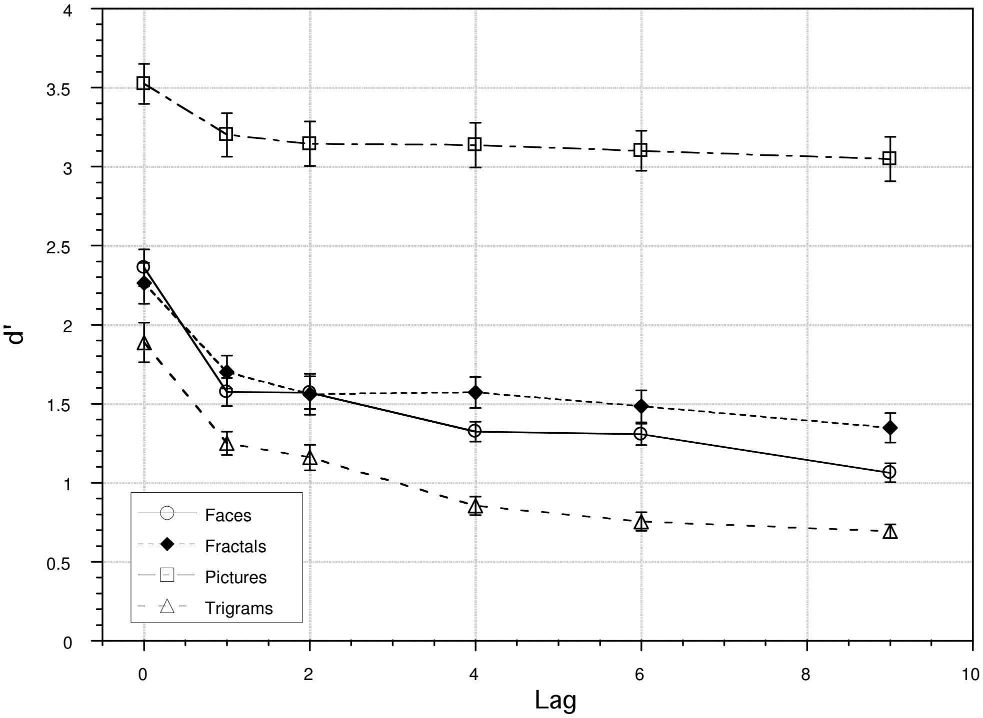

Figure 2.7: *d'* as a function of lag. Data = mean ± SEM.

There was a significant interaction between the stimulus type and lag (F(15, 807)=12.3, p<0.001), indicating that the effect of lag differed according to the stimulus set being recognised. Tukey's tests revealed that, whilst *d'* was significantly greater when there were no stimuli intervening between study and test (lag 0) for all stimulus types, further increases of lag did not affect retention of the picture stimuli. This differed from the pattern seen for other stimulus types where significant decline of *d'* was observed between lags 1 and 9 (faces, fractals and trigrams: p<0.01), and between lags 2 and 9 (faces and trigrams: p<0.001).

Differences were also observed between lags 1 and 4 ($p<0.01$), lags 1 and 6 ($p<0.001$) , lags 2 and 4 ($p<0.05$), and lags 2 and 6 ($p<0.001$) in the trigrams condition. It appeared that increasing lag had little effect on recognition of pictures, but resulted in a large and progressive decline in performance for recognition of trigrams. Face and fractal retention curves were similar in form, but shallower.

2.1.3.2 False alarm rates

False alarm rates were analysed in order to make inferences about how difficult discrimination between individual items within sets was. The more difficult stimuli were for participants to tell apart, the higher the number of false alarms they would be expected to make. The effect of stimulus type on false alarm rates (see Figure 2.8) was compared using a one-way ANOVA, and a significant main effect was revealed ($F(3,196)=59.3$, $MSe=0.012$, $p<0.001$). Tukey's post-hoc tests revealed that participants recognising pictures made significantly fewer false alarms than those recognising all other stimulus types ($p<0.001$), and that participants recognising trigrams made significantly more false alarms than those recognising all other stimulus types ($p<0.001$). False alarm rates for faces and fractals were not significantly different.

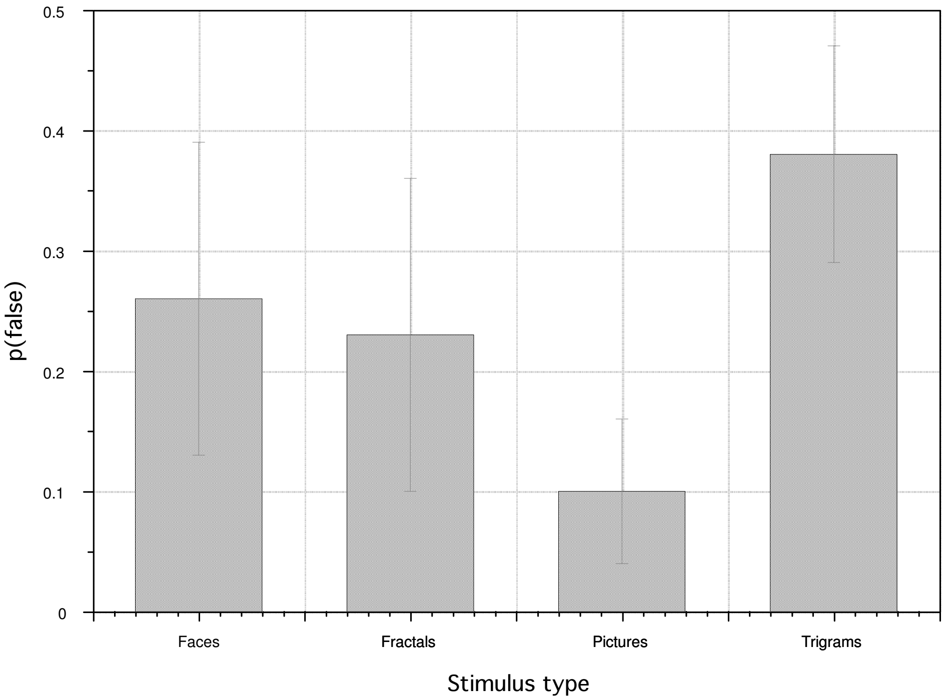

Figure 2.8: False alarm rates associated with recognition of different stimuli. Data = mean ± SEM.

2.1.3.3 Reaction times

The reaction time data for hits across the four conditions is shown in Figure 2.9. Mauchly's test of sphericity was significant for lag, and Greenhouse-Geisser epsilon corrected results are reported below. A similar 4 x 6 ANOVA to that described for *d'* data was carried out on these data, and revealed a significant main effect of stimulus type (F(3,196)=3.90, MSe=105000, p<0.05). Tukey's post-hoc analysis revealed significant differences between pictures and faces (p<0.05), and between pictures and fractals (p<0.05) but not between any other groups. Pictures were recognised faster than the faces and fractals. Once again, a significant main effect of lag was revealed (F(4.56, 161)=90.5, MSe=21400, p<0.001). Post-hoc Tukey's tests revealed significantly faster reaction times for hits at lag 0 than all other lags (p<0.001). There were also faster reaction times at lag 1 than lag 4 (p<0.05), 6, and 9 (p<0.001), and at lag 2 compared with lags 6 (p<0.01) and lag 9 (p<0.05). The interaction between stimulus type and lag was significant (F(13.7, 893)=3.54, p<0.001). Tukey's post-hoc tests revealed significantly faster reaction times at lag 0 than all other lags for all stimuli, faster reaction times at lag 1 compared with lag 6 and 9 for faces (p<0.05) and trigrams (p<0.01, p<0.05), and at lag 2 compared with lag 6 for trigrams only (0.05). Whilst recognition latency was significantly slower when stimuli intervened between study and test for all stimuli, further significant increases in reaction time with increasing lag were only incurred for recognition of certain stimuli (faces and trigrams).

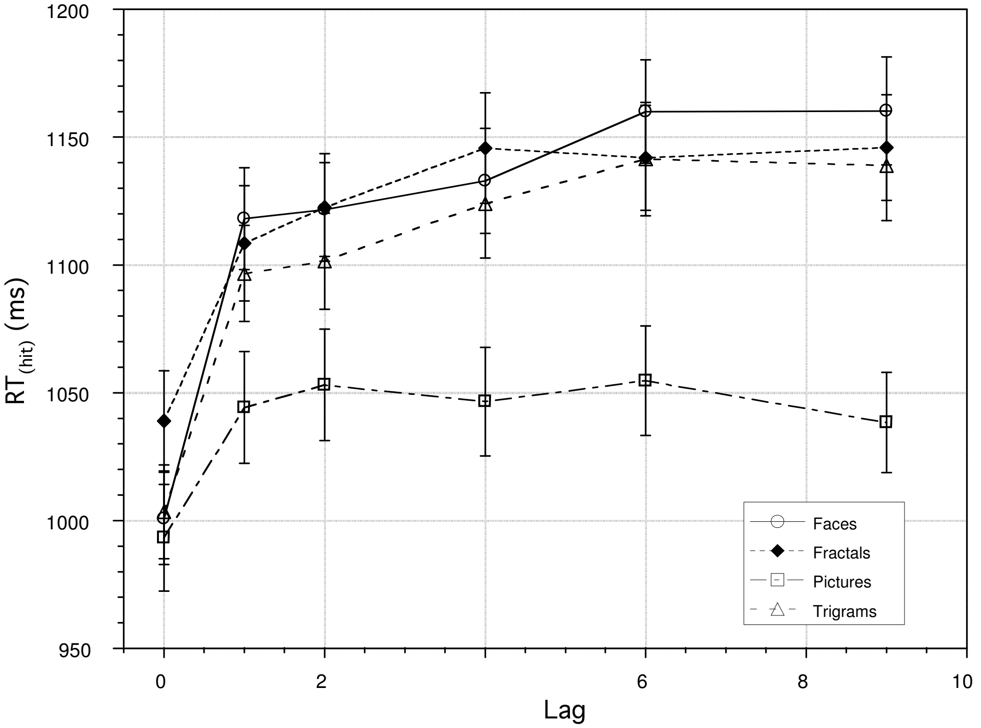

Figure 2.9: Reaction times for correct recognition (ms) as a function of lag. Data = mean ± SEM.

2.1.3.4 *Hit and false alarm rates by epoch*

Whilst the overall performance of participants during the experiment are of interest, inferences about stimulus sets can be made from analysing changes in the hit (p(hit)) and false alarm (p(false)) rates during the course of the experiment. It is possible to determine when (or if) these rates reach a ‘steady state’, and whether there are differences between different types of stimulus. By dividing data into epochs of 10 consecutive trials, a p(hit) and p(false) for each epoch was determined. These data are shown in Figure 2.10 and Figure 2.12, respectively. What is clear from these graphs is that both p(hit) and p(false) were low at the start of the experiment, and then rose to reach stable levels during the course of the experiment. The rate at which these changes occurred appeared to differ between stimulus types. To further explore these phenomena the p(hit) data for the first 10 epochs (see Figure 2.11) were compared in a 4 (stimulus type) x 10 (epoch) ANOVA. Mauchly’s test of sphericity was significant and Greenhouse-Geisser epsilon corrected data are reported. A significant effect of stimulus type was found ($F(3,193)=53.6$, $MSe=0.071$, $p<0.001$) with Tukey’s post-hoc tests revealing significantly higher p(hit) values for pictures than all other stimuli during this period ($p<0.001$). Values for faces were also greater than those for fractals ($p<0.05$). An effect of epoch was also found ($F(6.87,133)=16.1$, $MSe=0.053$, $p<0.001$) and there was a significant interaction between epoch and stimulus type ($F(20.6,133)=2.63$, $MSe=0.053$, $p<0.001$). Using Tukey’s post-hoc tests it was determined that p(hit) reached a stable level (defined as being the first epoch for which significant differences between p(hit) at that epoch at later epochs were not found) at epoch 3 for faces, epoch 4 for trigrams, and epoch 5 for fractals. It was stable from the start of the experiment for pictures.

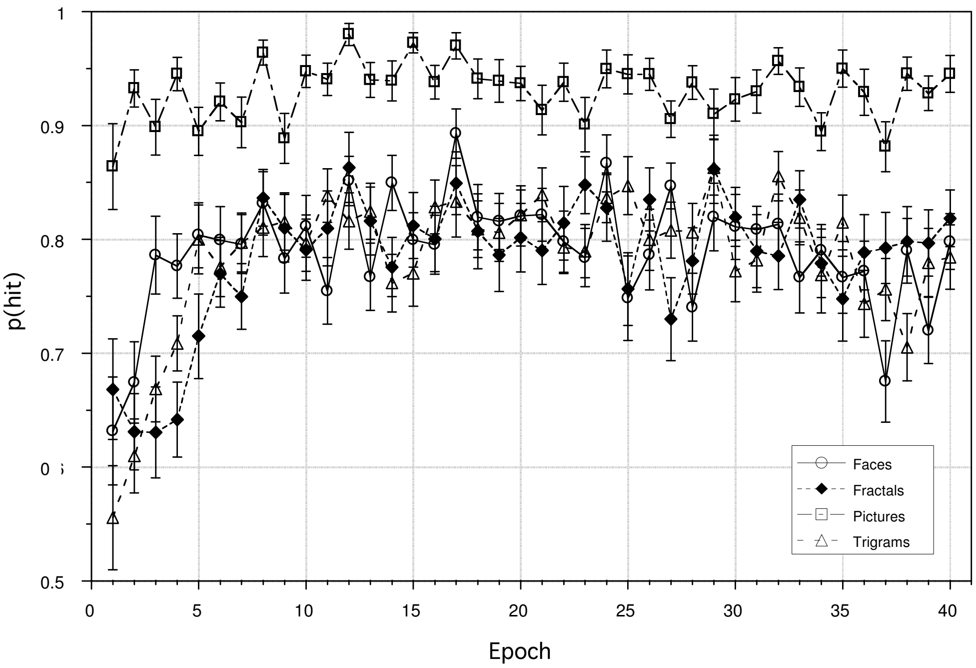

Figure 2.10: Hit rate variability during the course of the experiment. Hit rate was calculated for 10 trial epochs. Data = mean ± SEM.

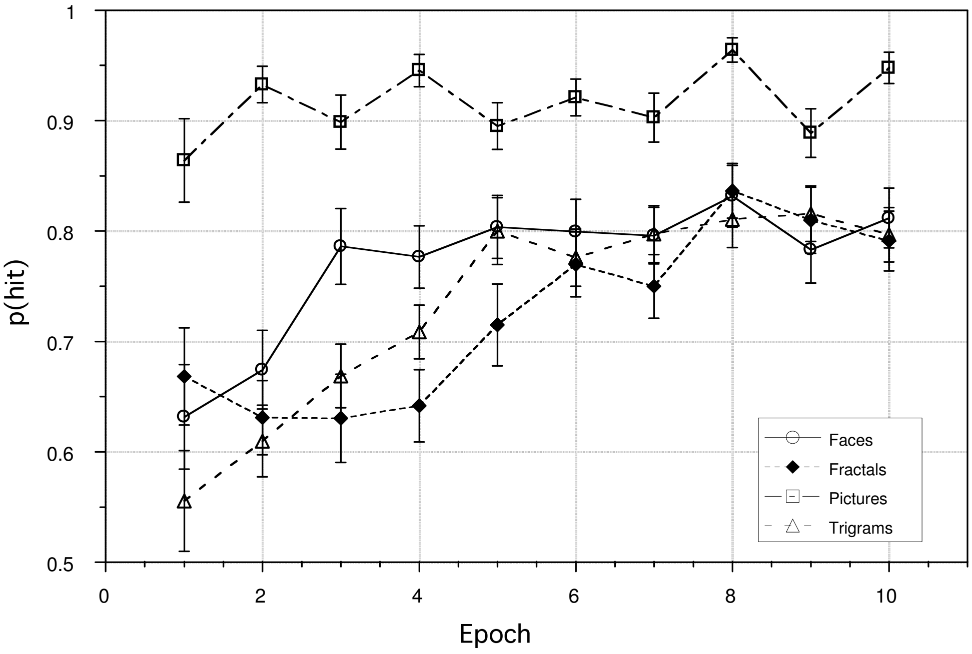

Figure 2.11: Hit rate data for the first 10 epochs. Data = mean ± SEM.

False alarm rate data was put into a similar 4 x 10 ANOVA as that described for the p(hit) data above (see Figure 2.13). Mauchly's test of sphericity was significant, and Greenhouse-Geisser epsilon corrected degrees of freedom are reported. Stimulus type was significant (F(3,179)=75.9, MSe=0.101, p<0.001), and Tukey's post-hoc analysis revealed significant differences between all stimulus types (all p<0.001, except faces vs. fractals, p<0.01). Trigram p(false) values were the highest, followed by faces, fractals, and pictures respectively. Epoch was significant (F(6.89,123)=19.2, MSe=0.057, p<0.001) as was the interaction between stimulus type and epoch (F(20.7,123)=3.37, p<0.001). In a similar manner as described above for p(hit), stabilisation of p(false) was determined to occur at epoch 3 for trigrams, and epoch 5 for faces and fractals. Again, values were stable for pictures throughout this period.

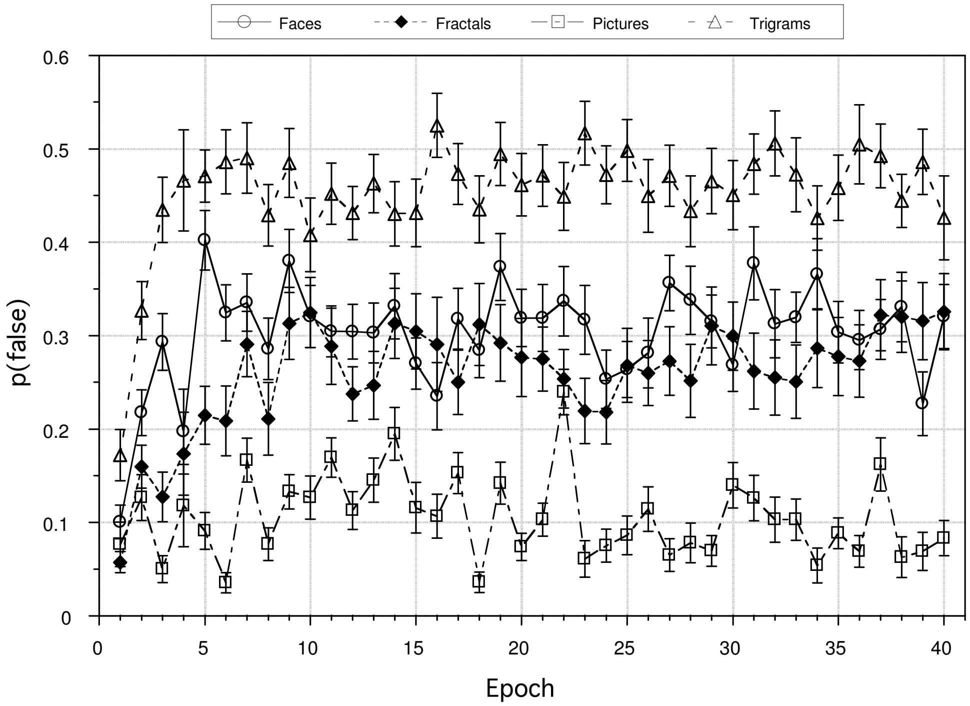

Figure 2.12: False alarm rate variability during the course of the experiment. False alarm rate was calculated for 10 trial epochs. Data = mean ± SEM.

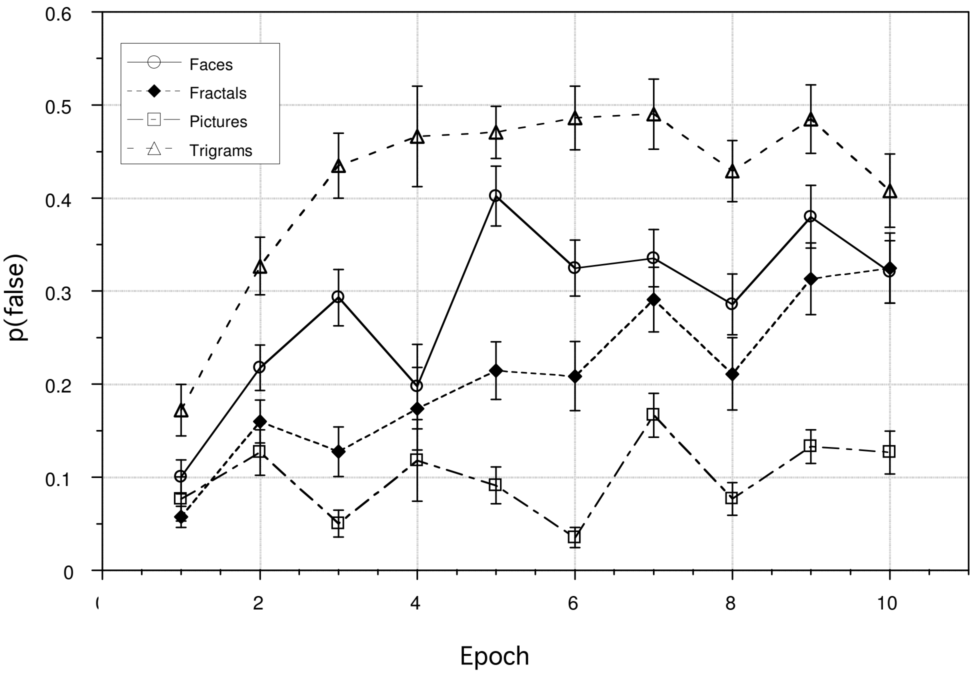

Figure 2.13: False alarm rate data for the first 10 epochs. Data = mean ± SEM.

2.1.3.5 Summary

To summarise, the results can broadly be said to divide the stimuli into three classes, based on participants' ability to recognise them. Recognition of pictures was far superior to that for all other stimuli, on all measures reported (*d'*, reaction times, hit rates and false alarm rates). Trigram recognition was worse in terms of accuracy (*d'*), although was no slower than that of faces and fractals. False recognition of the trigrams was very high. Faces and fractals were largely indistinguishable, although hit rate data from the early part of the experiment suggest that it took participants' longer to achieve a stable hit rate with the fractal stimuli than any other stimulus type.

2.1.4 Discussion

The results clearly demonstrate that the manipulation of the stimulus set had marked effects on recognition. The most robust of these effects was the clear superiority of participants' recognition for picture stimuli compared to that for all other stimulus sets. Pictures were recognised more accurately and rapidly than other stimulus types, and were associated with very low false alarm rates. No differences on any of these measures were observed between recognition of faces and recognition of fractals, but recognition of trigrams was less accurate, apparently as a result of a greater false alarm rate as opposed to a reduced hit rate. Whilst no detailed curve fitting was carried out for these data[1], the statistical and graphical comparisons of different stimulus types at different lags reveal that the retention curves asymptoted at different points for different stimuli. Stimuli that were less accurately recognised overall were also associated with retention curves that asymptoted at higher lags.

The differences in performance that were observed must be attributable to some aspects of the composition of the stimulus sets. Indeed the picture set was composed in a manner that was very different to the others. Whereas stimuli in

---

[1] Although it would be possible to fit an arbitrary function to the data relating recognition performance (e.g. *d'*) to lag this was not done for a number of reasons. Firstly, although previous studies (e.g. Rubin et al., 1999) have shown that performance on continuous recognition memory experiments may be reasonably well *described* by a function composed of a series of exponentials, its psychological validity in terms of the underlying memory process(es) involved is currently indeterminate. Secondly, without an established *a priori* reason for preferring one permitted function over another, caution must be expressed when deciding, meaningfully, what is the best fitting function to a set of data, especially when the candidate functions are permitted to differ in terms of the number of free parameters available. All else being equal, more free parameters will inevitably lead to better fits. Indeed, the best fitting curve of Rubin et al. (1999) has a relatively large number (6) of free parameters, when compared with the number of lag values used in the

the other sets were relatively homogeneous and differed only in their features, stimuli in the pictures set differed from each other in two different ways, requiring two different types of discrimination. Within-category discriminations, between two different stimuli from the same category (e.g. two oranges), would be assumed to rely on similar processes to those required for the discrimination of two faces or two fractals, whereas cross-category discriminations (e.g. an orange and a dog) do not require such subtle abilities. Two items from different categories can be discriminated by verbal label, perhaps supplemented by visual discrimination of gross physical differences. The interference generated by stimuli intervening between the learning and test presentations is likely to be considerably less if these stimuli are from different categories, than if they are different items from the same category. Bearing in mind that within the total set of 200 pictures, pictures from 20 different categories were represented, a simple calculation (200/20) reveals that items from each category occurred once every 10 stimuli on average. In an experiment where the longest lag was 10, in an average run of 10 presentations the number of within-category discriminations required would be less than 1.

Recognition of the pictures set and recognition of the faces and fractals sets can be considered to be mediated by two different cognitive processes. Recognition of specific items from others of the same category (as was the case for faces and fractals), is different to the participants' ability to categorise (as seen for the pictures condition). As Goldstein and Chance (1970) argued, in criticism of Shepard's (1967) and Nickerson's (1968) studies of picture recognition, participants' ability to recognise pictorial and other visual stimuli cannot be assessed where these items vary widely within a set (as they also did

---

present study, and thus it would be surprising if it did not fit the recognition memory data well. The usefulness of such a function in the present context may be limited.

in the picture set featured in this study). The experimenter cannot tell whether the recognition of items within such sets is reliant on semantic factors, such as the gist of a scene, or is due to faithful memorisation of an item's visual representation. The difference between categorisation and 'true' recognition is revealed in the fact that the retention curve for the pictures appeared to be only minimally affected by lag. Apart from a difference between immediate recognition (lag 0) and delayed recognition (lags greater than 0) no further differences to either the *d'* or reaction time data were observed. Also, whilst there appeared to be some learning of the other stimuli categories during the first 40-60 trials, as observed in an increasing hit rate, no such learning was observed for pictures. Likewise, no build up of interference, as inferred from the increase in false alarm rates at the start of the experiment, was found for picture stimuli, although it was observed for the other classes of stimuli. These results suggest that discrimination between the pictures, probably on the basis of basic level naming, was already optimal, and was not improved by experience.

A previous study investigating both cross- and within-category discrimination of colours and facial expressions (Roberson & Davidoff, 2000), examined the effect of both verbal and visual interference between the study and test phases of a 2 alternative forced choice procedure. Participants were tachistoscopically presented with target stimuli followed by a delay of 5 or 10 seconds, and were then asked to choose which of two test stimuli matched the target. During the delay, participants were presented with either a blank card (control), a multicoloured dot pattern or face-like features (visual interference), or a list of nonbasic colour words or adjectives relating to emotional expressions (verbal interference). The most important finding of the study was that verbal interference selectively affected cross-category discriminations. Visual interference produced the same decrement on both within- and cross-category judgements. These findings provide a useful insight into the possible basis of the recognition advantage for cross-category discriminations, suggesting that it is entirely reliant on verbal processes. Visual discrimination was shown by the study to play some part in both types of discrimination, but whether the discrimination was categorical or not had no effect on the decrement in recognition caused by visual interference. One should be careful in applying these findings to the current results, as the stimuli employed here were considerably more visually complex than dot patterns, but the effect of naming is of undoubted important for discriminating between items from different categories.

Roberson and Davidoff suggested that the greater accuracy of recognition they observed for cross-category discriminations was due to implicit naming on the part of the participants. Implicit naming is the term given to the propensity for humans to automatically name items where this is possible, and has been proposed to account for the finding that memory span for pictures of objects does not exceed that for their labels (Schiano & Watkins, 1981). Naming is proposed by Paivio (1989) to be initiated by object recognition, which leads to referential processing whereby activation of verbal representations connected to the appropriate imagen (nonverbal representation of the object) occurs. The most appropriate verbal representation (logogen) (i.e. the one that exceeds the threshold of activation first) is selected to name the item. This dual coding theory therefore proposes two systems of cognitive representations, logogens and imagens, which are connected by association, and that naming is the process whereby an imagen activates the most strongly associated logogen.

In the current study, recognition of the picture images could be considered to be most amenable to this process, as the different categories have commonly used labels at the basic level of generality (e.g. apple, dog, etc.) The faces and fractals could only be distinguished with much more complex verbal labels, with correspondingly more complex connections to the initial imagen. It can be reasonably assumed that, given the assumptions of the dual coding model, implicit naming would have occurred most rapidly and with most certainty for items in the picture set where a single strong connection between the image of an item and the corresponding category label would have generated an unambiguous name. For trigrams, naming involved a longer more complex name (e.g. “three-X-two”), perhaps explaining why there was no reaction time advantage for correct recognition of these stimuli. The names generated for these items were also considerably more similar to one another, given that they were composed of different combinations of a limited set of component digits and letters. This may be the reason why trigram recognition gave rise to a much higher false alarm rate than other stimulus sets. As regards face and fractal recognition, it could be hypothesised that assigning names would have been useless in discriminating between the individual items and the naming process would take considerably longer in generating unique verbal labels. In this scenario a more efficient procedure would be to perform discrimination between the imagens themselves without ever relying on the referential process of naming.

There is evidence from brain imaging studies to support the hypothesis that different neural systems are involved in the identification of familiar objects and unfamiliar, unnamed items. Verbal encoding and retrieval of items in memory tests has been associated with activation of prefrontal and medial temporal regions of the left cerebral hemisphere, whereas nonverbal encoding and retrieval is associated with analogous regions in the right hemisphere (Kelley et al., 1998; J. J. Kim et al., 1999; A. C. H. Lee, Robbins, Pickard, & Owen, 2000; McDermott, Buckner, Peterson, Kelley, & Sanders, 1999; Wagner, Desmond, Glover, & Gabrieli, 1998), although quite often bilateral activation is found and has been explained with reference to dual coding theory. A recent PET study of object and face recognition (Simons, Graham, Owen, Patterson, & Hodges, 2001) found that recognition of familiar items was associated with activation of left ventrolateral prefrontal cortex (associated with verbal memory), whereas unfamiliar item recognition caused activation of occipital regions, associated with visual perception, most marked in the right hemisphere. These data provide strong evidence that familiar items are memorised by association with verbal labels, otherwise a process of perceptual discrimination occurs. Given this increasing utilisation of brain areas associated with verbal material for recognition as items become more familiar, it can be suggested that naming is a more efficient process for discrimination than visual discrimination. This would also account for the findings of this study – that nameable items are remembered more accurately and rapidly than unnameable items.

An alternative explanation for the finding that the picture stimuli were easier to recognise than the other types of stimuli, is that, due to the greater variation in visually detectable features present in this heterogeneous set, such as the shape and configuration of elements within the image, perceptual representations of pictures were more distinct from one another. Murdock (1960) defined the distinctiveness of a given stimulus as the extent to which it “stands out” relative to others within a set. Whilst the phenomenon was initially studied for cases where a few ‘distinctive’ items were recognised considerably better than the majority of ‘typical’ items within a set, more recent attempts to map distinctiveness to co-ordinates in multidimensional similarity space (e.g. Busey & Tunnicliff, 1999) demonstrate the possibility of defining distinctive and typical sets. Sets where the stimuli are much less constrained (such as the pictures) can be considered as having a high average distinctiveness between pairs of stimuli, whereas sets of stimuli generated with highly constrained parameters (such as the faces) give rise to stimuli that are very similar and less visually distinctive. Busey and Tunnicliff found that items defined as distinctive in terms of their position in multidimensional similiarity space, were associated with higher hit rates and lower false alarms in tests of recognition, and this is a similar profile to that seen for pictures in this study. However, given the finding of Roberson and Davidoff (2000) that visual interference causes a comparable decrement to measures of recognition in both within- and cross-category discriminations, it would appear that distinctiveness does not facilitate visual discrimination *per se*, but may make an item easier to label verbally. Indeed, this is borne out by the findings of Busey and Tunnicliff who noted that items distinct in multidimensional similarity space were also more likely to have qualitatively different features to other more typical faces (e.g. beards), allowing differential classification.

The question that remains is why discrimination of verbal labels is more accurate and faster than visual discrimination. It would appear to contradict intuition about the visual system and the order in which visual stimuli are processed. Information entering the human visual system passes through stages of elemental feature recognition in areas concerned with visual perception in the occipital cortex, before objects are recognised by networks of cells in the inferotemporal (IT) cortex. These stages can be thought of as encompassing the initial ‘object recognition’ stage of naming proposed in the dual encoding model (Paivio, 1986), through the production of object perception. Once this is achieved, information about object identity can be linked with verbal labels in other brain regions. Why is it, then, that naming, and the discrimination of names, can occur so much more rapidly than visual discrimination of the perceived object from others in memory?

Some light may be shed on the question by authors investigating a rather different phenomenon – perceptual learning (Ahissar & Hochstein, 2002). These authors proposed that learning begins at higher levels in the perceptual hierarchy, and progresses downwards in reverse to the visual process. Higher levels, such as populations of neurones involved in categorisation and recognition (e.g. IT cortex), are the default location for learning because these are tuned to global entities, including objects, rather than local features, and can generalise over the spatial dimensions of stimuli. These neuronal populations provide the information necessary for the performance of simple visual discriminations. However, where more specific spatial discriminations are required populations of neurones at lower levels in the visual system are better suited for the task. Ahissar and Hochstein propose a search tree seeking the best neuronal population to yield the best signal:noise ratio according to the task demands, starting at high levels and continuing down through the hierarchy.

Whilst the task demands required of participants in the current study are unlikely to have involved perceptual learning of the kind studied by Ahissar and Hochstein, it seems possible that a similar reverse hierarchy may be operating. Once visual processing has led to perception of the stimulus, higher level cognitive processes may then begin to compare the item to others in memory. If the high-level visual representation outputted by the visual system is sufficient to discriminate it from objects from other categories the task may be completed. Otherwise, a search process may commence looking for lower levels of visual representation to distinguish the specific visual features of the object from others of the same class, taking correspondingly longer to generate a response.

Whether the recognition advantage for pictures is due to verbal labelling of the stimuli or simply faster and more accurate visual discrimination, it has been hypothesised that the form of memory retention is the same for recognition of both verbal and visual stimuli. The findings of Ward et al. (2005) using lists of stimuli in a 2-alternative forced choice memory task, and those of Doty and Savakis (1997) with continuous recognition, support this theory. Ward and colleagues (2005) compared recognition for unfamiliar faces and nonwords and found similar profiles, with limited recency and no primacy. Doty and Savakis (1997) compared the recognition of common words and random abstract stimuli during continuous recognition, and found evidence that, whilst recognition of the words was more accurate, the shape of the retention curves were similar (see

Figure 1.5). Whether the better recognition profile observed for pictures here was the result of a switch to simple verbal coding, or of the detection of more salient visual features, the result might have been the same. Only further experimentation will shed light on how categories are encoded for pictures.

Whilst the discussion thus far has focussed on the overall distinctions between stimulus types, averaged across the whole experiment, sorting the data into epochs allowed the examination of changes during the course of the experiment. Interestingly, whilst initial increases in p(false) were observed during the first 100 trials, after this p(false) stabilised at a roughly steady state. This is different to the observation of Shepard and Teghtsoonian (1961) that p(false) was still increasing after several hundred trials. There were also differences between stimulus types in how quickly this occurred. These increases are presumed to be the result of a build-up of proactive interference from stimuli in previous trials. Perhaps the differences between the current data and those of Shepard and Teghtsoonian can be attributed to the nature of the stimuli employed. Whilst their three digit number stimuli were similar to the trigrams used here, there would appear to be yet more interstimulus similarity between items in their set than in the trigrams used in the present study. This might be expected to result in a considerably longer period before a steady state could be achieved, as was the case in their experiment. Amongst the stimuli used in the current experiment, a steady state of interference appeared to be reached during epoch 3 for trigrams, and then by faces and fractals during epoch 5. There appeared to be no build-up of interference at all for pictures. These findings are similar to those concerning changes in p(hit) during the first 10 epochs. Again, whilst there was no systematic change in p(hit) during this period for the pictures, increases in p(hit) stabilised at epoch 3 for faces, 4 for trigrams, and 5 for fractals; a similar pattern to that found for p(false). The findings indicate that for faces, fractals and trigrams, there is a period during which participants learn to optimally discriminate the stimuli. The fact that this does not occur for pictures indicates that it is not the result of learning how to carry out the task, but is the result of learning about the stimuli themselves. The differences between stimulus types can be explained with reference to participants' familiarity with discriminating them. Whilst the pictures were of common objects that participants were likely to have already been highly familiar with, the other stimuli were not likely to have been familiar to participants. This might explain not just the differences in learning to discriminate the stimuli, but also the differences in the level of interference caused by stimuli. Because faces, fractals and trigrams were initially novel, individual items may have seemed more distinctive at the start of the experiment than they did once participants became familiar with these stimulus types. Conversely, because the category of objects was constantly changing for participants in the pictures condition, novelty would have been maintained, to some extent, throughout the experiment.

In conclusion, the significantly more accurate and faster recognition of stimuli from the pictures set is likely to be attributable to the relative ease with which participants discriminate between categories as opposed to individual exemplars of a single category. Such categorisation, whether related to verbal labelling or not, was not feasible when comparing faces or fractals that varied only in the dimensions of their local features. The recognition of trigrams was seriously hindered by a higher rate of false recognition than observed for recognition of the other sets, and this was assumed to be due to the much greater interitem similarity for this set. An important question that remains to be answered, is whether or not pictures of cartoon objects are recognised better than faces and fractals even when selected from the same category. In order to test this, a future experiment could include a condition in which all of the pictorial stimuli are drawn from a single basic level category. This would be an excellent way of testing whether restricting the category from which recognised objects are drawn has any effect on the recognition of those items. In addition, the suggestion that interitem similarity had a significant effect on recognition of trigrams leads to the question of exactly what factors affect recognition in this paradigm. The current results suggest that an interference- or distinctiveness-based account might be most appropriate, although the effect of decay is yet to be tested.

2.2 Experiment 2: The recognition advantage for pictures is not solely the result of cross-category discriminations

2.2.1 Introduction

In Experiment 1, a robust recognition advantage for a set of picture stimuli over face-like, fractal, and trigram stimuli was observed. Recognition accuracy and speed were considerably superior for this set. The set was composed of clipart images selected from a number of different basic level categories (e.g. apples, umbrellas, animals) in such a manner that it was highly probable that all stimuli in a retention period would be drawn from different categories. This was different from the other stimulus sets used, which were composed of stimuli that all had a similar configuration but differed in their features. As such it was not possible to definitively state whether the advantage was a result of this different set composition, or whether it was due to the nature of the stimuli themselves. Was better recognition the result of an ability to rely on basic level category naming or the visual distinctiveness of items from different categories, or was it that images of familiar objects were better recognised than abstract images?

Whilst, in discussing the results of Experiment 1, evidence implicating the role of verbal naming and visual distinctiveness in determining recognition performance was heavily drawn upon, there is also a substantial literature suggesting that prior familiarity with a class of objects improves the ability to discriminate them. A general finding from the literature on such perceptual learning, is that stimuli that were previously psychologically “fused” can become differentiated with repeated exposure (Goldstone, 1998). Tarr and Gauthier (1998) found that familiarity with several members of a class of objects generalised to other objects of the same class, in a viewpoint-dependent manner. In other words, increased exposure to members of a class of objects results in learning about other objects from the same class. In the neurophysiological literature, Rainer and Miller (2000) have found evidence that monkey prefrontal cortex neurones that were responsive to visual objects became more finely tuned and resistant to image degradation with experience. A similar sharpening of tuning for IT neurones in monkeys was found to occur to a greater extent for diagnostic than non-diagnostic features learned during a categorisation task (Sigala & Logothetis, 2002). These results suggest that experience results in a heightened ability to discriminate between members of a category of objects. Indeed, Tarr and Gauthier suggest that the human ability to discriminate very perceptually similar faces is the result of our expertise with these stimuli, gained through repeated exposure to this category of objects. Could better recognition of pictures in comparison to other classes of stimuli result from expertise with these objects, in comparison to a lack of expertise with novel visual stimuli like schematic face-like stimuli and fractals?

In order to determine to what extent the advantage for recognition of pictures over other stimuli was the result of a reliance on categorical labels, the continuous recognition paradigm used in Experiment 1 was replicated with a set of picture stimuli drawn entirely from the ‘birds’[2] category. This made impossible the reliance of participants on basic-level verbal labels for the discrimination of items, and allowed the comparison of recognition of a class of familiar visual stimuli (pictures of different kinds of birds) with recognition of unfamiliar classes of visual stimuli (faces and fractals). It also made possible the comparison of

---

[2] Birds are popular with certain experimenters (e.g. Grill-Spector & Kanwisher, 2005; Rhodes & McLean, 1990), as an example of a relatively homogenous class of living objects, with which most humans have only limited expertise. They were also chosen for practical reasons, as the clipart collections from which we drew stimuli contained large numbers of bird pictures drawn from a similar viewpoint.

recognition of pictures from different categories (the pictures set from Experiment 1) with recognition of pictures from within a single category (the new birds set).

2.2.2 Methods

2.2.2.1 Participants

All of the 25 participants were students at the University of Nottingham, of which 16 were female and 9 male. The mean age of participants was 23 years. Inclusion criteria were the same as those for Experiment 1, and all participants met these criteria. The mean number of 'no responses' was 4.16, the mean probability of recognition for lags 0 and 1 combined was 0.91 (SD=0.08), and the mean false alarm rate was 0.22 (SD=0.09).

2.2.2.2 Stimuli

The pictures consisted of 200 unique clipart images of birds, converted to greyscale in order to make them more homogeneous. Examples are given in Figure 2.14.

Figure 2.14: Example bird stimuli.

2.2.2.3 Presentation, design, procedure and scoring

All stimuli were presented using the same hardware and software described for Experiment 1. Stimuli were presented in the centre of the screen using *Matlab* v5.2.1 (Mathworks UK), running the psychophysics toolbox. All images were converted to the same standard size of 59 x 59 pixels, as described previously. Session design, procedure, and scoring were all the same as described in the Methods section of Experiment 1.

2.2.3 Results

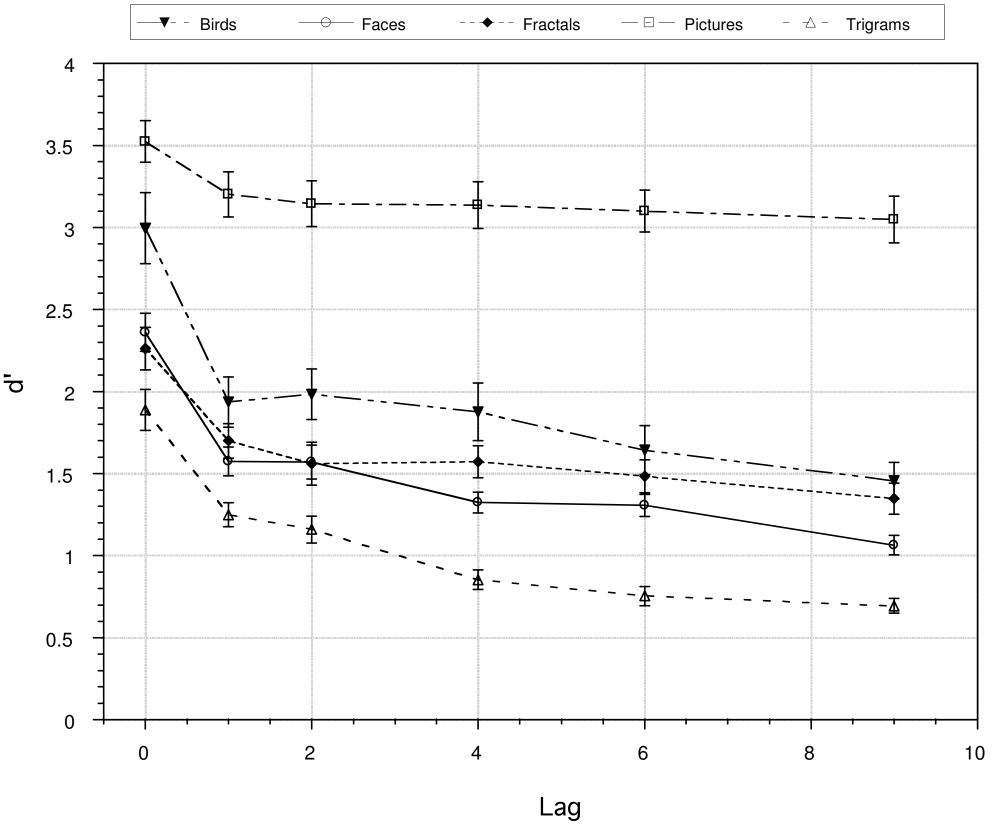

Figure 2.15: *d'* as a function of lag for birds and stimuli from Experiment 1. Data = mean ± SEM.

2.2.3.1 D-prime scores

The d' results for recognition of birds are shown alongside the results of Experiment 1 in Figure 2.15. Mauchly's test for sphericity was significant for lag, so Greenhouse-Geisser epsilon corrected ANOVA results are reported for this factor. A 5 (stimulus type) x 6 (lag) mixed ANOVA was performed on *d'* data from the birds set, and the 4 data sets from Experiment 1, and revealed a significant main effect of stimulus type (F(4,220)=85.5, MSe=2.20, p<0.001). Tukey HSD post-hoc tests revealed that birds were recognised more accurately than faces (p<0.01) and trigrams (p<0.001), but less accurately than the pictures (p<0.001). A significant main effect of lag was also revealed (F(4.15, 913)=239, MSe=0.284,

$p<0.001$) and there was a significant interaction between stimulus type and lag ($F(16.6, 913)=4.17$, $p<0.001$). Tukey tests revealed that, for bird stimuli *d'* scores were greater at lag 0 than all other lags ($p<0.001$), at lag 1 than lag 9 ($p<0.01$), at lag 2 than lag 9 ($p<0.01$), and at lag 4 than lag 9 ($p<0.05$).

2.2.3.2 False alarm rates

The mean false alarm rate for birds was 0.220 ± 0.094, and false alarm rate data were compared with data from Experiment 1 using a one-way ANOVA. A significant main effect was revealed ($F(4,220)=46.1$, $MSe=0.009$, $p<0.001$). Tukey's post-hoc tests revealed that recognition of birds was associated with greater false alarm rates than pictures ($p<0.001$), and lower false alarm rates than trigrams ($p<0.001$), in a similar pattern to the results observed for faces and fractals.

2.2.3.3 Reaction times

The reaction time data for correct recognition of birds accompanied by data from Experiment 1 is shown in Figure 2.16. Mauchly's test of sphericity was significant for lag, and Greenhouse-Geisser epsilon corrected results are reported below. A similar 5 x 6 ANOVA to that described for the *d'* data was carried out on these data, and revealed a significant main effect of stimulus type ($F(4,220)=2.93$, $MSe=109000$ $p<0.05$). Tukey post-hoc analysis failed to reveal any significant differences between birds and other stimulus types. Lag was significant ($F(4.56,1000)=110$, $MSe=4570$, $p<0.001$), and post-hoc Tukey tests revealed the same differences between lags as those found in analysis of the data from Experiment 1 alone. The interaction between stimulus type and lag was also significant ($F(18.2,1000)=3.02$, $p<0.001$) and Tukey post-hoc tests revealed significantly faster reaction times for recognition of birds at lag 0 than all other lags ($p<0.001$), as was observed for fractals and pictures. This was different to the pattern for faces and trigrams, which were associated with differences between reaction times at greater lags.

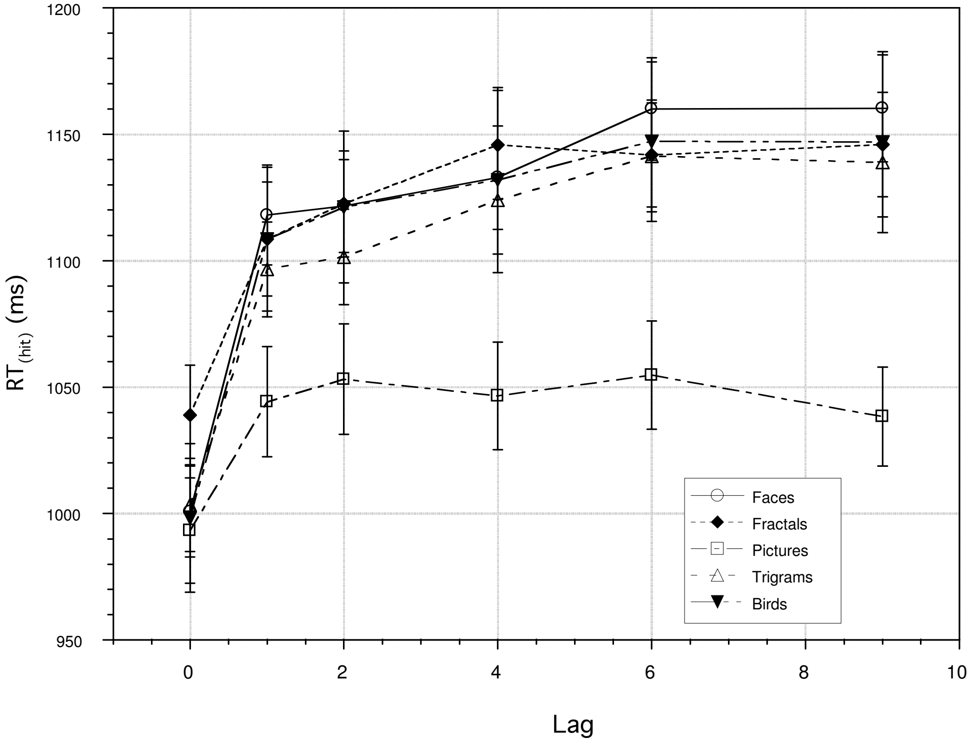

Figure 2.16: Reaction times for correct recognition (ms) as a function of lag for birds, with data from Experiment 1 for comparison. Data = mean ± SEM.

2.2.3.4 Hit and false alarm rates by epoch

Comparison of p(hit) and p(false) for birds by epoch (as described for Experiment 1) was carried out, and the data for the first 10 epochs are shown in Figure 2.17 and Figure 2.18 respectively. The p(hit) data for this period were compared in a 5 (stimulus type) x 10 (epoch) ANOVA. Mauchly's test of sphericity was significant for epoch and Greenhouse-Geisser epsilon corrected data are reported. A significant effect of stimulus type was found (F(4,220)=32.7, MSe=0.0764, $p<0.001$) with Tukey's post-hoc tests revealing significantly higher p(hit) values for birds than fractals ($p<0.001$) and trigrams ($p<0.01$), and lower values than pictures ($p<0.001$). A main effect of epoch was also found

(F(6.94,1500)=13.8, MSe=0.0518, p<0.001) and there was a significant interaction between epoch and stimulus type (F(27.8,1500)=2.39, MSe=0.0518, p<0.001). Using Tukey's post-hoc tests it was determined that p(hit) for birds did not significantly differ across these first 10 epochs, in a similar manner to the pattern observed for the pictures condition in Experiment 1.

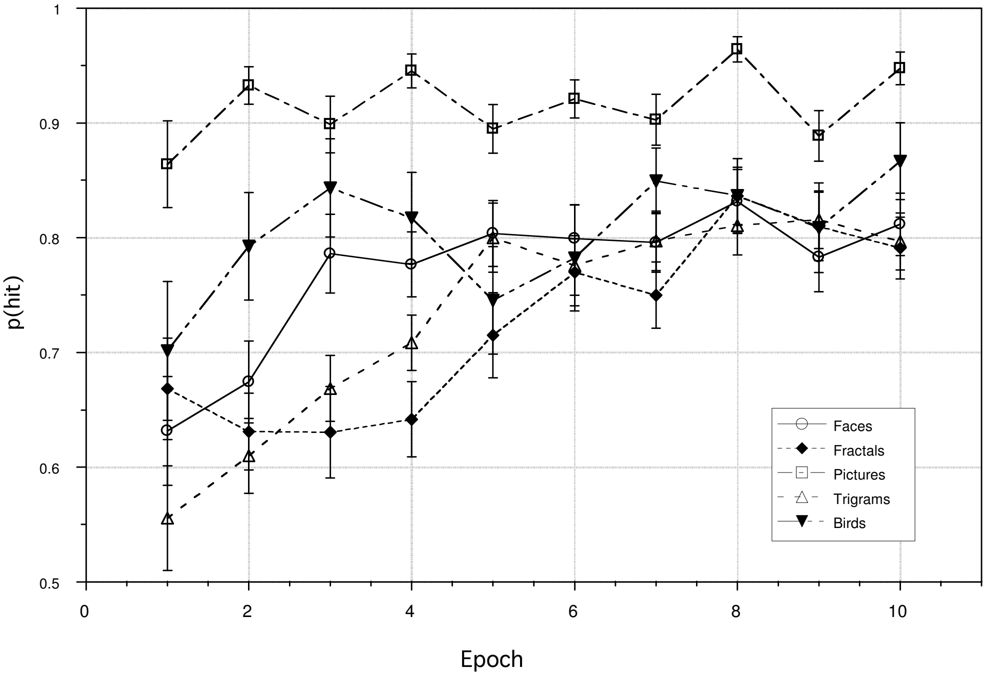

Figure 2.17: Hit rate data for birds for the first 10 epochs, compared with similar data from Experiment 1. Data = mean ± SEM.

False alarm rate data were analysed using a similar 5 x 10 ANOVA as that described for p(hit) data above. Mauchly's test of sphericity was significant for epoch, and Greenhouse-Geisser epsilon corrected degrees of freedom are reported. The main effect of stimulus type was significant (F(4,220)=53.5, MSe=0.0967, p<0.001), and Tukey's post-hoc analysis revealed significantly greater false alarm rates for birds compared with pictures (p<0.001) and lower false alarm rates compared with trigrams (p<0.001). The main effect of epoch was significant (F(7.12,1410)=16.8, MSe=0.0561, p<0.001) as was the interaction between stimulus type and epoch (F(28.5,1410)=2.65, MSe=0.0561, p<0.001). Post-hoc tests demonstrated that whilst p(false) for birds was significantly lower at epoch 1 compared with epoch 9 (p<0.05), no other differences were found for these stimuli. This was similar to the effect of epoch on false alarm rates to fractals and pictures, but different from the pattern for faces and trigrams, where significant differences were found at greater lags as well. As with the lack of change observed for p(hit), the relative lack of interference build-up measured by p(false) across this period suggests that birds are similar to the pictures set for these measures.

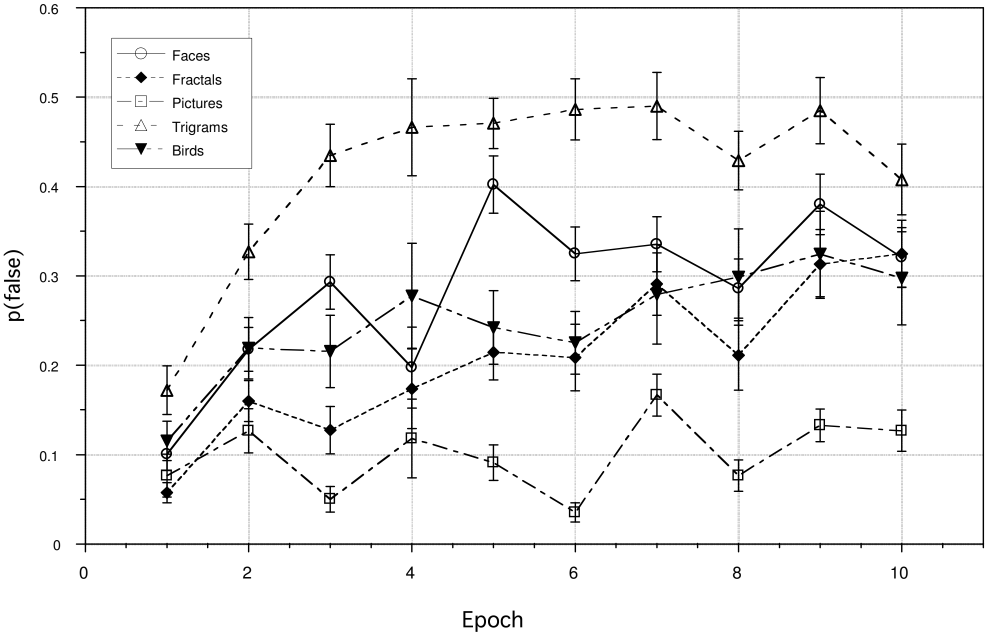

Figure 2.18: False alarm rate data for birds for the first 10 epochs, in comparison to data from Experiment 1. Data = mean ± SEM.

2.2.3.5 Summary

The comparison of recognition of the new birds set of stimuli with that of previously defined sets suggests that the birds set shares some characteristics with faces and fractals, and some with pictures. Unlike the relatively flat *d'* curve for pictures, the retention curve for birds continues to decrease at lags greater than 0, in a manner similar to those for other stimuli. However, whilst *d'* scores for birds were significantly lower than those for pictures, they were also significantly greater than the values for faces and trigrams, suggesting that there is a recognition advantage for birds over these stimuli. False alarm rates for birds were similar to those observed for faces and fractals, being significantly lower than those for picture recognition. However, analysis of p(hit) and p(false) during the first quarter of the experiment revealed no systematic effects of epoch on either measure for birds, whereas increases across early epochs were found for both measures with faces, fractals and trigrams. The lack of an effect related to epoch for birds is similar to that noted for picture stimuli in Experiment 1.

2.2.4 Discussion

By taking images from one form of stimulus (clip art pictures) and including items of only one category (birds) it was possible to show that most of the recognition advantage that was observed for pictures in Experiment 1 was related to a reliance on cross-category discriminations. Once pictures were drawn from only one category to make a stimulus set, and recognition required discrimination between individual exemplars of this bird category, the false alarm rates observed were similar to those previously seen for faces and fractals. The speed of recognition was also significantly slower than that for pictures, and was indistinguishable from the reaction time profiles for faces, fractals and trigrams. These data are evidence that the discrimination of items from different categories can be achieved more rapidly, and with less interstimulus interference, than discrimination of similar items selected from the same category.

The *d’* values calculated for recognition of the birds were significantly lower than those calculated for pictures, providing further evidence that the cross-category composition of the pictures set had a large effect in producing the recognition advantage for these stimuli. However, whilst *d’* scores for birds were not statistically different from those of fractals, they were higher than those measured for faces, suggesting that discrimination of birds was better than discrimination of faces. This may seem surprising as both sets consisted of visual stimuli with a similar configuration whose features differed from item to item. However, there are two possible explanations for the difference. Firstly, there is the difference of familiarity. The faces were unfamiliar in the sense that they were not realistic representations of faces, being composed of face-like configurations of ellipses that were probably unlike anything the participants had seen before. The bird pictures, however, were realistic depictions of birds. It can be argued that the bird stimuli, despite their similarities, would have been more familiar to participants than the face stimuli and that, as such, the participants had some expertise in their perception and discrimination. This is likely the result of some form of perceptual learning about stimuli and may be related to the reorganisation of higher visual perceptual and mnemonic areas with experience (C. A. Erickson, Jagadeesh, & Desimone, 2000; Rainer & Miller, 2000). In addition, participants might have used subordinate level categorisation of this familiar category to discriminate between certain individual birds (e.g. sparrow vs. blackbird). However, given that this familiarity-based explanation would also predict an advantage of bird recognition over fractal recognition, and no difference was found between these stimulus types, this explanation seems incomplete.

The second possibility is that there was more variation in physical characteristics between the individual bird stimuli than there was between individual face stimuli. As the faces were generated algorithmically with constraints on the level of permitted variation, whereas the birds were simply selected from a set of pictures of a natural category with no such regularities, it can be assumed that there was greater variation between individual birds than individual faces. As such, birds would be expected to be more distinct from one another than faces, and would have been recognised with greater accuracy. This might also explain the lack of difference between *d'* values for birds and fractals as, whilst fractals were also generated by computer algorithms with constraints on parameters, the nature of the algorithm meant that small variations in the parameters might lead to large variations in the outputted image. Whilst face parameters directly corresponded to feature dimensions, such as the height of a face or the width of an ‘eye’, fractal parameters corresponded to limits such as the maximum amount a line could ‘deflect’ (see the Methods section for Experiment 1 for more details about how these stimuli were generated). The net result of these differences was that fractal images could radically differ in their shape whilst faces retained the same basic shape throughout. Whilst this difference was not enough to generate significant differences in participants' ability to distinguish between faces and fractals, it may have been enough to make fractal recognition indistinguishable from bird recognition. The mean d' scores show a pattern in which fractal d' scores generally fall between the scores for birds and faces.

However, this explanation of difference is hard to square with the false alarm data, that show statistically similar scores for all three stimulus types. If greater visual distinctiveness was the reason that bird stimuli were better recognised than faces one would expect to see this reflected in both higher hit rates and lower false alarm rates for the bird stimuli. However, whilst the mean false alarm rates for each stimulus type were not statistically distinguishable, the order of the mean values was what one would expect in that false alarm rates were lowest for birds, followed by fractals and then faces. As with the *d'* scores, which did not show a dissociation between fractals and faces or fractals and birds, the differences are small and further research is needed to establish the reliability of this finding. It can certainly be concluded that, whilst small differences in recognition accuracy may be produced by these different stimuli, these pale in comparison to the much larger distinction between pictures and all other stimuli as a result of cross-category discrimination.

Interestingly, the results that show birds to be most similar to pictures are the epoch-based analyses of hit and false alarm rates. Whilst there are periods for face, fractal and trigram stimuli where participants' hit and false alarm rates increase significantly over the course of several epochs, this is not the case for either pictures or birds. In the discussion of the results of Experiment 1 it was suggested that the lack of interference and learning for the picture stimuli was the result of an already optimal perceptual representation of these familiar items. This would mean that participants could not learn to discriminate them any better than they were already doing. This result lends support to the familiarity-based explanation for the superiority of bird recognition over other stimulus types. If these stimuli were more familiar to participants than the novel computer-generated stimuli one would expect to see better discrimination of them accompanied by little to no learning during the course of the experiment, as was the case.

Whilst this experiment has clearly demonstrated that the large differences between picture stimuli and all other stimulus types observed in Experiment 1 were mainly the result of the cross-category composition of that set, it has also revealed important differences between recognition of pictures of familiar objects and recognition of artificial stimuli. Recognition was superior for the more familiar bird stimuli although it is unclear whether this was the result of an ability to use subordinate level verbal labels, or perceptual learning resulting in a better ability to visually discriminate between individual items. Indeed it may be the case that the bird stimulus set contained more interstimulus variety of physical characteristics than other sets as a result of its composition. Further experiments manipulating the physical similarity of stimuli resistant to naming are required, in order to determine the exact involvement of similarity in recognition performance.

Chapter 3 The nature of forgetting in continuous recognition

3.1 Experiment 3: Separating decay and interference in continuous recognition

3.1.1 Introduction

Experiments 1 and 2 made extensive use of the continuous recognition paradigm and plotted conventional curves of recognition performance as a function of lag. Whilst lag nominally denotes the number of items intervening between study and test presentations, because trials are presented at a constant rate, the time between study and test is also proportional to lag. In explaining the results obtained it is not possible to distinguish between the effect of time and the effect of intervening stimuli. Traditional accounts of loss of information from memory utilise two concepts: decay and interference. Decay-based theories propose a loss that is directly related to the time elapsed since exposure to the material to be memorised, whereas interference theory proposes that forgetting is caused by interference from other information. In order to discriminate between the two, a continuous recognition experiment was carried out in which interstimulus interval was manipulated.

The standard model of short-term memory (STM) is based upon the idea of activation, a mnemonic property that keeps information in an accessible form (Cowan, 2001; Shiffrin, 1999). Information may be activated from long-term stores to become accessible in STM. Two counter-acting processes are proposed to determine whether information is retained or forgotten: rehearsal and decay. The activated items are presumed to rapidly decay with time unless they are consciously rehearsed. This rehearsal is sufficient to ‘refresh’ the memory, and counteract decay of activated information.

Evidence for decay as the mechanism of forgetting from STM came from experiments which prevented rehearsal by requiring participants to count backwards, or read a list of digits, after presentation of material (lists of letters) to be retained (J. Brown, 1958; Peterson & Peterson, 1959). The finding that there was a rapid loss of information to be recalled, was considered striking because the distracting material (digits) was sufficiently different from the study material (letters) that interference could be considered to have been avoided (Wickens, Born, & Allen, 1963). The conclusion the authors arrived at was that, during the period that rehearsal had been prevented, the memory trace decayed over time. Support for this conclusion came from later findings that forgetting occurred in the absence of interference from previous trials (i.e. on the very first trial) when ceiling effects were avoided (A. D. Baddeley & Scott, 1971). It has been suggested that decay is an adaptation to the statistical structure of the environment (Anderson & Milson, 1989) as it allows “overwriting” of the memory with the most recent and relevant information about the individual’s environment.

Some authors (e.g. Nairne, 2002), however, have dispensed with the rehearsal mechanism in their theories pointing to evidence that item-based differences in memory remain when rehearsal is blocked and articulation rates are held constant. For example, high-frequency words have a greater memory span than low-frequency words when the articulation rate is held constant (Roodenrys, Hulme, Alban, Ellis, & Brown, 1994), as do words over nonwords (Hulme, Maughan, & Brown, 1991), and concrete words over abstract words (I. Walker & Hulme, 1999). Theories relying solely on decay ignore the effects of the retrieval environment and the type of activity between study and retrieval. In contrast, many authors have attempted to create models of STM without appealing to the process of decay, explaining forgetting by the various forms of interference that can occur (e.g. G. D. A. Brown & Hulme, 1995; G. D. A. Brown, Preece, & Hulme, 2000; Neath, 1993).

The idea that the interference of existing memories with new information is responsible for forgetting has a long history dating back to the work of McGeoch (1932). Interference can be either proactive (when associations learned prior to current learning interfere with current learning) or retroactive (when associations learned after the learning of material to be retained interfere with that learning). Keppel and Underwood (1962) proposed that proactive interference could account for the findings of Peterson and Peterson (1959). Rather than the memory trace for study material decaying whilst the participants were engaged in the distractor task, Keppel and Underwood argued that forgetting occurred due to the build up of interference from previously learnt material. Waugh and Norman (1965) found that varying the rate of presentation in a list-based memory task had little effect compared to the serial position of an item, suggesting that time was not as important as the number of intervening, interfering items between study and test. However, whilst interference-based explanations for forgetting are attractive, it is advisable to bear in mind the fact that any putative effect of time can be explained by various combinations of retroactive and proactive interference (Cowan, 2001).

Whilst interference and decay have historically been contrasted with one another as opposing explanations, a recent theory of Altmann and Gray (2002) combines decay- and interference-based explanations of forgetting. The functional decay model proposes that, when an attribute has to be constantly updated in memory, its value decays to prevent proactive interference. This rate of decay is theorised to adapt to the rate at which memory must be updated. Such a system would avoid memory becoming rapidly filled with items that would hinder recall of the most recent information (proactive interference). The authors found evidence for this in the gradual decline of performance on the current task set in a task-switching experiment. This decline was attributed to memory decay. Single-digit number stimuli were presented, and participants were required to label them either as odd or even, or greater or less than 5. The less frequently the task was updated, the more gradual the decline in performance became, suggesting an adaptive rate of forgetting.

Shepard and Teghtsoonian designed the continuous recognition paradigm with the intention of studying recognition under conditions in which “the possibility of rehearsal is minimized while the interference of preceding material is maximized” (Shepard & Teghtsoonian, 1961, p. 303, p. 303). Their aim was to create an experimental situation in which forgetting occurred at a constant rate, although their results demonstrate that the false alarm rate was still gradually increasing at the end of the experiment, suggesting that this steady state was never actually achieved. This continual increase in the false alarm rate suggested that memory stores became overloaded with previous stimuli, making novel stimuli progressively more difficult to distinguish from an increasingly large set in memory. However, this effect was not found in Experiments 1 and 2; rather the false alarm rate initially increased and then reached a plateau. This finding compliments the findings of other researchers (Doty & Savakis, 1997; Hockley, 1982) who noted that reaction times for correct responses in a continuous recognition procedure did not increase as a function of list position.

Whether or not overall list position of a stimulus affects old/new discrimination, a universally reported effect is the effect of lag. For example, Hockley (1982) noted that correct reaction times decreased with increasing lag, and that the function was approximately logarithmic. Doty and Savakis (1997) have studied old/new discrimination of completely novel stimuli from items tested in continuous recognition 1 or 2 weeks previously. Participants carried out continuous recognition of one set of complex, abstract images on week 1 (mean *d’*=1.82), and a second on week 2 (mean *d’*=1.26). On week 3 they carried out a similar continuous recognition experiment where half of the stimuli had been seen previously in either the first or second weeks. The effect was to greatly decrease discriminative ability to a mean *d'* of 0.52. The same participants carried out a similar procedure with common 4-letter words on weeks 4 and 5 and the resulting decrease in mean *d'* was from 2.97 to 0.89. Significant increases in reaction times were also observed. Nonetheless, the authors concluded that the "robust behavioral retention over a period of 1 week or more offers the likelihood that time *per se* is not critical" (Doty & Savakis, 1997, p. 292, p.292).

To the author's knowledge no study has attempted to dissociate the effects of number of intervening items and time under conditions of continuous recognition, with the same participants, within the same experimental session. In the current experiment the intertrial interval (the time elapsed between successive trials) varied systematically across three blocks of either "long" (8 sec), "medium" (3.5 sec) or "short" (1.25 sec) values. Within each block the frequencies of lags were controlled. As such the effect of intertrial interval and number of intervening items could be separated.

Most of the classic literature regarding forgetting is based on experiments using words and other verbalisable stimuli, and often refers to subvocal rehearsal as a mechanism for maintaining information (e.g. B. B. J. Murdock, 1961; Peterson & Peterson, 1959; Waugh & Norman, 1965). Given the assumption set out in Chapter 2 that certain visual stimuli (e.g. fractals) are not amenable to verbal labels, it was important to dissociate the effects of intertrial interval and number of intervening items for both verbalisable and non-verbalisable stimuli. Each participant, therefore, carried out two sessions: one where the stimuli were fractals and one where they were trigrams.

3.1.2 Methods

3.1.2.1 Design

The experiment had a three-way within subjects design. The independent variables were the type of stimulus tested, the time interval between successive stimuli (intertrial interval), and the number of intervening trials (lag). Stimulus type had 2 levels: fractals and trigrams. Intertrial interval had 3 levels: short (1.25 sec), medium (3.5 sec), and long (8 sec), giving total trial durations of 2.25 sec, 4.5 sec, and 9 sec, respectively. Lag had 3 levels: 1, 4 and 9 intervening items. The dependent variables were the participants' *d'* scores, and reaction times for correct recognition.

3.1.2.2 Participants

A total of 21 participants were tested, of which 16 were female and 5 male. All were students at the University of Nottingham with a mean age of 20±0.3 years. All participants had normal or corrected-to-normal vision.

In order to exclude data from participants who 'gave up' on the tasks, during one or other of the relatively long 35 min sessions, criteria were set for inclusion of the data in analyses. Each participant's data were entered into a one-tailed chi square, to determine that the number of correct responses were significantly above that expected by chance.

3.1.2.3 Stimuli

Two different categories of stimulus were employed: fractals and trigrams. These stimulus sets were the same as those described in the Methods section of Experiment 1.

3.1.2.4 Presentation

All stimuli were presented using an *Apple Macintosh G3* computer (300Mhz, 384Mb RAM) with an *ATI Radeon 7000* (32Mb) graphics card, on a 21" *Mitsubishi* colour display monitor (size: 1024 x 768 pixels, resolution: 72 x 72 dpi, refresh rate: 75Hz). Stimuli were presented in the centre of the screen using *Matlab* v5.2.1 (Mathworks UK), running the psychophysics toolbox. All bitmap images (fractals) were converted to a standard size of 59 x 59 pixels, and had an area of 2 cm x 2 cm when displayed at the screen resolution described above. Participants were seated at a distance of 57.5 cm from the screen so stimuli subtended an area of approximately 2° x 2° of visual angle.

3.1.2.5 Procedure

Each experiment took place in the form of two sessions, identical except for the type of stimuli (fractals and trigrams) and the order of intertrial interval blocks (see more info below). The order in which the participants carried out the fractals session and the trigrams session was counterbalanced across participants.

For each session a pseudorandomly determined frame of 120 trials was generated, providing 16 study and 16 test trials at each of the 3 lags. These lags (1, 4, and 9) were selected as early, mid, and late points from the retention curves obtained in Chapter 2. Twelve study-test pairs of unscored filler trials made up the spaces in the frame.

This frame, or block, was repeated three times, with novel stimuli each time. For each block the intertrial interval duration was different (1.25 sec, 3.5 sec, or 8 sec), with the order in which these occurred counterbalanced across participants. In addition, the experiment began with a buffer of 40 unscored filler trial pairs to prevent the occurrence of primacy effects. Intertrial interval for the buffer was always medium length (3.5 sec). This yielded a total of 400 trials.

Once the order of trials for the entirety of a session had been generated, the 200 stimuli were randomly assigned to the 200 pairs of trials, so that each participant experienced the stimuli in a different order. Once begun, the session progressed through the buffer, and the three blocks continuously without breaks in between, the aim being to maintain the steady state of performance encountered previously (see Chapter 2).

For each trial stimuli appeared on the screen for 1 sec. During this period participants were instructed to respond either ‘old’ or ‘new’ with the left or right mouse button respectively, ‘old’ to previously seen and ‘new’ to previously unseen items. In order that participants did not forget which button corresponded to which response during the experiment, a label of which was which was placed in clear view of the participant. After 0.5 sec the stimulus was replaced with a blank screen. One second later feedback was given for incorrect responses, or where no response was given. Feedback consisted of a low-pitched beep. No feedback was given if the response was correct. The total length of time that the screen remained blank varied according to the intertrial interval. For the short intertrial interval the total period was 1.25 sec (total trial length of 2.25 sec), for medium it was 3.5 sec (trial length = 4.5 sec), and for long 8 sec (trial length = 9 sec). The experiment lasted a total of 34.5 min.

3.1.2.6 Scoring

As mentioned previously, participants’ responses were only detected during the 1 sec presentation period of each stimulus. Once a response had been made it was final, and no opportunity for the correction of responses was allowed. Reaction time was measured as the latency from the start of the trial until the detection of the response.

3.1.3 Results

‘No responses’, when participants failed to respond within the timeframe, were not scored. The number of correct responses to test presentations for each lag, for each interstimulus interval, for each stimulus type, was divided by the total number of scored test presentations for that category, to obtain the hit rate, or proportion of correct responses.

From these results a one-tailed chi square test was conducted on each participant’s responses, to determine that the number of ‘old’ responses to repeated stimuli was significantly above that expected at chance. For four participants there was no significant difference, and those participants’ data were omitted from further analysis.

To obtain a hit rate independent of the participant’s bias to respond ‘old’, *d’* was calculated using hit rates and false alarm rates. The values for fractals and trigrams are presented as a function of the number of intervening items between study and test in Figure 3.1 and Figure 3.3, and as a function of study-test interval in Figure 3.2 and Figure 3.4. Study-test interval is the period of time, in seconds, elapsing between the end of the initial study presentation of a stimulus, and the onset of the second, test presentation.

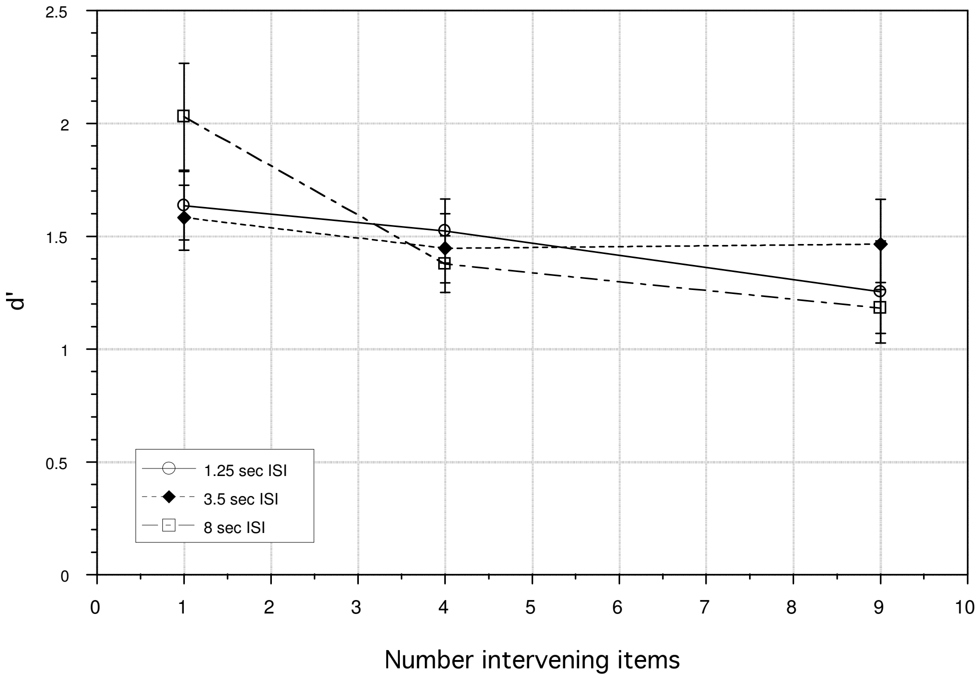

Figure 3.1: d' values for fractal recognition, as a function of the number of items intervening between study and test. ISI = interstimulus interval. Data = Mean ± SEM.

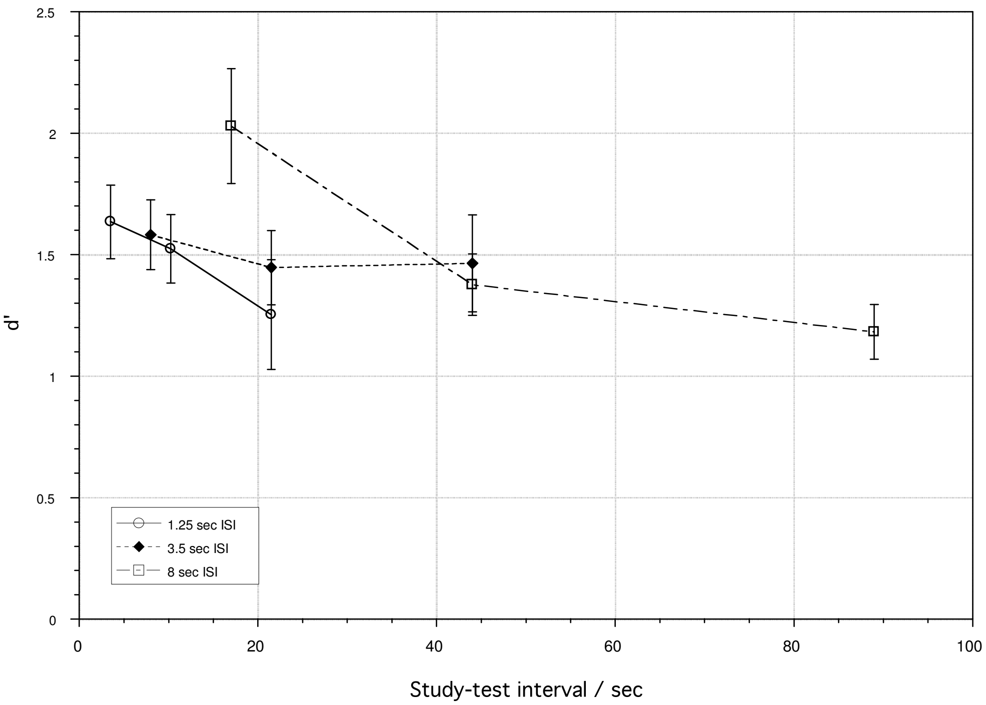

Figure 3.2: d' values for fractal recognition, as a function of study-test interval (sec). ISI = interstimulus interval. Data = Mean ± SEM.

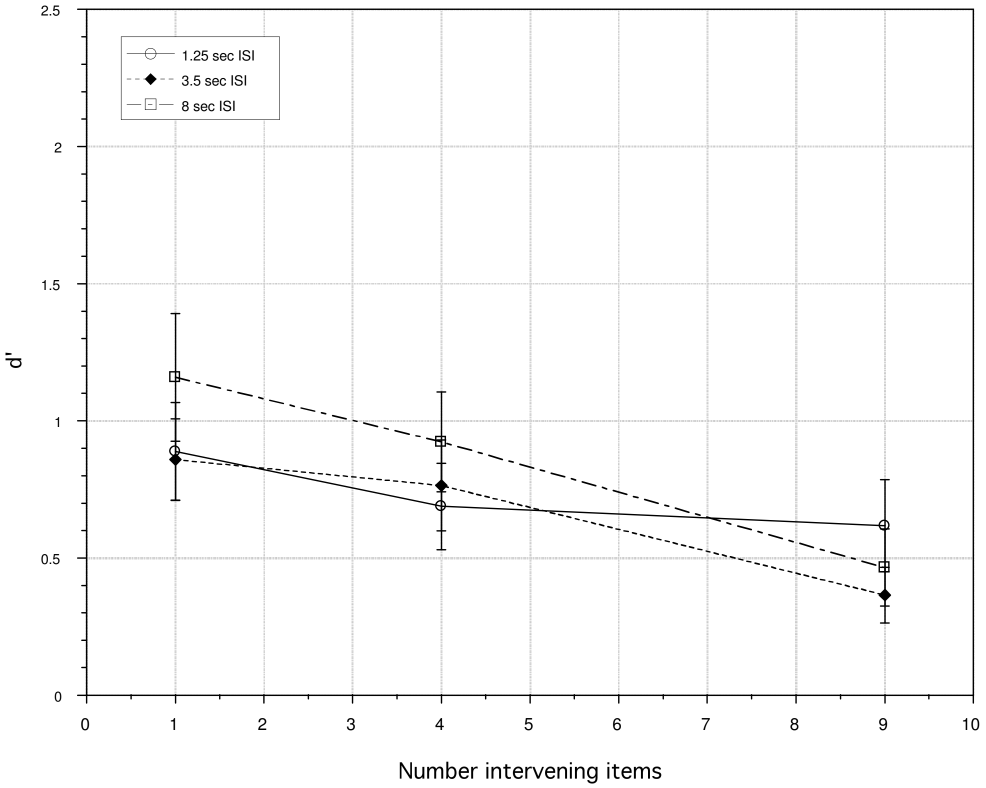

Figure 3.3: d’ values for trigram recognition, as a function of the number of items intervening between study and test. ISI = interstimulus interval. Data = Mean ± SEM.

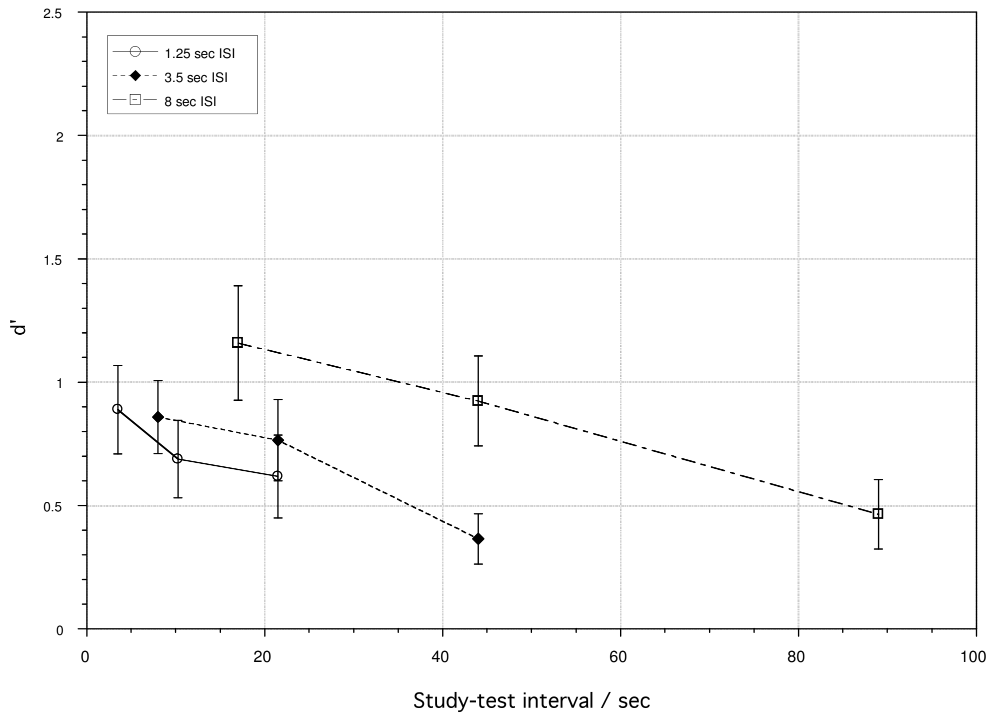

Figure 3.4: d' values for trigram recognition, as a function of study-test interval (sec). ISI = interstimulus interval. Data = Mean ± SEM.

3.1.3.1 D-prime scores

Mauchly's test of sphericity was significant for lag, and all subsequent results for lag are Greenhouse-Geisser epsilon corrected. A 2 (stimulus type) x 3 (interstimulus interval) x 3 (lag) repeated-measures ANOVA was performed on *d'* data and revealed a significant main effect of stimulus type (F(1,16)=77.5, MSe=0.558, p<0.001). Fractal *d'* scores were higher than those for trigrams, a finding consistent with the findings of Experiment 1. There was also a significant main effect of lag (F(1.33,21.2)=23.8, MSe=0.353, p<0.001), and Tukey's post-hoc tests revealed significantly higher scores at lag 1 than lag 4 (p<0.01), at lag 1 than lag 9 (p<0.001), and at lag 4 than lag 9 (p<0.01). This finding is unsurprising and consistent with the effects of lag reported in Chapter 2. There was no main effect of interstimulus interval. No interactions between any of the factors reached significance. The amount of time elapsing between study and test had no effect on the participants' performance, which was instead affected by the number of items intervening between study and test.

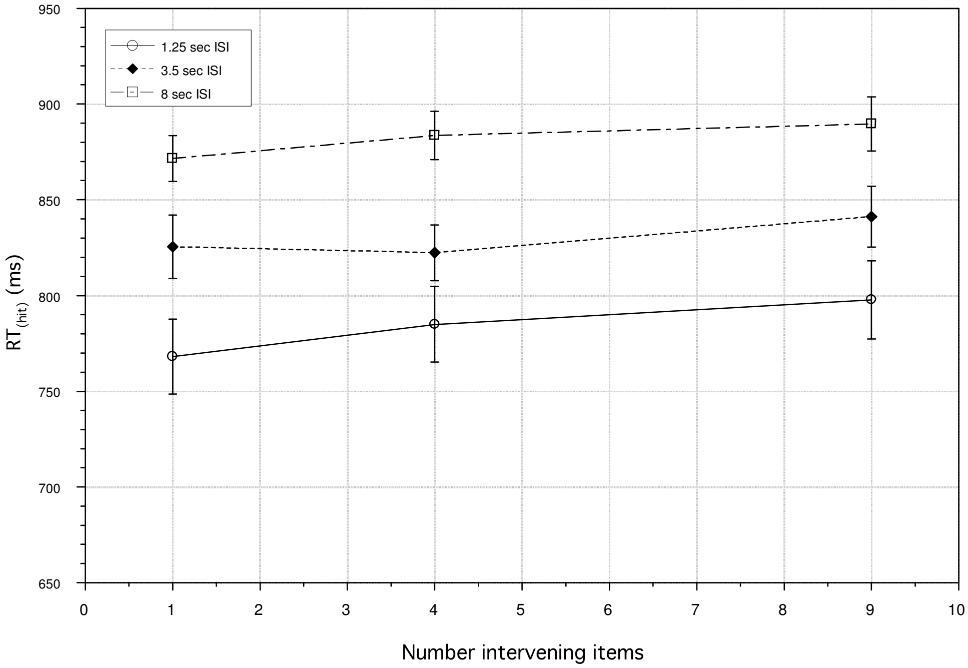

Figure 3.5: Reaction time values for correct recognition of fractals (ms). ISI = interstimulus interval. Data = Mean ± SEM.

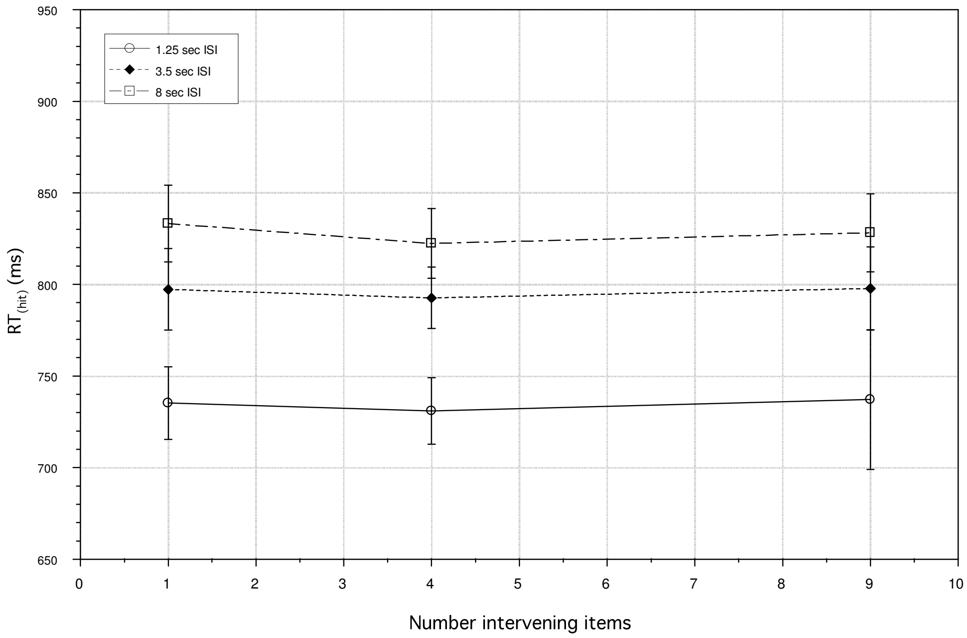

Figure 3.6: Reaction time values for correct recognition of trigrams (ms). ISI = interstimulus interval. Data = Mean ± SEM.

3.1.3.2 Reaction times

The reaction time data for hits (correct recognition) is shown in Figure 3.5 (fractals) and Figure 3.6 (trigrams). Mauchly's test of sphericity was significant for the stimulus type x intertrial interval interaction, the stimulus type x lag interaction, the intertrial interval x lag interaction, and the stimulus type x intertrial interval x lag interaction. All subsequent results pertaining to these interactions are quoted as Greenhouse-Geisser epsilon corrected values. A similar 2 x 3 x 3 repeated measures ANOVA to that described for *d'* results was carried out on these data, and revealed a significant main effect of stimulus type (F(1,16)=16.5, MSe=9640, p<0.01). Trigrams were recognised more rapidly than fractals, as was also the case in Experiment 1. There was also a significant main effect of intertrial interval (F(2,32)=64.9, MSe=3620, p<0.001). Short intertrial intervals were associated with faster reaction times for correct recognition than both medium and long intervals (p<0.001), and medium intervals were associated with longer reaction times than long intervals (p<0.001). There was no effect of lag on reaction times, in a departure from the results of Experiment 1, which showed significantly faster reaction times at lower lags for both stimulus sets. However, the reaction times in the current experiment were much faster on average than those recorded for Experiment 1, and this may have resulted in a ceiling effect. No significant interactions were observed.

3.1.3.3 Serial position

Data for p(hit) and p(false) was again calculated for 10 trial epochs and are shown in Figure 3.7. As was found in Chapter 2, both p(hit) and p(false) were initially low and then increased throughout the course of the experiment before reaching a steady state, and p(false) was higher for trigrams than fractals. Interestingly, the steady state for p(false) during trigram recognition appeared to have been achieved later in current experiment, suggesting that trigram recognition was more difficult when interstimulus interval varied. This same effect was not present in fractal recognition. The data were complicated by the counterbalancing of the order in which short, medium, and long interstimulus intervals occurred, and no consistent trends associated with the start of experimental blocks were observable.

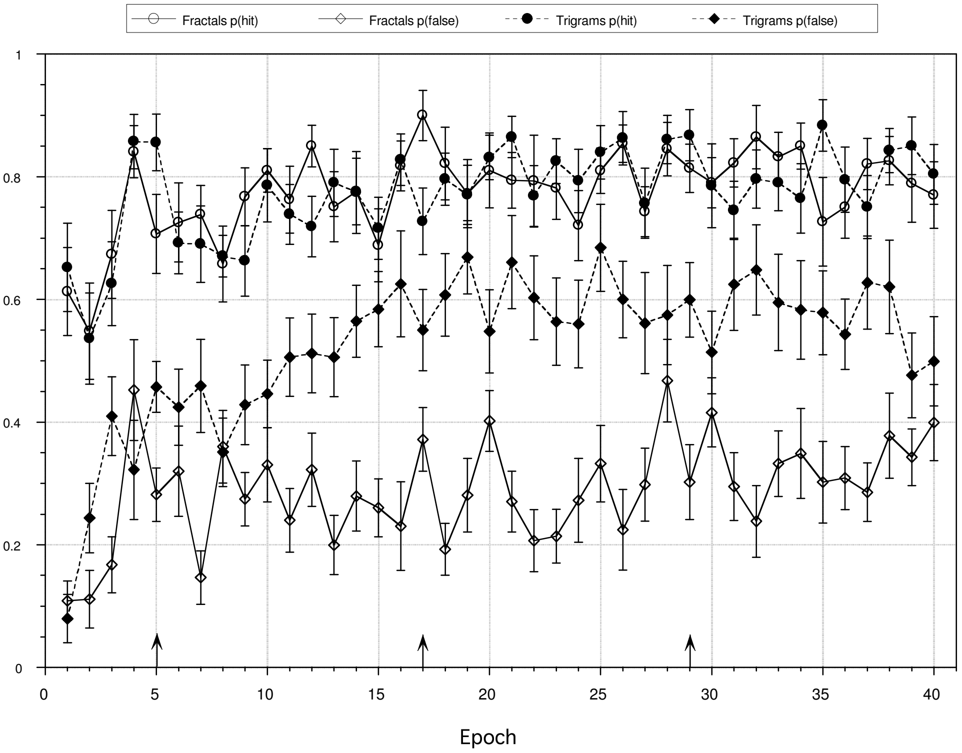

Figure 3.7: Hit and false alarm rates for recognition data by epoch. Epoch = 10 trials. Arrows indicate the start of each experimental block. Data = mean ± SEM.

3.1.3.4 Summary

Increasing lag decreased discriminative ability, as measured by the *d'* scores, but it had no effect on reaction times for hits. Conversely, whilst increasing interstimulus interval had no effect on *d'* scores, it led to significantly slower reaction times for correct recognition. That is, whilst accuracy of recognition was only affected by the number of items intervening between study and test, speed of recognition was affected only by the rate of presentation. No significant interactions were found between factors for any of the data, suggesting that recognition of both the fractals and the trigrams were similarly affected by interstimulus interval and lag. However, the serial position data suggest that trigram recognition was significantly adversely affected by the manipulation of interstimulus interval, resulting in higher false alarm rates, and a longer period of interference buildup.

3.1.4 Discussion

The results provided clear evidence that the number of items intervening between study and test was the major factor determining recognition accuracy, not the length of time elapsed. Even when the intertrial interval was increased 4-fold, *d'* values remained unchanged. Importantly, this suggests that interference is the major factor affecting memory performance in continuous recognition. Reaction times for correct responses to test stimuli (hits), however, were slower at greater interstimulus intervals. This suggests that it may take longer to retrieve the memory even though it has not decayed.

The classic decay-based model of memory is inadequate to explain the kind of memory processes involved in visual continuous recognition. The continual presentation of novel information in the paradigm would have prevented any hypothetical ‘visual rehearsal’, leaving the participant with only a decaying memory trace for each item according to the classic model. How then, are participants still able to recognise items with reasonable accuracy as long as 90 sec after their original presentation? Visual rehearsal, then, may be a redundant concept, as has been suggested previously (Washburn & Astur, 1998; Wright, Santiago, & Sands, 1984).

A more plausible explanation of the forgetting observed, is that participants’ memories for items are reasonably accurate until they are presented with further items that must also be held in memory. Indeed, this hypothesis is supported by the data presented in Chapter 2, in which recognition of items tested without any intervening items was associated with high *d'* values that fell sharply following one intervening item. Material appears to retroactively interfere with existing memory representations, so the more material intervening between study and test, the poorer recognition will be. In addition to a clear retroactive interference effect, proactive interference (PI) appears to have a limited effect in this particular paradigm. Serial position data from Experiment 1 clearly revealed that false alarm rates for novel stimuli rose continually during early periods of the experiment, although they asymptoted beyond this point. The obvious explanation for this primacy effect is that, as the total number of items that had been seen increased, PI from these items also increased, causing new items to be more difficult to distinguish from old.

Consistent with previous studies (Doty & Savakis, 1997; Korsnes, Magnussen, & Reinvang, 1996) recognition of both visual and verbalisable study materials appeared to be subject to the same lag effects. Nonetheless, correct recognition of trigrams was associated with significantly faster reaction times, perhaps a reflection of the relative speed at which simple verbal labels may be accessed, compared with visual feature identification. Also, the discrimination of trigrams was poorer at all levels, and dissociation of hit and false alarm rates revealed that this was the result of much higher false alarm rates associated with these stimuli. This is likely to have been the result of the trigrams' much greater interstimulus similarity resulting in greater confusion between items. The greater false alarm rates observed in the current experiment than those for the same set in Experiment 1 suggests that something about the manipulation of interstimulus interval makes trigrams harder to discriminate. One possibility is that by decreasing the time interval between trials less time was available for verbal rehearsal of the trigram 'name', making labelling a less efficient strategy. However, as the *d'* scores were unaffected by interstimulus interval, this seems unlikely. Perhaps then the very fact that the rate of presentation changed several (4) times during the experiment meant that it took longer for false alarm rates to settle at an approximately steady state.

Whilst the data do not show any obvious evidence of decay, this may be because the decay was 'functional' (Altmann & Gray, 2002). As there was a constant rate of presentation within blocks it is possible that the decay rate could have adapted in the manner proposed by functional decay theory, such that it remained appropriate in order to minimise the build-up of proactive interference. This would have differed according to which rate of presentation the participant was exposed to, and presumably would have changed throughout the experiment. However, the false alarm rates continued to rise for a relatively long period during the course of the experiment, indicative of a sustained period of proactive interference build-up. Also, one would expect there to be a continuous reduction in these effects during the course of an experimental block, followed by an abrupt increase at the start of a new block, as the decay rate would have to readapt to each new rate of presentation. In fact, the serial position data showed no evidence of this, with false alarm rates actually increasing overall during some blocks.

Little reference to what is known about the neural mechanisms thought to be involved in recognition has previously been made in the cognitive psychology literature. However, Doty and Savakis' (1997) comparative study of memory for words and pictures suggests that stimulus-specific adaptation (SSA) is a common mechanism underlying recognition of both types of stimuli. SSA is a property of neurones in the inferotemporal cortex (IT), a region of the brain central to visual object recognition. Certain cells of this area are known to respond to repetition of specific stimuli or object features with a significantly decreased firing rate (E. K. Miller, Li, & Desimone, 1991). Various subpopulations of IT neurones are thought to exist, with different types of repetition-sensitive responses. So-called 'novelty neurones' (Xiang & Brown, 1997) respond strongly to the first presentation of a novel object. There are also 'familiarity neurones' that respond significantly less to familiar than unfamiliar objects, and 'recency neurones' that respond differently according to whether an object is novel, unfamiliar, or familiar (Fahy, Riches, & Brown, 1993; Zhu, Brown, McCabe, & Aggleton, 1995). These repetition-sensitive responses, therefore, are capable of detecting relative recency in repetition of a stimulus or feature, and also the frequency of repetitions. Indeed, decreased blood flow to anterior IT was detected in a human PET study when participants made judgements based on feelings of familiarity rather than explicit recollection (Vandenberghe, Dupont, Bormans, Mortelmans, & Orban, 1995).

The altered response on the second presentation of a stimulus has been found at study-test intervals greater than 24hr and to outlast hundreds of intervening presentations for some cells (Fahy, Riches, & Brown, 1993; Xiang & Brown, 1997; , 1998), a span sufficient for memory at the longest lags tested in recognition experiments. Fahy et al. (1993) reported that 69% of neurones found to exhibit repetition sensitivity had a memory span of ≥5 intervening presentations, dropping to 60% for a span ≥20. The decreasing number of cells responsive to repetition at greater spans might explain why forgetting occurs. One might expect that, with decreasing numbers of neurones making the repetition-sensitive response, the signal that a stimulus is ‘old’ would become weaker and less discernable. It is not clear from the results of Fahy et al. whether time or intervening presentations is at the root of the drop-off in cell response rate, although one might expect that the system’s accuracy is based on the quantity of items to be remembered in a given period.

In conclusion, the results suggest that the effect of time on the accuracy of recognition in continuous recognition is minimal. The most important factor in determining whether information is retained or not is the number of intervening items between study and test, as seen in the effect of lag. There is also a build-up of proactive interference, reflected in an increasing false alarm rate, which reaches a steady state before the end of the experiment. This effect appears to be different for different stimuli, and is proposed to be dependent on the global similarity between stimuli within a set. Further experiments manipulating the global similarity within the constraints of a stimulus category are needed, to determine whether this is the case.

3.2 Experiment 4: The effect of similarity within stimulus sets on continuous recognition

3.2.1 Introduction

Experiment 3 found compelling evidence that loss of information regarding items in continuous recognition is largely determined by interference from other items. Increasing the number of items intervening between study and test had a robust effect of decreasing retention, independent of time delay, suggesting that retroactive interference from these items was responsible for the declining recognition accuracy. In addition, the increases in false alarm rates associated with serial position that occurred throughout Experiments 1, 2, and 3, have revealed the presence of proactive interference effects from items already seen. What factors determine how much interference is caused by the intervening items? Why do the items intervening between study and test of trigram stimuli have a greater interfering effect than a similar number of fractal items, for example?

An important factor assumed to be involved in determining the level of retroactive interference by items intervening between study and test, is their similarity to the studied item. Short-term forgetting is increased by both semantic (Dale & Gregory, 1966) and phonological (Wickelgren, 1965) similarity between presented items and distracters. Proactive interference, from accumulated items that are no longer relevant, is also affected by similarity relations between stimuli in humans (Wickens, Born, & Allen, 1963), as well as in nonhuman primates (Reynolds & Medin, 1981). Reynolds and Medin found that a multiplicative model of overall similarity (combining dimensions of colour, form, and position of stimuli, regardless of their relevance to the memory task) gave a good account of between-trials proactive interference.

Traditionally, the study of interference effects has been carried out using word, letter, and number stimuli. Lists of phonologically similar items, be they words, letters, or pictures with similar names, are recalled with less accuracy than lists of dissimilar items (e.g. A. D. Baddeley, 1966; Schiano & Watkins, 1981), and this effect is suppressed by concurrent, irrelevant articulation (e.g. counting backwards from 100 in steps of 3) (D. J. Murray, 1968; Schiano & Watkins, 1981). This phonological similarity effect has been demonstrated with the names of common objects at both short-term memory (STM) speeds (1 per second) and rapid serial visual presentation (RSVP) speeds (8 per second) (Coltheart, 1999). However, presentation of the pictures themselves was only affected by phonological similarity at the slower speed. In the continuous recognition literature Raser (1972) reported that lures that were phonologically and/or orthographically similar to studied words were associated with higher false alarm rates than dissimilar stimuli, and that this effect occurred as a function of lag.

Estes and Maddox (1995b) claimed that interitem similarity was the factor determining the increase in both false alarm and hit rates over trials in their continuous recognition study using word, letter, and number stimuli. The authors argued that there was a systematic increase in interitem similarity from their words stimulus set to the letter set to digits, and that this resulted in higher global similarities between items. They also noted that higher similarity would result in a greater retention loss between study and test. One of the problems of similarity research, touched upon by Estes and Maddox, is the definition of similarity itself. The authors defined similarity functionally as a measure of “confusability” between stimulus representations. One would certainly expect greater interitem similarity in sets composed of random strings of digits than in those composed of random strings of letters (there are only 10 possible digits, as opposed to 26 letters in the English alphabet). Whether words were recognised with greater accuracy because they could be chunked into discrete units that were less confusable than individual letters, or whether it was simply related to articulation time, was not explored. There is a further problem with this similarity-based definition, in that words are more familiar stimuli than nonwords, and the effects of familiarity cannot be separated from any putative effects of similarity.

Whilst these results gave some indication of the effect of similarity within continuous recognition, a more systematic exploration of similarity is needed. For example, rather than contrasting stimuli drawn from different (albeit closely related) categories, it would make more sense to manipulate the similarity of items from within the same category. This would avoid the ambiguity of results in which the effect of similarity is confounded with the category of stimulus. Verbal stimuli are also not ideally suited to the study of similarity, because similarity relations at orthographic, phonological, and semantic levels must be accounted for in the construction of suitable lists. The advantage of exploring similarity effects with parametrically defined visual stimuli, is that similarity relations are easy to assess and quantify.

Memory for complex visual stimuli has been investigated in a recent fMRI study in which participants were required to study particular works by famous painters with distinctive styles (Yago & Ishai, 2006). A subsequent recognition test, introducing novel paintings, found that novel paintings that were different to the prototypes studied were classified more rapidly than novel paintings that were similar to the prototypes. The behavioural effects of similarity were accompanied by effects on the activation of brain regions involved in memory, including the hippocampus, where responses to new exemplars were reduced with decreased similarity to the studied prototypes. Whilst this study provides compelling evidence that similarity has a key role in visual recognition, for a more detailed understanding of the relationship of similarity to recognition, quantifiable changes to the stimuli need to be examined. This is clearly not possible where the stimuli are not defined through the use of readily manipulable parameters.

The similarity of stimuli is intimately related to the concepts of typicality and distinctiveness. Distinctiveness can be thought of as the extent to which an item stands out within a set, whereas typical items are those which do not. In a set with high global similarity, there will be a preponderance of typical items and few distinctive ones. The distinctiveness of dissimilar stimuli is associated with a greater accuracy of recognition (e.g. Hunt, 1995; Lockhead, 1970).

In terms of current models, Nosofsky's generalised context model (GCM) of recognition (Nosofsky, 1988, 1991) assumes that individual exemplars of studied items are represented as points in multidimensional perceptual space, and that the similarity of exemplars is defined by their distance in that space. Items more distant from one another are assumed to be less similar. By summing the similarities between test items and the exemplars held in memory, information that forms the basis for recognition judgements is obtained. Whilst this model has proved to be a good predictor of how the similarity of stimuli affects false alarm rates, it also predicts that hit rates for typical stimuli should be higher than those for distinctive items. There is evidence from face-recognition literature that the opposite is often the case (Bartlett, Hurry, & Thorley, 1984; Light, Kayra-Stewart, & Hollander, 1979; Valentine & Ferrara, 1991; Vokey & Read, 1992). Indeed, a study by Busey and Tunnicliff (1999) found that faces determined as distinctive according to their co-ordinates in a multidimensional perceptual space were associated with greater hit rates than typical faces. This was perhaps a result of the presence of discrete, individuating features in these particular faces (e.g. beards, scars).

Nosofksy and Zaki (2003) investigated the effectiveness of two different definitions of distinctiveness in predicting the results of experiments involving recognition of colour patches. When the parameters of stimuli were mapped in continuous-dimension similarity space, and distinctive items were defined as those lying in isolated regions of this space, the effects of distinctiveness were weak. However, when distinctive items were defined as those containing certain discrete, individuating features, large old item distinctiveness effects were observed. By specifically including discrete, individuating properties in stimuli, a strong impact on hit rates was observed. These results suggest that when superior hit rates are associated with distinctive stimuli, it is because these stimuli contain individuating features, rather than because they lack perceptual similarity to other stimuli in a set.

As well as the similarity between studied and target items, the similarity of items intervening between study and test affects recognition. This has been demonstrated systematically by Kahana and Sekuler (2002) using sinusoidal grating stimuli that varied in their spatial frequency. Recognition of items from short lists was successfully modelled with a noisy exemplar model (NEMO) of recognition, that assumes that memory is affected by the similarities among list items, as well as similarities between list items and the test probe. Earlier array models (Estes, 1986; Hintzman, 1986, 1988) and the exemplar-retrieval model GCM (Nosofsky, 1986, 1992) theorised that recognition decisions were made on the basis of summed pairwise similarity between the probe and each stimulus representation. Kahana and Sekuler's model takes into account the similarity relationships between items in the list as well.

Whilst the similarity relationships between items in list-based memory have been extensively explored (Kahana & Sekuler, 2002; Nosofsky & Zaki, 2003; Zhou, Kahana, & Sekuler, 2004), the continuous recognition literature lacks a systematic examination of interitem similarity effects. Whilst Estes and Maddox (1995b) alluded to similarity in the explanation for the cross-category differences they observed, no parametrically-based examination of the effects of similarity within a category under conditions of continuous recognition has been reported. This experiment systematically varied the global similarity of stimuli within experimental blocks, during continuous recognition.

3.2.2 Methods

3.2.2.1 Design

The experiment had a three-way within subjects design. The independent variables were the type of stimulus tested, the global similarity, and the number of intervening trials (lag). Stimulus type had 2 levels: fractals and trigrams. Similarity had 3 levels: similar, medium, and dissimilar, reflecting the level of constraints on stimulus generation. Lag had 3 levels: 1, 4 and 9 intervening items. The dependent variables were the participants' *d'* scores, and reaction times for correct recognition.

3.2.2.2 Participants

A total of 19 participants were tested (15 female and 4 male). All were students at the University of Nottingham with a mean age of 20.5 ± 0.2 years. All participants had normal or corrected-to-normal vision.

In order to exclude data from participants who 'gave up' on the tasks, during one or other of the relatively long 35 min sessions, criteria were set for inclusion of the data in analyses. Each participant's data were entered into a one-tailed chi square, to determine that the number of correct responses were significantly above that expected by chance.

3.2.2.3 Stimuli

Two different categories of stimulus were employed: fractals and trigrams. These stimuli were generated according to the rules described in the Methods section of Experiment 1, although constraints on variation were manipulated, in order to produce three sets of each stimulus type: similar, medium, and dissimilar. In addition, the routine producing 'irregular' fractals was removed.

3.2.2.3.1 Fractals

The minimum and maximum limits for the following variables were manipulated: number of superpositions, number of sides of original polygon, depth of recursions, size of recursion. These were manipulated in such a manner that the mean value for each variable for each set remained the same, whilst the amount of variation from the mean varied according to ‘similarity’. The limits for each set are shown in Table 3.1. Examples of the stimuli produced are shown in Figure 3.8.

Table 3.1: Limits for variation of variables for fractal generation, for sets of different global similarities

	Number of super-positions	***Number of sides of polygon***	***Recursion depth***	***Size of recursion***
Similar	5	6	6	0.3-0.7
Medium	4-6	4-8	4-8	0.15-0.85
Dissimilar	3-7	3-9	3-9	0-1

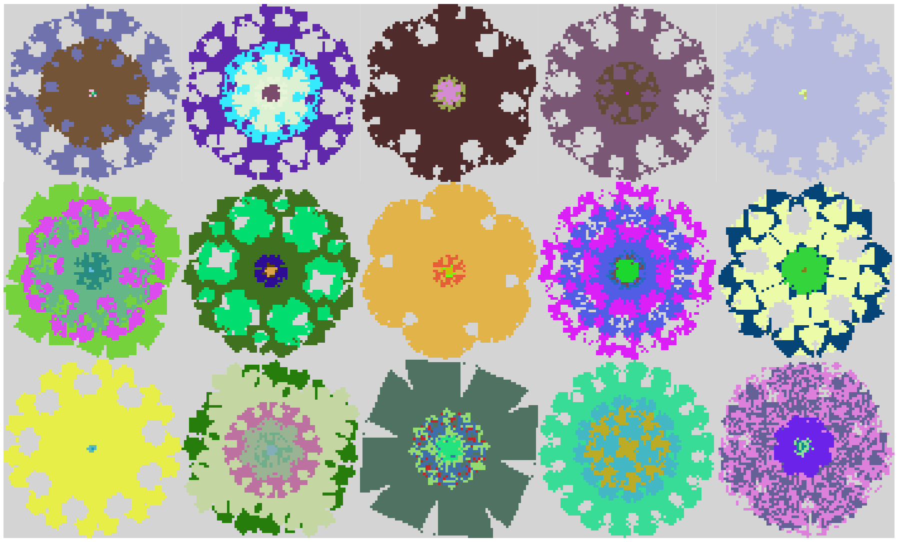

Figure 3.8: Example fractal stimuli. Top row = similar, middle row = medium, bottom row = different.

3.2.2.3.2 Trigrams

Trigrams were of a similar format to those used previously, in that they were digit-letter-digit. Numbers could vary from 1-9 in all conditions. Similarity was regulated by constraints on the letter. In the similar condition, letters could be any of the five letters allowed in previous experiments (K, V, W, Y, or Z), e.g. 4V9, 2W2, 9Z1. In the medium similarity condition it could be any of the first 10 letters of the alphabet, e.g. 3F7, 8A7, 5C8. In the dissimilar condition it could be any letter of the alphabet, e.g. 8J6, 3Y5, 1P7.

3.2.2.4 Presentation

All stimuli were presented using the same equipment described for Experiment 3. Stimuli were presented in the centre of the screen and subtended approximately 2° x 2° of visual angle.

3.2.2.5 Procedure

Each experiment took place in the form of two sessions, identical except for the type of stimuli (fractals and trigrams) and the order of similarity blocks. The order in which the participants carried out the fractals session and the trigrams session was counterbalanced across participants.

For each session a pseudorandomly determined frame of 120 trials was generated, providing 16 study and 16 test trials at each of the 3 lags. Twelve study-test pairs of unscored filler trials made up the spaces in the frame.

This frame, or block, was repeated three times, with novel stimuli each time. For each block the global stimulus similarity was different: either similar, medium, or dissimilar. The order in which these occurred was counterbalanced across participants. The experiment began with a buffer of 40 (medium similarity) unscored filler trial pairs to prevent the occurrence of primacy effects. This yielded a total of 400 trials. The order of stimuli within each block was always randomly determined. There were no breaks between blocks.

For each trial stimuli appeared on the screen for 2 sec. During this period participants were instructed to respond either ‘old’ or ‘new’ with the left or right mouse button respectively, ‘old’ to previously seen and ‘new’ to previously unseen items. The instructions stressed that, in order to be classified as old, items had to be identical to a previous item. In order that participants did not forget which button corresponded to which response during the long experiment, a label of which was which was placed in clear view of the participant. After 0.5 sec the stimulus was replaced with a blank screen. One second later feedback was given for incorrect responses, or where no response was given. Feedback consisted of a low-pitched beep. No feedback was given if the response was correct. The screen remained blank for a further 1 sec, before the start of the next trial.

3.2.2.6 Scoring

Participants’ responses were only detected during the 2 sec presentation period of each stimulus. Once a response had been made it was final, and no opportunity for the correction of responses was allowed. Reaction time was measured as the latency from the start of the trial until the detection of the response.

3.2.3 Results

From the results a one-tailed chi square test was conducted on each participant's responses, to ensure that the number of 'same' responses to old stimuli was significantly above that expected by chance. This was the case for all of the participants, and all data was entered into subsequent analyses.

3.2.3.1 D-prime scores

The *d'* values for fractals and trigrams are presented as a function of lag in Figure 3.9 and Figure 3.10. Mauchly's test of sphericity was significant for the interactions of stimulus type x lag, similarity x lag, and stimulus type x similarity x lag. All reported results for these interactions are Greenhouse-Geisser epsilon corrected. A 2 (stimulus type) x 3 (similarity) x 3 (lag) repeated-measures ANOVA was performed on *d'* data and revealed a significant main effect of stimulus type (F(1,18)=11.3, MSe=0.853, p<0.01). Fractal *d'* scores were higher than those for trigrams. There was also a significant main effect of similarity (F(2,36)=12.5, MSe=0.588, p<0.001), and Tukey's post-hoc tests revealed significantly higher scores for dissimilar blocks than both similar (p<0.001) and medium (p<0.01) blocks. Scores for high and medium similarity blocks were not significantly different. In addition, a significant main effect of lag was reported (F(2,36)=41.4, MSe=0.331, p<0.001), with significant differences between all lags. There was also a significant interaction between stimulus type and lag (F(1.48,26.7)=4.74, MSe=0.354, p<0.05). Post-hoc tests revealed significant differences between fractals and trigrams only at lags greater than 1. Also, whilst scores for trigrams at lags 1 and 4 were significantly different, they were not for fractals. Increasing the similarity between stimuli from medium to dissimilar made performance on the task less accurate.

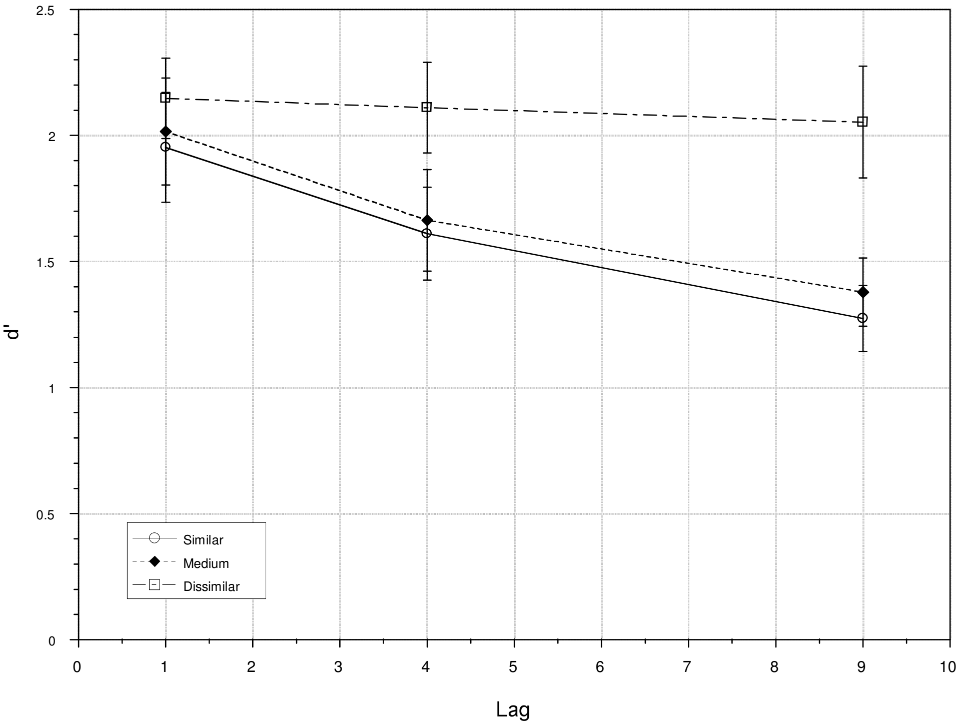

Figure 3.9: d’ values for fractal recognition, as a function of lag. Data = Mean ± SEM.

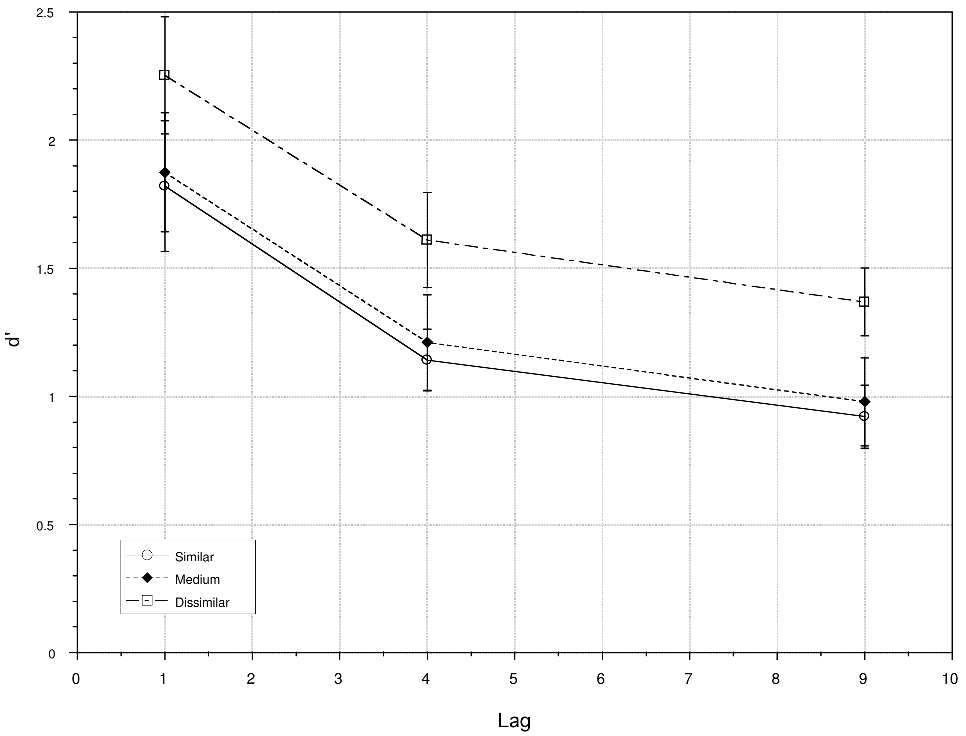

Figure 3.10: d' values for trigram recognition, as a function of lag. Data = Mean ± SEM.

3.2.3.2 Reaction times

The reaction time data for hits (correct recognition) is shown in Figure 3.11 (fractals) and Figure 3.12 (trigrams). Mauchly's test of sphericity was not significant for any of the measures reported. A similar 2 x 3 x 3 repeated measures ANOVA to that described for *d'* results was carried out on these data, and no significant effect of either stimulus type or similarity was found. There was a significant main effect of lag on reaction times (F(2,36)=18.1, MSe=8370, $p<0.001$). No significant interactions were observed.

In contrast to the effects on *d'* scores, similarity between stimuli appeared to have little or no effect on reaction times for hits.

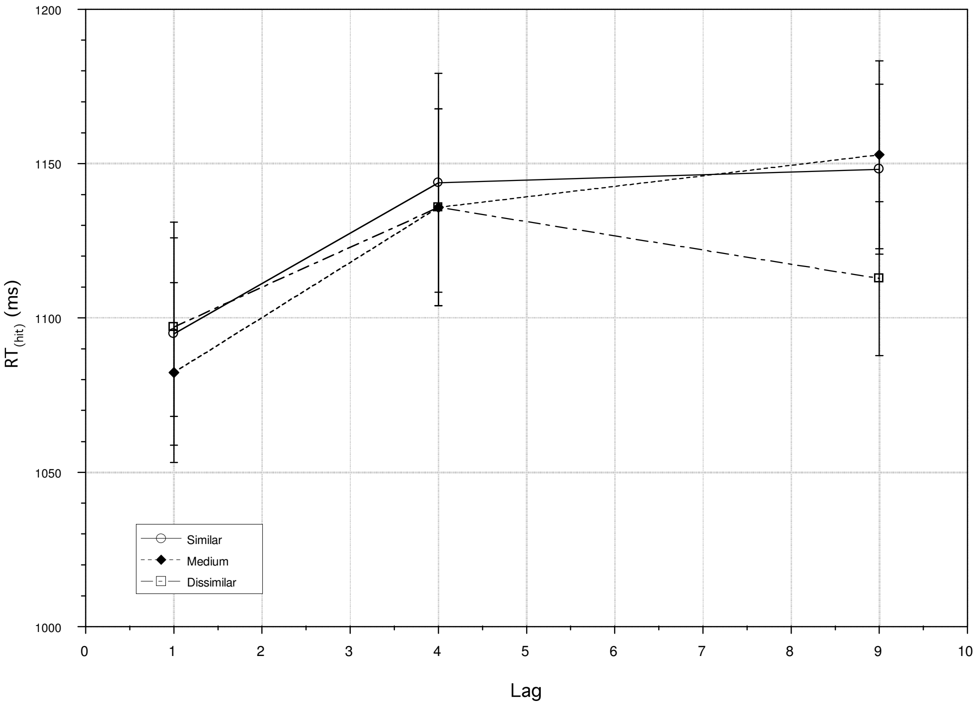

Figure 3.11: Reaction time values for correct recognition of fractals (ms). Data = Mean ± SEM.

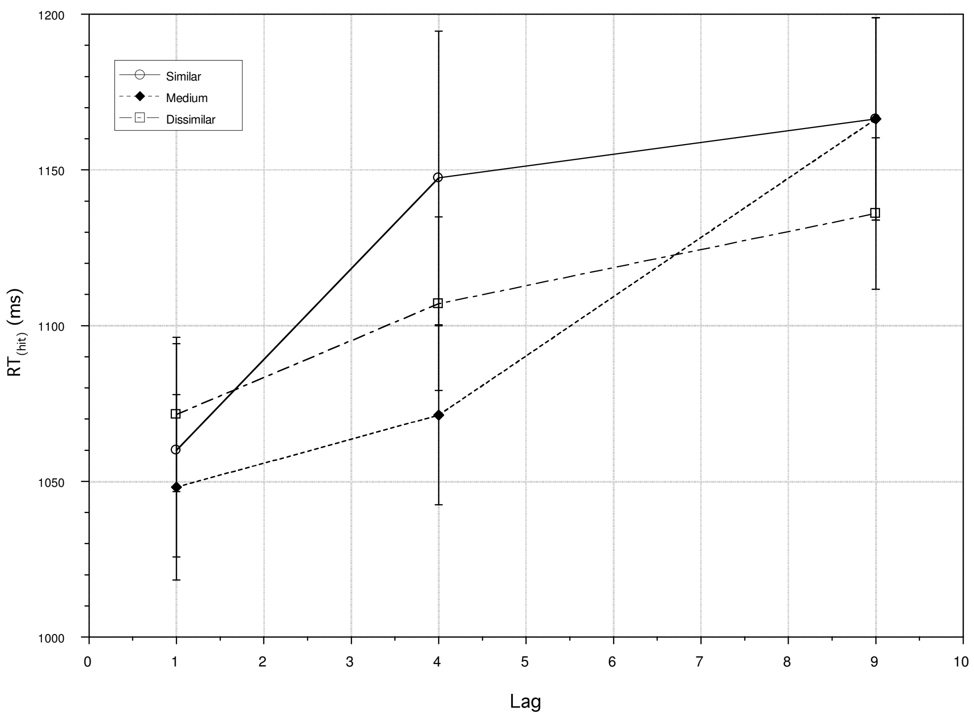

Figure 3.12: Reaction time values for correct recognition of trigrams (ms). Data = Mean ± SEM.

3.2.3.3 Hit and false alarm rates

In a similar manner to that described in Experiment 3, trials were analysed in epochs of 10 subsequent trials, and hit (p(hit)) and false alarm (p(false)) rates were calculated for each epoch. Data were then reorganised according to the level of similarity of the block, such that the changes in p(hit) and p(false) during the course of blocks for each similarity level could be examined.

Data for p(hit) can be seen in Figure 3.13 (fractals) and Figure 3.14 (trigrams). The data were entered into a 2 (stimulus type) x 3 (level of similarity) x 12 (epoch) repeated measures ANOVA. Mauchley's test for sphericity was not significant for any of the variables studied. There were no significant effects of any of the variables tested, indicating that p(hit) was similar for both fractals and trigrams, was unaffected by the level of similarity, and remained constant with regards to serial position within blocks.

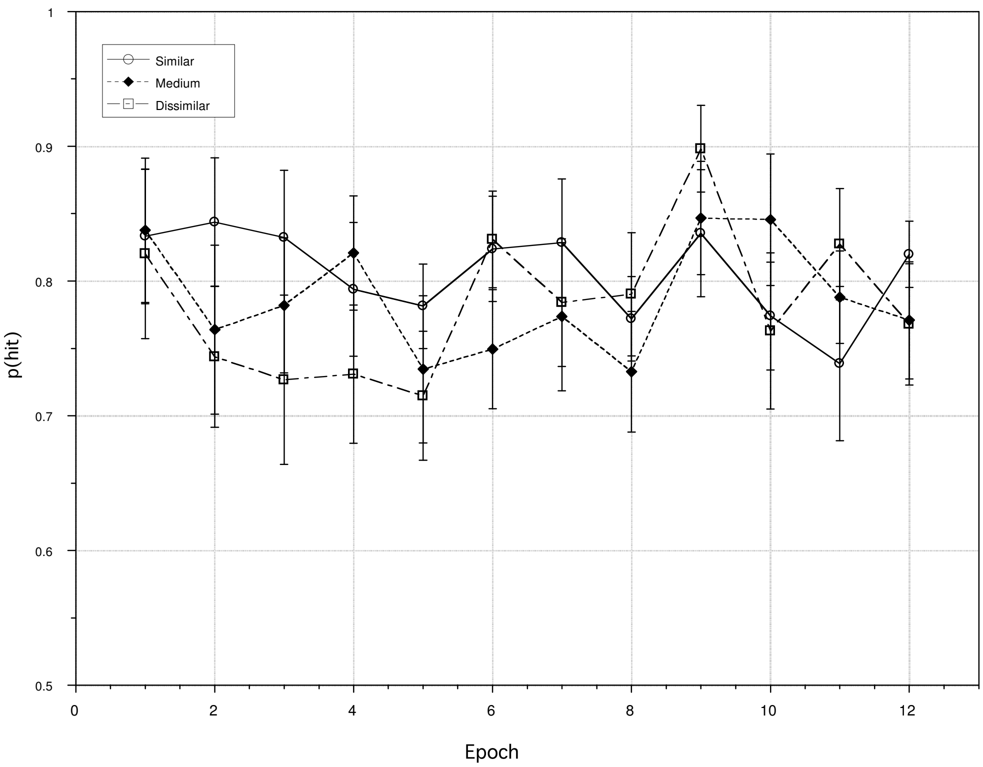

Figure 3.13: The effect of interitem similarity on hit rates for trigram recognition. Each epoch represents 10 consecutive trials. Data = mean ± SEM.

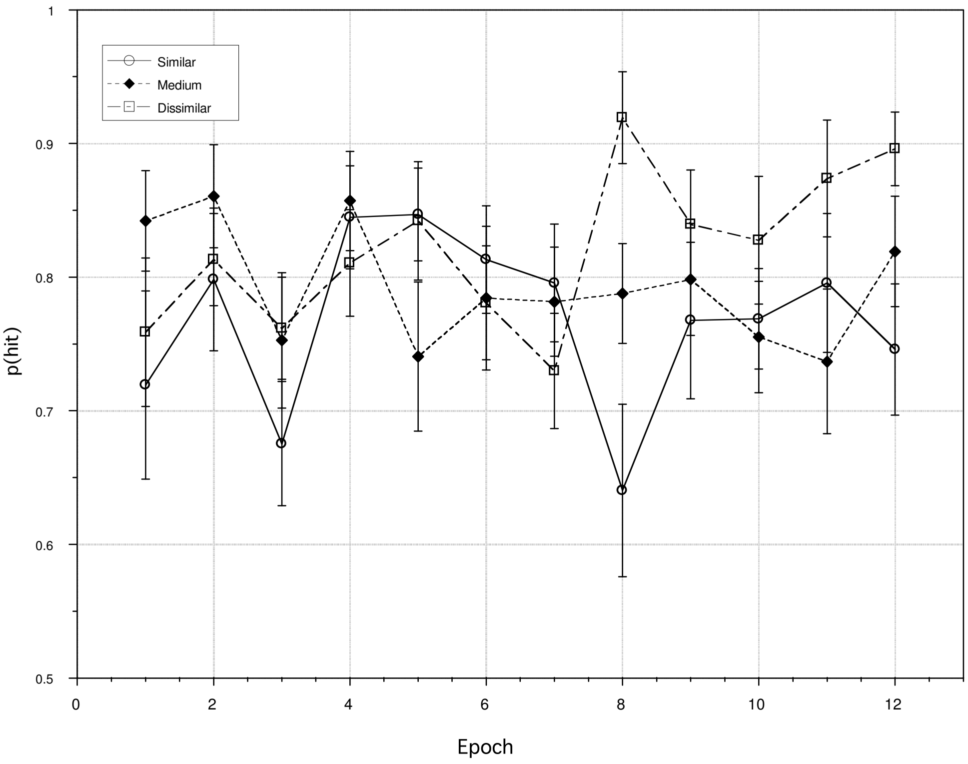

Figure 3.14: The effect of interitem similarity on hit rates for trigram recognition. Each epoch represents 10 consecutive trials. Data = mean ± SEM.

Data for p(false) is shown in Figure 3.15 (fractals) and Figure 3.16 (trigrams). The data were entered into a 2 (stimulus type) x 3 (level of similarity) x 12 (epoch) repeated measures ANOVA. Mauchley's test for sphericity was not significant for any of the variables studied. There was a significant main effect of stimulus type (F(1,17)=6.88; MSe=0.487, p<0.05), with p(false) for trigram recognition being significantly greater than p(false) for fractals. There was also a significant main effect of similarity (F(2,34)=23.6, MSe=0.820, p<0.001). Blocks of dissimilar stimuli were associated with lower p(false) than were both medium and similar stimuli (Tukeys, p<0.001). Epoch had a significant main effect (F(11,187)=4.52, MSe=0.0344, p<0.001). Post-hoc tests revealed significantly lower p(false) at epochs 1, 2, 3, and 4 than at lag 8 (p<0.05), and at epoch 1 than epoch 10 (p<0.05). No interactions were significant.

These results indicate that there was significant build-up of proactive interference, as observed in increased p(false), during experimental blocks of the same similarity level, and possibly the experiment as a whole. High and medium similarity blocks were associated with greater p(false) than dissimilar stimuli. The lack of any changes in p(hit) suggests that the differences in *d'* reported for different similarity conditions are largely the result of variation in p(false).

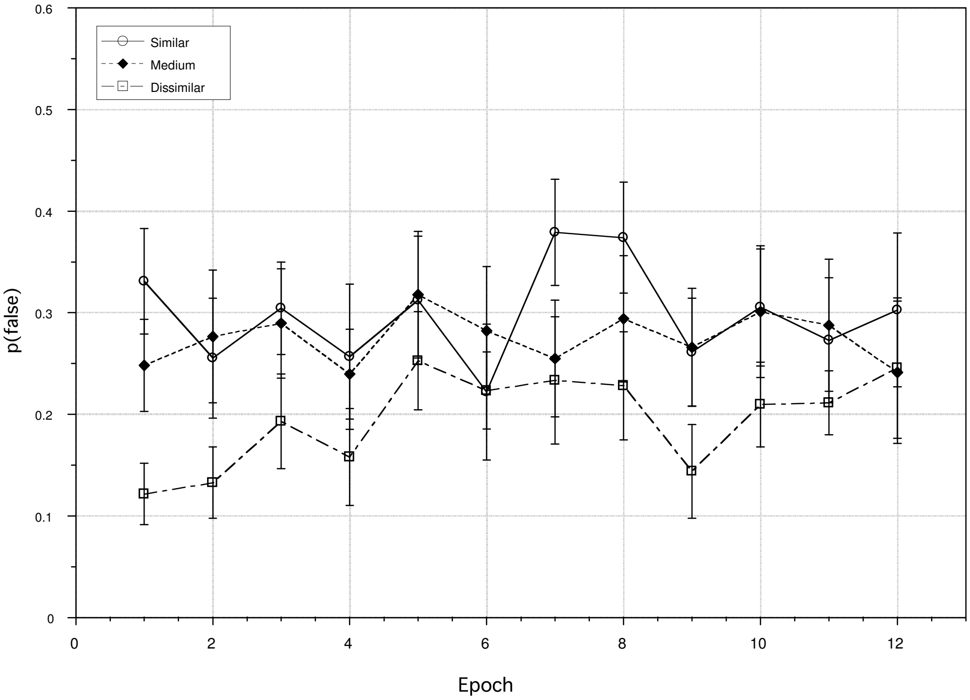

Figure 3.15: The effect of interitem similarity on false alarm rates for fractal recognition. Each epoch represents 10 consecutive trials. Data = mean ± SEM.

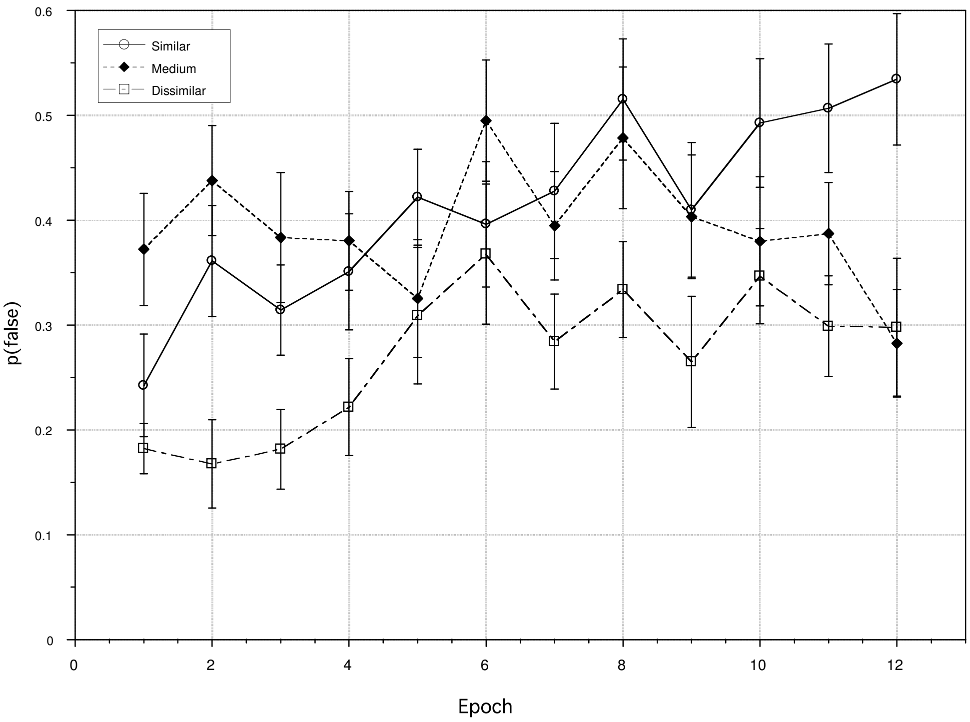

Figure 3.16: The effect of interitem similarity on false alarm rates for trigram recognition. Each epoch represents 10 consecutive trials. Data = mean ± SEM.

3.2.4 Discussion

The manipulation of interstimulus similarity within a continuous recognition experiment resulted in significant effects on measures of recognition accuracy. By constraining stimuli so that they were more similar to one another, recognition was poorer, and new stimuli were more frequently mistaken for old (increased false alarm rates), although hit rates appeared to remain constant. This result suggested that the effect of similarity only affected the correct discrimination of new trials from old, with performance on old trials unaffected.

The results provide further evidence that retroactive interference between study items and items occurring between study and test, is a major factor in determining the effect of lag. If a simple decay of the studied stimulus' memory trace was the sole cause of lag effects on recognition, manipulating the properties of intervening stimuli would not alter the lag function. This was clearly not the case.

Constraints on similarity had comparable effects regardless of stimulus type, indicating that similarity has an equivalent effect on stimuli that are amenable to verbal labeling and those that are not. Further research is needed to determine whether this trend can be generalised from fractals and trigrams to other types of stimuli.

A point of interest is the failure to observe any differences between recognition of high and medium similarity blocks. No differences were observed between the two similarity levels for either fractal or trigram recognition, whereas the transition from medium similarity to low similarity (dissimilar condition) was associated with much better recognition. There are a number of factors that may have contributed to this result.

Firstly, the parametric similarity that was manipulated in the experiments did not necessarily correlate with the subjective experience of similarity (e.g. as measured by similarity ratings), which may involve factors more complex than simple geometric coordinates (e.g. saliency of features). Also, it is important to note that the changes to constraints for both stimulus sets resulted in non-linear effects on the number of possible output stimuli. For fractals, changing the possible number of superpositions of fractal stimuli from 5 to 4-6 to 3-7 (for similar to medium to dissimilar stimuli) changes the number of possible values of number of superpositions from 1 to 3 to 5 (linear). However, when coupled with an increase of possible sides of polygons from 6 to 4-8 to 3-9, and an increase of possible recursion depths of 6 to 4-8 to 3-9, the number of possible combinations of these three variables increases from 1 to 75 to 245. There were over 3 times as many possible combinations of these variables at the dissimilar level (245) than there were at the medium level (75) of similarity, and these figures do not take size of recursion into account. For trigrams, again, the transition from one level of similarity to another was not linear in terms of the number of possible outputted stimuli. By changing the number of possible letters from 5 to 10 to 26, the number of possible trigrams was increased from 405 to 810 to 2106. Again, the number of possible stimuli increased exponentially.

Perhaps parametric similarity only affected recognition performance when objects varied dramatically in terms of their parameters. Putative thresholds in the detection of variation, perhaps the boundaries of perceptual categories, might explain why similar and medium similarity fractals and trigrams interfere with one another in a comparable fashion, whereas the exponentially greater variation at dissimilar levels reduces this interference.

When parameters are allowed to vary to a greater extent, it is more likely that distinctive stimuli that stand out from the others will be produced. This distinctiveness might form the basis for discrimination between items in memory. Whether distinctive items are better defined as those lying in isolated regions of parametric space, or those containing discrete, individuating features in continuous recognition (see Nosofsky & Zaki, 2003), is a question for future investigations. This implies that distinctiveness and categorisation may be intimately connected, and that a distinctive item might be considered to be one which falls outside of existing categories, or falls into a separate category from the majority of stimuli. This might explain the finding of Nosofsky and Zaki (2003) that the presence of certain individuating features made items more distinctive and, easier to remember. Those features might form the basis for inclusion in a separate perceptual category to the majority of stimuli that do not contain the feature. Discrimination between items that fall into this smaller group of stimuli might be expected to be easier than discrimination between the larger number of items that do not contain the distinguishing feature.

Surprisingly, the results for hit and false alarm rates revealed that the effect of similarity on overall recognition performance was mediated solely through changes in the false alarm rate. Hit rates did not vary greatly with the degree of similarity, suggesting that participants' ability to correctly recognise old items was unaffected by the manipulation. This may explain why no significant effects on reaction times for hits were observed. False alarm rates rose significantly during each experimental block suggesting that changes to similarity disrupted the achievement of a steady state, although this could simply be because serial position was confounded with position within a block in this analysis. Comparison of these results with the serial position curve obtained for results from Experiment 3 (Figure 3.7) suggests that hit rates did not vary between fractals and trigrams for that experiment either. Again, the difference in discrimination can be ascribed to the false alarm rate.

The results confirm the suggestions of Estes and Maddox (1995b) and Raser (1972) that interitem similarity is an important factor in determining recognition performance on the continuous recognition task. Through the use of verbal stimuli that varied only in one parameter of their construction (trigrams) the role of similarity was examined in a manner that was more systematic than Estes and Maddox’s cross-category comparisons. The results also confirm that Kahana and Sekuler’s (2002) finding that recognition of probes is affected by similarity relations between items in a preceding list, applies to the more complex relationships between items in continuous recognition as well.

Chapter 4 Translation invariance of immediate recognition with respect to retinal position

4.1 Experiment 5: Positional effects in the recognition of face-like stimuli

4.1.1 Introduction

One of the most important functions of object memory is the recognition of familiar objects regardless of transformations of size, viewpoint, location in the visual field, etc. Changes in the distance between an observer and an object change the size of the object on the retina. In a similar manner, the same object can be viewed from different 3D positions in space, or at different retinal locations due to either a change in the position of the object, or a change in the observer's fixation. Whilst changes in these conditions may result in very different patterns of activation on the retina, it is necessary for the visual system to recognise that an object's identity does not change in order to produce a coherent perceptual representation of objects in the observer's environment. How this is achieved is a fundamentally important issue and the subject of much debate. For example, an object moving from one visual field location to another is thought to enable feature analysers at different stages of cortical processing to become associated with one another through the strengthening of mutual connections. How much experience, and of what kind, are important questions that remain to be answered. By investigating the conditions under which translation invariance is achieved, and when it breaks down, may reveal important information about how object identity is represented in memory.

An aspect of transformational invariance, about which there are conflicting reports, is translational invariance of retinal location. Evidence from both human psychophysical studies (Dill & Edelman, 2001; Dill & Fahle, 1997, 1998; Gratton, Corballis, & Jain, 1997; Hornak, Duncan, & Gaffan, 2002; Kahn & Foster, 1981; Nazir & O'Regan, 1990), and electrophysiological studies in monkeys (Chelazzi, Duncan, Miller, & Desimone, 1998; DiCarlo & Maunsell, 2003; Lueschow, Miller, & Desimone, 1994), have revealed that, under certain conditions, translation invariance breaks down, and previously learned objects are not recognised in their new locations. Much of the evidence is discussed in more depth in 0, but an outline of the most important findings is given below.

Neuronal populations in inferior temporal (IT) cortex involved in object perception are thought to learn to produce translation invariant representations of objects (i.e. representations of objects that are not specific to a particular retinal location) by experience of objects changing within the visual field (Wallis & Rolls, 1997). Stimulus-selectivity in IT neurones is often assumed to be constant throughout their receptive fields, and as these receptive fields are large, throughout the majority of the visual field (Desimone, Albright, Gross, & Bruce, 1984). This view is supported by the finding that there was no decrement to the firing of IT neurones when an effective object was shifted up to 4º from fixation (Tovee, Rolls, & Azzopardi, 1994). However, in a study of object recognition in which monkeys were required to discriminate between target images and distracters, changes in the location of achromatic forms of 1.5° from fixation resulted in a mean 60% decrease in the responses of IT neurones responsive to the objects (DiCarlo & Maunsell, 2003). In a delayed-matching-to-sample (DMS) task involving shifts of 5°, Lueschow et al. (1994) found that, whilst the order of neurones' stimulus preferences did not change, 69% of cells recorded preferred a given retinal location. Only 7% of cells showing repetition-sensitive responses were invariant for location. Interestingly, the nature of the result in these experiments appears to be related to the nature of the task. Both Desimone et al. (1984) and Tovee et al. (1994) measured passive responses to stimuli that were not related to any task. In both Lueschow et al. (1994) and DiCarlo and Maunsell's (2003) studies, animals were actively engaged in making responses based on the stimulus identity in either form recognition or DMS tasks. Also, the neuronal response did not always match the behavioural response. DiCarlo and Maunsell found that, whilst neuronal responses were decreased at 1.5°, behavioural measures of recognition were unaffected.

Psychophysical studies demonstrate considerable invariance for recognition on certain tasks, with certain stimuli. Biederman and Cooper (1991) carried out an experiment in which pictures of readily nameable objects (e.g. pianos) were presented to participants twice in two separate blocks for identification by naming. The interval between study and test presentations was approximately 7 min. The first presentation had a priming effect on an object's subsequent naming, measured as both a faster reaction time and a reduced error rate, and this effect was independent of changes in the object's location. In different experiments items could occur in either the left or the right visual hemifields, or the upper and lower visual hemifields, but whether the object was in the same or a different location to the initial presentation, the effect of priming was the same. Likewise, Bricolo and Bülthoff (1993) found no effect of 2.5° shifts on the recognition of wire-like objects from their training positions.

There is, however, plenty of evidence from pattern recognition literature demonstrating the breakdown of translation invariance. For example, Kahn and Foster's (1981) study of participants' ability to discriminate sequentially presented dot patterns found that *d'* scores for recognition were highest at identical positions, and decreased with increasing distance from training position.

Nazir and O'Regan (1990) trained participants to discriminate dot patterns at a fixed location to the left or right of fixation, from two non-target distracters.

Subsequent recognition rates dropped when stimuli were presented at novel positions, regardless of whether the retinal eccentricity was 2.4°, 0.86°, or 0.49°. Most recently, the experiments of Dill and Fahle (1997; , 1998) confirmed the findings of Nazir and O’Regan, that improvement in discrimination performance at one retinal location does not necessarily transfer to new locations. In addition, recognition of dot clouds and checkerboard stimuli were examined at a range of eccentricities up to 2° from an initial position, and recognition performance decreased with increasing distance from study position (Dill & Fahle, 1998). The positional specificity (lack of translation invariance) observed in these studies seems to be related to the use of novel patterns, and the difficulty of discriminations. However, as recognition accuracy following transfer from the learned to the unlearned position was significantly above chance in this study, positional specificity was incomplete.

One limitation of several of these studies (Dill & Fahle, 1997; Kahn & Foster, 1981; Nazir & O'Regan, 1990) is that it is impossible to discern whether their results reflect a global positional specificity *per se* or whether this phenomenon is linked to the presentation of items at particular locations in the visual field (i.e. foveal vs. parafoveal/peripheral presentation, within- vs. between-hemifield shifts). Some studies have suggested that recognition accuracy is related to whether the test presentation occurs within the same left/right hemifield as the studied item, regardless of the size of the shift (Gratton, Corballis, & Jain, 1997; Hornak, Duncan, & Gaffan, 2002). Gratton and colleagues found significantly better recognition for symmetrical line patterns shifted vertically within the same visual hemifield, compared with recognition of stimuli shifted horizontally across the midline. Hornak et al. (2002) used a design in which stimuli could appear at any of the four corners of an imaginary square centred on the fixation point. Each point was equidistant from fixation, allowing the comparison of the effects of vertical and horizontal shifts at a constant retinal eccentricity. During learning, stimuli (pictures of nameable objects, containing different exemplars of the same categories) were presented in pairs in diametrically opposite positions. In a later test phase, pairs contained either two novel pictures or one learned and one novel picture, and participants were required to discriminate between the two. The results showed a significant decrement to recognition caused by horizontal shifts, whereas vertical shifts were comparable to the no shift controls. Both studies were cited by their authors as evidence for the hemispheric organisation of visual memory.

Table 4.1: Comparison of previous human studies of positional effects

Study	Task	Stimuli	Shift(s)	Effect of translation?
Kahn & Foster (1981)	Same/different discrimination	Dot patterns	0.5°, 1.0° Fixation, left and right	Yes
Nazir & O'Regan (1990)	Training followed by recognition	Dot patterns, Columns of grey squares	Learned at +/- 2.4°/0.86°/0.49°, Tested at left, right, up and down from fixation	Yes
Biederman & Cooper (1991)	Picture naming	Line drawings of common objects	2.4° to left and right of fixation	No
Bricolo & Bülthoff (1993)	Training followed by recognition	Wire-like objects	2.5° left, right, up, and down from fixation	No
Dill & Fahle (1997)	Training followed by recognition	Checker-boards	2.4° to left and right of fixation	Yes
Gratton et al. (1997)	Training followed by recogntion	Line patterns	Square around fixation: 8.4° vertical/horizontal; 6° to left and right of fixation	Yes; different hemifield > same hemifield
Dill & Fahle (1998)	Same/different discrimination	Dot patterns	0°/0.5°/1°/1.5°/2 left, right, above and below study	Yes; for *same* trials

			at 1°	
Dill & Edelman (2001)	Same/different discrimination	Animal-like shapes	Square around fixation: 5.5° horizontal/vertical, 8° diagonal	No; except when configuration scrambled
Hornak et al. (2002)	Recognition of objects in old-new, and new-new pairs	Pictures of common objects	Square around fixation: 6.7° horizontal/vertical	Yes; different hemifield > same hemifield

A comparison of the important psychophysical studies of stimulus translation (Table 4.1) reveals that an important factor in determining whether or not translation invariance is found seems to be the complexity of the stimuli used. When stimuli were objects with which participants had some prior familiarity (e.g. pictures of common objects), translation invariance was generally found, whereas studies using novel patterns with a complex configuration (e.g. dot clouds, checkerboards) found evidence of positional effects. This suggests that the cause of these effects may lie in the abstract and/or highly similar nature of the stimuli employed. Dot clouds and checkerboards are unlikely to have been seen by participants before engaging in the experiment, and participants are unlikely to have experience with similar objects. The patterns are also likely to be more similar to one another than items from a set of common objects, or geometric shapes. This suggests that it is necessary to have some experience with a category of objects before translation invariant representations of those objects can be formed.

A study that sheds light on which properties of objects are important for translation invariance is that of Dill and Edelman (2001). In same/different experiments there was no effect of translation for their ‘animal-like’ stimuli, regardless of their interstimulus similarity. The animal-like stimuli were composed of a standard number of computer-generated features such as legs and heads, and different ‘animals’ were defined both by the identity of their component features and the features' positional configurations. Translation invariance of recognition was found for these stimuli, and was maintained when stimuli were 'scrambled' by randomising the identity of the component features whilst maintaining the global configuration of the animal. However, when the stimuli were made to differ in the locations but not the shapes of corresponding parts an effect of position emerged. This suggests that the representation of the identity of an object's features may be position invariant, although the representation of their configuration may be specific to the position in which it was learned.

In an attempt to ascertain whether the incomplete translation invariance encountered in the Dill and Edelman (2001) study can be generalised to other complex stimuli, the present study examined the effect of positional shifts on same/different discrimination of face-like stimuli. The face-like stimuli (previously described in Chapter 2) were like the animal-like stimuli in that they were defined both by the identity and configuration of their features. In a departure from previous approaches to translation invariance, precise and consistent changes to certain parameters of the stimuli occurred between each new stimulus and the next, in an attempt to insure that interitem similarity was constant. The effects of changes of feature size and location were examined systematically.

The purported hemispheric organisation of memories was also studied, by comparing the effects of horizontal and vertical shifts in a manner similar to Hornak et al. (2002). The previously reported poorer performance for trials in which the stimulus is shifted between visual hemifields has only been observed for abstract line-patterns (Gratton, Corballis, & Jain, 1997) and pictures for which naming is possible (Hornak, Duncan, & Gaffan, 2002). This experiment intended to determine the extent to which representations of complex visual stimuli that are not readily amenable to naming are translation invariant.

4.1.2 Methods

4.1.2.1 Design

The experiment had a two-way mixed measures design. The between-subjects variable was the type of change between different stimuli, which had two levels: size of features only, or size and location of features. The within-subjects variable was the nature of the shift (or lack thereof) between stimulus location on initial viewing (study presentation) and subsequent viewing (test presentation). This factor had five levels: horizontal shift (a shift of 5.6° horizontally, between the right and left visual hemifields), vertical shift (a shift of 5.6° vertically, between the lower and upper visual fields), centre-periphery shift (a shift of 4.0°, from fixation to one of the four visual quadrants), and conditions in which the stimuli remained either at fixation (centre (no shift)) or in the same peripheral location (periphery (no shift)) (see Figure 4.1). The shift sizes were chosen to be consistent with those used by Dill and Edelman (2001).

The dependent variables were the participants' *d'* score, reaction times for correct recognition, and hit and false alarm rates.

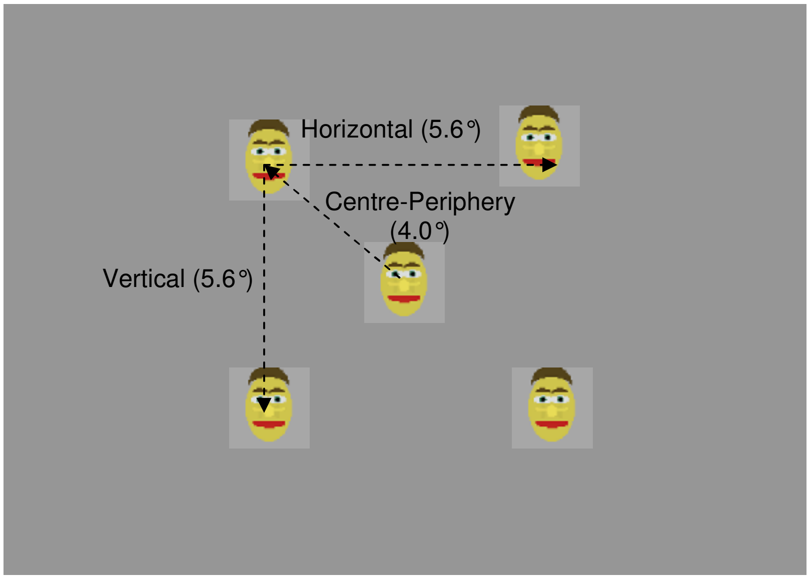

Figure 4.1: The five possible stimulus positions (top left, top right, centre, bottom left and bottom right) as well as examples of the three shift types (horizontal shift, vertical shift and centre-periphery shift). Actual stimuli were in full colour.

4.1.2.2 Participants

All of the 43 participants were students at the University of Nottingham (27 female, 16 male), taken from an opportunity sample. All participants were aged between 20-22 years, and had normal or corrected-to-normal vision. None had previously been exposed to the stimuli. Of these participants, 21 were assigned to the size only change condition, and 22 were assigned to the size + location condition.

4.1.2.3 Apparatus and materials

The experiment was run using an *Apple Macintosh G3* computer (300Mhz, 384Mb RAM) with an *ATI Radeon* 7000 (32Mb) graphics card, on a 21” *Mitsubishi* colour display monitor (size: 1024 x 768 pixels, resolution: 72 x 72 dpi, refresh rate: 75Hz). Stimuli were presented in the centre of the screen using *Matlab* v5.2.1 (Mathworks UK), running the psychophysics toolbox. All bitmap images were converted to a standard size of 59 x 59 pixels, and had an area of 2 cm x 2 cm when displayed at the screen resolution described above. Participants were seated at a distance of 57.5 cm from the screen so individual stimuli subtended an area of approximately 2° x 2° of visual angle.

4.1.2.4 Stimuli

The stimuli used were two series of 253 abstract ‘faces’ composed of modified ellipses. The basic procedure for generating faces is described in the Methods section for Experiment 1.

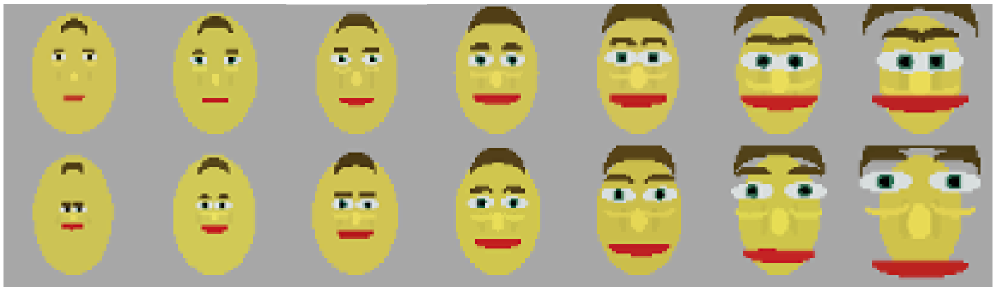

Figure 4.2: The seven steps of feature size. The top row shows faces from the size condition. Each face has features 20% larger than the face to its left. The bottom row shows faces from the size + location condition. Each face has features 20% larger than the face to its left and some features differ in location by 20% between subsequent faces. Actual stimuli in full colour.

The two series differed slightly in the differences that occurred from item to item. In the “size” condition each new face in the series differed from the previous by a 20% increase or decrease in the size of certain features. In the “size + location” condition each face differed from the previous both in the size of those features and 20% changes in both their horizontal and vertical co-ordinates. In addition, small changes to feature form and colour occurred at random throughout the series. Seven steps of feature size (and location) were permissible (see Figure 4.2). Apart from at the minimum and maximum feature sizes in the set, feature size would increase or decrease at random. Because the faces were exemplars of the same category, the stimuli were considered to be resistant to verbal encoding.

Of the 253 faces, 200 occurred twice forming shift pairs of study and test presentations. These 200 pairs were balanced for horizontal, vertical, centre-periphery, centre (no shift), and periphery (no shift) conditions with 40 of each type. Within each shift type, pairs were balanced for the location of the shift. So, for example, of the 40 horizontal shift pairs, 10 were from top left to top right, 10 from top right to top left, 10 from bottom left to bottom right, and 10 from bottom right to bottom left. This avoided the confounding of results for a particular shift type (e.g. horizontal) with the effects of shift direction and shift locations.

In order to prevent the emergence of obvious patterns in presentation (i.e. ‘different’-‘same’-‘different’-‘same’), two further conditions were included. There were 50 ‘foils’: faces that occurred once and did not recur (10 in each location), and 25 ‘repeats’: faces identical to the preceding pair (5 in each location). Both were balanced across all five locations, but the shift was not specified.

The order of the 253 faces was always the same, but the order of the conditions within the sequence was always determined randomly. Together with an initial unscored buffer of 3 trials, the experiment comprised 478 trials.

Faces were displayed in a continuous series. Participants were required to respond according to whether they thought each image was the same or different to the previously presented one, without distinct learning and test phases. This modification of the continuous recognition paradigm has a number of advantages over designs with distinct phases, minimising the opportunity for rehearsal and covert regrouping of stimuli (Shepard & Teghtsoonian, 1961).

A grey cross was displayed in the centre of the screen as a fixation point. It was presented at low contrast in order to avoid after-images that might interfere with the processing of stimuli presented at the centre of the screen. The screen background was a mid-grey and remained the same throughout the experiment.

4.1.2.5 Procedure

Instructions were presented to participants in written form. The instructions informed them that they would see a series of faces, and that for each they were required to make a response indicating whether they thought it was the same or different from the last image. The instructions stated that an old item was one that was identical to any previous item in all aspects except spatial location. As such, an item could be classed as ‘old’ even if it appeared in a different location to its previous appearance. Participants were instructed to fixate on the central cross throughout the experiment. The cross was present throughout except when a stimulus was presented in the central position.

Participants carried out a 15-trial practice session, with face stimuli that were different to those used in the main experiment. If the experimenter was convinced that they had understood the instructions of the experiment the main task was begun. If the participant seemed unsure as to the correct procedure further verbal instruction was given before they commenced with the experiment.

Stimuli were displayed in the appropriate position for 100 ms, a presentation time too short to allow the participant to fixate a peripherally-presented stimulus via a saccade (Saslow, 1967). Following the presentation participants were required to make a response with the click of a mouse button. If they thought the stimulus was the ‘same’ as the previous stimulus, they were to click the left button. If they thought it was ‘different’ they were to press the right. The next trial would not start until both a response had been given, allowing participants to pace the experiment according to their own ability. There was a brief interval between the stimulus presentation and the next stimulus during which a feedback tone was played. A high-pitched tone indicated a correct response and a low-pitched tone indicated an incorrect response. This provided some motivation for participants to maintain attention throughout the duration of the experiment.

4.1.3 Results

Hit rates were calculated for each shift type in both change conditions. A false alarm rate was also calculated for each change condition. It was clear that some participants' found the task difficult and had performed very poorly. In order to exclude such data a one-tailed chi square was conducted on each participant's responses, to determine that the number of 'same' responses to 'old' stimuli was significantly above that expected by chance. In two cases there was no significant difference, and those participants' data were omitted from further analyses (both of these participants had taken part in the 'size only' change condition).

The included data were converted into d' scores according to both shift type (Figure 4.3) and by test location (Figure 4.4).

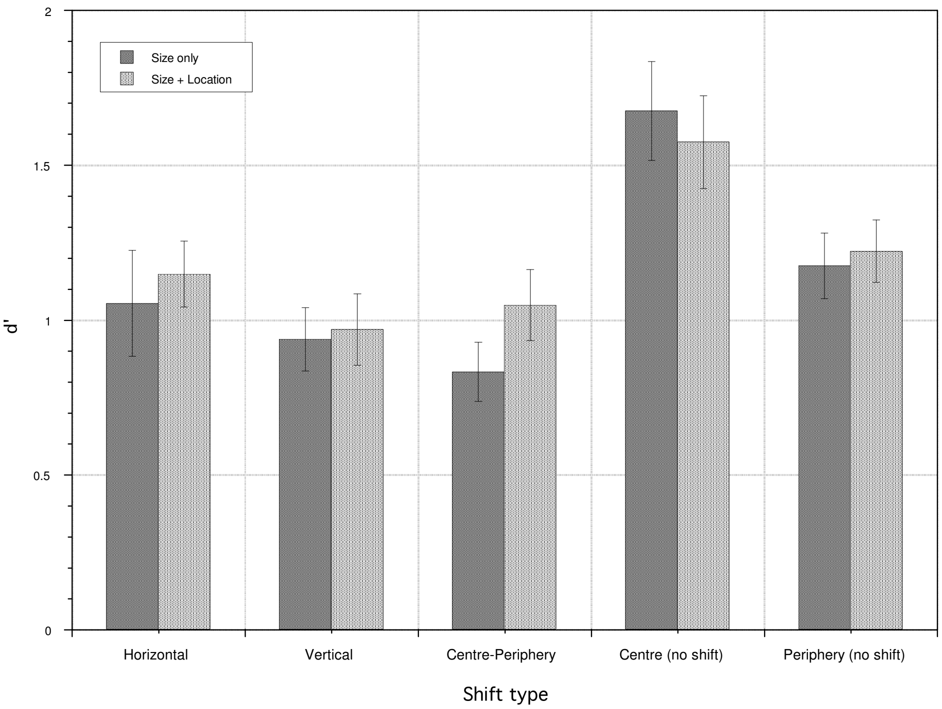

Figure 4.3: The effect of positional shift on d' scores. Data = mean ± SEM.

4.1.3.1 D-prime scores

Data were entered into a 2 (change conditions) x 5 (shift type) mixed ANOVA. Mauchley's test of sphericity was significant and consequently Greenhouse-Geisser epsilon corrected degrees of freedom are quoted for shift type. There was a significant main effect of shift type (F(2.80,109)=34.2, MSe=0.137, $p<0.001$). Tukey's post-hoc tests revealed that recognition in the centre (no shift) condition was significantly better than in all other conditions (all $p<0.001$). In addition, recognition in the periphery (no shift) condition was better than that associated with vertical ($p<0.01$) and centre-periphery ($p<0.001$) shifts. The main effect of change condition was not significant, and neither was the interaction between shift type and change type.

Performance was also analysed by test location. Hit rates for each of the five stimulus positions were used to calculate *d'* scores. A 2 (change condition) x 5 (test position) ANOVA was performed on the data. Mauchley's test of sphericity was not significant. There was a significant main effect of location (F(4,156)=35.8, MSe=0.0619, p<0.001). Tukey's post-hoc tests revealed recognition at the central position was significantly better than at all four peripheral locations (all p<0.001). Again, no significant main effect of change condition was found and there was no interaction between the two factors.

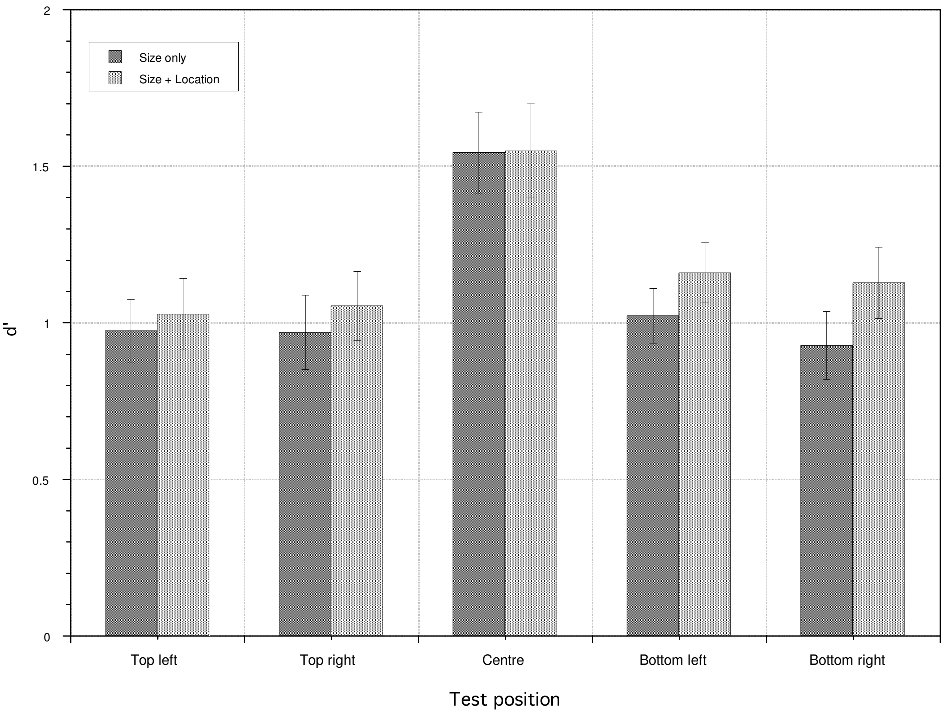

Figure 4.4: The effect of test position on d' scores. Data = mean ± SEM.

4.1.3.2 Reaction times for correct recognition

Reaction times were also measured for hits, and the mean and SEM values for hits of each shift type are shown in Figure 4.5. The data were also analysed by test position (Figure 4.6).

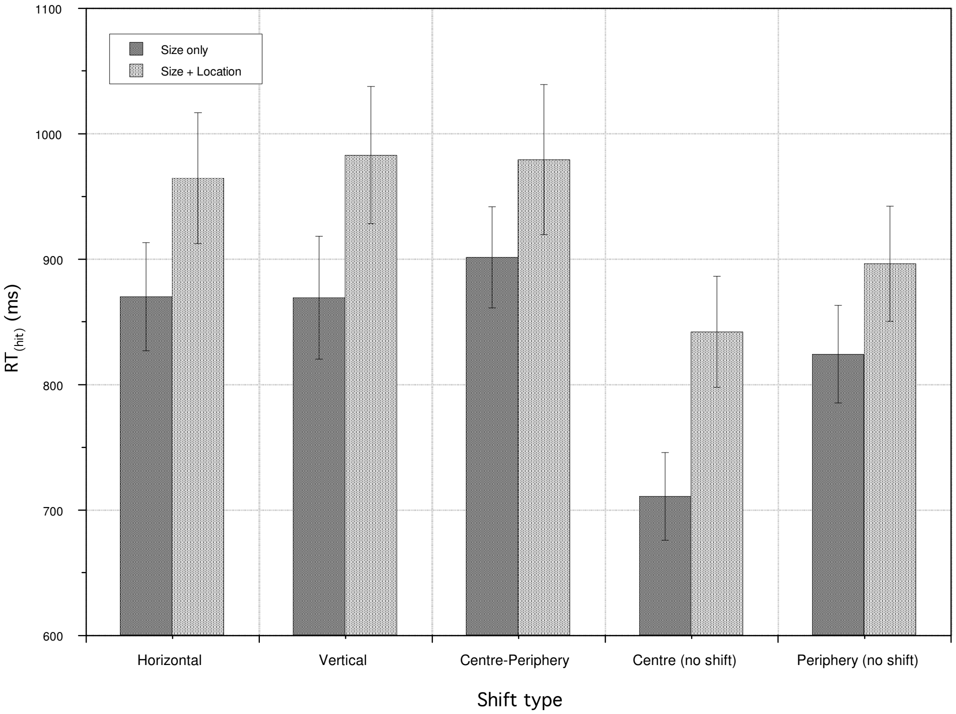

Figure 4.5: The effect of positional shift on reaction times for correct recognition (ms). Data = mean ± SEM.

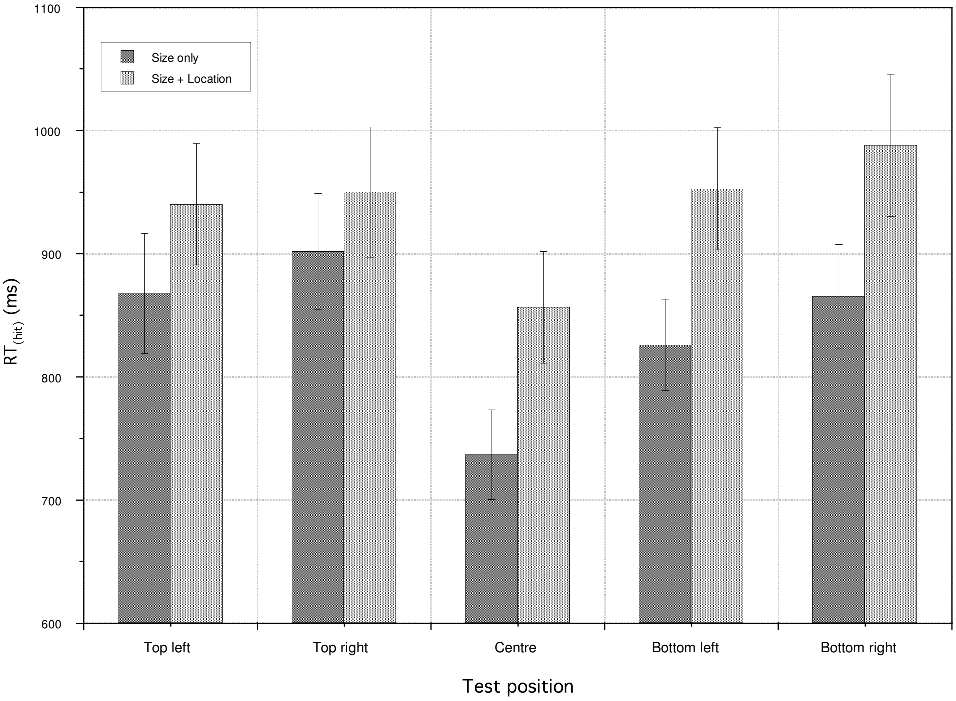

Figure 4.6: The effect of test position on reaction times for correct recognition (ms). Data = mean ± SEM.

A similar 2 x 5 ANOVA as that detailed for the effect of positional shift on *d'* scores was carried out for the reaction time data. There was a significant main effect of shift type (F(4,156)=26.0, MSe=7120, p<0.001). Tukey's post-hoc tests revealed significantly faster reaction times associated with centre (no shift) than all other conditions (p<0.001), and faster reaction times for periphery (no shift) than for horizontal (p<0.05), vertical (p<0.01), and centre-periphery (p<0.001) shifts. Despite the fact that, for each shift condition, size only reaction times were faster than those for size + location, no significant effect of change condition was found. There was also no interaction between change condition and test position.

The reaction time results for test position were also analysed with a similar 2 x 5 ANOVA as that described previously for *d'* scores. Mauchley's test of sphericity was not significant. A significant main effect of test position was found (F(4,156)=15.8, MSe=7440, p<0.001). Tukey's post-hoc tests revealed significantly faster reaction times associated with recognition at the central position than at all peripheral locations (all p<0.001). Neither the main effect of change condition nor the interaction between change condition or test position reached significance.

4.1.3.3 Serial position

In order to compare this hybrid same/different-continuous recognition paradigm with the continuous recognition results obtained previously (see Chapter 2 and Chapter 3), the hit (p(hit)) and false alarm (p(false)) rates were determined for 10 trial epochs that were plotted in sequence (Figure 4.7). Unlike the results described in Chapters 2 and 3, these data do not show any consistent trends of p(hit) or p(false) associated with serial position, suggesting that recognition was performed at a steady state.

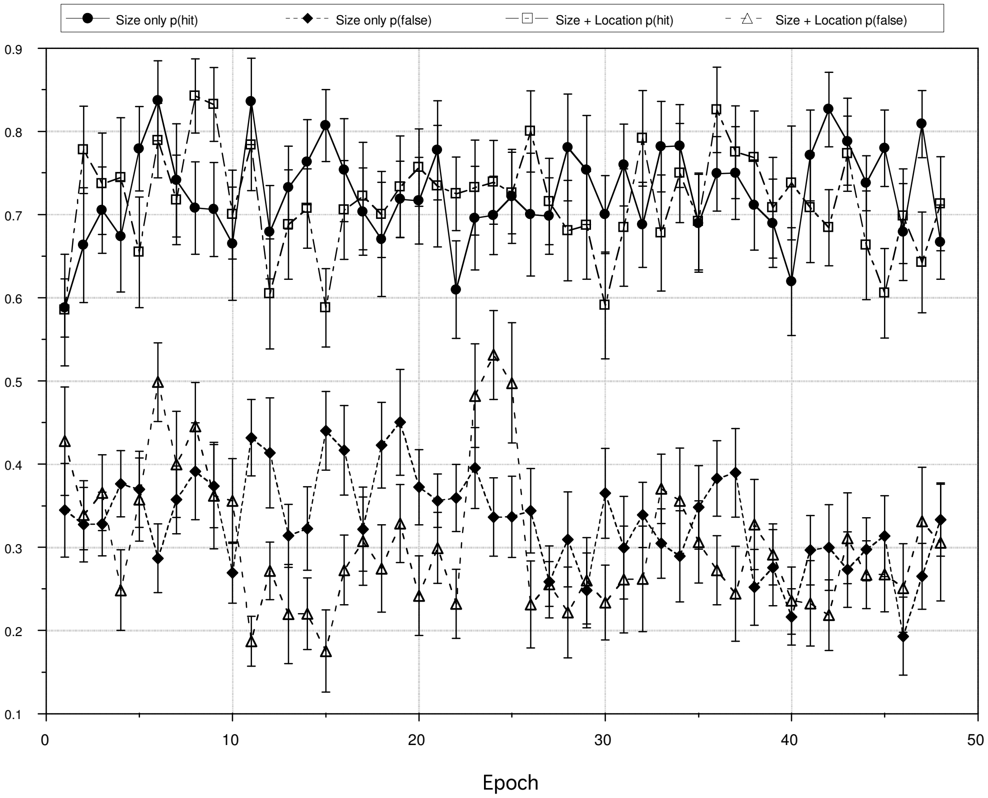

Figure 4.7: The effect of serial position on hit and false alarm rates. Epoch = 10 trials. Data = mean ± SEM.

4.1.4 Discussion

Positional translation impaired recognition accuracy and speed. Whilst, perhaps unsurprisingly, recognition was best when a stimulus was studied and tested at fixation, there was also an advantage for stimuli that remained in the same peripheral location compared to shifted stimuli. In the case of *d'* scores this advantage was not complete. The periphery (no shift) condition was associated with higher scores than both vertical and centre-periphery shifts, but was not significantly different to horizontal shifts. However, reaction times were faster for the periphery (no shift) condition than all shift conditions.

It is unsurprising that recognition was best at fixation, because one would expect the better acuity of foveal vision, compared with peripheral vision, to be more suited to the fine discrimination of features that the task required. Interestingly, however, this better acuity was of no help in forming an accurate representation of items subsequently tested in the periphery. When stimuli were studied centrally and then tested at a peripheral location (centre-periphery shift) *d'* scores were lowest, and reaction times slowest.

The results suggest that horizontal and vertical shifts were comparable in terms of their effects on recognition, in contrast to the findings of Gratton et al. (1997) and Hornak et al. (2002), but in agreement with those of Dill and Edelman (2001) who used a more similar stimulus set and task to the present study. Hornak et al.'s study presented stimuli in pairs on opposite sides of the fixation point, requiring the participant to make recognition judgements based on the identity of both stimuli. This task may have been more demanding than a same/different decision on a single stimulus (as in the current study), as it required division of visual attention across both left and right, or upper and lower, visual hemifields. This could conceivably have resulted in participants using verbal labels for the pictures of objects that were used as stimuli, as an augmentation strategy for memorising items. Language abilities are, of course, to some degree hemispherically-organised, and the verbal labeling strategy might be expected to be disrupted more by horizontal shifts than vertical ones. The authors argue that, as they included multiple exemplars of each category of object used, naming would not have been an efficient strategy for object discrimination. However, as has been demonstrated in Chapter 2, this precaution is not sufficient to prevent a category-based advantage in recognition from occurring. Indeed, a recent study by Grill-Spector and Kanwisher (2005) found that participants presented with natural images performed as quickly and accurately on a categorisation task as they did on a task requiring object detection. This suggests that object categorisation is automatic and instantaneous with conscious realisation of an object's presence and, as such, one might infer that object categorisation would have played a part in the Hornak et al. study.

Such explanations, however, cannot account for the similar findings of Gratton et al. (1997) who used line patterns that were, presumably, not amenable to verbal labelling. However, the kind of discrimination required in Gratton and colleagues' study was quite different to that required in the current experiment. In Gratton et al.'s experiment, stimuli were very similar to one another being composed of lines of a limited number of orientations (horizontal, vertical, 45º left and right), and also with the requirement that they be either horizontally or vertically symmetrical. Due to the limited number of component features (the lines) 'new' stimuli could conceivably be considered to be rearrangements of the preceding stimuli, thus making the changes between items similar to those encountered in the configurational 'scrambling' described by Dill and Edelman (2001). These authors found that when stimuli retained the same features but differed in their spatial configuration, they were discriminated in a manner different to that for stimuli whose configuration remained the same, but whose features differed. Dill and Edelman did not, however, find any evidence of differences between vertical and horizontal shifts, suggesting that these differences may be specific to the type of stimuli being tested.

Interestingly, the nature of variation in the stimuli (change condition) had no significant effect on any of the results of the current experiment. Whether the change between stimuli consisted of a simple change in the size of features, or in both their size and location, the effect was the same. However, as noted in the Results section, reaction times for size only were always faster than those for size + location, and *d'* scores were almost always lower for this condition. Further research would be needed to establish the reliability of this finding. Based on the findings of Dill and Edelman, one might have expected to see a difference in the level of positional specificity of representations when the configuration of features was manipulated compared to when only the properties of the features themselves were altered. However, the configurational change in the current study was relatively mild compared to that employed by Dill and Edelman, where features could be present anywhere within the stimulus. In the current experiment, features retained the same basic relationship to one another, even if their proportional positions altered from item to item.

The results can be interpreted as evidence that memory representations of complex novel visual stimuli have some positional specificity, as has previously been determined for simple patterns, but this specificity is not complete. This suggests that the dichotomy between total specificity and complete invariance may be false, and that the accuracy of perception of different objects under different conditions lies along a continuous dimension between these two extremes. Indeed, it might be more sensible to consider these results as evidence for same positional advantage rather than positional specificity, as the superiority of periphery (no shift) trials over trials shifted within the periphery was not clear-cut. Future work is needed to determine whether this advantage is persistent or whether certain experimental conditions eradicate it.

4.2 Experiment 6: The putative effect of position on recognition is not observed when eye position is monitored

4.2.1 Introduction

The accuracy of results obtained in experiments examining the effects of retinal position on recognition, is crucially dependent on the assumption that participants are fixating where they should be. Whilst many studies rely on participants following instructions to maintain stable, voluntary fixation (Biederman & Cooper, 1991; Dill & Edelman, 2001; Dill & Fahle, 1997, 1998; Gratton, Corballis, & Jain, 1997; Hornak, Duncan, & Gaffan, 2002; Kahn & Foster, 1981), very few have verified fixation using eye-tracking equipment. One might expect that, as participants are normally unable to predict the location at which the next stimulus will appear, the central fixation point will be fixated. However, there is the possibility that fixation will tend towards the location of the last stimulus producing an advantage for discrimination at that location. For example, if a stimulus is presented to the left of fixation and subsequently a participant's fixation tends towards the left, if the next stimulus is presented at that same location it will be experienced at a more central region of the retina than the experimenter assumes. If the next stimulus is presented to the right of fixation, it will be experienced more peripherally than assumed. If, as has been suggested in Experiment 5, there is a recognition advantage associated with central vision, the former case (when discrimination was between two presentations at the same location) might be expected to be associated with better recognition, and the latter (where the two presentations are at different locations) might be expected to be associated with poorer recognition. Unless eye position is monitored, fixation can only be assumed.

To the author’s knowledge eye-tracking has only been included in one previous study (Nazir & O'Regan, 1990). In their first experiment, a photoelectric device was used to measure eye movements, and if a participant’s gaze deviated more than 0.3° from fixation the target was immediately masked. There is, however, no data on the accuracy of the device used in the Nazir and O’Regan study, and no way of knowing how rapidly their apparatus could mask the stimuli once deviations from fixation had been detected. Given these potential sources of error it is possible that a participant’s eye movements could have taken their gaze considerably closer to the target than the reported methods describe.

In order to address this issue, the current experiment replicated Experiment 5, with the addition of eye-tracking to verify fixation at the start of each trial. No trial would start until fixation of a cross at the central location was detected. Whilst this was often associated with longer intertrial intervals at the start of the experiment, the effect was one of training the participants to fixate the cross between trials, such that later trials were performed without delay. The use of eye-tracking ensured that stimuli were presented at their assumed retinal locations.

4.2.2 Methods

4.2.2.1 Design

The experiment had a one-way within subjects design. The independent variable was the nature of the positional shift (or lack thereof) between stimulus location on initial viewing (study presentation) and subsequent viewing (test presentation). The shift types studied were the same as those described for Experiment 5. The dependent variables were the participants' *d'* scores, reaction times, and hit and false alarm rates.

4.2.2.2 Participants

All of the participants were students at the University of Nottingham (10 female, 10 male), taken from an opportunity sample. The mean age of participants was 20.4 (± 0.3) years of age. All participants had normal or corrected-to-normal vision. None had previously been exposed to the stimuli.

4.2.2.3 Apparatus and materials

The experiment was run using the same computer and software described for Experiment 5. The images were displayed at the same size (approximately 2° x 2° when viewed at a distance of 57.5 cm).

Eye movements were monitored using an *ISCAN Inc. RK-726PCI* Pupil/Corneal reflection tracking system. The system divided the video signal into a 512H x 265V pixel matrix. Eye position data was refreshed every 16.7 ms with a resolution of approximately 0.3°.

4.2.2.4 Stimuli

The stimulus set employed was the "size only" set of face-like stimuli previously described for Experiment 5.

4.2.2.5 Procedure

Instructions were presented to participants in written form. These were similar to those given for Experiment 5, with the additional information that each trial would only begin once participants were fixating on the cross at the centre of the screen. The low-contrast fixation cross (approximately 1° in each direction from the centre of the screen) was present throughout the experiment, except when a stimulus was presented in the central position.

Participants were seated in front of the screen, and the eye-tracking apparatus was calibrated. Participants were asked to fixate each of 5 crosses, presented on the screen, in turn. One was in the centre, and then there was one in each corner of the screen. The information from each was used subsequently to estimate eye position. Participants' fixation of each cross was tested a second time to determine whether the estimations were accurate, and the process was repeated until accuracy was achieved. Participants first participated in a 15-trial practice session, with similar but different stimuli to those used in the main experiment. If the experimenter was convinced that they understood the instructions of the experiment the main task was begun. If the participant seemed unsure as to the correct procedure further verbal instruction was given before they began the experiment.

Before both the practice and main sequences the experimenter gave the instruction to fixate on the central cross. Once the participant was fixating, the experimenter pressed a key to set the current eye-tracker co-ordinates as those for future fixation assessments. Once this was done the participant was free to start the sequence with a mouse click.

At the start of each trial the program waited until the participant was fixating within a 1° x 1° area around the centre of the fixation cross. Once correct fixation was detected the stimulus would be displayed in the appropriate position for 100 ms. Following the presentation, participants were required to make a response with the click of a mouse button. If they thought the stimulus was the ‘same’ as the previous stimulus, they were to click the left button. If they thought it was ‘different’ they were to press the right. The next trial would not start until a response had been given, allowing participants to pace the experiment according to their own ability. There was a brief interval between the stimulus presentation and the next stimulus in which a feedback tone was played. A high pitched tone indicated a correct response and a low pitched tone indicated an incorrect response. This provided some motivation for participants to maintain attention throughout the duration of the experiment.

4.2.3 Results

Both hit rates and false alarm rates were calculated in a similar manner to that described for Experiment 5. In order to exclude data from participants who performed very poorly on the task, a one-tailed chi square was conducted on each participant's responses, to determine that the number of 'same' responses to old stimuli was significantly above that expected at chance. In one case there was no significant difference, and that participant's data was omitted from further analysis. The resulting data were converted into *d'* scores (Figure 4.8).

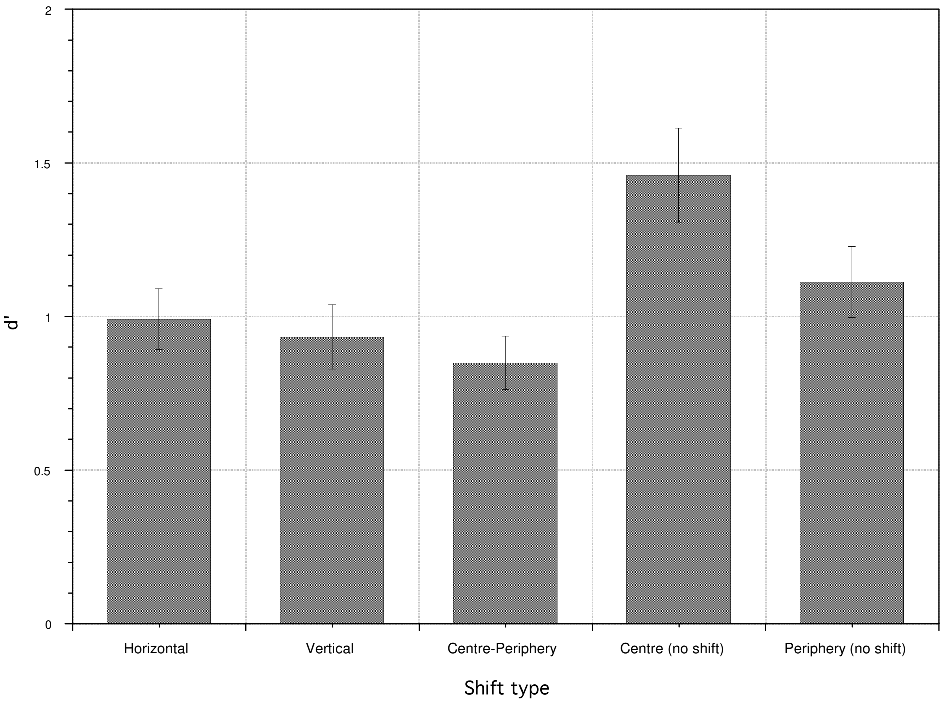

Figure 4.8: The effect of shift type on d' scores. Data = mean ± SEM.

4.2.3.1 D-prime scores

Mauchley's test of sphericity was not significant for these data. A one-way within subjects ANOVA with 5 levels was performed on the data with shift type as the within-groups factor. There was a significant main effect of shift type (F(4,72)=10.9, MSe=0.0994, p<0.001). Tukey's post-hoc tests revealed that scores for the centre (no shift) condition were significantly higher than those for all other conditions (at least p<0.01). In a departure from the results of Experiment 5, no differences between the periphery (no shift) condition and any of the shift conditions were found.

Performance was also analysed by location. Hit rates and false alarm rates for each of the five stimulus positions were used to calculate *d'* scores (Figure 4.9).

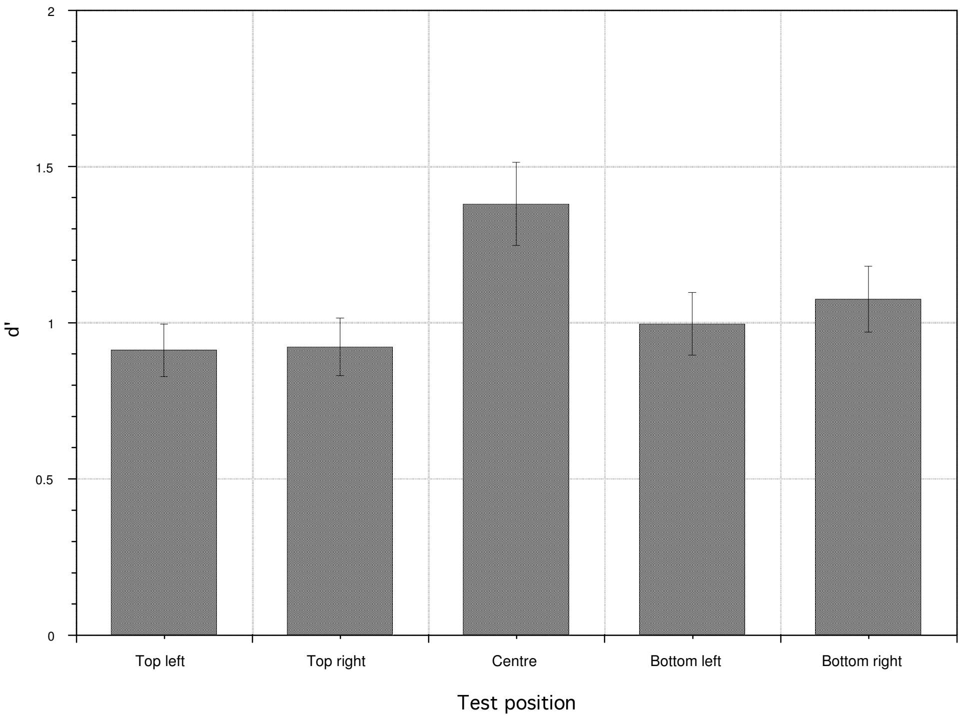

Figure 4.9: The effect of test position on d' scores. Data = mean ± SEM.

Mauchley's test for sphericity was not significant. A one-way within subjects ANOVA with 5 levels was performed on the data with test position as the within-groups factor. There was a significant main effect of location (F(4,72)=11.1, MSe=0.0631, p<0.001). Tukey's post-hoc tests revealed that scores for the central location were significantly higher than for all other locations (at least p<0.01). No other differences reached significance.

To summarise, *d'* scores showed a recognition advantage when test occurred at fixation (i.e. the centre (no shift) condition). However, whether stimuli in the periphery were shifted or not made no difference to *d'* scores. The addition of eye tracking to the experimental setup effectively eradicated the advantage for periphery (no shift) over some shift types that was observed in Experiment 5.

4.2.3.2 Reaction times for correct recognition

Reaction times for hits were also measured (Figure 4.10). Mauchley's test of sphericity was not significant. A one-way repeated measures ANOVA was carried out on the data for shift type. There was a significant main effect of shift type on reaction times for hits (F(4,72)=14.9, MSe=7040, p<0.001). Tukey's post-hoc tests revealed significantly faster reaction times for the centre (no shift) condition than horizontal, vertical, and centre-periphery shifts (p<0.001). In addition, the periphery (no shift) condition was associated with faster reaction times than both vertical (p<0.05) and centre-periphery (p<0.001) shift conditions. These results are important in demonstrating that the superior *d'* scores for the centre (no shift) shift type were not the result of a speed-accuracy trade off.

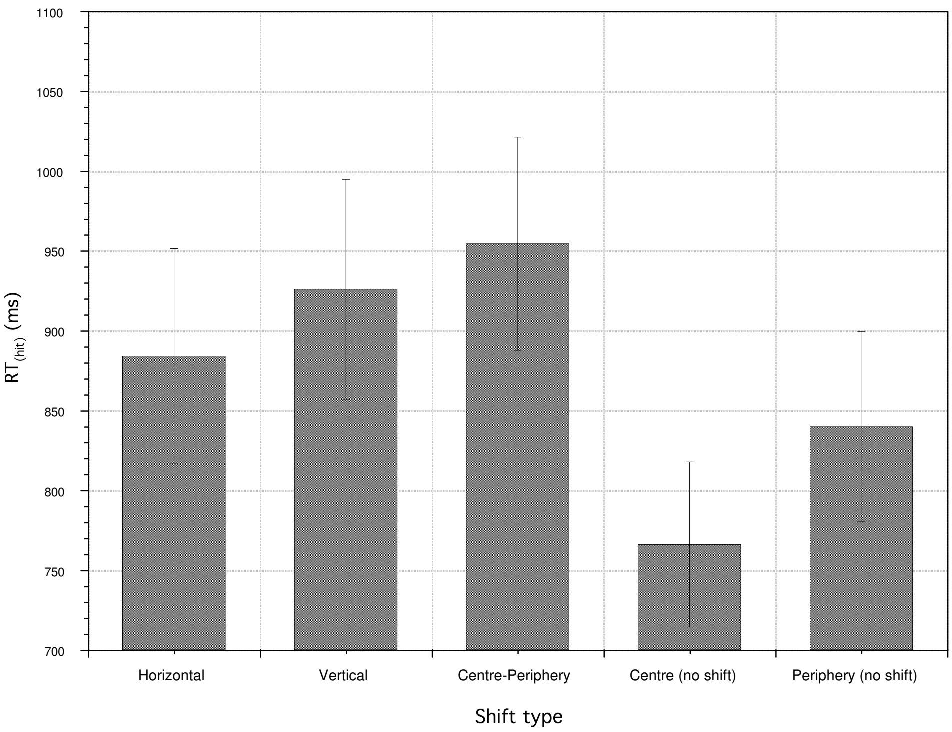

Figure 4.10: The effect of shift type on reaction times for hits (ms). Data = mean ± SEM.

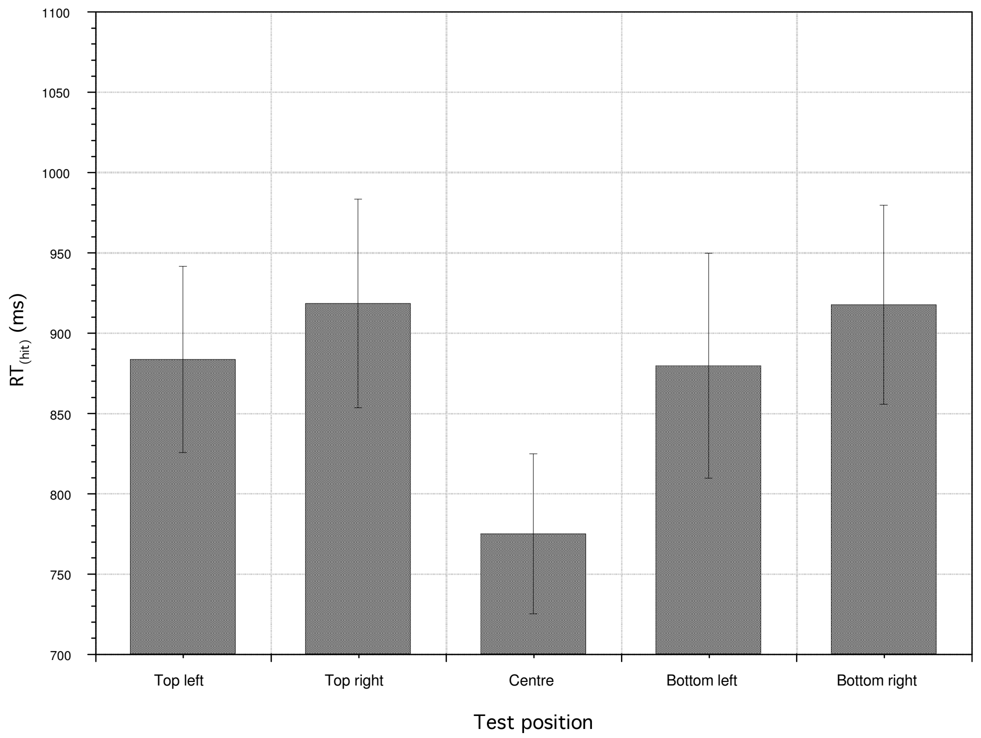

Figure 4.11: The effect of test position on reaction times for hits (ms). Data = mean ± SEM.

Reaction times were also calculated by test position (Figure 4.11). A similar one-way ANOVA to that carried out for the shift type data was performed on test position data. Mauchley's test of sphericity was not significant. There was a significant main effect of test position on reaction times for correct recognition (F(4,72)=10.3, MSe=6360, $p<0.001$). Reaction times were faster at the central position than top left, top right, bottom right ($p<0.001$) and bottom left ($p<0.01$), again demonstrating that the central advantage was not the product of a speed-accuracy trade off.

Interestingly, whilst the superior *d'* scores associated with the periphery (no shift) condition in Experiment 5 were not replicated in the current experiment, faster reaction times for this condition were present.

4.2.3.3 Serial position

The data for p(hit) and p(false) were calculated for epochs of 10 trials (Figure 4.12). The p(false) values decreased consistently throughout the experiment. This is in contrast to the effect of serial position on p(false) observed for the continuous recognition task, in Experiments 1-4. In these experiments, p(false) increased during the course of the experiment before reaching a steady state. In addition p(hit) values were approximately constant throughout this experiment.

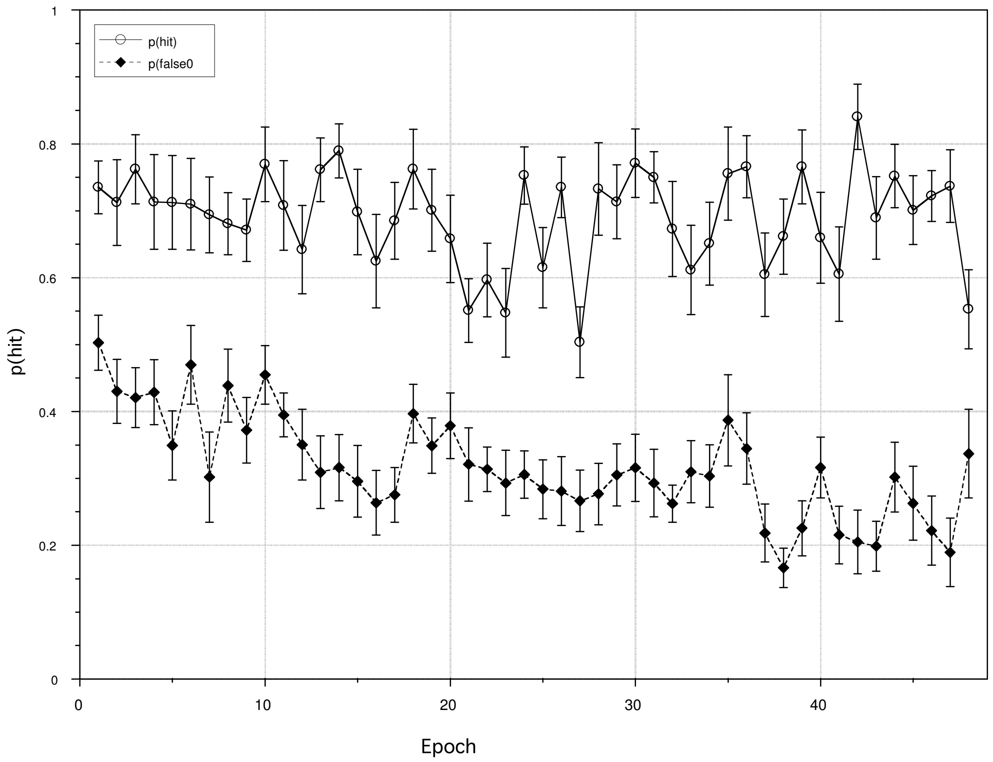

Figure 4.12: The effect of serial position on p(hit) and p(false). Epoch = 10 trials. Data = mean ± SEM.

4.2.4 Discussion

The results demonstrated that shifting a stimulus between study and test presentations had no significant effect on recognition accuracy when fixation was verified with an eye-tracker. A significant effect of ‘shift type’ was demonstrated but post-hoc analysis revealed that this was attributable to a difference between the centre (no shift) condition and all other conditions. In contrast to the findings of Experiment 5, recognition accuracy, as measured by *d’* scores, was no better when study stimuli presented in the periphery were subsequently tested in the same location, compared to when they were shifted. It would seem that the apparent positional effects of Experiment 5 (i.e. some degree of positional specificity in the peripheral visual field) were artefacts of changes in fixation. When participants are trained to return their eyes to the central cross between trials, *d’* scores for the periphery (no shift) condition were no different to the scores for horizontal, vertical, or centre-periphery shifts. The remaining valid differences between all four of these conditions and the centre (no shift) condition can be explained with reference to the better acuity of vision in the foveal region, which would be expected to produce more detailed mnemonic representations of stimuli, and facilitate discrimination. Also, whilst the 100 ms presentation time was too rapid for a saccade to peripherally presented items to be carried out, there is the possibility that the initial stages of saccade initiation may have begun during this time (e.g. Holt, 1903; Ross, Morrone, & Burr, 1997; Volkmann, 1962), and this may have disrupted perception of the stimulus. Obviously this would not have been the case for items presented at fixation.

The most likely reason for the differences observed between Experiments 5 and the current experiment, is that when fixation was not explicitly controlled, the gaze of participants remained directed closer to the location of the last stimulus observed than the central location (when presentation occurred in the periphery). This would have meant that if such a stimulus was the study stimulus of a periphery (no shift) trial, the subsequent test, in the same location, would have been experienced more centrally. As vision has greater acuity nearer the central fovea, this might have aided discrimination in the manner suggested previously. Alternatively (regardless of acuity) position specificity may be a property of central vision not the periphery. Vertically or horizontally shifted presentations would likely have been further away from this putative drifted point of fixation, with a corresponding deficit to discrimination. Such advantages and disadvantages would have been eradicated by the task demand of training the gaze on the fixation cross at the start of each trial.

Interestingly, whilst the discrimination advantage for periphery (no shift) was eradicated by the introduction of eye-tracking, the reaction time advantage for this condition over shift conditions was not entirely dissipated. Recognition was faster in the periphery (no shift) condition than in both the vertical and centre-periphery shift conditions, although it was no different to the horizontal shift condition (a difference that was present in the results of Experiment 5). This finding gives an indication that there remains some recognition advantage, in terms of speed of processing (but not necessarily accuracy), for items presented in the same location at study and test (albeit a weaker one than that found with free fixation).

The demonstrated importance of verifying fixation location has far-reaching implications for research in this field, which has largely neglected this precaution. None of the previous positional translation studies, with the exception of one of Nazir and O'Regan's (1990) experiments, has measured the direction of the gaze, and merely assume that fixation remains at the designated marker. This casts previous findings of positional effects by a number of authors into serious doubt (Dill & Edelman, 2001; Dill & Fahle, 1997, 1998; Kahn & Foster, 1981). However, Nazir and O'Regan's (1990) Experiment 1 did find positional effects, despite monitoring the direction of gaze. Could it be that differences of task and/or stimulus type were responsible for the breakdown of translation invariance in this case? A possible explanation for the discrepancy between the current results and those of Nazir and O'Regan, is that the two experiments followed very different procedures, used different stimuli at different retinal locations, and even reported a different measures of recognition. In Nazir and O'Regan's study, participants were given unlimited time to familiarise themselves with targets at their 'study' location (±2.4° from fixation) and were then trained extensively in discrimination of targets from distracters at that location. This is a very different kind of learning to that involved in rapidly memorising an image seen only once. Also, the stimuli were dot patterns, so recognition could only be achieved through memorisation of their spatial configurations (as all dots were the same size, and all configurations consisted of the same number of dots). With many repetitions of stimuli at the same retinal location one might expect a certain degree of perceptual learning at that location that would aid subsequent recognition there. Indeed, Dill (2002) reviews evidence from a number of studies suggesting that position-specific perceptual learning of such patterns occurs over hundreds of trials.

The findings of the current experiment make clear the importance of using eye-tracking equipment to objectively verify the fixation location, in order to ensure that participants are fixating where they are supposed to. Despite the methodology of this experiment differing from Experiment 5 only in the controlling of fixation, entirely different results were produced. They suggest that the finding of Experiment 5 that same-different discrimination of face-like stimuli is subject to positional specificity effects, is potentially erroneous, and that, in fact, such discriminative ability is translation invariant. Further experiments monitoring the direction of gaze without controlling it would be useful in determining whether the advantages of discrimination accuracy are the product of drifting fixation, or whether other factors are involved.

Future work is also required to determine whether this effect is specific to particular tasks and stimuli, or whether a reevaluation of the vast majority of positional translation literature is needed. Given the results of Nazir and O'Regan (1990), whose study did take into account the direction of gaze, further work needs to establish what features of stimuli, and what kinds of task, cause a breakdown of translation invariance. Are the effects of position related to configurational versus featural discrimination? Are they the product of perceptual learning at a single location as opposed to a single exposure? These are questions that will need to be answered in order to better understand how the translation invariance observed in the current experiment can, under certain circumstances, breakdown.

4.3 Experiment 7: Translation invariance for location transfers to fractal stimuli

4.3.1 Introduction

The results of Experiment 6 demonstrated that the sensitivity of same/different discrimination of face-like stimuli was invariant with regards to positional translation. However, what is unclear is whether this finding can be generalised to performance in this particular task, or whether it is specific to the recognition of the face-like stimuli employed.

There is some evidence in the literature that the degree of positional specificity associated with object representations in memory is related to participants' familiarity with those objects. For example, Biederman and Cooper (1991) found no evidence of positional specificity with their experiments using pictures of common objects, and neither did Dill and Edelman (2001; Experiments 1-3) with animal-like objects. However, experiments using abstract patterns (e.g. Dill & Fahle, 1997, 1998; Kahn & Foster, 1981; Nazir & O'Regan, 1990) have found evidence of incomplete invariance. Indeed, perhaps most compelling are the findings of Dill and Edelman's (2001) later experiments (Experiment 4) in which they scrambled the configuration of the component features of their animal-like objects to make unfamiliar constellations, resulting in the breakdown of translation invariance.

It can be argued that, like the animal-like stimuli used by Dill and Edelman, the face-like stimuli employed in Experiments 5 and 6, whilst not particularly realistic, are familiar to participants to some extent. They emulate the basic configuration of human faces. Without examining the generality of the findings of Experiment 5 to other, more abstract/novel stimuli, it is impossible to tell whether the positional translation observed is common to the recognition of other objects in this task. In order to test, in a limited way, the generality of the findings the same experiment was carried out using abstract fractal patterns, that differed from one another by regular changes to their features in a manner similar to that employed in Experiment 5.

4.3.2 Methods

4.3.2.1 Design

The design of the experiment replicated that described for Experiment 6.

4.3.2.2 Participants

All of the participants were students at the University of Nottingham (11 female, 9 male), taken from an opportunity sample. The mean age of participants was 21.0 (± 0.9) years of age. All participants had normal or corrected-to-normal vision. None had previously been exposed to the stimuli.

4.3.2.3 Stimuli

The stimuli used were a series of 254 fractals, the basic construction of which is given in the Methods section of Experiment 1. In order that this series was in some way comparable to the series of face stimuli used previously (i.e. that measured changes took place between subsequent stimuli) 254 fractals were generated, each differing from the previous by measured changes in each superposition of the stimulus. For each superposition the variable GA, specifying the depth of deflection, was either increased or decreased by 30% at random. Also for each superposition a randomly selected value from the RGB value was either increased or decreased by 0.08 (within the minimum and maximum limits). An example of a series of different stimuli generated in this way can be seen in Figure 4.13.

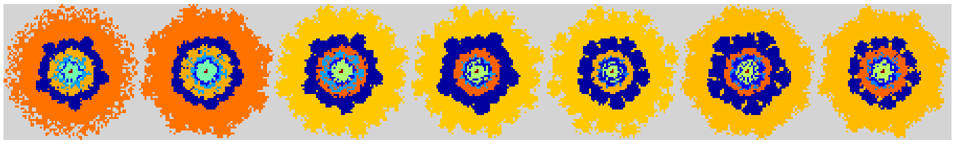

Figure 4.13: An example of seven subsequent fractals. Each fractal varies in the depth of deflection for each superposition, and in changes to variables defining their colours.

4.3.2.4 Procedure

The procedure was the same as Experiment 6, and again, eye-tracking equipment was used to monitor the location of fixation.

4.3.3 Results

A similar preliminary one-tailed chi-square test to that carried out for Experiment 5 was carried out on the results, and all participants were shown to have performed better than chance. Consequently, all participants' data were entered into subsequent analyses. D-prime scores were calculated from the hit and false alarm rates in the manner previously described for Experiment 5.

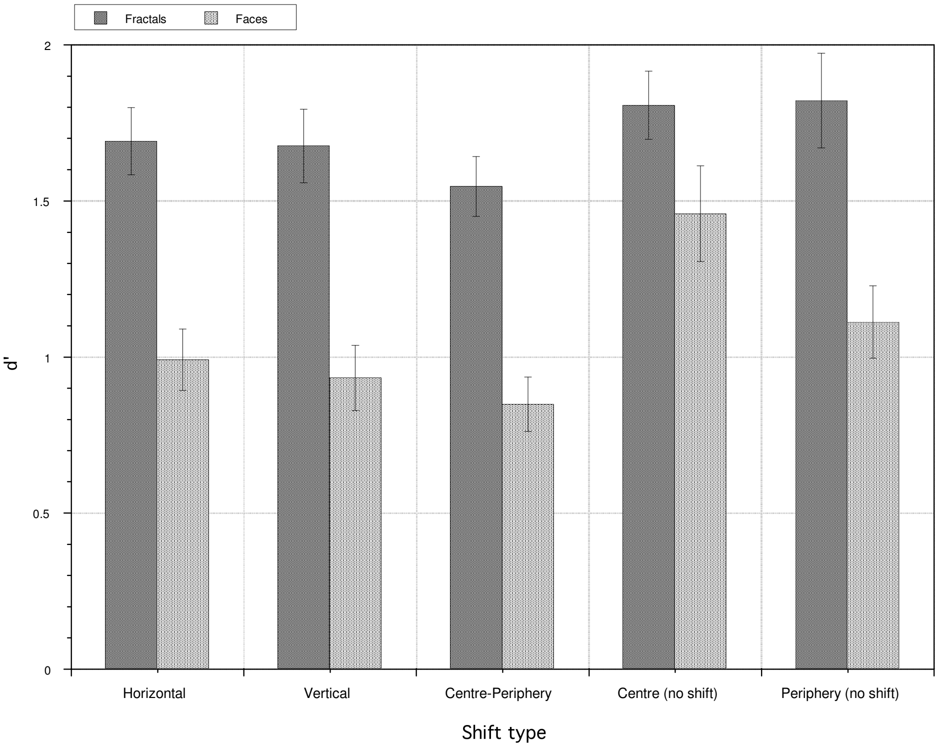

Figure 4.14: The effect of shift type on *d'* scores for face (replotted from Experiment 6 for comparison) and fractal recognition. Data = mean ± SEM.

4.3.3.1 D-prime

D-prime scores were calculated for shift type and test position, separately, as described for Experiment 5. The scores from the current experiment, and those from Experiment 6 were compared (Figure 4.14). The data were entered into a 2 (stimulus type) x 5 (shift condition) mixed ANOVA. Mauchley's test for sphericity was significant for shift type, and Greenhouse-Geisser epsilon corrected values are quoted for these results. There was a significant main effect of stimulus type (F(1,37)=21.0, MSe=0.950, p<0.001). Fractals were recognised with more accuracy than were faces. There was also a significant main effect of shift condition (F(2.51,93.0)=8.60, MSe=0.144, p<0.001). Tukey's post-hoc tests revealed that the centre (no shift) condition was associated with significantly higher scores than horizontal, vertical, and centre-periphery shifts (all p<0.001). In additon, the periphery (no shift) condition was associated with higher scores than the centre-periphery shift (p<0.01). A significant interaction between stimulus type and shift condition was detected (F(2.51,93.0)=6.32, MSe=0.144, p<0.01). Post-hoc tests revealed that for horizontal, vertical, and centre-periphery shifts, and periphery (no shift), recognition of fractals was significantly better than that of faces (all p<0.05), but in the centre (no shift) condition there was no difference. In addition, for face recognition there were significant advantages for centre (no shift) over horizontal, vertical and centre-periphery shifts (p<0.001) and periphery (no shift) (p<0.01). Recognition of fractals was associated with only one significant difference: periphery (no shift) was associated with higher scores than centre-periphery shift (p<0.05).

These results appear to show some effect of position, as there was an advantage of the periphery (no shift) condition over one shift condition (centre-periphery shift). Interestingly, the otherwise consistent advantage of fractal over face recognition was not replicated for the centre (no shift) condition, perhaps suggesting a ceiling effect, or a loss of central advantage for fractal recognition.

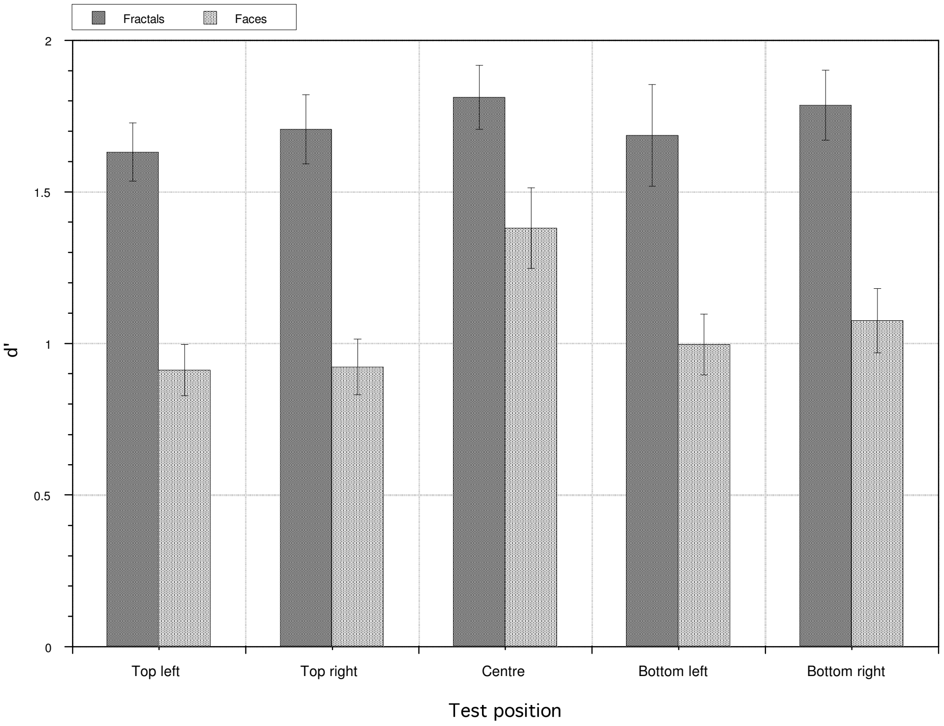

Figure 4.15: The effect of test location on *d'* scores for fractal and face recognition. Data = mean ± SEM.

D-prime data was also analysed by test location (Figure 4.15) in a 2 (stimulus type) x 5 (test location) ANOVA. Mauchley's test of sphericity was not significant and sphericity was assumed. The data showed a significant main effect of stimulus type (F(1,37)=23.1, MSe=0.940, p<0.001) (fractals associated with higher scores than faces). There was also a significant main effect of test location (F(4,148)=7.71, MSe=0.0835, p<0.001). Post-hoc tests revealed a significant advantage for discrimination at the central location over top left and top right (p<0.001), as well as bottom left (p<0.01) locations. No significant interactions were detected.

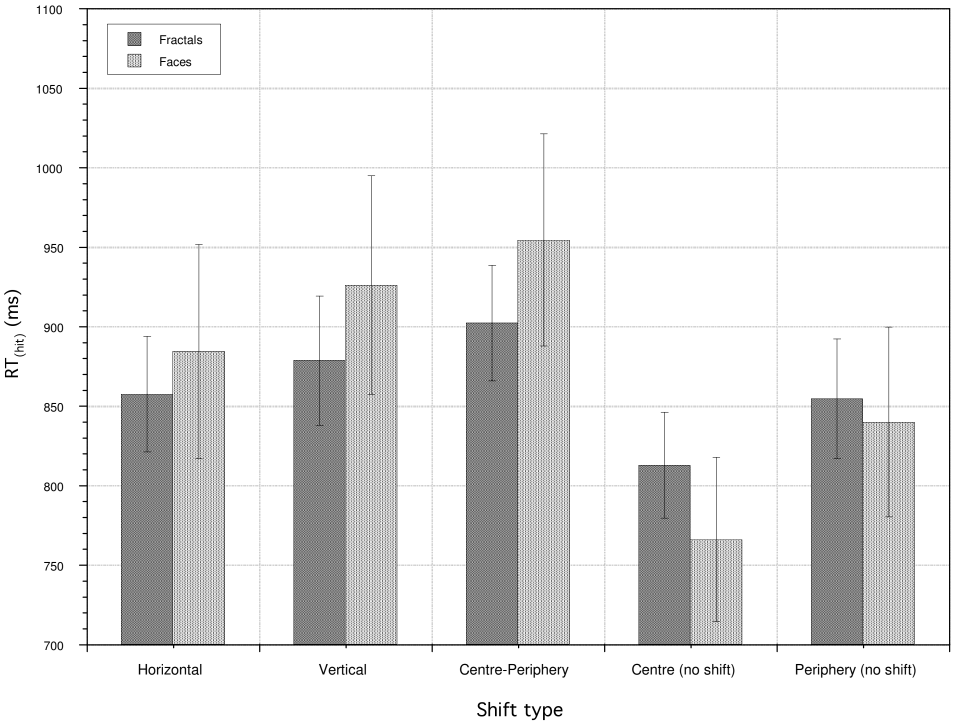

Figure 4.16: The effect of shift type on reaction times for hits (ms) for fractal and face recognition. Data = mean ± SEM.

4.3.3.2 Reaction times for correct recognition

Reaction time data for hits was also analysed by shift type (Figure 4.16) and test location (Figure 4.17). Data analysed by shift type were entered into a 2 (stimulus type) x 5 (shift condition) ANOVA. Mauchley's test of sphericity was not significant. Stimulus type was not significant, but there was a significant main effect of shift (F(4,148)=8.45, MSe=5490, p<0.001). Tukey's post-hoc tests revealed significantly faster reaction times for the centre (no shift) condition than horizontal, vertical, and centre-periphery shifts (p<0.001), and also the periphery (no shift) condition (p<0.01). The periphery (no shift) condition was associated with faster reaction times than centre-periphery (p<0.001) and vertical (p<0.01) shifts. The horizontal shift condition was associated with faster reaction times than centre-periphery shift (p<0.01). Furthermore, there was a significant interaction between stimulus type and shift type (F(4,148)=15.0, MSe=5490, p<0.001). Fractal recognition was associated with faster reaction times for centre (no shift) compared with both centre-periphery (p<0.01) and vertical shifts (p<0.05). Face recognition was associated with significantly faster reaction times for centre (no shift) compared with horizontal, vertical, and centre-periphery shifts (p<0.001) and periphery (no shift) (p<0.05), as well as advantages for periphery (no shift) over centre-periphery (p<0.001) and vertical (p<0.01) shifts, and for horizontal shift over centre-periphery shift (p<0.05).

These results suggest that the effect of translation on reaction times is more moderate for fractal recognition than it is for recognition of faces, although the rank order of the shift conditions (centre (no shift) < periphery (no shift) < horizontal shift < vertical shift < centre-periphery shift) remains the same for both stimulus types.

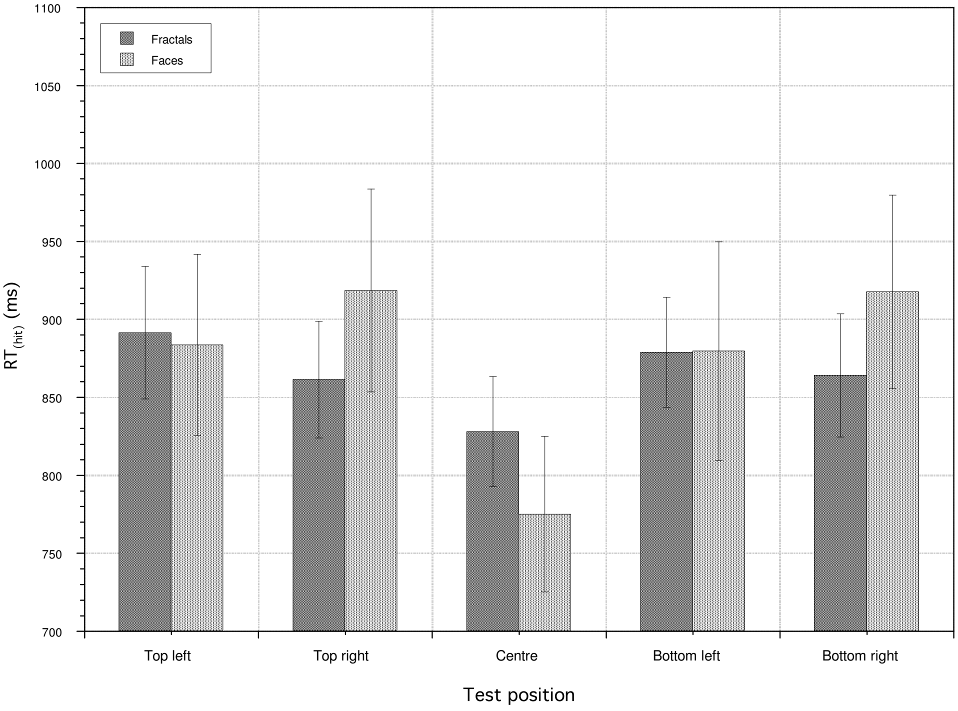

Figure 4.17: The effect of test location on reaction times for hits (ms) for fractal and face recognition. Data = mean ± SEM.

Reaction time data were also analysed by test location in a 2 (stimulus type) x 5 (test location) ANOVA. Mauchley's test of sphericity was not significant. There was no main effect of stimulus type, although a significant main effect of test location was revealed (F(4,148)=9.66, MSe=5970, p<0.001). The central location was associated with significantly faster reaction times than were any of the four peripheral locations (p<0.001). There was a significant interaction between stimulus type and test position (F(4,148)=3.46, MSe=5970, p<0.05). Whilst there was a significant reaction time advantage for discrimination of faces in the central location, no significant differences between reaction times for recognition of fractals by location were identified. This is further evidence that the central advantage observed for face recognition is not as pronounced in the recognition of fractals.

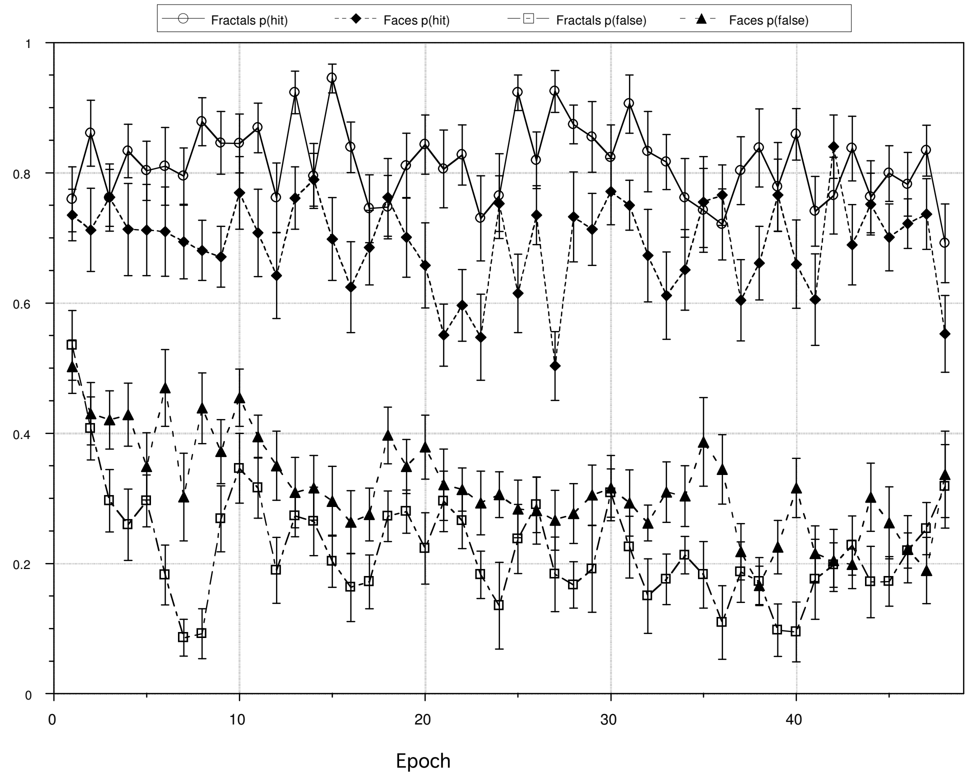

Figure 4.18: The effect of serial position on p(hit) and p(false) for recognition of fractals and faces. Epoch = 10 trials. Data = mean ± SEM.

4.3.3.3 Serial position

Data were divided into 10 trial epochs based on the serial position of trials within the experiment. For each epoch p(hit) and p(false) were calculated and compared with data from Experiment 6 (Figure 4.18). The effect of serial position on p(false) for fractal recognition was similar to that observed for faces, decreasing throughout the course of the experiment. Furthermore, p(hit) values were approximately constant throughout the course of the experiment for both types of stimuli.

4.3.4 Discussion

Importantly, the results demonstrate that the pattern of results observed for face-like stimuli in Experiment 6 is largely replicated for fractals. Fractals were recognised with greater ease than were the faces, perhaps indicating that the changes from stimulus to stimulus were easier to detect than were the changes between faces, but there were few indicators that recognition was more accurate when study and test stimuli were in the same location. The only evidence to suggest any effect of translation was the finding that the periphery (no shift) condition was associated with higher *d'* scores than was the centre-periphery shift for the fractal stimuli. This was, perhaps, a result of the fact that stimuli were shifting from a retinal location where acuity was high to a test location where it was lower, in the centre-periphery shift. Features encoded at high resolution may have been more difficult to recognise at low resolution in peripheral vision. Certainly, the centre-periphery shift condition was associated with the lowest *d'* scores and slowest reaction times for recognition in both Experiment 6 and the current experiment. The speed of recognition for fractal recognition also demonstrated translation invariance, displaying no central test location advantage, unlike the results for faces. Nonetheless, the overall pattern of results for fractal recognition were remarkably similar to those for face recognition, albeit with fewer significant differences between conditions.

These results provide support for the idea that the modified same/different task, in which there is no training at given retinal locations, is not sensitive to the location of stimuli. This is true whether the stimuli are naturalistic face-like images with a familiar configuration, or whether they are abstract fractal patterns. The translation invariance, therefore, does not appear to be limited to ‘familiar’ configurations (faces), as it is also found for novel abstract patterns (fractals).

This finding supports the finding of Dill and Edelman (2001) that animal-like stimuli are similarly unaffected by translation, unless the configuration of their parts is scrambled. However, it contrasts with the findings of Dill and Fahle (1998) in which same/different discrimination of dot patterns was shown to be linearly affected by shifts in position. This may have been due to the fact that the shifts they used were not controlled for eccentricity. Shifts could occur horizontally or vertically from a study position 1º from fixation. Therefore, a variety of eccentricities of stimulus were compared in the results. As is shown in the results of this chapter, shifts between locations of different eccentricities (and, hence, retinal acuity) are associated with poorer recognition. However, the differences between Dill and Fahle's results and those of the experiments in this chapter may also be the result of radically different stimuli. Dill and Fahle's dot patterns are considerably less visually rich than the stimuli used in these experiments, and their recognition relies on the accurate encoding of their spatial configuration, rather than on component features. The same is true of the scrambled animals of Dill and Edelman. On the other hand, both faces and fractals could be discriminated from one another on the basis of changes to their component features. The evidence seems to suggest that different processes are involved in the encoding and recognition of configural and featural information, and that the latter is more resistant to changes of location than the former.

An important question for future experiments is whether other changes to stimuli, barring the previously explored avenue of scrambling the configuration of object features, can cause a breakdown of invariance with respect to object location. For example, do additional translations, such as rotation, cause a deficit in the recognition of objects that have been shifted from one position to another? Indeed, are the mnemonic representations of objects that can be discriminated on the basis of their component features invariant with respect to rotation, in the same way that they are with respect to their location?

Chapter 5 The effects of rotational translation, in addition to positional translation, on same/different discrimination

5.1 Experiment 8: Rotation has no effect on translation invariance of same/different recognition

5.1.1 Introduction

The experiments in Chapter 4 demonstrated that same/different recognition of complex visual objects was relatively invariant with regards to positional translation. Whether this was the case because adult visual memory is capable of instantly recognising novel objects at novel locations in the visual field, or whether it was the result of learning during the course of the experiment was not determined. It is conceivable that the visual system readily achieves invariance with regard to translation along a single dimension, but that this might break down when the task becomes more difficult, e.g. with additional transformations of the stimulus. It would also be of interest to determine whether the visual system is invariant with regards to different types of transformation (e.g. object rotation) in their own right.

Recognition of familiar items (e.g. letters, digits) (Corballis, Zbrodoff, Shetzer, & Butler, 1978; Simion, Bagnara, Roncato, & Umilta, 1982) and items with pronounced diagnostic features (e.g. line drawings of objects) (Eley, 1982; Jolicoeur, 1985) rotated in the image plane are associated with relatively small error rate and reaction time costs. These costs are usually eliminated with practice (Jolicoeur, 1985). However, if the objects are novel (e.g. letter-like shapes) greater costs are incurred (Tarr & Pinker, 1989), and practice does not eliminate these effects. This finding was not replicated in a study of macaque monkey recognition, where an initial orientation dependency in the identification of novel objects gave way to an ability to generalise across rotations (Logothetis, Pauls, & Poggio, 1995). The effect of familiarity on invariance with regards to rotation appears to depend on the nature of the stimuli employed, and the nature of the task.

There is other evidence to suggest that the effects of rotation on recognition are different for different classes of objects. Cooper and Brooks (2004) found that reaction times for the recognition of animals, as a function of angle of rotation, formed an inverted-U shaped curve (i.e. that reaction times were slowest for recognition at 180º). The curve for recognition of objects at the basic-level (e.g. piano, phone, tricycle) was M-shaped, similar to the inverted-U but with faster reaction times at 180º. This finding is similar to that of a number of psychophysical studies of pattern recognition, which have shown that angles close to 90º reduce accuracy more than rotations of 180º (Dearborn, 1899; Foster, 1978; Rock, 1973). Cooper and Brooks also found a right hemisphere advantage for the recognition of animals but none for other objects.

The differences described by Cooper and Brooks have been ascribed to the involvement of different brain regions operating on different spatial relations. The right fusiform gyrus is a region involved in face recognition, and may also be involved in the recognition of other biological forms. The homologous region in the left hemisphere is suggested to be involved in more general, bilateral object recognition (Rossion et al., 2000). The system mediating face recognition codes coordinate spatial relations specified relative to a fixed location, whereas that mediating basic level object recognition codes categorical spatial relations (Cooper & Wojan, 2000). It could be the case that these differences are also responsible for the different effects of rotation on familiar and novel objects, especially when rotations of 180º are used.

There is considerable evidence that faces are recognised by different mechanisms to most other object classes, and the effects of rotation on the recognition of faces are quite different to its effects on recognition of other objects. One effect that seems to be unique to faces is the disproportionate disruption to recognition caused by inversion, or 180º rotation. Inversion of faces causes a disruption of the normal pattern of facial features, also known as facial syntax (Ellis, 1986). When recognition of upright and inverted objects from a range of classes was compared, upright faces were the easiest to recognise, whereas inverted faces were the most difficult (Yin, 1969). Diamond and Carey (1986) explained the effect with reference to expertise, as a similar inversion effect was found for dog experts' recognition of the faces of dogs. However, the finding of a right hemisphere advantage for recognition of unfamiliar faces (i.e. those seen only once previously) (Bruyer, 1986; Ellis, 1983; Rhodes, 1985), coupled with no lateralisation for inverted faces, is evidence that differential processing of faces occurs. Indeed, more recently Kanwisher et al. (1997) have found evidence for the involvement of the right fusiform gyrus in the recognition of faces.

The finding of an 'inversion effect' for faces begs the question of why this is the case. Facial identification relies predominantly on configural information, such as the spatial relations between different parts of the face, due to the unusual homogeneity of faces as a visual stimulus class. The configurational cues that aid identification of an individual face are disrupted more by inversion than isolated feature cues (Rhodes, 1993). This may explain why recognition of other objects is affected less by inversion than the recognition of faces, because their discrimination is more often carried out on the basis of featural information. Collishaw and Hole (2002) used blurring to disrupt the featural processing in the recognition of both upright and rotated faces. Whilst upright blurred faces were recognised well above chance, blurred inverted faces were not recognised above chance. The authors ascribed the difference to the ability to process configural information in upright and inverted faces, suggesting a linear relationship between the degree of rotation and the degree of disruption to configural cues. At 180º configural cues were assumed to have been totally disrupted hence the chance level of performance.

Further to the reliance on configural cues in the processing of faces, it has been suggested that upright faces are processed in a holistic, global manner, whereas the recognition of inverted faces relies on the recognition of isolated features. Robbins and McKone (2003) trained participants extensively in the discrimination of faces of identical twins, at both upright and inverted orientations, and from multiple views. They found that upright discrimination was supported by holistic processing whereas in the inverted orientation discrimination was dependent on the learning of local feature differences specific to certain images of the faces. Participants did not learn to holistically process the faces, suggesting that the advantage for upright faces is not related to expertise. This supports the theory of an innately driven component in face recognition (de Gelder & Rouw, 2000; Farah, Rabinowitz, Quinn, & Liu, 2000; Morton & Johnson, 1991) perhaps involving a critical period in infancy for the development of holistic processing (Le Grand, Mondloch, Maurer, & Brent, 2001).

Electrophysiological study of IT neurones in monkeys demonstrates that they are more sensitive to changes in nonaccidental properties (NAPs) than changes in metric properties (MPs) (Vogels, Biederman, Bar, & Lorincz, 2001). Whilst MPs, such as the degree of curvature of a contour, are view-dependent, NAPs, such as the linearity of a contour or the coterminal of pairs of contours, are relatively invariant over rotations in depth. The relative sensitivity of IT neurones to these properties, in spite of the fact that the image variation produced by changes in MPs is greater than that produced by changes in NAPs, suggest that IT is more sensitive to the transformation invariant properties of objects than view-dependent properties. Vogels and colleagues suggest that this may enable the immediate recognition of novel objects at new views. Whilst these results are more pertinent to the study of 3D rotations, rotation in the image plane might be expected to be even more dependent on NAPs as there is less view-dependent variation in the image. Visual areas dedicated to extracting NAPs from visual information would be essential for achieving invariance of recognition for objects rotated in 2D. Given the theory that IT is organised according to experience (C. A. Erickson, Jagadeesh, & Desimone, 2000), it can be predicted that invariance with regards to rotation may be dependent on how familiar objects and their components are.

In addition to the effects of rotation *per se*, one might expect this transformation to effect positional translation. A number of studies have combined the two transformations in the investigation of pattern matching. Kahn and Foster's (1981) same/different experiments with random dot patterns employed both rotations of 90º and 180º, as well as positional shifts of 0.5º and 1.0º of visual angle. The results indicated that rotation had a lesser effect on recognition accuracy than the distance between study and test locations. Whilst a significant effect of positional shift was found for objects rotated 90º (performance was lower for objects shifted between the left and right peripheral locations than other conditions) the effect of shift was not significant for point inversion (180º rotation). Dill and Fahle's (1998) very similar study also found an effect of translation for shifts of up to 2º. Recognition accuracy was inversely related to the distance of the shift, but again this effect was eradicated after 180º rotation of the pattern. The authors argued that this indicated that positional specificity occurred at relatively low levels of processing. When the 'higher' levels of cognitive processing needed to mentally rotate an object were required in the task, the displacement effect vanished. Patterns rotated as well as displaced showed no variation in accuracy or speed of recognition.

These results are similar to those of Larsen and Bundesen (1998) who studied same/different discrimination of simultaneously presented patterns, located on the periphery of a circle with a radius of approximately 3º, centred on fixation. They found no effect of spatial separation on *d'* scores in the absence of rotation, but when patterns forming a 'same' pair could differ from one another by both a positional translation and a rotation, a different effect emerged. When these task demands were in place, but the rotational component of a 'same' pair was 0º, *d'* was a decreasing function of the distance between the two patterns. The effect of spatial separation was less pronounced at rotations of 10º, 20º and 30º. The authors explained this result by proposing that 'same' judgements were made by the mental translation of one pattern to the location and orientation of the other, and testing for a match. When a rotation was required, reaction times were a function of the angle of rotation.

Having previously discovered that same/different discrimination of face-like stimuli differs from that of patterns when stimuli were translated between locations in the visual field, it was of interest to determine whether this effect would change when the requirement of mental rotation was added to the task. In the current experiment the positional translations were accompanied by rotations of some stimuli by 180º. It has been suggested by Dill and Edelman (2001) that translation invariance breaks down when the configuration of the stimuli is disrupted by scrambling. The current experiment aimed at determining whether this was the case when the configuration of face-like stimuli was disrupted by inversion.

5.1.2 Methods

5.1.2.1 Design

The experiment had a two-way within subjects design. The independent variables were the shift type, and the rotation type of the stimuli. Three positional shift types were studied (see Figure 5.1): periphery-periphery (a shift of 5.6°, between the right and left peripheral positions), centre-periphery (a shift of 2.8°, between the central position and either right or left position), and periphery-centre (a shift of 2.8°, from either left or right to centre). In addition, conditions in which the stimuli remained either at fixation (centre (no shift)) or in the same peripheral location (periphery (no shift)) were studied. Four rotation types were studied across two consecutive blocks (see Figure 5.2): upright-upright, inverted-inverted, normal-inverted and inverted-normal. Inverted stimuli differed from upright stimuli by a 180° rotation.

The dependent variables were the participants' *d'* scores, and the reaction times for correct recognition (hits).

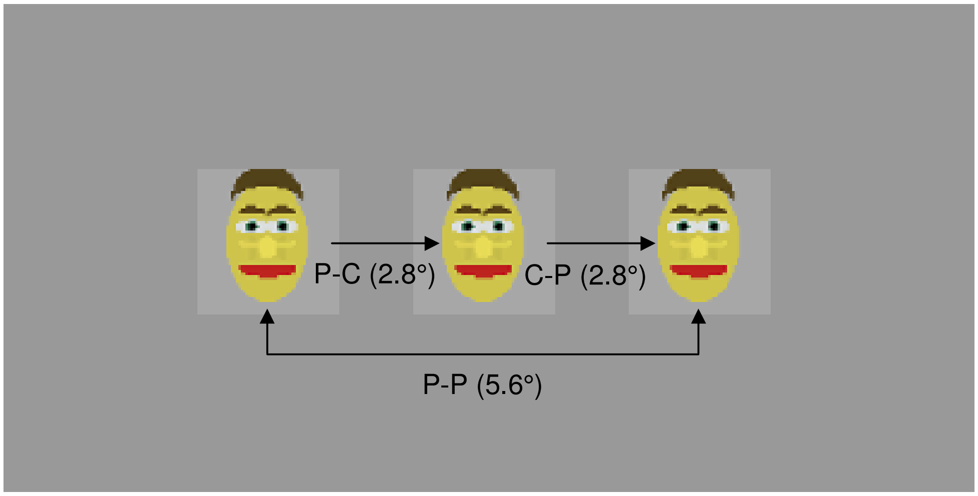

Figure 5.1: The three possible stimulus positions (left, centre, right) as well as the three shift types (P-P = periphery-periphery shift, C-P = centre-periphery shift, and P-C = periphery-centre shift). Actual stimuli were in full colour.

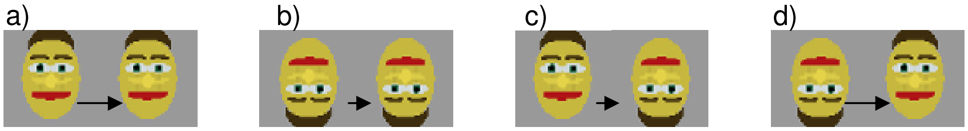

Figure 5.2: The four rotation types for a 'same' study-test pair: a) upright-upright, b) inverted-inverted, c) upright-inverted, and d) inverted-upright. Actual stimuli were in full colour.

5.1.2.2 Participants

The 24 participants (15 female, 9 male), were students of Nottingham University, taken from an opportunity sample. The mean age of participants was 24.9 (±0.8) years of age. All participants had normal or corrected-to-normal vision. None had previously been exposed to the stimuli.

5.1.2.3 Apparatus and materials

The apparatus and experimental set-up were the same as those described for Experiment 6.

5.1.2.4 Stimuli

The stimuli were generated in a similar manner to the series used in Experiment 6. Two sets of 253 faces were generated: one for each experimental block. When an inverted face was specified by the program, the stimulus' co-ordinates were rotated 180° before display. Participants were required to respond according to whether they thought each image was the same or different to the previous one, regardless of whether the image was in a different location or orientation to the previous image.

Within each series of 253 faces, 200 occurred twice forming shift pairs of study and test presentations. These 200 pairs were balanced for periphery-periphery, centre-periphery, periphery-centre, centre (no shift), and periphery (no shift) conditions with 40 of each type. Within each shift type, pairs were balanced for the location of the shift. So, for example, of the 40 horizontal shift pairs, 20 were from left to right and 20 from right to left. This avoided the confounding of results for a particular shift type with the effects of shift direction and shift locations.

In order to prevent the emergence of obvious patterns in presentation (i.e. 'different'-'same'-'different'-'same'), two further conditions were included. There were 50 'foils': faces that occurred once and did not recur (10 in each location), and 25 'repeats': faces identical to the preceding pair (5 in each location). These were balanced such that 20 of the foils and 10 of the repeats occurred on both the left and the right, and that 10 of the foils and 5 of the repeats were in the centre.

The order of each set of 253 faces was always the same, but the order of the conditions within the sequence was always determined randomly for each participant. Together with an initial unscored buffer of 3 trials, the experiment comprised 478 trials.

In block *A* trials were subdivided into two sections of 100 study-test pairs. Section *a* was composed entirely of upright stimuli and section *b* entirely of inverted stimuli. The order of these sections was counterbalanced across participants. The buffer trials were composed of stimuli of the orientation of whichever section occurred first. The reason for this division was to prevent the ease with which upright orientation study-test pairs could be discriminated from inverted pairs, were the two to be intermixed. In block *B* 100 study-test pairs were upright-inverted and 100 were inverted-upright. Foils were of the opposite orientation to the preceding trial and repeats were of the same orientation. In this block the two orientation types were intermixed, as the distinction between different pairs was harder to detect.

5.1.2.5 Procedure

The procedure was largely same as for Experiment 5, and was repeated so that each participant completed both blocks. The order of blocks *A* and *B* was counterbalanced. The only other difference in the procedure was that, in addition to stressing that items differing from a previous item in their spatial location but being identical in all other respects would be classified as old, the instructions for the current experiment specified that items differing by a rotation of 180º, but otherwise identical, should also be classified as ‘old’.

5.1.3 Results

All trials in which participants took longer than 3 sec to respond were rejected. From the remaining trials hit rates were calculated from test trial accuracy data for each shift type and rotation type. False alarm rates were calculated for each rotation type. In order to exclude data from any participants who performed exceptionally poorly, a one-tailed chi square was conducted on each participant's responses, to determine that the number of 'same' responses to old stimuli was significantly above that expected by chance. In one case there was no significant difference, and that participant's data were omitted from further analysis.

5.1.3.1 D-prime scores

The resulting data were converted into *d'* scores, to give a measure of discrimination (Figure 5.3).

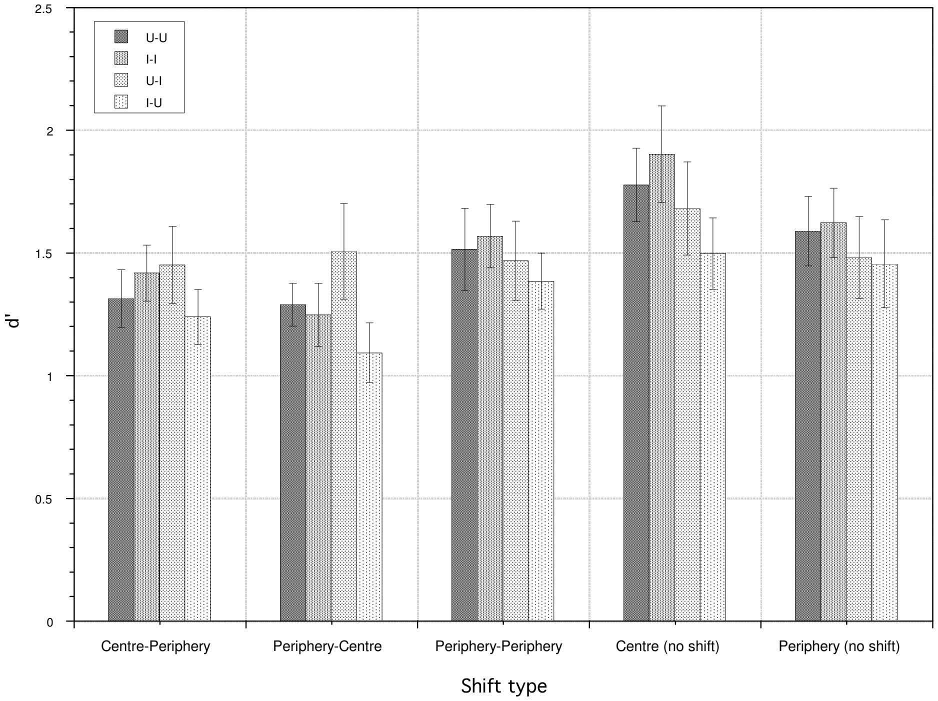

Figure 5.3: The effect of rotation type and shift type on d' scores. U-U = upright-upright, I-I = inverted-inverted, U-I = upright-inverted, I-U = inverted-upright. Data = mean ± SEM.

A two-way within subjects ANOVA with 4 (rotation type) x 5 (shift type) levels was performed on the data. Mauchley's test of sphericity was not significant for rotation type or shift type. There was no significant main effect of rotation type, but there was a significant main effect of shift type ($F(4,92)=9.77$, $MSe=0.275$, $p<0.001$). Tukey's post-hoc tests revealed significantly higher scores for the centre (no shift) condition than all shift types (at least $p<0.05$). In addition, periphery (no shift) was associated with higher scores than periphery-centre shift ($p<0.05$). The interaction between rotation type and shift type was not significant.

The manipulation of rotating stimuli had no effect on *d'* scores. There was evidence of a positional effect in the data (centre (no shift) advantage, periphery (no shift) recognition better than periphery-centre), but as in Experiment 7, it was limited to circumstances in which stimuli shifted between central vision and the visual periphery.

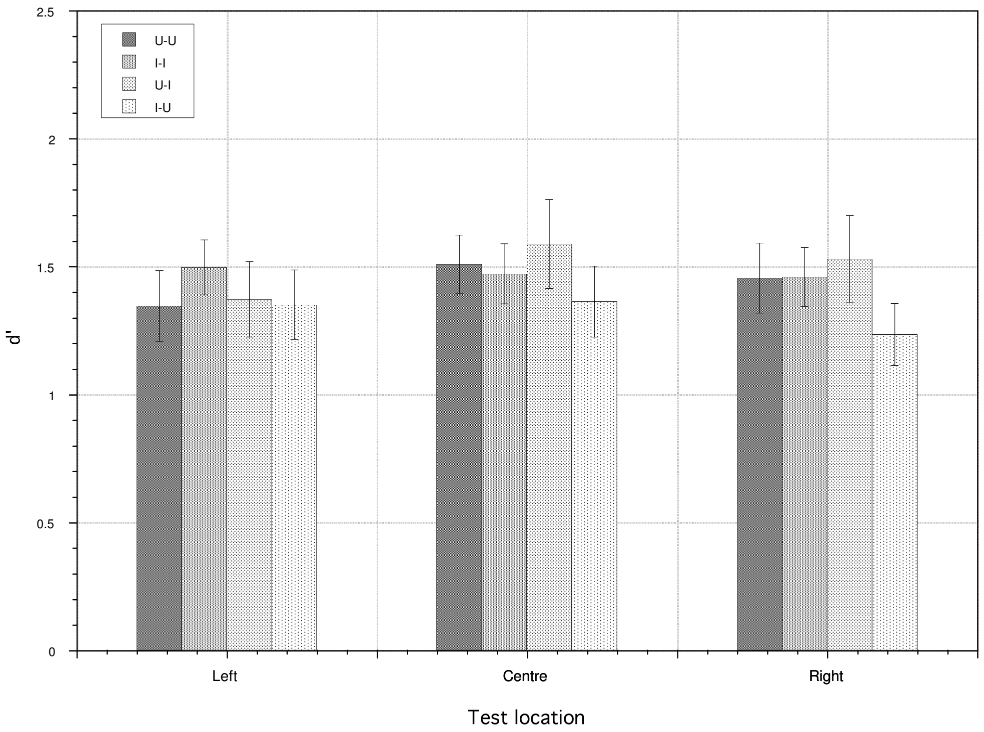

Figure 5.4: The effect of test position on d' scores. Data = mean ± SEM.

D-prime scores were also analysed by test position (Figure 5.4) in a 4 (rotation type) x 3 (test position) ANOVA. Mauchley's test of sphericity was not significant for any of the results. No significant effects of either rotation type or test position were found.

In order to determine whether the orientation of either the study or test stimuli were significant in affecting the results, *d'* scores were calculated for shift type for both of these factors (Figure 5.5 and Figure 5.6). The study orientation scores were entered into a 2 (study orientation) x 5 (shift type) repeated measures ANOVA. There was no significant main effect of either study orientation or shift type, and the interaction between the two factors was not significant.

Test orientation scores were entered into a 2 (test orientation) x 5 (shift type) ANOVA. Again, no significant effect of orientation on *d'* scores was found, and neither the main effect of shift type nor the interaction was significant.

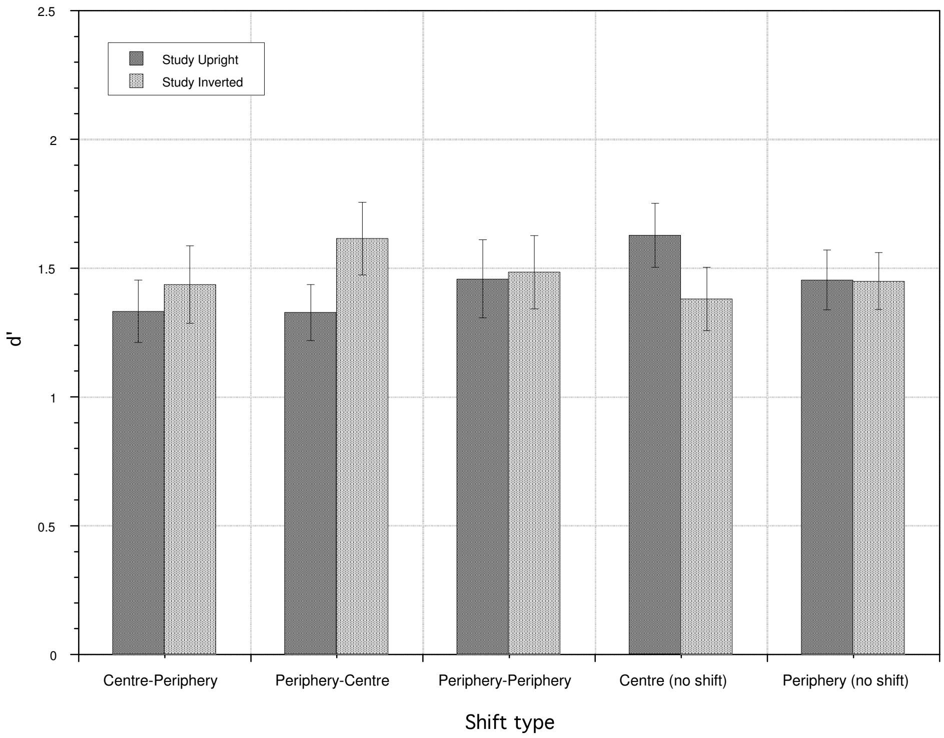

Figure 5.5: The effect of study orientation on d' scores. Data = mean ± SEM.

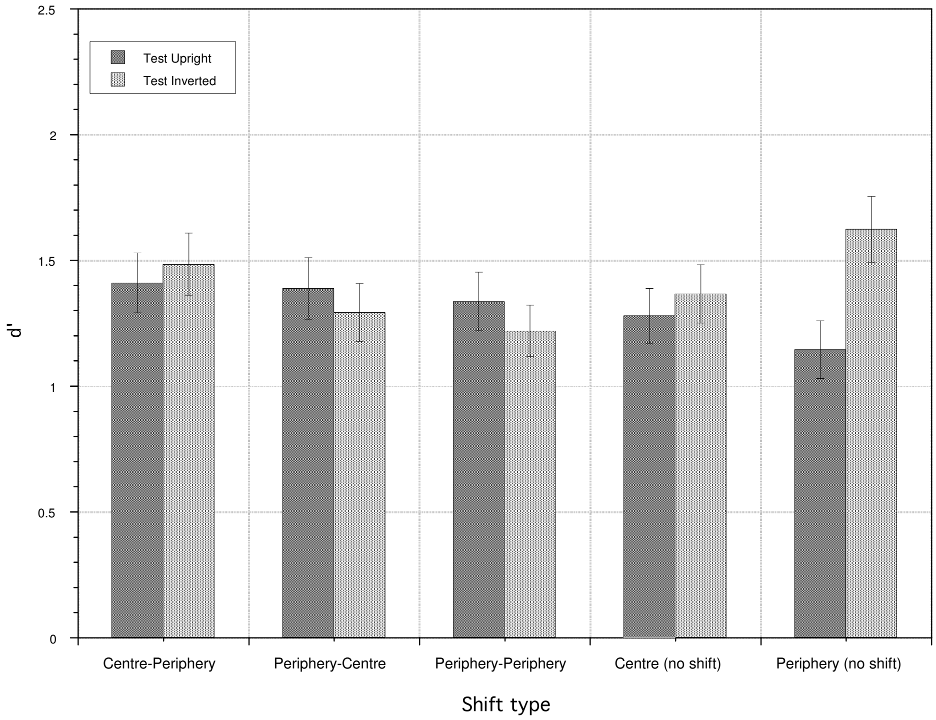

Figure 5.6: The effect of test orientation on d' scores. Data = mean ± SEM.

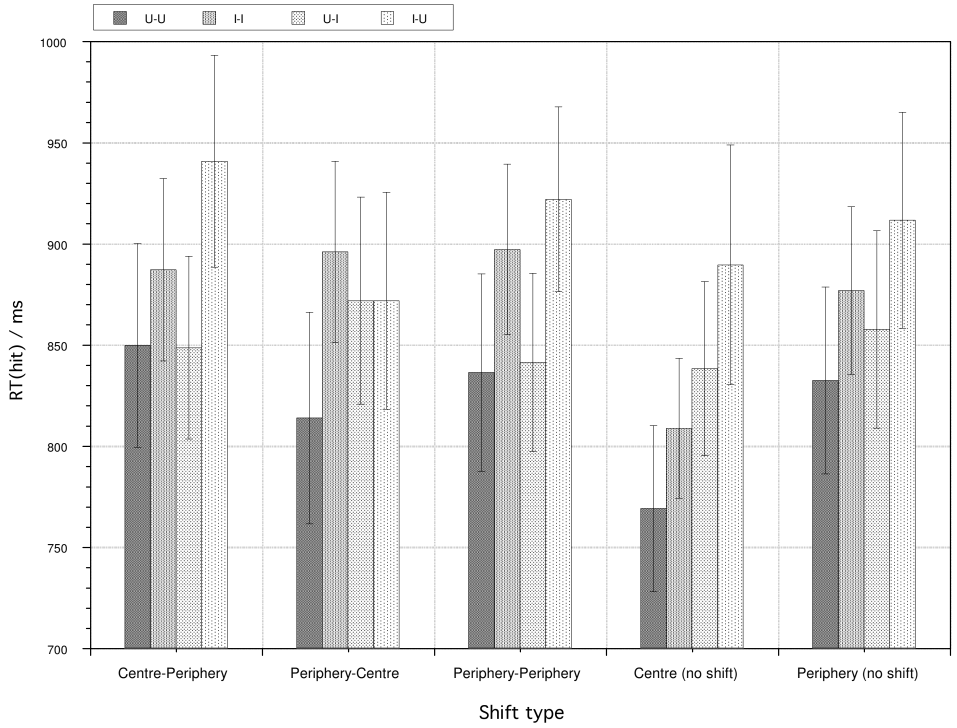

Figure 5.7: The effect of rotation type and shift type on reaction times for hits (ms). Data = mean ± SEM.

5.1.3.2 Reaction times for correct recognition

A similar two-way repeated measures ANOVA to that previously described for *d'* scores was carried out on reaction time data. Mauchley's test of sphericity was significant for both rotation type and the rotation type x shift type interaction, although neither of these were significant. There was a significant main effect of shift type (F(4,92)=4.25, MSe=10500, p<0.01). Tukey's post-hoc tests revealed significantly faster reaction times for the centre (no shift) condition than the centre-periphery shift (p<0.01), periphery-periphery shift, and periphery (no shift) (p<0.05).

Rotation had no effect on the speed of recognition. There was a centre (no shift) advantage for the speed of recognition, but no other differences were observed.

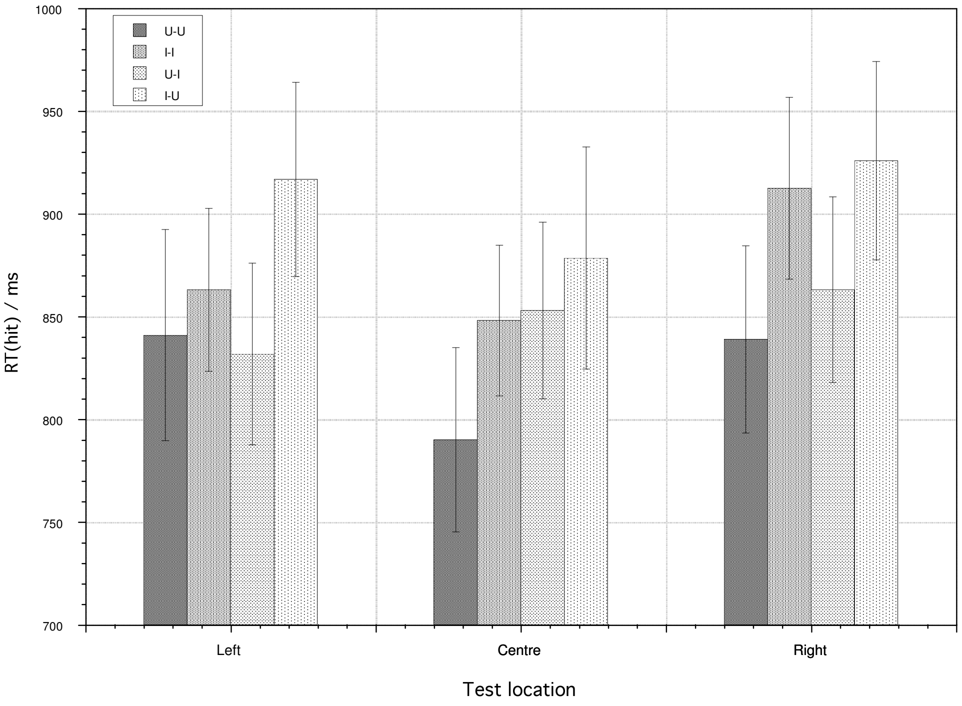

Figure 5.8: The effect of test location on reaction times for hits (ms). Data = mean ± SEM.

Reaction time data were also analysed according to test location with a similar ANOVA to that previously described for the analysis of *d'* scores. Mauchley's test of sphericity was significant for rotation type. Rotation type was not significant, but there was a significant effect of test position (F(2,46)=6.55, MSe=6670, p<0.01). Post-hoc tests revealed significantly faster recognition of items tested at the central location in comparison to on the right (p<0.01). The interaction between rotation type and test position was not significant.

Reaction time data were also calculated for study orientation (Figure 5.9) and test orientation (Figure 5.10). The data for study orientation were entered into a 2 (study orientation) x 5 (shift type) ANOVA. Mauchley's test of sphericity was not significant for any of these data. There was a significant main effect of study orientation (F(1,23)=4.25, MSe=21600, p<0.01): items studied in the upright orientation were recognised significantly faster than those studied in an inverted orientation. The main effect of shift type was not significant, and there was no significant interaction between study orientation and shift type.

The test orientation results were analysed in a 2 (test orientation) x 5 (shift type) ANOVA which revealed no significant main effects, and no interaction between the two factors.

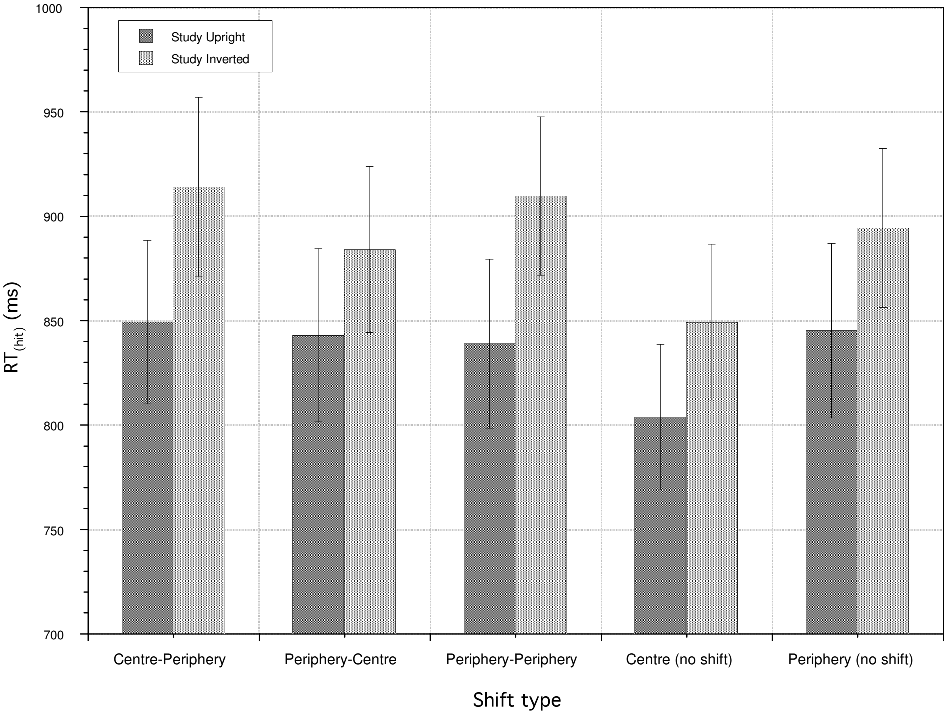

Figure 5.9: The effect of study orientation on reaction times for correct recognition. Data = mean ± SEM.

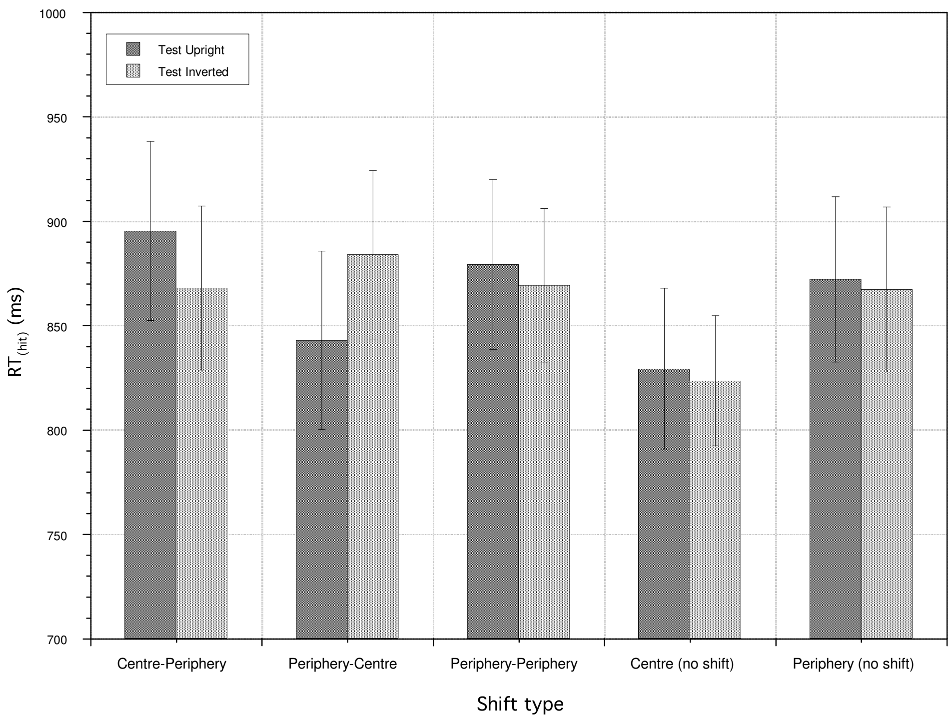

Figure 5.10: The effect of test orientation on reaction times for correct recognition. Data = mean ± SEM.

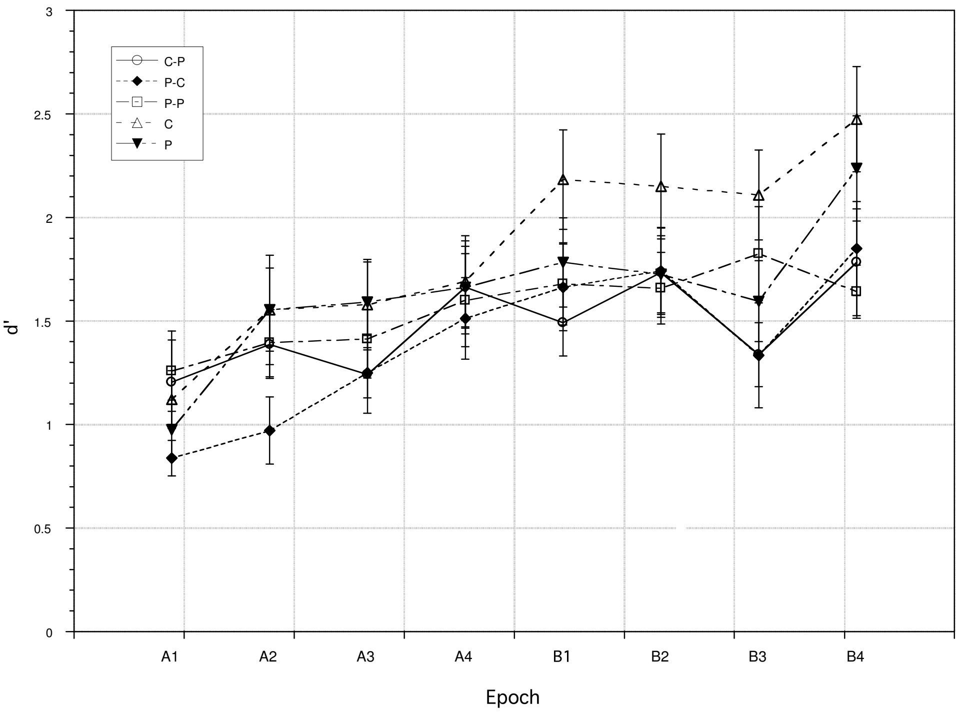

Figure 5.11: The effect of learning throughout the course of the experiment for each shift type. Epoch = 121 trials. A = first block, B = second block. C-P = centre-periphery shift, P-C = periphery-centre shift, P-P = periphery-periphery shift, C = centre (no shift), P = periphery (no shift). Data = mean ± SEM.

5.1.3.3 Serial position

In order to assess the effect of learning throughout the experiment, trials from both experimental blocks were divided into four parts. The three buffer trials from each block were discarded and then the remaining 484 trials were divided in four such that an epoch consisted of data from 121 trials. Blocks were then arranged in the order that participants carried them out, such that data from 8 consecutive epochs could be compared. For each epoch p(hit) and p(false) were used to calculate *d'* scores for each shift type (Figure 5.11). These scores were then analysed in a 5 (shift type) x 8 (epoch) repeated-measures ANOVA. Mauchley's test of sphericity was significant for both shift type and epoch, and the results for these factors are Greenhouse-Geisser epsilon corrected. There was a significant main effect of shift type (F(2.82,64,8)=7.63, MSe=1.12, p<0.001) and of epoch (F(3.31,76.2)=4.51, MSe=4.67 p<0.01). Tukey post-hoc tests revealed significantly lower scores at epoch 1 compared to those at epochs 5, 6 and 8 (all p<0.05), and lower scores at epoch 2 than at epoch 8 (p<0.05). There was no significant interaction between shift type and epoch.

What these results reveal is a significant learning effect resulting in improvement of performance over the course of the experiment. The lack of interaction between shift type and epoch, however, suggests that the learning is not dependent on positional shift, but occurs across all trials.

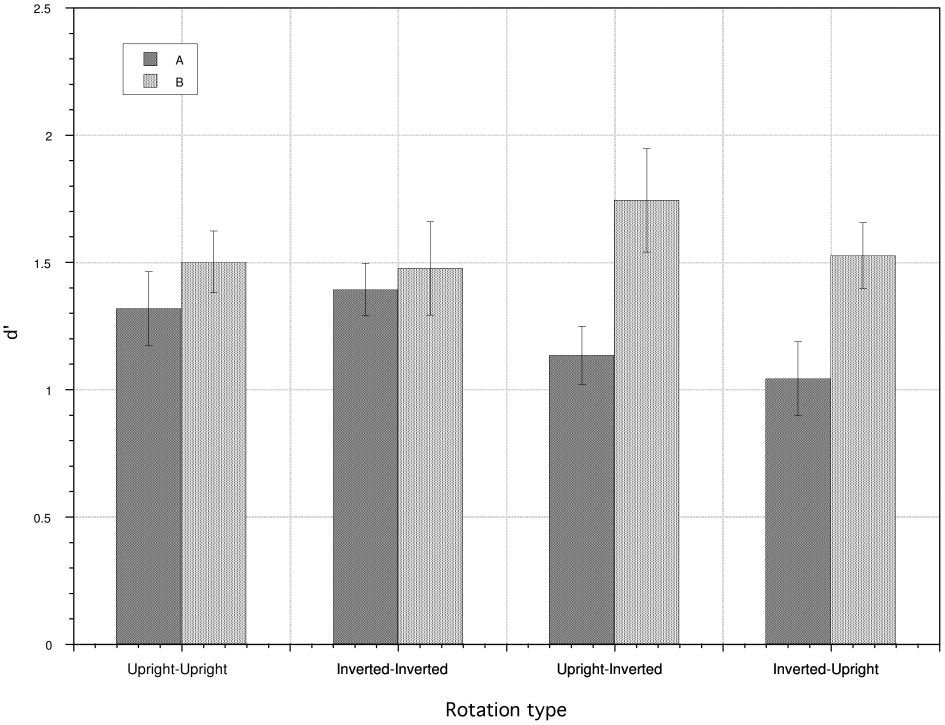

Figure 5.12: The effect of learning throughout the course of the experiment for each rotation type. A = first block, B = second block. Data = mean ± SEM.

In addition to analysing the effect of learning by shift type, the data were also analysed by rotation type. Because only two rotation types were present in each experimental block (upright-upright and inverted-inverted in the separated block, and upright-inverted and inverted-upright in the mixed block) it seemed that comparing the performance of those participants who carried out that block first (A) should be compared with that of those who carried it out second (B), rather than attempting to break the results down further by epoch. D-prime scores for each rotation type and each block were calculated and organised in order that this comparison could be made (Figure 5.12). The data were subsequently analysed in two ANOVAs. The first compared the performance of participants who performed the separated block first with those who participated in it second. The ANOVA was a 2 (order) x 2 (rotation type) mixed measures design. Neither factor was significant and there was no interaction between the two. The second ANOVA was similar but compared performance on the mixed block by the order of the experiment. Again, there was no effect of rotation type, but there was a significant main effect of order ($F(1,22)=7.55$, $MSe=0.473$, $p<0.05$). Participants who performed the mixed block as the second experimental block performed better than those who carried it out as the first block.

Again, these results demonstrate an effect of learning and, interestingly, also suggest a difference between conditions. It seems that participants' prior experience of the separated block improved their performance on the mixed block, although the opposite effect did not occur. This can be tentatively assumed to suggest that familiarity with the stimuli facilitated participants' subsequent ability to recognise rotated stimuli, but that performance for unrotated stimuli remained constant.

5.1.4 Discussion

Rotation of the stimuli had no effect on recognition. Rotation type had no effect on either the accuracy or the speed of recognition, although there was evidence that upright studied items were recognised faster than those studied in the inverted orientation. The disruption of the configuration of face-like stimuli by 180° did not appear to have any major effect on the participants' ability to recognise the stimuli, providing further evidence that these stimuli may not be recognised in the same manner as real faces. No 'inversion effect' was observed. The finding of slower reaction times associated with recognition of stimuli initially viewed in the inverted orientation, suggests that representations of inverted configurations were more difficult to retrieve than those for upright configurations, even though their accuracy was unaffected by orientation. Studies using items both studied and tested in the upright orientation or studied and tested in the inverted orientation, have previously found evidence for an orientation effect (e.g. Yin, 1969), but the current experiment extends that finding further to suggest that it is the study orientation that is critical.

Previous studies also indicate a right hemisphere advantage for recognition of unfamiliar faces (Bruyer, 1986; Ellis, 1983; Rhodes, 1985), for which there is some evidence in the present study. Face-like stimuli tested in the right visual field (processed in the left hemisphere) were recognised significantly more slowly than those tested at the central location, whereas those tested in the left visual field (right hemisphere) were not. Previous comparisons of upright and inverted faces, however, have found that this lateralisation breaks down when the face is inverted (Rapaczynski & Ehrlichman, 1979; Yin, 1970; Young & Bion, 1980), whereas the current results show a left visual-field (right hemisphere) advantage for stimuli irrespective of the orientation. Indeed, the involvement of the right fusiform gyrus in recognition of these stimuli seems unlikely, given the relative ineffectiveness of rotation at impairing recognition of the face-like images. This pattern is more suggestive of recognition on the basis of categorical spatial relations, as is found in basic level object recognition, which is impaired more by rotation of 90º than 180º (e.g. Foster, 1978). This would suggest that the face-like stimuli are recognised in a manner more similar to that of patterns than that of real faces.

The effect of positional translation was similar to that seen in Experiments 6 and 7. There was a recognition advantage in terms of both speed and discriminative ability for items that were studied and tested at fixation, although in a departure from previous experiments, the central test location was not associated with more sensitive discrimination than left and right locations. This was likely because of the inclusion of a periphery-centre shift in the current experiment, associated with much lower *d'* scores for items tested in the central location. Previously only centre (no shift) trials were tested in the central location, and this condition was associated with high *d'* scores. The results also showed higher *d'* scores associated with recognition of periphery (no shift) trials than with the periphery-centre shift – the condition in which performance was poorest overall. The inclusion of this positional shift demonstrated that study in the periphery and test at the central location was as effective as, if not better than, the opposite pattern of locations in disrupting recognition. The shift of objects' positions between retinal locations of different visual acuity seemed to be the cause of the positional effect, as changes between peripheral locations of similar acuity (periphery-periphery shift) did not have the same deleterious effect on recognition. As in Experiments 6 and 7 there were effects of translation, although these can be more readily explained with reference to retinal acuity changes than as a function of spatial separation *per se*.

Interestingly, given the previous results in studies combining positional translation and rotation (Dill & Fahle, 1998; Kahn & Foster, 1981; Larsen &

Bundesen, 1998), rotation of stimuli had no effect on the effect of positional translation. In previous studies, the differences of recognition performance associated with retinal translation in the upright orientation were eradicated by 180° rotation. In the current experiment, a comparison of the *d'* scores for both the upright-upright and upright-inverted conditions does show some evidence of this. Figure 5.3 shows considerable differences between the mean scores for different shift types in the upright-upright condition, whereas the results for upright-inverted are remarkably similar across all shift conditions. However, this interaction was not statistically significant, so it is necessary to be cautious in making any inferences based on it. Indeed, there were no significant interactions between shift type and rotation type or orientation in any of the statistical analyses, suggesting that the two transformations are processed independently.

As the current experiment was longer than previous experiments in this thesis examining translation alone, it was possible to divide trials into large epochs, for which it was possible to obtain *d'* scores for each positional shift type, and examine learning of each independently. From these results it was possible to determine a clear effect of learning over the course of the experiment, resulting in better performance on the second experimental block than in the first. The effect was not related to shift type, indicating a general increase in discriminative ability affecting all locations equally. However, an examination of changes based on rotation type revealed a differential effect of experience on discrimination of rotated and unrotated stimuli. Performing the experimental block containing unrotated stimuli before that containing rotated stimuli significantly improved performance on the latter, but the same was not true in the opposite order of testing. Whilst the mean scores for performance of the unrotated block were slightly greater if it was carried out second, this was not significant. These results suggest that discrimination of unrotated stimuli was not affected by the additional familiarity with the stimuli gained by carrying out a previous block of trials.

However, prior experience with the block of unrotated stimuli did aid subsequent discrimination of rotated ‘same’ stimulus pairs. This suggests that transformations of rotation and positional shift are processed separately, as learning appeared to be unaffected by shift, or lack thereof, whereas the effect of learning seems to differ according to whether or not stimuli were rotated. Due to the low number of participants in each group for these analyses (N=12), however, the reliability of this finding needs to be confirmed in future work.

The current experiment provides further evidence that the face-like stimuli are, paradoxically, not that much like faces, in terms of participants’ recognition performance with these stimuli. There was little effect of 180º rotation on recognition of these stimuli. Even though stimuli studied in the upright orientation were marginally better recognised than those studied in the inverted orientation, those studied and tested in the inverted orientation were associated with higher scores than those studied and tested in the upright orientation, in marked contrast to the pattern with photographs of faces (Moscovitch, Winocur, & Behrmann, 1997; Rhodes, Brake, & Atkinson, 1993; Valentine, 1988; Yin, 1969). Indeed, the results suggest that these stimuli are recognised in a manner more similar to that of basic-level objects and patterns. As such, inversion of the stimuli did not greatly impair participants’ perception of their spatial configuration, and consequently did not breakdown translation invariance in the manner described by Dill and Edelman (2001) for configurally ‘scrambled’ objects.

The addition of a periphery-centre shift condition added to the existing information from Experiments 6 and 7 about the nature of positional effects observed in recognition. The eradication of the central test location advantage when stimuli have been studied in the periphery makes it clear that better spatial acuity at the central location cannot be the sole cause of the differences observed between shift conditions. Rather, this finding makes it clear that detrimental effects to recognition are caused by shifts between retinal regions with different acuities. Further experiments employing spatial blurring of different levels, to control for the greater acuity of central vision, must be carried out in order to determine whether this hypothesis is tenable. If recognition of stimuli shifted between areas of different eccentricity is impaired by different spatial acuities at those eccentricities, one might expect blurring at areas of greater acuity to reduce this deficit.

Chapter 6 The effect of visual field position on continuous recognition of complex visual stimuli

6.1 Experiment 9: Intervening stimuli induce an effect of translation on recognition

6.1.1 Introduction

The preceding chapters have examined the effects of both temporal separation (in the case of continuous recognition experiments) and spatial separation (in the case of positional translation experiments) between study and test stimuli, in isolation from one another. The current experiment combined variation in both dimensions.

The results of previous continuous recognition experiments (e.g. Rubin, Hinton, & Wenzel, 1999; see also Chapters 1 and 2; Shepard & Teghtsoonian, 1961) have demonstrated that the presence of other items between study and test presentations is detrimental to recognition, and that this detriment increases as a function of the number of items that intervene. In particular, there is a marked difference between a lag of 0, when there are no items between study and test, and a lag of just 1 intervening item. This effect appears to be independent of the type of stimuli employed, e.g. 3-digit numbers (Estes & Maddox, 1995b; Shepard & Teghtsoonian, 1961), words (Hintzman, 1969; Hockley, 1982; Okada, 1971; Reder et al., 2000), photographs (Nickerson, 1965), pictures of common objects (current thesis, Experiment 1), or abstract pictures (Doty & Savakis, 1997; current thesis, Experiment 1). Experiments 3 and 4 demonstrated that the effect is primarily one of retroactive interference, as it was found to be unaffected by study-test interval, but was related to the level of interstimulus similarity.

The effects of spatial shifts between study and test presentation are less clear, especially considering certain methodological deficiencies of past experiments (i.e. the lack of fixation verification with eye-tracking equipment), but may differ according to the type of information encoded. For example, Dill and Edelman (2001) suggest that recognition of objects is normally invariant with respect to translation in the visual field, but that positional changes are detrimental when configural cues for recognition are disrupted. The former finding is consistent with the findings of Chapters 4 and 5, which did not show any effect of translation *per se* on same/different discrimination of face-like stimuli and fractals, although disruption of the configural cues of face-like stimuli by their inversion did not breakdown translation invariance. The pattern of results for the recognition of objects is not replicated in studies of pattern recognition, which is sensitive to positional changes in the absence of the disruption of configural cues (Dill & Fahle, 1998; Kahn & Foster, 1981; Nazir & O'Regan, 1990). This may be because when recognising patterns one cannot make use of the translation invariant configural information that aids the recognition of natural objects, but instead one must rely on other (perhaps feature-based) information that is sensitive to position. However, there is also the possibility that discrimination of very similar abstract patterns is more difficult to achieve than that of objects, and that positional translation exerts greater cognitive costs as a result. Indeed, whilst there were no significant effects of translation *per se* in Experiments 6, 7 and 8, the mean *d'* values for unshifted peripheral stimuli were greater than those for peripherally-shifted stimuli (horizontal, vertical, or periphery-periphery shift conditions). It is possible that experiments of these sizes (N=19-24) had insufficient statistical power to reveal small but consistent differences between performance in these conditions.

The effects of interference from intervening trials on recognition of translated stimuli have not been systematically examined. Some experiments have tested recognition of translated stimuli after intervening stimuli, despite that not being the aim of those experiments. For example, Biederman and Cooper (1991) carried out an experiment in which participants named 48 objects in a ‘priming’ phase, and then the same objects again in a ‘primed’ phase, with an average 7 min between presentations. Objects could occur to the left or right of fixation, and half of the primed items were presented in the same location, and half in a different location, to the initial priming position. Biederman and Cooper did not find significant effects of translation on priming effects. However, the measure they used, priming of the latency of object naming, is quite different to recognition. Recognition requires more than simply giving a basic-level name to an object, including the accurate identification of the object as one that has been seen previously.

Nazir and O’Regan’s (1990) study of pattern recognition included presentation, learning, and test phases. In the presentation phase, the participants passively viewed a target stimulus at a peripheral location. The learning phase required the participants to discriminate the target from distracters at the same location in blocks of 90, until a criterion of 95% accuracy was reached. There then followed a 3-block test phase, in which participants were required to discriminate target stimuli from distracters at the trained location (to the left or right of fixation), at the opposite location, and at fixation. The error rate was significantly higher for both non-trained locations in comparison with that for the trained location. In this experiment, however, whilst there were stimuli intervening between the study and test presentations of stimuli, the fact that there were multiple repetitions of the same stimuli served to reinforce learning of the discrimination, not to interfere with it.

Whilst there is little evidence in the literature to guide theories about what effect interference by intervening stimuli might have on translated stimuli, there is evidence concerning the effects of increasing interference by increasing interstimulus similarity. Dill and Fahle (1998) specifically manipulated the ‘difficulty’ of their same/different task by introducing dot cloud stimuli that differed from same trials only in the position of 1 of 6 dots. The dot in question was always displaced horizontally by 80% of the stimulus size. This condition was compared with a ‘horizontal’ condition, in which stimuli varied randomly. The effect of the manipulation was to greatly increase the similarity between stimuli and, thus, the level of interference between such stimuli. The ‘difficult’ condition was associated with a larger detrimental effect of positional translation on recognition accuracy ($p<0.001$) than the ‘horizontal’ condition ($p<0.05$). In another experiment (Dill & Fahle, 1998, Experiment 6), the authors compared randomly varying but horizontally symmetrical checkerboard patterns with those that were different only by one square. In this experiment there was a significant difference between *d’* scores for the more difficult than the easier condition, but positional translation had a similar effect on both. A similar effect was found in Dill and Edelman’s (2001) Experiment 2, in which they manipulated the interstimulus similarity of animal-like stimuli. Again, there was an overall effect of increased similarity on *d’* scores, but no interaction with positional translation. The effect of increasing similarity and interference between stimuli seems to differ according to the stimuli employed.

In the absence of any previous studies, the current experiment was an attempt to systematically study the effects of ‘lag’ on the recognition of items subject to positional translation. A continuous recognition design was used to examine the effects of a number of different ‘lags’ (numbers of intervening stimuli) in combination with the effects of the positional translation of items. Fractal stimuli were used in the current experiment, as previous experiments had determined the effects of interference and similarity (Experiments 3 and 4), as well as those of positional translation (Experiment 7), on their recognition. The position of the stimuli intervening between study and test was also manipulated, to determine whether their location with reference to study and test affected the level of interference observed. It was hypothesised that, with the increased cognitive demands required of participants during continuous recognition, stimulus position would have a greater effect on recognition. As such, it was predicted that longer lags would be associated with a breakdown of translation invariance.

6.1.2 Methods

6.1.2.1 Design

The experiment had a three-way within subjects design. The independent variables were lag, shift type, and the position of intervening stimuli. Four lags were used: 0, 1, 4, and 8 intervening stimuli. Stimuli were either shifted (5.6º between left and right visual field locations, see Figure 6.1) or not shifted. Stimuli intervening between study and test (at lags greater than 0) were either all in the 'same' position as the study trial, all in the 'different' position to the study trial, or (for lags greater than 1) 'half' were in the same location and half were in the different location (Figure 6.2).

The dependent variables were the participants' *d'* scores, and the reaction times for correct recognition (hits).

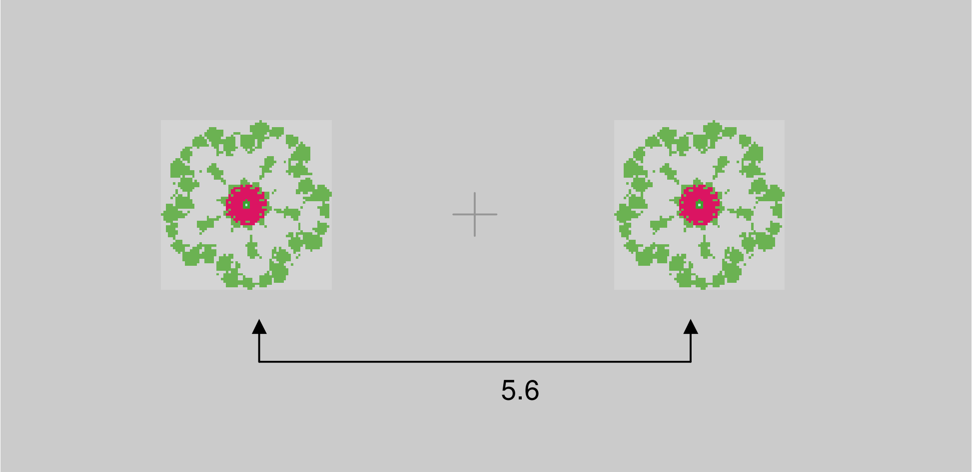

Figure 6.1: The two possible stimulus locations (left and right) and the size of the shift between the locations (5.6º). The cross denotes the point for fixation. Actual stimuli were in full colour.

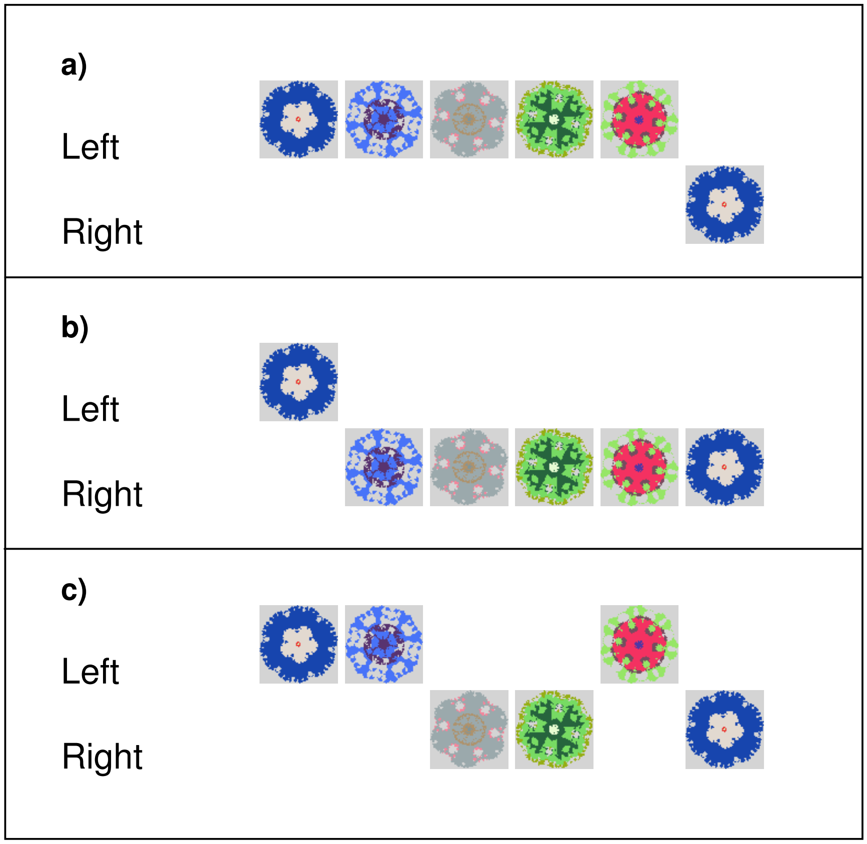

Figure 6.2: Examples of the different intervening items conditions. Left and Right denote the left and right visual field positions. The example given is a sequence of trials forming a shifted study-test pair, with a lag of 4. The locations of intervening stimuli are given for a) the ‘same’ condition b) the the ‘different’ condition and c) an example of locations for the ‘half’ condition (half of intervening trials on one side and half on the other). Actual stimuli were in full colour.

6.1.2.2 Participants

The 20 participants (13 female, 7 male), were students of Nottingham University, taken from an opportunity sample. The mean age of participants was 19.7 (±0.5) years of age. All participants had normal or corrected-to-normal vision. None had previously been exposed to the stimuli.

6.1.2.3 Apparatus and materials

The apparatus and experimental set-up were the same as those described for Experiment 6. Fixation was measured throughout the experiment with an eye-tracker.

6.1.2.4 Stimuli

The stimuli were a series of 340 fractals generated using the ‘medium’ similarity rules described in the Methods section for Experiment 4.

6.1.2.5 Session design

A pseudorandomly determined frame of 650 trials was generated for each participant, providing 10 study and 10 test trials for each lag, shift, and position of intervening stimuli combination (e.g. there were 10 study-test pairs for lag 8, no shift, intervening stimuli in the ‘same’ position). There were 60 study-test pairs each for lag 8 and lag 4 (shifted and no shift x same, half and different intervening stimuli), 40 for lag 1 (shifted and no shift x same and different intervening stimuli), and 20 for lag 0 combinations (shifted and no shift). Each subset of 10 study-test pairs consisted of 5 pairs with the study location in the left position and 5 pairs with the study location in the right position. Study-test pairs comprised 360 of the trials. The remaining spaces in the 650 trial frame were composed of 145 study-test pairs of unscored filler trials. In addition, the experiment began with a buffer of 30 unscored filler trial pairs to prevent the occurrence of primacy effects, yielding a total of 680 trials. Once the order of trials for the entirety of a session had been generated, the 340 stimuli were randomly assigned to the 340 pairs of trials, so that each participant experienced the stimuli in a different order.

Instructions were presented to participants in written form. The instructions informed them that they would see a series of fractals, and that for each they were required to make a response indicating whether they thought it was completely novel (‘new’) or had been seen previously (‘old’). They were informed that items differing from one another in spatial location but identical in all other aspects should be classified as ‘old’. Participants were instructed to fixate on a central, mid-grey cross which was present throughout the experiment.

Stimuli were displayed in the appropriate position for 100 ms. Following the presentation, participants were required to make a response with the click of a mouse button. If they thought the stimulus was ‘new’, they were to click the left button. If they thought it was ‘old’ they were to press the right. The next trial would not start until both a response had been given, and fixation on the central cross was detected, allowing participants to pace the experiment according to their own ability. There was a brief interval between the stimulus presentation and the next stimulus during which a low feedback tone was played if the previous response was incorrect. This provided some motivation for participants to maintain attention throughout the duration of the experiment.

6.1.3 Results

All trials in which participants took longer than 3 sec to respond, or those in which it took longer than 5 sec to achieve fixation on the central location, were discarded. From the remaining trials hit rates were calculated from test trial accuracy data for each shift type. A single false alarm rate was calculated for the entire experiment.

In order to exclude data from participants who performed particularly poorly, a one-tailed chi square was conducted on each participant's responses, to determine whether the number of 'old' responses to repeated stimuli was significantly above that expected by chance. This was the case for all of the participants, and consequently all data were included in subsequent analyses.

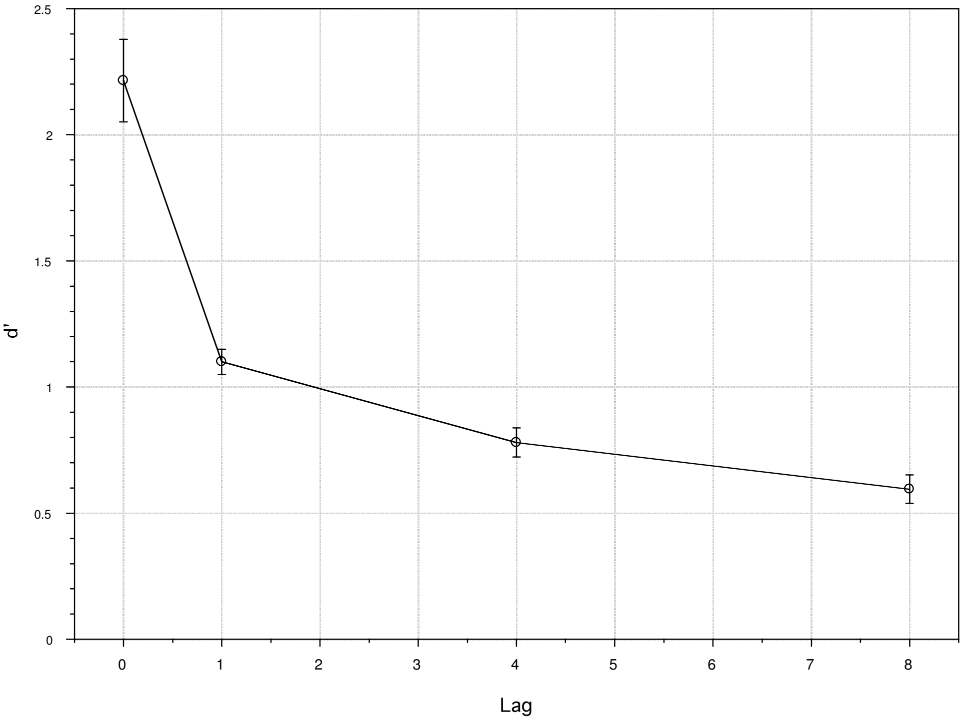

Figure 6.3: The effect of lag on *d'* scores. Data = mean ± SEM.

6.1.3.1 D-prime scores

Hit and false alarm rates were used to calculate *d'* scores at each of the four lags used in the experiment (Figure 6.3). There was a clear decline in *d'* with increasing lag, similar to that found in Chapters 2 and 3.

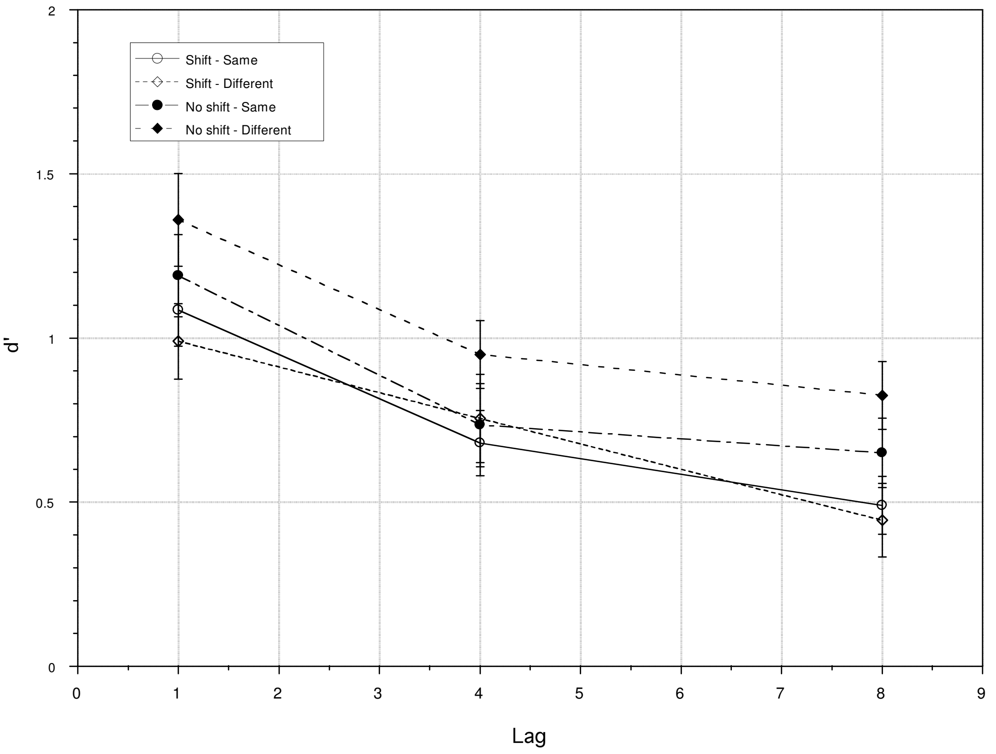

Figure 6.4: The effect of shift and same vs. different position of intervening stimuli, at lags 1-8. Data = mean ± SEM.

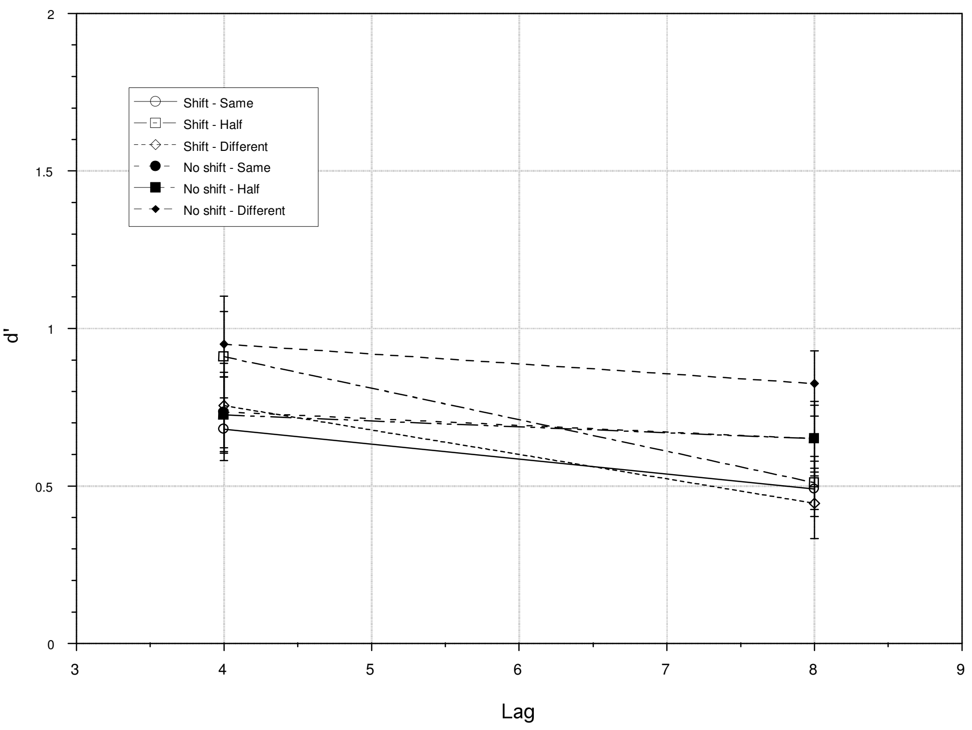

Figure 6.5: The effect of shift type and intervening stimuli position on d' scores, at lags 4 and 8. Data = mean ± SEM.

In order to determine the effects of the three manipulated variables, *d'* scores were calculated by lag, shift type, and position of intervening stimuli, using hit rates for each condition and false alarm rates calculated according to shift type. From these scores two comparisons were made: the effects of lag, shift type, and same vs. different position of intervening stimuli for lags 1-8 (Figure 6.4), and the effect of lag, shift type, and same vs. half vs. different position of intervening stimuli for lags 4 and 8 (Figure 6.5).

The data from the first comparison were entered into a 3 (lag) x 2 (shift type) x 2 (intervening stimuli position) repeated measures ANOVA. Mauchley's test of sphericity was significant for lag and the interaction between lag and shift type, and these results are quoted with Greenhouse-Geisser epsilon corrected degrees of freedom. There was a significant main effect of lag (F(1.55,29.5)=30.3, MSe=0.272, p<0.001). Post-hoc analyses revealed significantly better scores for lag 1 than lags 4 and 8 (p<0.001), and better scores for lag 4 than lag 8 (p<0.05). There was also a significant main effect of shift type (F(1,19)=8.62, MSe=0.309, p<0.01). Stimuli that were not shifted between study and test presentations were recognised significantly better than those that were shifted (i.e. changed location in the visual field). The main effect of intervening stimuli position was not significant and neither were any of the interactions between the factors.

The effects of lag are similar to those encountered in previous experiments (Experiments 1, 2, 3 and 4), but, interestingly, there was an effect of translation between regions of the same retinal acuity. This was different to the results found for previous positional translation experiments in which no stimuli intervened between study and test (Experiments 6, 7 and 8).

The second comparison was analysed with a 2 (lag) x 2 (shift type) x 3 (intervening stimuli position) repeated measures ANOVA. Mauchley's test of sphericity was not significant for any of the factors. Again, there was a significant main effect of lag (F(1,19)=13.0, MSe=0.180, p<0.01), with *d'* scores significantly higher at lag 4 than lag 8. There were no main effects of either shift type or intervening stimuli position. The only significant interaction was that between lag and shift type (F(1,19)=7.68, MSe=0.0821, p<0.05). Post-hoc analyses revealed significantly better recognition at lag 4 than lag 8 when stimuli were shifted, but no difference when there was no shift.

From the graphs (Figure 6.4 and Figure 6.5), it would appear that the superiority for non-shifted stimuli over shifted stimuli at lags 1 and 8 is less marked at lag 4. Perhaps this fact, in combination with additional data from the 'half' intervening stimuli position conditions, meant that there was no overall effect of shift in the second analysis. Indeed, the interaction demonstrated an effect of shift at lag 8, but not at lag 4.

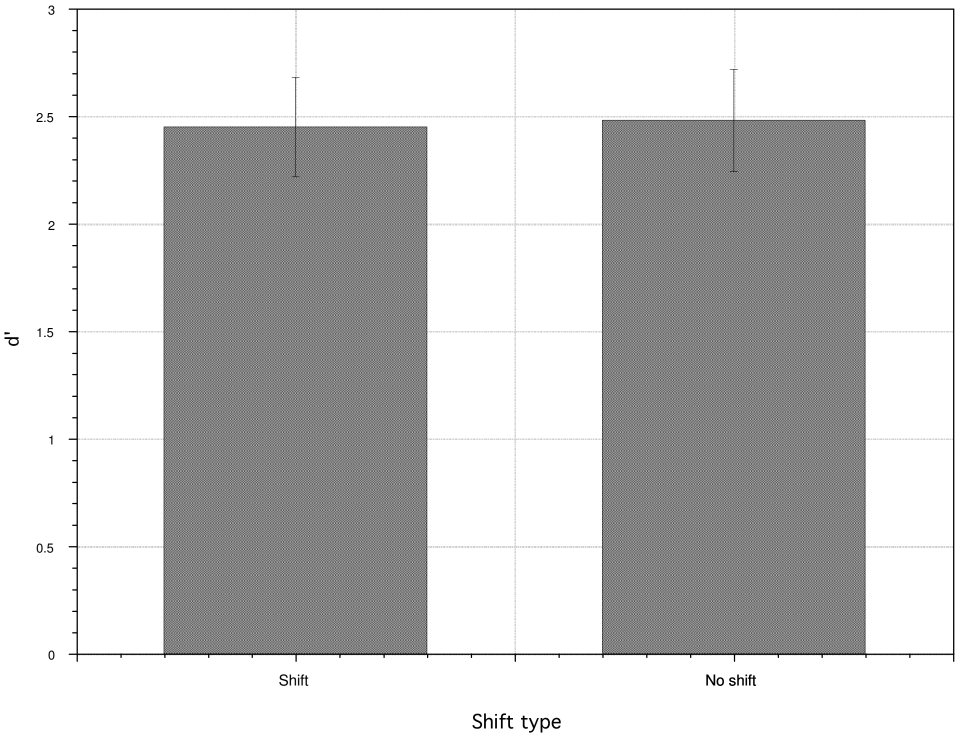

Figure 6.6: The effect of shift on d' scores at lag 0. Data = mean ± SEM.

The previous results demonstrate a partial effect of shift at lags greater than 0. In order to determine whether this was the case at lag 0 as well, *d'* scores were calculated for shift and no shift conditions at lag 0 (Figure 6.6). These data were entered into a paired t-test and no significant difference between the two conditions was found. This is consistent with the results of Experiment 8, in which the periphery (no shift) condition was not associated with significantly different scores to those in the periphery-periphery shift condition (equivalent to the no shift and shift conditions of the present experiment). It is also consistent with the results of Experiments 6 and 7, which did not find any advantages of same peripheral location over different peripheral locations of the same eccentricity. It appears to demonstrate that the effects of shift differ when stimuli intervene between study and test, as effects of shift were only found for lags greater than 0.

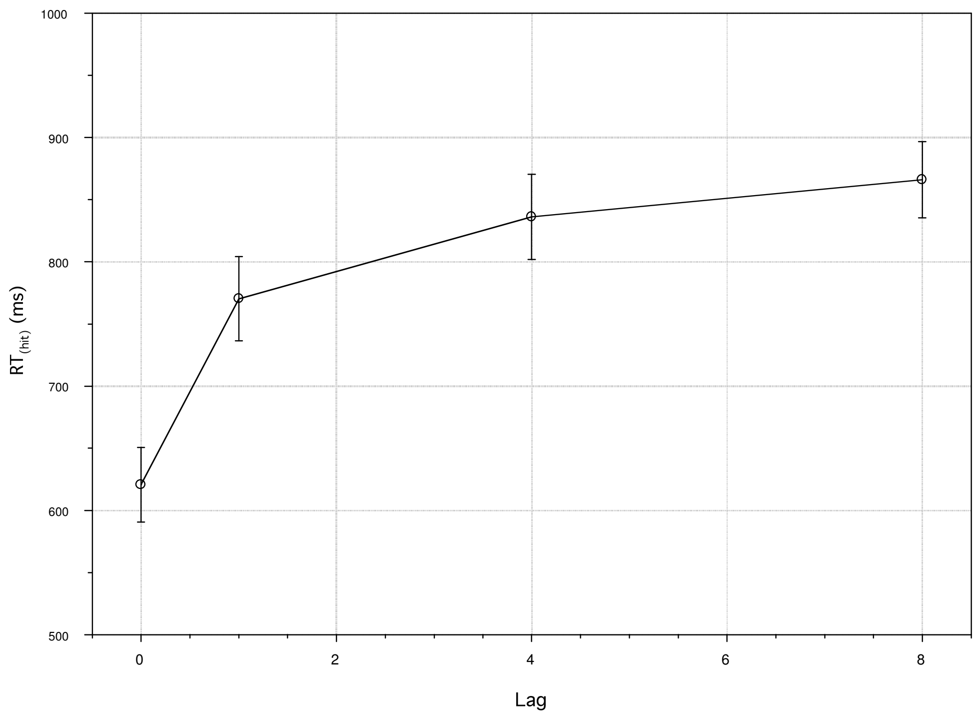

Figure 6.7: The effect of lag on reaction times for hits (ms). Data = mean ± SEM.

6.1.3.2 Reaction times for correct recognition

Reaction time (RT) data for correctly recognised stimuli were analysed by lag (Figure 6.7). A similar pattern of results to those of Chapters 2 and 3 was found, with RTs increasing with increasing lag.

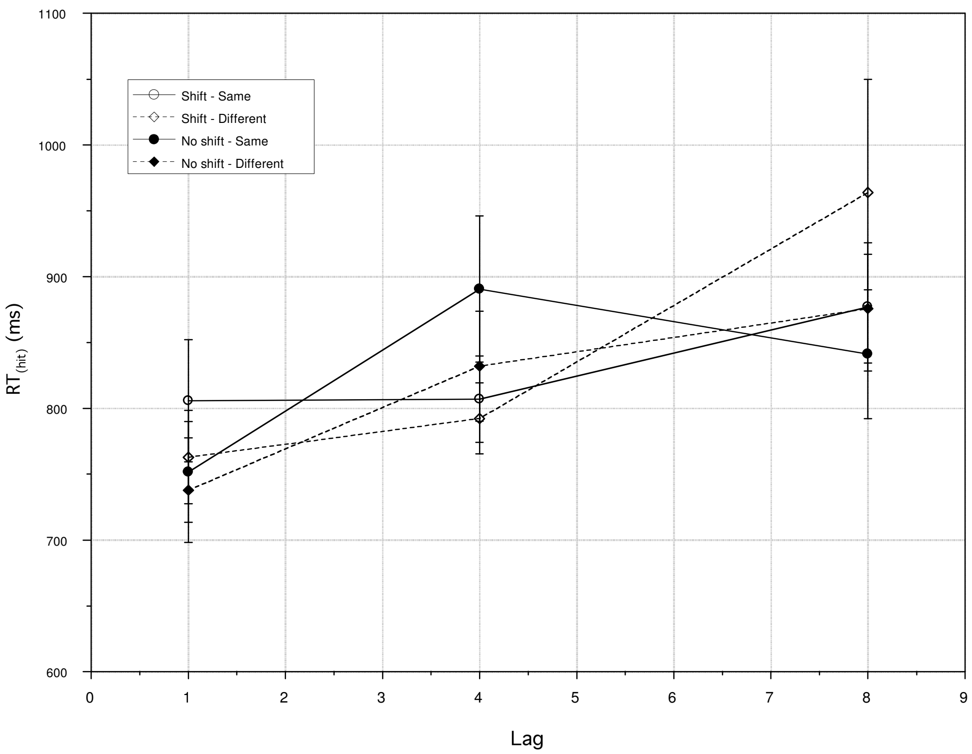

Figure 6.8: The effect of shift type and same vs. different intervening stimuli positions on reaction time data for correct recognition (ms) at lags 1-8. Data = mean ± SEM.

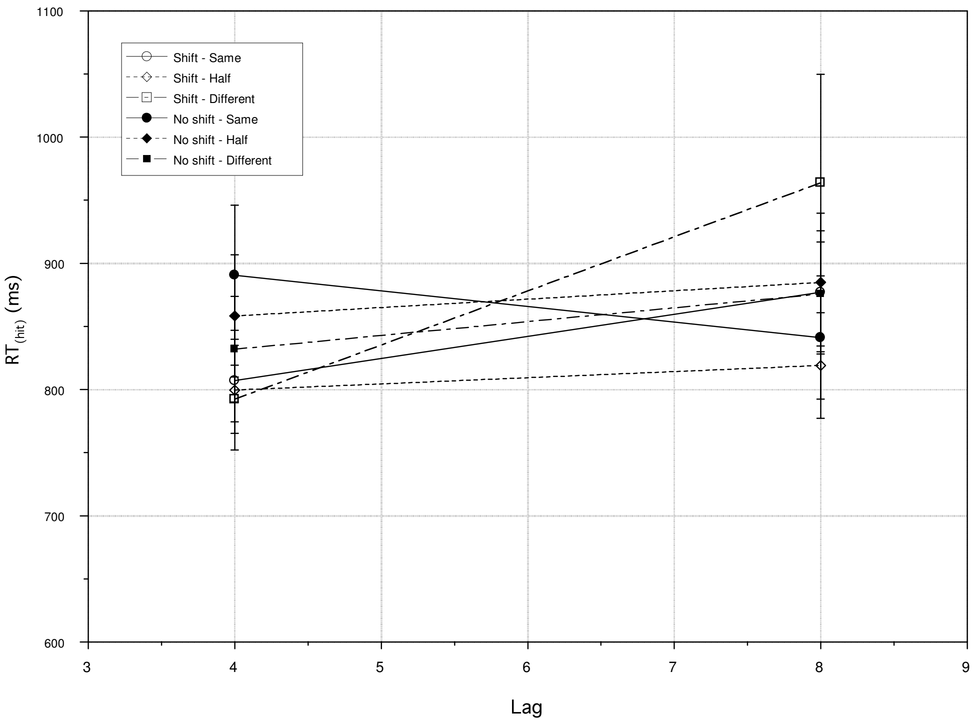

Figure 6.9: The effect of shift type and intervening stimuli position on reaction times for correct recognition (ms) at lags 4 and 8. Data = mean ± SEM.

The effects of lag, shift type, and intervening stimuli position on RT data for hits are illustrated in Figure 6.8 and Figure 6.9. Two ANOVAs, similar to those carried out for *d'* data, were carried out on the RT data. An ANOVA examined the effects of lag, shift type, and intervening stimuli position at lags 1-8. Mauchley's test of sphericity was significant for lag and Greenhouse-Geisser epsilon corrected degrees of freedom are quoted. There was a significant main effect of lag (F(1.35,25.7)=8.52, MSe=54300, p<0.01) and post-hoc tests revealed faster RTs at lag 1 than lag 8 (p<0.001). There were no significant main effects of shift type or intervening stimuli position, and none of the interactions reached significance. The ANOVA examining data for lags 4 and 8 found no significant main effects and no significant interactions.

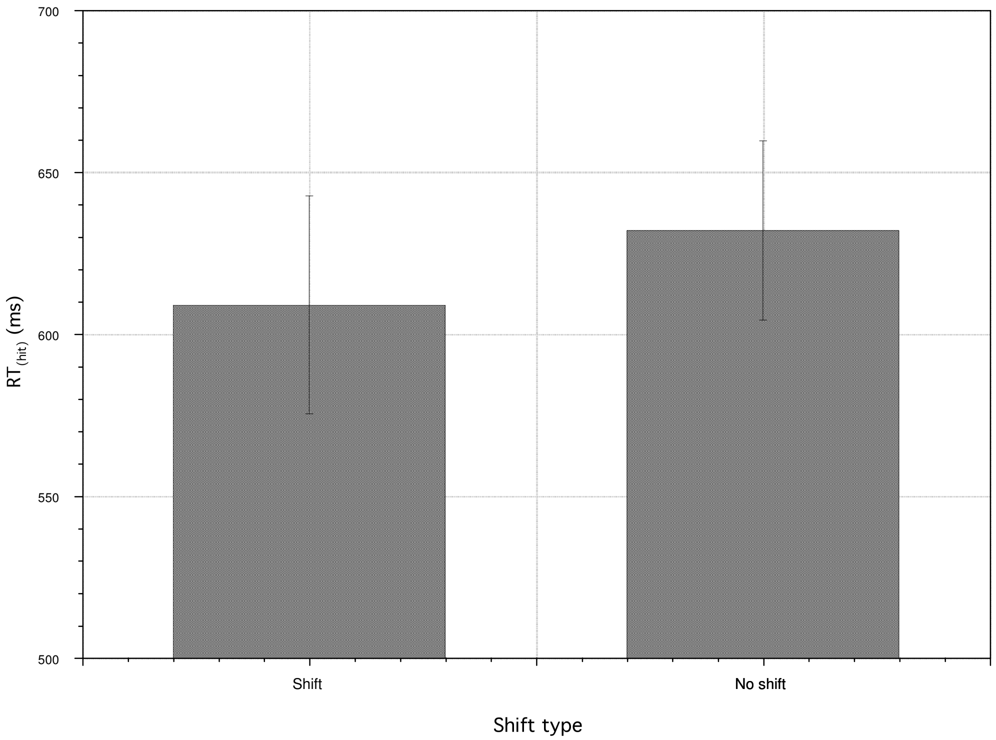

Figure 6.10: The effect of shift on reaction times for hits (ms) at lag 0. Data = mean ± SEM.

The RT data for lag 0 were divided into shifted and unshifted conditions (Figure 6.10) and compared with a paired t-test. Again, there was no significant effect of shift type. RTs did not appear to be affected by shift in this experiment.

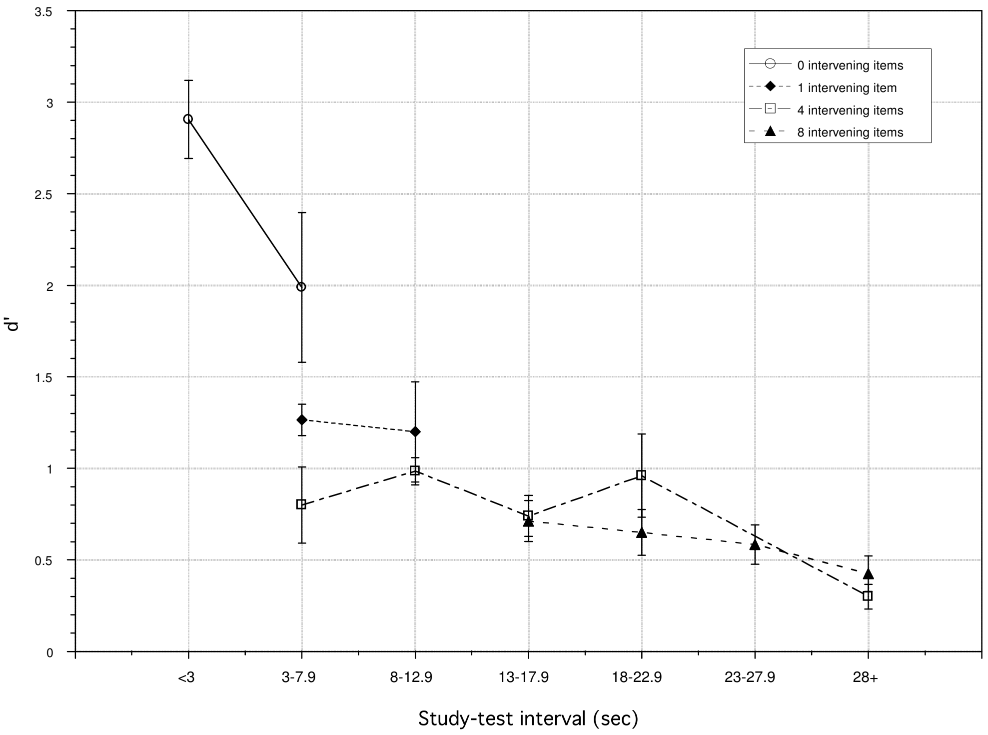

Figure 6.11: The effect of study-test interval (sec) and number of intervening items between study and test on *d'* scores. Data = mean ± SEM.

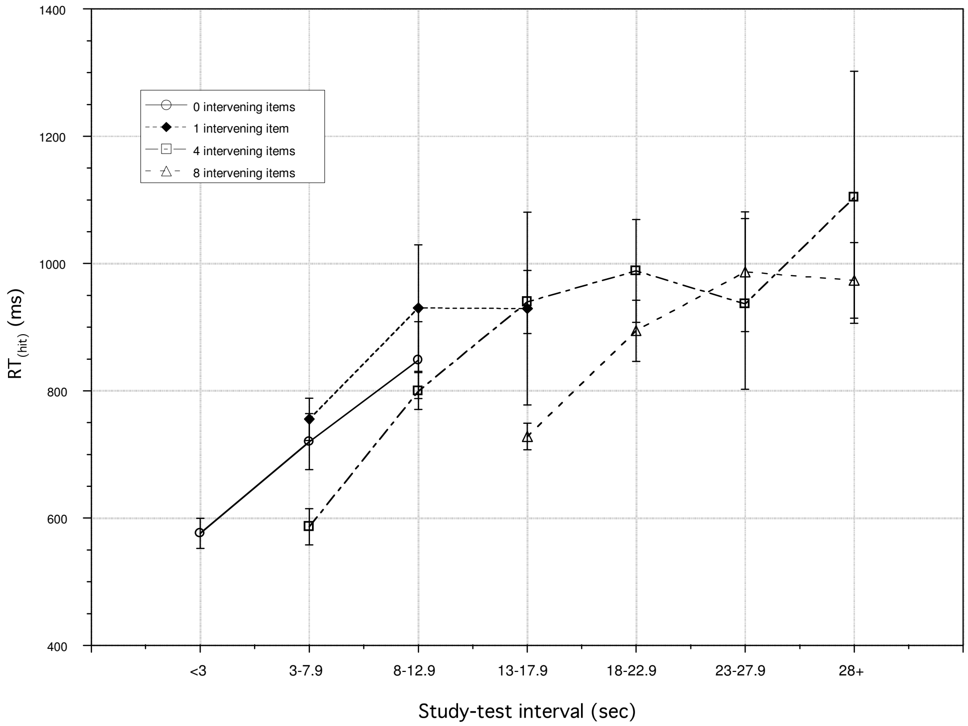

Figure 6.12: The effect of study-test interval (sec) and number of intervening items between study and test on RTs for hits (ms). Data = mean ± SEM.

6.1.3.3 Lag vs. study-test interval

In contrast to previous continuous recognition experiments in this thesis (Chapters 2 and 3), the current experiment was self-paced. Consequently it was possible to analyse the results by both number of intervening items and study-test interval. As has been discussed previously in greater depth (Experiment 3), continuous recognition normally confounds the number of intervening items with the study-test interval in the single variable, lag. Because the design of the current experiment did not impose a fixed relationship between the number of intervening items and time elapsed between study and test, the data were analysed by both factors in order to determine their relative importance.

Various preliminary analyses were used to determine appropriate study-test interval time bins for the final analyses. On the basis of these analyses it was decided that 5 sec intervals, starting from 3 sec and going up to 28 sec, would capture the majority of the data. Study-test intervals below 3 sec and over 28 sec were included in separate bins. Data were assigned to the appropriate time bin for which *d'* scores and RTs for hits were calculated for each number of intervening items.

The *d'* data (Figure 6.11) show that, overall, the effect of increasing study-test interval was to decrease *d'* scores. Indeed, this was also the case when the number of intervening items remained constant. However, examination of time bins for which there was data from more than one number of intervening items demonstrates that this factor stratifies the data. For example, for the 3-7.9 sec study-test interval data point, the differences between 0, 1, and 4 intervening items is large. This effect seems to be lessened with increasing study-test interval. It would appear that study-test interval and the number of intervening items affect the sensitivity of participants' discriminations.

The results for RTs for hits appear to show a stronger effect of study-test interval (Figure 6.12). The recognition latency increases with increasing study-test interval overall and for each number of intervening items. Again, though, there does appear to be a stratification between data points for different numbers of intervening items at the same point, e.g. between 8 and 4 intervening items in the 13-17.9 sec time bin.

In conclusion, both study-test interval and number of intervening items seem to be important in determining recognition performance, under the conditions of this experiment. However, the current experiment was not explicitly designed to separate out these two factors, and it is quite possible that individual differences may be at the root of the differences observed. Furthermore, the data are potentially confounded because the longest lags (e.g. 8 intervening items) are necessarily correlated with the longest study-test intervals so the pattern of results should be viewed tentatively. Further research is needed to confirm the reliability of these findings.

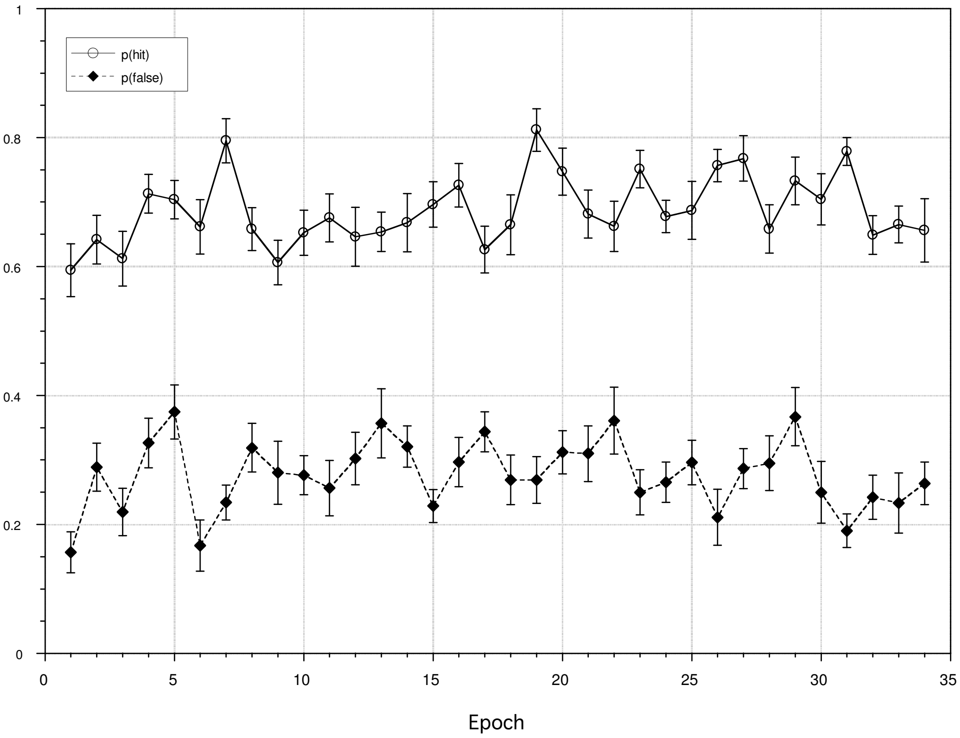

Figure 6.13: The effect of serial position on p(hit) and p(false). Epoch = 20 trials. Data = mean ± SEM.

6.1.3.4 Serial position

As the current experiment combined aspects of both timed continuous recognition (for which an increase in p(false) over the course of the experiment has previously been found), and self-paced positional shift experiments (for which a decrease in p(false) over the course of the experiment has previously been found) it was of interest to determine the effect of serial position on both p(hit) and p(false) over the course of the current experiment.

The data were divided into 20 trial epochs and p(hit) and p(false) were calculated for each epoch (Figure 6.13). Interestingly, p(false) remained comparatively constant over the course of the experiment, presumably neither increasing as a result of proactive interference nor decreasing as a result of a learnt improvement in discrimination. It is possible that participants' ability to set the pace of the experiment diminished the effects of proactive interference on their ability to correctly recognise novel stimuli. As in previous experiments, p(hit) remained relatively constant for the duration of the experiment.

6.1.4 Discussion

The results of the current experiment confirm the hypothesis that the effects of object position are more pronounced when items intervene between the study and test presentations. d’ scores were significantly higher for unshifted items than shifted items at lags 1-8, regardless of the position of the intervening trials. This difference was not present between scores for shifted and non-shifted trials when no trials intervened between study and test (lag 0), consistent with the results of previous studies (Experiments 6, 7 and 8). It is already well established that memory for items in continuous recognition decreases as a function of the number of items intervening between study and test, and Experiment 3 suggests that this is principally due to interference from the intervening items rather than decay of the memory trace. Presumably, with the increased demands on memory imposed by having to retain information over as many as 9 trials, the ability of visual memory to transfer information across retinal locations is impaired, resulting in the positional effect (i.e. a failure of translation invariance).

This qualitative change in the effect of translation does not appear to be caused by an increased difficulty in the task *per se*. Both Dill and Fahle (1998; Experiment 6, checkerboard stimuli) and Dill and Edelman (2001) found that increasing the difficulty of discrimination by increasing the similarity of stimuli resulted in an overall reduction of performance, but no change in the effect of positional translation. Indeed, in the Dill and Edelman study the global similarity between stimuli was increased, which would have been expected to increase retroactive interference between stimuli in a similar matter to that demonstrated in Experiment 4. However, to appreciate the effect of this change, one must take into account the different pattern of false alarm rates encountered by Dill and Edelman. In a similar manner to the current study, different false alarm rates for different shift conditions of study stimuli were used to calculate *d’*. They found that false alarm rates differed among shift conditions in a similar manner to the hit rates, such that shifted stimuli were associated with lower hit rates *and* lower false alarm rates. As such, overall, there was no change in *d'* by shift condition, in spite of the significantly different hit rates. Essentially, shifting stimuli made them more likely to be judged as novel. This effect was greater when the similarity of the stimuli was increased. The current experiment found the opposite effect on false alarm rates – recognition of shifted stimuli were associated with lower hit rates and *higher* false alarm rates. Shifting the stimuli decreased sensitivity without altering response bias. Evidently the effect of increasing interference by increasing the global similarity of items is different to that of increasing interference by presenting items between study and test. Increases in global similarity decrease the overall sensitivity of recognition with a decrease in hit rate and a decrease in false alarm rate. They also change the response bias for shifted stimuli towards ‘new’. Increasing the number of items intervening between study and test decreases overall sensitivity by a decrease in hit rate. It also results in reduced recognition sensitivity to shifted stimuli.

Surprisingly, the position of the intervening items between study and test had no effect on recognition. It was expected that, given previous evidence that memory is lost primarily through retroactive interference in this paradigm, if the intervening stimuli appeared in the same location as the studied item, they would have a greater interfering effect on memory for that item. The finding that this is not the case indicates that information about the spatial location of items is relatively unimportant in terms of interference. The fact that subsequent items were present was sufficient to interfere with memory for the study stimulus. It would seem that the task of retaining information about several similar stimuli in memory becomes more difficult solely as a function of the number of additional items, regardless of where that item appears. Interestingly, this suggests a position invariant process of interference, even though recognition may have some dependency on stimulus position under these conditions.

An important difference between the current results and those of Experiment 3, was that the link between intervening items and accuracy was less marked in the current experiment. Due to the fact that the current experiment was self-paced, there was considerable variation in the study-test interval between trials of the same lag, both within and between participants. As such, it was possible to plot *d'* and RT data for each lag (number of intervening items) against study-test interval. From this it was possible to infer that, whilst there was a clear effect of lag on *d'* at lower study-test intervals, the data for different lags became indistinguishable at intervals of 8 sec and greater. D-prime values for lags 4 and 8 also showed a downward trend at longer study-test intervals. This might be because the experiment was self-paced, and did not proceed at a constant and, from the participants' perspective, uncontrollable rate. Perhaps when participants had control over the rate of presentation they could adapt it in some manner that minimised interference, making the effects of decay more prominent. Speed of recognition showed a much clearer sensitivity to study-test interval, with RTs for each lag increasing as a function of time elapsed. This effect was consistent with the findings of Experiment 3. These results suggest that, in addition to the effects of interference, there may be variation in the decay of memory traces. Perhaps memory traces decay slowest in visual regions associated with a stimulus' location, resulting in an increased positional specificity of recognition as time passes. Of course, these effects may not necessarily indicate a spontaneous decay in the memory trace, but may be the result of interference from sources other than intervening stimuli (e.g. conscious efforts to maintain fixation during periods of tiredness, which would also contribute to longer trial durations).

The current results add to the existing literature on positional translation by demonstrating that the intervention of similar but different stimuli between study and test produces an effect of shift on recognition. Recognition where there are no intervening items between study and test is not affected by positional shifts. This is not the same pattern of results as might be expected based on other manipulations of task difficulty, such as the manipulation of similarity between stimuli, which have found no change in the effects of translation. An interesting future experiment to determine how these factors interact, if at all, would be a comparison of different levels of global stimulus similarity on the effects of positional translation in continuous recognition. These results also suggest that interference from intervening stimuli may be less important than the passage of time *per se* in determining information retention at longer lags. In order to examine this phenomenon more systematically further work should be done using timed, rather than self-paced, experiments, which vary the speed of presentation. With the aid of such future experiments, the factors causing the breakdown of translation invariance observed in the current experiment will be further elucidated.

Chapter 7 General discussion

7.1 Summary

The current thesis investigated the effects of a variety of spatial and temporal factors on visual recognition memory in human adults. Temporal factors examined included the presence of variable numbers of stimuli intervening between the study and test presentations of a novel stimulus, and the amount of time elapsed between study and test. These factors were studied using continuous recognition, a paradigm that interleaves study-test pairs in such a manner that memory performance reaches a relatively steady state, and is affected to a much lesser degree by serial position effects than list-based tests of memory. The spatial factors that were examined included the composition of the stimulus set (i.e. whether stimuli were drawn from a single category or many different categories, and the nature of the categories), the similarity between items, changes in the spatial location of items between study and test, and rotations of items between study and test. Whilst the effects of all of these temporal and spatial factors on recognition have been examined before in isolation, the current thesis has refined and extended previous findings with the use of parametrically defined stimuli that could be quantifiably manipulated. For example, the effects of interitem similarity in continuous recognition had previously only been determined non-systematically with the use of different classes of verbal stimuli (Estes & Maddox, 1995b). Experiment 4 extended these findings to readily quantifiable changes in the parameters used to define fractal and trigram stimuli. The use of stimuli that were both unfamiliar and not amenable to verbal labelling has also extended some of the findings of the extensive body of work concerning the recognition of words and other verbally encoded materials into the study of perceptually-based recognition e.g. the well-characterised effects of decay and interference on the recognition of words were investigated in the recognition of fractals (Experiment 3). In addition, the interrelations of various combinations of factors that have hitherto only been examined in isolation have been studied systematically in the experiments described here. Whilst both the effects of items intervening between study and test and of positional translations between study and test have been studied extensively previously, Experiment 9 found evidence of an important relationship between the two factors. In addition to these empirical advances, the work contained within this thesis has made possible an important discovery with methodological implications for the study of positional translation. Experiment 6 found that fixation verification with eye-tracking apparatus in a common positional translation experiment design resulted in a very different pattern of results to that found in its absence (e.g. Experiment 5). This has far-reaching implications for the existing body of literature on positional translation, most of which is based on studies that do not objectively verify fixation.

Experiments 1 and 2 used continuous recognition to investigated the effect of lag (the number of items intervening between study and test) on recognition of a variety of stimulus sets (face-like stimuli = ‘faces’, fractals, pictures of common objects = ‘pictures’, and digit-letter-digit trigrams) with large numbers of participants (50 per group). This was done in order to obtain a relatively precise comparison of the recognition profiles for these different classes of stimulus. Recognition of stimuli amenable to verbal encoding in terms of both simple basic-level object names (pictures), and more complex labels (trigrams), could be compared with those for which verbal recoding was a less useful strategy (faces, fractals). The experiment determined that recognition of pictures was superior to that of all other stimulus types. Recognition of this set was associated with significantly greater *d’* scores and faster recognition latencies. In addition, trigram recognition was associated with poorer *d'* scores than all other stimulus sets. These changes were associated with lower false alarm rates (p(false)) for pictures and greater p(false) for trigrams. There were also interactions between stimulus type and lag. For pictures, *d'* scores and reaction times indicated significantly better recognition at a lag of 0 than at all other lags, but after this stage their values remained relatively constant. Recognition of other stimulus sets were associated with a more gradual decline of performance with increasing lag. Significant increases in both hit rate (p(hit)) and p(false) during the first ten 10-trial epochs of the experiment were found for all stimulus types except the pictures.

Continuous recognition confounds the number of intervening items between study and test presentations of a stimulus with the time elapsed. In order to probe further how these factors affect recognition under conditions of continuous recognition they were separated in an experiment comparing recognition performance for stimuli presented at different rates (Experiment 3). Both trigrams and fractals were studied with a number of different numbers of intervening items, and at three different presentation rates. The study-test interval had no effect on *d'* scores for both fractal and trigram stimuli, which were affected solely by the number of intervening items.

The role of interference was investigated further in Experiment 4, which examined the effect of interitem similarity on recognition. Interitem similarity had previously been assumed to affect continuous recognition (Estes and Maddox, 1995), but no systematic tests of this hypothesis had been carried out. The global similarities of blocks of fractals and trigrams were controlled through constraints on the variability of certain stimulus parameters. These affected the range of letters that could be present in a trigram, or the variability of a number of parameters involved in the generation of fractals. The lowest level of global similarity amongst stimuli (‘dissimilar’) was associated with lower p(false) and higher *d'* scores than both 'medium' and 'similar' stimuli. Hit rates were not significantly changed by the manipulation. Similarity had no effect on reaction times.

Experiment 5 studied the effects of spatial separation between study and test using same/different recognition of face-like stimuli. Different stimuli differed from one another by standard changes of 20% in either the size of certain features, or both the size and location of those features. Horizontal, vertical and centre-periphery positional shifts between study and test were compared with control conditions in which the stimuli remained at the same location. The study found a recognition advantage, in *d'* and reaction time data, for stimuli that were studied and tested in the central location over all other conditions. There was also an advantage in both measures for stimuli that were studied and tested in the same peripheral location compared with those that were shifted. These effects occurred whether the changes were of feature size only or both feature size and location, with no significant differences between the two conditions.

As most previous studies of positional translation had made unverified assumptions about participants' fixation, Experiment 6 used eye-tracking to objectively verify central fixation with a similar design to Experiment 5. In this version, the next trial would only commence once the participant's fixation was detected in a small area around the fixation cross. The result was that the previously observed recognition advantage in *d'* scores for peripheral unshifted stimuli was eradicated. The advantage in reaction times was also reduced.

In order to determine whether the findings of Experiment 6 were applicable to other complex, abstract stimuli, or whether they were limited to the perception of face-like configurations, Experiment 7 replicated the former experiment using fractals. When the data from the two experiments were compared a familiar pattern of results emerged. The centre (no shift) condition was associated with greater *d'* scores than all other conditions, and the periphery (no shift) condition was associated with higher scores than the centre-periphery shift condition. The interaction results revealed that the former effect was limited to recognition of faces whilst the latter was limited to recognition of fractals. However, the relative ordering of mean values for recognition of both sets was largely unchanged. It was possible that the much better recognition of fractals than faces had resulted in a ceiling effect, such that recognition of fractals in the centre (no shift) condition did not significantly differ from that of peripheral stimuli. Reaction time data revealed a greater effect of shift on recognition latency for faces than that for fractals, but the ranking of conditions remained the same. Serial position analyses demonstrated decreases in p(false) for both sets throughout the duration of the experiment.

It was hypothesised that the translation invariance of recognition might breakdown when recognition was made more difficult by the addition of rotational transformations, especially given the well-described effects of inversion on recognition of faces (Yin, 1969). In Experiment 8, upright and inverted stimuli were tested in both rotated and unrotated conditions, and with and without positional translations. Rotation was found to have no effect on recognition, but stimuli studied in the upright orientation were recognised faster than those studied in the inverted orientation. The effect of shift was similar to that observed in Experiments 6 and 7, with a centre (no shift) advantage in *d'* scores and recognition latencies, and a periphery (no shift) advantage over periphery-centre shift in *d'*. The reaction time results also demonstrated a central test position advantage. Serial position data were analysed in epochs of 100 and revealed significantly more sensitive discrimination in the second experimental block (of 2) than the first. This effect was similar for all shift types.

In the final experiment (Experiment 9) spatial and temporal factors were combined in a continuous recognition experiment with positional shifts. It was thought that interference from items intervening between study and test presentations of stimuli would increase the difficulty of the task sufficiently that translation invariance would breakdown. The results revealed that the presence of items intervening between study and test did indeed induce an effect of shift on *d’* scores. Items shifted at a constant eccentricity were associated with significantly poorer recognition than those that remained in the same location. No effect of translation on reaction times was found for this experiment. The position of the intervening items was manipulated so that they could occur on the same side as the study presentation, the opposite side, or (for longer lags) half on the same side and half on the other. However, the position of intervening items had no effect on measures of recognition. Post-hoc analysis of the results by both number of intervening items and by study-test interval showed an effect of both factors.

7.2 Categorisation vs. recognition

Whilst it has been documented that the discrimination of visual stimuli on the basis of their inclusion in different categories is superior to that of recognition of individual exemplars from the same category (e.g. Roberson & Davidoff, 2000), Experiment 1 further characterised the superiority of categorisation as a function of lag, and of serial position. The superiority of recognition of the pictures set was assumed to be largely the result of its composition. As it was composed of items from many different categories it was highly probable that most discriminations could be achieved on the basis of simple categorisation as opposed to ‘true’ recognition (Goldstein & Chance, 1970). Discrimination of items within this set was associated with a very shallow memory retention curve, and no learning effects during the course of the experiment. Compared to other stimulus sets, for which discrimination relied on the identification of individual exemplars from a single category, discrimination based on categorisation showed only a very small decline in sensitivity as a function of lag, and performance did not change significantly during the course of the experiment. Discrimination based on recognition from a single category was much more adversely affected by items intervening between study and test, and showed improvement during the early stages of the experiment.

A previous study comparing cross-category and within-category recognition has suggested that the recognition advantage for cross-category decisions is eliminated by verbal interference, suggesting that the advantage is based on verbal labels (Roberson & Davidoff, 2000). The use of verbal labels is associated with efficient recognition of familiar items, whereas visual discrimination is employed for less familiar stimuli (Simons, Graham, Owen, Patterson, & Hodges, 2001). The use of basic-level object names to label stimuli would have been sufficient for a high discriminative sensitivity for the pictures set. This interpretation is supported by the finding that there was no increase in p(hit) or decrease in p(false) during the course of the experiment, indicating that discrimination of the stimuli was already optimal. As the familiarity of participants' with basic-level object names was assumed to be high at the start of the experiment one would not expect any change in participants' ability to use them as the experiment progressed.

An alternative explanation for good performance with the pictures is that, as they were not drawn from a homogeneous category, they were more likely to be visually distinct from one another, and that this relative distinctiveness was sufficient to improve recognition performance. However, a study of distinctiveness effects in face recognition has demonstrated that faces that are highly distinct in terms of their spatial characteristics are likely to have certain individuating features (e.g. beards) that make them amenable to categorisation (i.e. with verbal labels) (Busey & Tunnicliff, 1999). The implicit naming that was assumed to be involved in the recognition of pictures is also highly likely to have been employed in the recognition of trigrams. However, because these items were highly similar to one another they were assumed to form a homogeneous category, and their recognition was correspondingly poorer than more heterogeneous stimuli.

7.3 Changes in recognition with experience

Through the use of both novel and familiar stimuli (e.g. those that participants were likely to already have some experience with, such as pictures of common objects) it was possible to determine that recognition changed as a result of increased familiarity with stimuli. A comparison of the results of Experiment 1 and Experiment 2 revealed an advantage in discriminative ability for the birds set over that of similarly composed sets of abstract, computer-generated stimuli. This was thought to be the result of a greater initial familiarity with birds, a suggestion supported by the lack of subsequent improvement in discriminatory ability for these stimuli throughout the duration of the experiment. It was assumed that, as participants were already highly familiar with the bird configuration, little further learning of this category was possible. However, it was possible that subordinate level categorisation (e.g. using subcategories of birds such as blackbird, pigeon, etc.) would have been sufficient to increase recognition accuracy for this set. It can certainly be concluded that recognition was better for natural classes of objects, with which participants may have had prior experience, than for artificial categories created for the purposes of the experiment (e.g. face-like stimuli, fractals). Indeed, it is perhaps nonsensical to attempt to distinguish between perceptual familiarity and the ability to categorise at the subordinate level, as the two may develop simultaneously, as a result of expertise.

In addition to finding that classes of stimuli already familiar to participants were recognised better than novel stimuli, there was evidence from serial position analyses of the data that recognition performance with previously novel stimuli improved throughout the duration of experiments. In the case of continuous recognition experiments this entailed an increase in p(hit) (e.g. Experiments 1, 2, and 3). This indicated that participants were improving their ability to detect ‘old’ stimuli. However, this was often accompanied by an increase in p(false) suggesting that the improvement in p(hit) may simply have been a change in response bias towards responding ‘old’. In the same/different experiments, p(false) often decreased during the course of the experiment (Experiment 6, 7, 8), whilst p(hit) remained relatively constant, indicating an improvement in participants’ ability to detect novelty in stimuli irrespective of bias. This was most pronounced in the longest experiment, Experiment 8 (956 trials), for which there was a clear and significant increase in *d’* during the course of the experiment, although significant changes occurred in the first 100 trials in several experiments. It would seem that relatively little experience is required to improve discriminative ability, and that this improvement is immediate.

There was also evidence from Experiment 8 that the type of experience is important in determining the extent of learning. This experiment was divided into two blocks of stimuli, stimuli that were rotated between study and test, and those that were not. The order in which participants carried out the two blocks was counterbalanced, and post-hoc analyses revealed that performance on the rotated block was improved significantly if preceded by the unrotated block, whereas the reverse effect was not significant. It would seem that the ability to mentally rotate stimuli was improved by prior experience of discrimination of upright and inverted stimuli, whereas the ability to discriminate the stimuli without rotation was unaffected by experience. This analysis is to be validated by future experiments.

7.4 The importance of interference in continuous recognition

Considerable evidence for the involvement of both proactive interference (PI) and retroactive interference (RI) in affecting continuous recognition memory performance was found. With the exception of highly familiar stimulus sets (pictures, birds) increases in p(false) during the early stages of Experiments 1-4 were found. This effect suggested that participants' ability to detect novel stimuli was initially good, but was impaired as the number of previously seen stimuli increased. This was presumably because novel arrangements of features became increasingly hard to distinguish from the enlarging array of previous stimuli. This effect occurred only up to a point as p(false) did reach an approximate steady state for each stimulus type in each experiment, in a departure from the results of Shepard and Teghtsoonian (1961). The number of trials taken to achieve the steady state appears to be linked to the difficulty of stimulus discrimination, as p(false) for the hardest stimuli to discriminate (e.g. digit-letter-digit trigrams in Experiment 1, and Shepard and Teghtsoonian's more difficult three digit numbers) reached asymptote over the longest numbers of trials.

The effect of lag was shown to be a function of the number of stimuli intervening between study and test, rather than the time elapsed, in Experiment 3. This finding supports an interference-based explanation for the effect of lag i.e. that memory for items remains relatively intact unless further different items must be remembered before test. Retroactive interference, interference with a memory trace by items occurring *after* participants' experience of the remembered item, would appear to be the most important factor determining *d'* with this task. This contradicts theories based on a decay over time, in which visual memory is thought to be maintained through a process of 'visual rehearsal' (A. Baddeley, 2000; A. D. Baddeley & Hitch, 1974; Cowan, 2005). It is also inconsistent with the functional decay theory of Altmann and Gray (2002), which would suggest a decline in p(false) during the course of experimental blocks with different rates of presentation, as the decay rate adapted to the rate of interference. No evidence for such a change was found.

Interestingly, the recognition latency was affected by the time elapsed between study and test, but not by the number of intervening items – the opposite pattern to these factors' effects on *d'*. This implied that the speed of recognition was a function of the speed of presentation, and was not affected by lag. However, the results of Experiments 1 and 4 did demonstrate an effect of lag on recognition latency when the rate of presentation was held constant. It would seem that, when the rate of presentation is allowed to vary, the speed of recognition adapts to this rate, in a manner that greatly weakens the effect of lag.

Experiment 4 provided further evidence that interference between stimuli is the primary cause of memory deterioration in continuous recognition, by determining that parametric similarity between stimuli within a set affected recognition performance. Stimulus sets that were composed of stimuli that could vary widely in their defining parameters were recognised considerably better than those with parameters that were relatively constrained. Importantly, this effect was similar for stimuli that were amenable to verbal labelling (trigrams) and those that were not (fractals), supporting the theory that perceptual and verbal memories are processed similarly (e.g. Doty & Savakis, 1997; Ward, Avons, & Melling, 2005). The lack of difference between high and medium similarity levels suggests that there may be perceptual thresholds for the detection of difference. Indeed, these thresholds may map onto category boundaries and change with experience. There is evidence in the neurophysiological literature that monkey perirhinal cortex neurones involved in recognition are organised according to sensory experience (C. A. Erickson, Jagadeesh, & Desimone, 2000), suggesting a possible neural substrate for altered sensitivity to perceptual similarity. If the detection of similarity is stratified rather than continuous, parametric similarity may not directly correlate with participants' subjective experience of similarity. It may, however, explain why certain stimuli are perceived as being more distinctive than others.

7.5 Translation invariance and positional effects in recognition

The experiments in which positional translation was examined determined that recognition memory for stimuli was translation invariant under some conditions, but that there was an effect of shift in others. When stimuli were examined under conditions of same/different recognition, in both upright and inverted orientations, and in both rotated and unrotated conditions (Experiments 6, 7, and 8), recognition was invariant with regards to translation, at a constant eccentricity. However, *d'* scores were decreased by shifts between fixation and peripheral locations. Recognition performance was also better when stimuli were both studied and tested at fixation than when study and test were both at the same peripheral location. It was hypothesised that the central advantage in recognition was the product of better visual spatial acuity at that region, allowing finer discrimination of features at that location than that possible in peripheral vision. This factor alone, however, would suggest that stimuli that were experienced at fixation for either study (centre-periphery shift) or test (periphery-centre shift) should be recognised better than stimuli that were both studied and tested in the periphery. In fact, the opposite effect occurred. This demonstrates that a constant eccentricity, and, therefore, acuity of vision, is important for the detection of sameness. Presumably representations of the same stimulus at different spatial scales are more likely to be judged as representing different objects, than representations of the same object at the same spatial scale.

The hypothesis that changes in eccentricity, rather than positional shift *per se*, are responsible for some of the translational effects previously reported, is supported to a certain degree by Dill and Fahle's (1998) study of pattern recognition. That study found that pattern recognition performance decreased as a function of increasing spatial separation. However, the studied patterns were located at an eccentricity of 1° in the periphery and were then shifted horizontally or vertically by multiples of 0.5°. Whilst the effect of translation was to decrease performance, the effect of translation was lesser at 2°, when half of the trials would have been at the same eccentricity as the study trial (e.g. when a pattern presented at 1° to the right of fixation was shifted 2° to the left, to a location 1° to the left of fixation).

The results were at variance with those of Gratton et al. (1997) and Hornak et al. (2002) who had used similar positional shifts and found evidence that horizontal between-hemifield shifts affected recognition more adversely than equivalent vertical within-hemifield shifts. There were no significant differences between the two conditions in the experiments reported in this thesis, suggesting that reported hemispheric differences in recognition may be dependent on the type of stimuli used (Gratton et al.: symmetric patterns, Hornak et al.: common objects), or the experimental design. Hornak et al. used a design in which pairs of objects were presented in which either one or neither of the stimuli had been seen previously. This required the division of attention across visual hemifields which may have revealed biases in attention not present in the experiments of this thesis.

In Experiment 8 stimuli were presented in two different orientations: upright, or inverted (rotated 180°), and 'same' stimulus pairs could differ in a rotation as well as a positional translation. Despite the very weak inversion effect observed (that 'faces' studied in the upright orientation were recognised better than those studied in the inverted orientation), and a slight right hemisphere advantage ('faces' were recognised more slowly in the right visual hemifield than at centre, whereas this was not the case for those recognised in the left hemifield), the face-like stimuli were not recognised in a similar manner to photographs of real faces. The pattern (e.g. that recognition was unimpaired by 180° rotation) suggested that the spatial relations of the face-like stimuli were likely to have been encoded categorically, in a manner similar to that for memorisation of common objects and patterns. As such, there was little disruption to perception of the stimuli's configuration, and the hypothesis of Dill and Edelman (2001) that such disruption causes a breakdown of translation invariance, could not be validated.

Whilst there was no effect of shift *per se*, items shifted between visual field locations of different eccentricities were associated with poorer recognition than those at a constant eccentricity, in a similar manner to that observed in Experiment 7. In a contrast to the findings of previous experiments (Dill & Fahle, 1998; Kahn & Foster, 1981; Larsen & Bundesen, 1998), there was no clear evidence that 180° rotation eradicated the effect of position on recognition, although mean *d'* scores for recognition of rotated stimuli appeared to be more similar than those for unrotated stimuli.

Serial position analyses revealed a clear learning effect over the course of the experiment, resulting in better discrimination at later epochs than earlier ones. This effect was independent of shift type, suggesting that an increased familiarity with the stimuli improved overall performance, but did not improve the ability to mentally rotate or recognise shifted stimuli. However, these analyses relied on a reduced sample size, and their reliability must be ascertained with future work.

Interestingly, a rather different pattern of results to those in which fixation was objectively verified was obtained in an experiment (Experiment 5) that, like many previous experiments in the positional translation literature, did not verify fixation. The results indicated that there was a same location advantage for recognition, resulting in partial position specificity. These results suggested that the face-like stimuli were recognised in a manner more similar to that of abstract dot-cloud and checkerboard patterns (Dill & Fahle, 1998) than that of animal-like stimuli (Dill & Edelman, 2001). Dill and Fahle (1998) had previously found evidence for positional effects in same/different pattern recognition, whereas recognition of animal-like stimuli was demonstrated to be invariant with regards to position (Dill & Edelman, 2001).

It was suggested that the important differences between Experiment 5 and Experiment 6 were caused by drifts in fixation when fixation was not controlled. If eye position drifted towards the location of the last stimulus then subsequent stimuli occurring at the same location would be perceived more centrally than was assumed, whereas subsequent stimuli occurring at different locations would be perceived more peripherally than was assumed. Given the hypothesis that the central advantage in recognition was the product of better spatial acuity at that region, one would expect the better acuity with which stimuli in the same location were perceived to result in better recognition than that for more peripheral stimuli at other locations. These findings pose important questions about the validity of the many past studies of positional effects in recognition, hat have made similar assumptions about fixation to those of Experiment 5. These previous findings may, perhaps, be similarly misleading in suggesting effects of translation that are mere artefacts of drifting fixation.

Experiment 9 demonstrated that a genuine breakdown of the ‘translation invariance’ (i.e. translation invariance at a constant eccentricity), observed in Experiments 6, 7, and 8, was possible. Recognition of items shifted between visual field locations of a constant eccentricity was impaired relative to those that stayed at a constant location. It is likely that the greater demands imposed on visual memory by continuous recognition meant that the ability to translate information between different retinal locations was impaired, resulting in the effect of shift at lags greater than 0. This was not assumed to be a general effect of task difficulty *per se*, as previous experimenters found no effects of increased interstimulus similarity on positional translation (Dill & Edelman, 2001; Dill & Fahle, 1998), but was thought to be specific to the presence of items intervening between study and test.

Whilst the effect of increasing similarity was to decrease both p(hit) and p(false) resulting in no overall change in *d'*, the effect of interference from intervening items was to decrease p(hit) and *increase* p(false) with a corresponding decrease in *d'*. The interfering effect of these intervening items was not dependent on their occurrence in the same location as either study or test. This suggests that the effects of interference are invariant across retinal locations, but increase the participants' reliance on positional information when making recognition judgements.

The analysis of recognition by both the number of intervening items and the time elapsed between study and test suggested that there was a greater 'decay' effect in Experiment 9 than that observed without positional shifts in Experiment 3. However, this analysis was carried out post-hoc, and attempted to address questions that the experiment was not specifically designed to address. Individual differences between participants and the effects of fatigue are just two factors that could have confounded this analysis. Only future experiments specifically designed to examine this issue will determine the reliability of this finding.

7.6 Neural mechanisms

As discussed in Chapter 1, the repetition suppression (RS) responses of IT neurones to effective stimuli are considered to be of primary important in visual recognition. Studies of the perirhinal cortex have identified certain subtypes of cells exhibiting such responses, including familiarity cells that encode the absolute novelty of a stimulus, and recency cells that encode the relative recency of exposure to a stimulus (Xiang & Brown, 1998). These responses persist when stimuli intervene between study and test (E. K. Miller, Li, & Desimone, 1993) although the robustness of this persistence is variable, with fewer cells demonstrating RS as time and the number of intervening items increase (Fahy, Riches, & Brown, 1993). What is not clear from previous studies is whether the effects of time and intervening items on the number of cells exhibiting RS can be dissociated. The results of Experiment 3 demonstrating that, on behavioural measures, intervening items are more important than time *per se*, suggest that such an effect may exist at the neural level also. Certainly experiments designed specifically to disentangle these two factors at a neural level would be useful in determining how closely behavioural measures of recognition are associated with RS at the neuronal level.

The responses of IT neurones have previously been described as invariant with respect to transformations of the stimulus dimensions, such as the size and position within the receptive field, of an object (Desimone, Albright, Gross, & Bruce, 1984). However, more recent studies of RS have found large changes to the response for items moved only 1.5º (DiCarlo & Maunsell, 2003), and cell ‘preferences’ for certain retinal locations (Lueschow, Miller, & Desimone, 1994). Lueschow and colleagues suggest that location is treated like an object feature for the purposes of recognition. It is interesting to note that neurophysiological studies of positional translation, like their human psychophysical counterparts, refer to shifts in terms of their distance in degrees of visual angle, rather than in terms of eccentricity. The findings of Experiments 6, 7, and 8 suggest that eccentricity is more important than spatial separation in determining the effects of positional translation. Future studies are required to determine whether eccentricity and spatial separation can be disentangled at the neuronal level. It would also be of interest to determine whether RS becomes more sensitive to shifts of the stimulus under conditions of continuous recognition, in the manner predicted by behavioural data in Experiment 9.

7.7 Future experiments

The work reported within this thesis has answered important experimental questions regarding human visual recognition. However, in the course of these experiments, many new questions have arisen that may provide impetus for future research. This section details some of those questions, and the further work that is required to address them.

The findings of Chapter 2 demonstrated that stimulus sets composed of items from different categories were recognised better than those drawn from just one of those categories. Whilst it was speculated that the primary mechanism for this advantage was verbal labelling (naming), there also exists the possibility that visual discrimination between members of different perceptual categories was responsible. In order to determine the extent to which verbal labels were important future experiments could be carried out in which verbal interference, e.g. the visual presentation of words likely to interfere with labelling, could be used in the intertrial period to interfere with any verbal labels generated in connection with the pictures. Whilst it was demonstrated that testing picture stimuli from only one category greatly reduced recognition performance, it would be of interest to determine whether the reverse effect is possible, with other types of stimuli. For example, what would the recognition profile be like for a stimulus set composed of 10 exemplars from each of 20 categories (the same numbers as those for the mixed pictures set) of complex abstract stimuli, e.g. fractals, face-like stimuli, Gauthier and Tarr's (1997) Greebles? A third method for determining the nature of the difference between discrimination of pictures and birds, is the use of neuroimaging. Through the use of fMRI or other neuroimaging methods, it would be possible to compare the activation of different brain regions during recognition of both sets. This would give a clear indication of whether areas primarily involved in verbal processing or visual processing were involved in the different types of recognition, and how these areas interacted. In addition, ERP studies might be able to examine the timing of the evoked potentials in these brain regions, and thus provide more information regarding differences in recognition latencies, and the temporal sequencing of recognition strategies.

The finding that the birds stimulus set was still associated with a recognition advantage over comparable sets of abstract stimuli suggested that prior familiarity with natural classes of stimuli was important in determining recognition performance. Whilst it would be difficult to mimic the level of exposure participants have over the course of their lives with these naturally occurring objects, it would be interesting to discover what the effects of greater exposure to previously novel stimuli would have in this paradigm. Would extensive training in the discrimination of different items from the same novel classes (e.g. face-like stimuli, fractals), perhaps over several days or weeks prior to continuous recognition, improve subsequent recognition performance?

Chapter 3 demonstrated the importance of interference from stimuli intervening between study and test in continuous recognition, and suggested that it was dependent on both physical and phonological similarity between items. However, the exact relationship between stimulus properties and the level of interference remains unknown. Certainly the number of stimuli is important, as is their parametric similarity to a certain degree. However, Experiment 9 demonstrated that the spatial location of an intervening stimulus is irrelevant to its interfering effect. The definition of 'similarity', as regards interference, must be expanded from one based on geometric parameters defining a stimulus' makeup, to an understanding of the subjective experience of similarity. For example, Experiment 4 used three arbitrary levels of interstimulus 'similarity', based on the variability of parameters used to define the properties of fractal and trigram stimuli. In both cases only the most dissimilar stimuli were recognised in a significantly different manner to the most similar stimuli. Stimuli of an intermediate level of global similarity were associated with the same pattern of recognition as that for similar stimuli. Why did the one change (from ‘medium’ to ‘dissimilar’ stimuli) create a significant change in recognition, whereas the other (from ‘similar’ to ‘medium’) did not? It was hypothesised that certain “thresholds” exist in the perception of difference, perhaps marking the boundaries of learnt perceptual categories. Further research is required to determine how changes to different stimulus features and properties are perceived, and what determines sensitivity for the detection of these changes. Whilst there is a significant body of literature elucidating such detection thresholds for simple stimuli, e.g. sinusoidal gratings, colour patches, etc., this research needs to be extended to more complex stimuli, as the basic properties of objects interact in combination to produce features and whole objects. It would be of interest to learn how robust memory for the perception of these changes is, e.g. whether the changes detectable in same/different recognition are also detectable after interference from other stimuli. Another interesting avenue that could be pursued is the investigation of how experience with artificial categories alters their perception. Do subcategories emerge with extensive experience of a set of stimuli, for example, ‘sad’ vs. ‘happy’ face-like stimuli? And are these differences reliant on perceptual discrimination or verbal labellling? Also, what kind of experience is necessary for such learning? Does it occur spontaneously through passive observation of stimuli or is active engagement in a relevant task (e.g. one involving perceptual discrimination) required? It is only through answering these questions with carefully designed recognition and perceptual learning experiments that a better understanding of ‘similarity’ can be achieved.

The finding that the use of eye-tracking to verify fixation in positional translation studies significantly altered the findings of those studies must cast serious doubt on the validity of previous work on positional translation that had neglected to verify fixation (e.g. Dill & Edelman, 2001; Dill & Fahle, 1998; Gratton, Corballis, & Jain, 1997; Hornak, Duncan, & Gaffan, 2002; Kahn & Foster, 1981). Considering that these studies form the vast majority of previous work in this field, future re-evaluation of these previously accepted findings may radically alter knowledge about the effects of translation. The major areas that must be re-examined are whether the breakdown of translation invariance for simple patterns is genuine, whether training participants with stimuli at fixed locations still produces positional learning, and whether the apparent breakdown of translation invariance when stimuli are configurally scrambled is still the case.

In relation to the stimuli used in Experiment 5, the apparently consistent but non-significant differences between faces that changed only in the size of features, and faces whose features changed in both size and location, should be examined in more depth. Would exaggerating the location changes (e.g. by making them larger) make the differences easier to detect? The intriguing finding of Dill and Edelman (2001) that the disruption of the learned spatial configuration of stimuli, but not the replacement of learned features with novel features, causes a breakdown of translation invariance in recognition, should be extended in order to ascertain what kind of disruption to the configuration is required. Experiments contained within this thesis have disrupted the configuration of stimuli through both featural size and location changes, and through rotation, without producing such an effect. Is the effect limited to well-learned stimuli or is a more radical disruption of configural cues required to breakdown translation invariance?

It is also important that the hypothesis proposed here, that differences in recognition performance evoked by positional shifts are the result of changes in eccentricity, and therefore visual acuity, and do not normally occur at a constant eccentricity, should be tested by subsequent work. One experiment that would be useful in testing the validity of this hypothesis would be to apply variable amounts of spatial blurring to stimuli, in order to disrupt any advantage due to the greater acuity of central vision. If pairs of stimuli are more likely to be considered to be different when they are processed with different levels of acuity, then this effect should be eradicated by sufficient blurring to equalise acuity at the different regions. In addition to trying to understand this phenomenon in more depth with the stimuli used in this thesis, it would be of value to determine whether it can account for some of the purported effects of positional shift described previously for simpler patterns (e.g. Dill & Fahle, 1998).

The results of Experiment 8 suggested that the face-like stimuli were not recognised in a similar manner to that of actual faces, in terms of not producing an ‘inversion effect’ or a strong right hemisphere advantage of recognition, in spite of containing a similar arrangement of simplified features. This begs the question, what properties of a stimulus make it face-like for the purposes of recognition? This is of particular importance considering the assumptions of previous investigators that face-like stimuli are a close approximation of actual faces (e.g. Goren, Sarty, & Wu, 1975; Lewis & Johnston, 1998).

There were no obvious effects of rotation on recognition in Experiment 8. However, previous experiments have suggested that 180º rotation eradicates positional shift-induced differences in recognition performance (Dill & Fahle, 1998; Kahn & Foster, 1981; Larsen & Bundesen, 1998). This effect might become more evident in an experiment like Experiment 9 where there is a clearer effect of shift. In addition, future work should address the effects of other degrees of rotation on positional shifts, especially those in the region of 90º that are more deleterious to accurate recognition. It would also be of interest to determine the effects of rotation of specific features in conjunction with positional changes, given evidence from the study of photographs of faces with rotated features that this manipulation severely disrupts configural cues for recognition (e.g. Robbins & McKone, 2003).

When the effects of positional translation were examined under conditions of continuous recognition, the previously reported translation invariance at constant retinal eccentricity broke down. This was only examined for only one type of stimulus: fractals. It would be useful to determine whether this result can be replicated for other types of stimuli, including stimuli amenable to verbal labelling, and of different levels of familiarity. It may be the case that phonological encoded or well-known stimuli are resistant to this positional effect, and are recognised by different processes to the fractals. In addition, whilst manipulations of global similarity have previously been documented to have no differential effect on shifted and unshifted stimuli (Dill & Edelman, 2001; Dill & Fahle, 1998), their effect may be different in continuous recognition. As the demands of continuous recognition are theorised to have impeded the ability to translate information about objects across the visual field, increasing the interference by increasing interstimulus similarity might be expected to further impair translation. A detailed analysis of how these different factors interact is needed to establish whether this is the case.

The current thesis has contributed to the understanding of human visual recognition memory through important and novel findings regarding the nature of interference in recognition and the processing of object location. It is hoped that the work reported here will not only add to the literature on these factors, but will also provide impetus for further research in this area.

References

Aggleton, J. P., & Brown, M. W. (1999). Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. *Behavioral & Brain Sciences, 22*(3), 425-444; discussion 444-489.

Aggleton, J. P., McMackin, D., Carpenter, K., Hornak, J., Kapur, N., Halpin, S., et al. (2000). Differential cognitive effects of colloid cysts in the third ventricle that spare or compromise the fornix. *Brain, 123*(4), 800-815.

Aggleton, J. P., & Shaw, C. (1996). Amnesia and recognition memory: a re-analysis of psychometric data. *Neuropsychologia, 34*(1), 51-62.

Ahissar, M., & Hochstein, S. (2002). The role of attention in learning simple visual tasks. In M. Fahle (Ed.), *Perceptual Learning* (pp. 253-272). Cambridge, MA: MIT Press.

Altmann, E. M., & Gray, W. D. (2002). Forgetting to remember: The functional relationship of decay and interference. *Psychological Science, 13*(1), 27-33.

Alvarez, P., Zola-Morgan, S., & Squire, L. R. (1995). Damage limited to the hippocampal region produces long-lasting memory impairment in monkeys. *Journal of Neuroscience, 15*(5 Pt 2), 3796-3807.

Amaral, D. G., & Insausti, R. (1990). Hippocampal formation. In G. Paxinos (Ed.), *The Human Nervous System* (pp. 711-755). London: Academic Press.

Amaral, D. G., & Price, J. L. (1984). Amygdalo-cortical projections in the monkey (*macaca fascicularis*). *Journal of Comparative Neurology, 230*, 465-496.

Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review, 96*, 703-719.

Atkinson, R. C., & Juola, J. F. (1974). Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, R. D. Luce & P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (pp. 243-293). San Francisco: Freeman.

Bachevalier, J., & Mishkin, M. (1986). Visual recognition impairment follows ventromedial but not dorsolateral prefrontal lesions in monkeys. *Behavioural Brain Research, 20*(3), 249-261.

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

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology, 18*(4), 362-365.

Baddeley, A. D. (1997). *Human Memory: Theory and Practice*. Philadelphia: Psychology Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47-90). New York: Academic Press.

Baddeley, A. D., & Scott, D. (1971). Short term forgetting in the absence of proactive interference. *Quarterly Journal of Experimental Psychology, 23*, 275-283.

Bahrick, H. P. (1965). Semantic Memory Content in Permastore: Fifty Years of Memory for Spanish Learned in School. *Journal of Experimental Psychology: General, 113*, 1-27.

Barbarotto, R., Laiacona, M., Macchi, V., & Capitani, E. (2002). Picture reality decision, semantic categories and gender: A new set of pictures, with norms and an experimental study. *Neuropsychologia, 40*, 1637-1653.

Bartlett, J. C., Hurry, S., & Thorley, W. (1984). Typicality and familiarity of faces. *Memory & Cognition, 12*(3), 219-228.

Beason-Held, L. L., Rosene, D. L., Killiany, R. J., & Moss, M. B. (1999). Hippocampal formation lesions produce memory impairment in the rhesus monkey. *Hippocampus, 9*(5), 562-574.

Benedek, L., & Juba, A. (1940). Weitere Beitraege zur Frage des anatomischen Substrates des Korsakowchen Symptomen Komplexes. *Archive fuer Psychiatrie und Nervenkrankheiten, 112*, 506-516.

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

Blake, R., Cepeda, N. J., & Hiris, E. (1997). Memory for Visual Motion. *Journal of Experimental Psychology: Human Perception and Performance, 23*(2), 353-369.

Boussaoud, D., Desimone, R., & Ungerleider, L. G. (1991). Visual topography of area TEO in the macaque. *Journal of Comparative Neurology, 306*(4), 554-575.

Bower, G. H. (1972). Stimulus-sampling theory of encoding variability. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 85-123). Washington, DC: V. H. Winston.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433-436.

Brewer, J. B., Zhao, Z., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1998). Making memories: brain activity that predicts how well visual experience will be remembered. *Science, 281*(5380), 1185-1187.

Bricolo, E., & Bulthoff, H. H. (1993). Rotation, translation, size and illumination invariances in 3D object recognition. *Investigative Ophthalmology & Visual Science, 34*(4), 1081.

Broadbent, D. E., & Broadbent, M. H. P. (1981). Recency effects in visual memory. *Quarterly Journal of Experimental Psychology, 33*, 1-5.

Brooks, D. N. (1972). Memory and head injury. *Journal of Nervous and Mental Disorders, 155*(5), 350-355.

Brooks, D. N. (1974a). Recognition memory after head injury: a signal detection analysis. *Cortex, 10*(3), 224-230.

Brooks, D. N. (1974b). Recognition memory, and head injury. *Journal of Neurology, Neurosurgery & Psychiatry, 37*(7), 794-801.

Brown, A. S., & Mitchell, D. B. (1994). A reevaluation of semantic versus nonsemantic processing in implicit memory. *Memory & Cognition, 22*, 533-541.

Brown, G. D. A., & Hulme, C. (1995). Modeling item length effects in memory span: no rehearsal needed? *Journal of Memory & Language, 34*, 594-621.

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review, 107*, 127-181.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology, 10*, 12-21.

Brown, M. W., & Aggleton, J. P. (2001). Recognition memory: What are the roles of the perirhinal cortex and hippocampus? *Nature Reviews, 2*, 51-61.

Brown, M. W., Wilson, F. A., & Riches, I. P. (1987). Neuronal evidence that inferomedial temporal cortex is more important than hippocampus in certain processes underlying recognition memory. *Brain Research, 409*(1), 158-162.

Brown, M. W., & Xiang, J.-Z. (1998). Recognition memory: Neuronal substrates of the judgement of prior occurrence. *Progress in Neurobiology, 55*, 149-189.

Brozinsky, C. J., Yonelinas, A. P., Kroll, N. E., & Ranganath, C. (2005). Lag-sensitive repetition suppression effects in the anterior parahippocampal gyrus. *Hippocampus, 15*(5), 557-561.

Bruyer, R. (1986). Lateral differences in visual processing: relative vs exclusive hemispheric specialization. *Human Neurobiology, 5*(2), 83-86.

Buckley, M. J. (2005). The role of the perirhinal cortex and hippocampus in learning, memory, and perception. *Quarterly Journal of Experimental Psychology B, 58*(3-4), 246-268.

Buckley, M. J., & Gaffan, D. (1997). Impairment of visual object-discrimination learning after perirhinal cortex ablation. *Behavioral Neuroscience, 111*(3), 467-475.

Buckley, M. J., & Gaffan, D. (1998a). Learning and transfer of object-reward associations and the role of perirhinal cortex. *Behavioral Neuroscience, 112*(1), 15-23.

Buckley, M. J., & Gaffan, D. (1998b). Perirhinal cortex ablation impairs configural learning and paired-associate learning equally. *Neuropsychologia, 36*(6), 535-546.

Buckley, M. J., & Gaffan, D. (1998c). Perirhinal cortex ablation impairs visual object identification. *Journal of Neuroscience, 18*(6), 2268-2275.

Buckner, R. L., Petersen, S. E., Ojemann, J. G., Miezin, F. M., Squire, L. R., & Raichle, M. E. (1995). Functional anatomical studies of explicit and implicit memory retrieval tasks. *Journal of Neuroscience, 15*(1 Pt 1), 12-29.

Buffalo, E. A., Ramus, S. J., Clark, R. E., Teng, E., Squire, L. R., & Zola, S. M. (1999). Dissociation between the effects of damage to the perirhinal cortex and area TE. *Learning & Memory, 6*, 572-599.

Buffalo, E. A., Stefanacci, L., Squire, L. R., & Zola, S. M. (1998). A reexamination of the concurrent discrimination learning task: the importance of anterior inferotemporal cortex, area TE. *Behavioral Neuroscience, 112*(1), 3-14.

Busey, T. A., & Tunnicliff, J. L. (1999). Accounts of blending, distinctiveness and typicality in the false recognition of faces. *Journal of Experimental Psychology: Learning, Memory & Cognition, 25*, 1210-1235.

Bussey, T. J., Saksida, L. M., & Murray, E. A. (2002). Perirhinal cortex resolves feature ambiguity in complex visual discriminations. *European Journal of Neuroscience, 15*(2), 365-374.

Bussey, T. J., Saksida, L. M., & Murray, E. A. (2003). Impairments in visual discrimination after perirhinal cortex lesions: testing 'declarative' vs. 'perceptual-mnemonic' views of perirhinal cortex function. *European Journal of Neuroscience, 17*(3), 649-660.

Bussey, T. J., Saksida, L. M., & Murray, E. A. (2005). The perceptual-mnemonic/feature conjunction model of perirhinal cortex function. *Quarterly Journal of Experimental Psychology B, 58*(3-4), 269-282.

Challis, B. H., & Brodbeck, D. R. (1992). Level of processing effects priming in word-fragment completion. *Journal of Experimental Psychology: Learning, Memory & Cognition, 18*, 595-607.

Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurones in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology, 80*(6), 2918-2940.

Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature, 363*(6427), 345-347.

Clarys, D. (2001). The psychology of human memory: Theoretical and methodological progress. *Annee Psychologique, 101*(3), 495-519.

Clayton, N. S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature, 395*(6699), 272-274.

Cohen, G. (1972). Serial position effects in the recall of picture sequences. *Quarterly Journal of Experimental Psychology, 24*(1), 41-47.

Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: dissociation of knowing how and knowing that. *Science, 210*(4466), 207-210.

Collishaw, S. M., & Hole, G. J. (2002). Is there a linear or a nonlinear relationship between rotation and configural processing of faces? *Perception, 31*(3), 287-296.

Coltheart, V. (1999). Comparing short-term memory and memory for rapidly presented visual stimuli. *International Journal of Psychology, 34*, 293-300.

Coltheart, V., Mondy, S., Dux, P. E., & Stephenson, L. (2004). Effects of orthographic and phonological word length on memory for lists shown at RSVP and STM rates. *Journal of Experimental Psychology: Learning, Memory & Cognition, 30*(4), 815-826.

Coney, J., & MacDonald, S. (1988). The Effect of Retention Interval Upon Hemispheric Processes in Recognition Memory. *Neuropsychologia, 26*(2), 287-295.

Cooper, E. E., & Brooks, B. E. (2004). Qualitative differences in the representation of spatial relations for different object classes. *Journal of Experimental Psychology: Human Perception & Performance, 30*(2), 243-256.

Cooper, E. E., & Wojan, T. J. (2000). Differences in the coding of spatial relations in face identification and basic-level object recognition. *Journal of Experimental Psychology: Learning, Memory & Cognition, 26*(2), 470-488.

Corballis, M. C., Zbrodoff, N. J., Shetzer, L. I., & Butler, P. B. (1978). Decisions about identity and orientation of rotated letters and digits. *Memory & Cognition, 6*(2), 98-107.

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

Cowan, N. (2005). *Working memory capacity*. New York: Psychology Press.

Crick, F. (1984). Function of the thalamic reticular complex: the searchlight hypothesis. *Proceedings of the National Academy of Sciences of the USA, 81*(14), 4586-4590.

Curran, T. (1999). The electrophysiology of incidental and intentional retrieval: ERP old/new effects in lexical decision and recognition memory. *Neuropsychologia, 37*(7), 771-785.

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition, 28*(6), 923-938.

Cutting, J. (1978). Patterns of performance in amnesic subjects. *Journal of Neurology, Neurosurgery & Psychiatry, 41*(3), 278-282.

Dale, H. C., & Gregory, M. (1966). Evidence of semantic coding in short-term memory. *Psychological Science, 5*, 153-154.

Daum, I., & Ackermann, H. (1994). Frontal-type memory impairment associated with thalamic damage. *International Journal of Neuroscience, 77*(3-4), 187-198.

de Gelder, B., & Rouw, R. (2000). Paradoxical configuration effects for faces and objects in prosopagnosia. *Neuropsychologia, 38*(9), 1271-1279.

De Renzi, E. (1968). Nonverbal memory and hemispheric side of lesion. *Neuropsychologia, 6*, 181-189.

De Valois, R. L., & De Valois, K. K. (1988). *Spatial Vision*. New York: Oxford University Press.

Dean, P. (1976). Effects of inferotemporal lesions on the behavior of monkeys. *Psychological Bulletin, 83*(1), 41-71.

Dearborn, G. V. N. (1899). Recognition under objective reversal. *Psychological Review, 6*, 395-406.

Delay, J., & Brion, S. (1969). *Le syndrome de Korsakoff*. Paris: Masson.

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience, 4*(5), 2051-2062.

Desimone, R., Fleming, J., & Gross, C. G. (1980). Prestriate afferents to inferior temporal cortex: an HRP study. *Brain Research, 184*(1), 41-55.

Desimone, R., & Gross, C. G. (1979). Visual areas in the temporal cortex of the macaque. *Brain Research, 178*, 363-380.

Desimone, R., & Ungerleider, L. G. (1989). Neural mechanisms of visual perception in monkeys. In F. Boller & J. Grafman (Eds.), *Handbook of Neuropsychology* (Vol. 2, pp. 267-299). Amsterdam: Elsevier.

Devinsky, O., & Luciano, D. (1993). The contribution of cingulate cortex to human behavior. In B. A. Vogt & M. Gabriel (Eds.), *Neurobiology of Cingulate Cortex and Limbic Thalamus* (pp. 528-556). Boston: Birkhauser.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115*, 107-117.

DiCarlo, J. J., & Maunsell, J. H. R. (2003). Anterior inferotemporal neurons of monkeys engaged in object recognition can be highly sensitive to object retinal position. *Journal of Neurophysiology, 89*, 3264-3278.

Dill, M. (2002). Specificity versus invariance of perceptual learning: The example of position. In M. Fahle (Ed.), *Perceptual Learning* (pp. 219-231). Cambridge, MA: MIT Press.

Dill, M., & Edelman, S. (2001). Imperfect invariance to object translation in the discrimination of complex shapes. *Perception, 30*(6), 707-724.

Dill, M., & Fahle, M. (1997). The role of visual field position in pattern-discrimination learning. *Proceedings of the Royal Society, London B, 264*, 1031-1036.

Dill, M., & Fahle, M. (1998). Limited translation invariance of human visual pattern recognition. *Perception & Psychophysics, 60*(1), 65-81.

Distler, C., Boussaoud, D., Desimone, R., & Ungerleider, L. G. (1993). Cortical connections of inferior temporal area TEO in macaque monkeys. *Journal of Comparative Neurology, 334*(1), 125-150.

Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition, 24*(4), 523-533.

Doty, R. W., & Savakis, A. E. (1997). Commonality of processes underlying visual and verbal recognition memory. *Cognitive Brain Research, 5*, 283-294.

Drake, A. I., & Hannay, H. J. (1992). Continuous recognition memory tests: are the assumptions of the theory of signal detection met? *Journal of Clinical & Experimental Neuropsychology, 14*(4), 539-544.

Dudai, Y. (2002). *Memory: From A to Z*. Oxford: Oxford University Press.

Dudai, Y. (2004). The neurobiology of consolidations, or, How stable is the engram? *Annual Review of Psychology, 55*, 51-86.

Duzel, E., Yonelinas, A. P., Mangun, G. R., Heinze, H. J., & Tulving, E. (1997). Event-related brain potential correlates of two states of conscious awareness in memory. *Proceedings of the National Academy of Sciences of the USA, 94*(11), 5973-5978.

Eacott, M. J., Gaffan, D., & Murray, E. A. (1994). Preserved recognition memory for small sets, and impaired stimulus identification for large sets, following rhinal cortex ablations in monkeys. *European Journal of Neuroscience, 6*(9), 1466-1478.

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology* (H. A. Ruger & C. E. Bussenius, Trans.). New York: Dover.

Eichenbaum, H. (2000). A cortical-hippocampal system for declarative memory. *Nature Reviews: Neuroscience, 1*(1), 41-50.

Eldridge, L. L., Knowlton, B. J., Furmanski, C. S., Bookheimer, S. Y., & Engel, S. A. (2000). Remembering episodes: a selective role for the hippocampus during retrieval. *Nature Neuroscience, 3*(11), 1149-1152.

Eley, M. G. (1982). Identifying rotated letter-like symbols. *Memory & Cognition, 10*(1), 25-32.

Ellis, H. D. (1983). The role of the right hemisphere in face perception. In A. W. Young (Ed.), *Functions of the Right Hemisphere*. London: Academic Press.

Ellis, H. D. (1986). Processes underlying face recognition. In R. Bruyer (Ed.), *The Neuropsychology of Face Perception and Facial Expression*. Hillsdale, NJ: Erlbaum.

Engel, A. K., Konig, P., Kreiter, A. K., Schillen, T. B., & Singer, W. (1992). Temporal coding in the visual cortex: new vistas on integration in the nervous system. *Trends in Neuroscience, 15*(6), 218-226.

Erickson, C. A., Jagadeesh, B., & Desimone, R. (2000). Clustering of perirhinal neurons with similar properties following visual experience in adult monkeys. *Nature Neuroscience, 3*(11), 1143-1148.

Erickson, R. P. (1968). Stimulus coding in topographic and nontopographic afferent modalities: on the significance of the activity of individual sensory neurons. *Psychological Review, 75*(6), 447-465.

Ericsson, K. A., & Kintsch, W. (1995). Long-Term Working Memory. *Psychological Review, 102*, 211-245.

Estes, W. K. (1986). Array models for category learning. *Cognitive Psychology, 18*(4), 500-549.

Estes, W. K., & Maddox, W. T. (1995a). Interactions of Stimulus Attributes, Base Rates, and Feedback in Recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*(5), 1075-1095.

Estes, W. K., & Maddox, W. T. (1995b). Interactions of stimulus attributes, base rates, and feedback in recognition. *Journal of Experimental Psychology: Learning, Memory & Cognition, 21*(5), 1075-1095.

Fahy, F. L., Riches, I. P., & Brown, M. W. (1993). Neuronal activity related to visual recognition memory: long-term memory and the encoding of recency and familiarity information in the primate anterior and medial inferior temporal and rhinal cortex. *Experimental Brain Research, 96*, 457-472.

Farah, M. J., Rabinowitz, C., Quinn, G. E., & Liu, G. T. (2000). Early commitment of neural substrates for face recognition. *Cognitive Neuropsychology, 17*, 117-124.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1*(1), 1-47.

Fine, I., & Jacobs, R. A. (2002). Comparing Perceptual Learning Across Tasks: A Review. *Journal of Vision, 2*, 190-203.

Foster, D. H. (1978). Visual comparison of random-dot patterns: evidence concerning a fixed visual association between features and feature-relations. *Quarterly Journal of Experimental Psychology, 30*(4), 637-654.

Fuchs, K. L., Hannay, H. J., Huckeba, W. M., & Espy, K. A. (1999). Construct validity of the Continuous Recognition Memory test. *Clinical Neuropsychology, 13*(1), 54-65.

Fujita, I., Tanaka, K., Ito, M., & Cheng, K. (1992). Columns for visual features of objects in monkey inferotemporal cortex. *Nature, 360*, 343-346.

Gabrieli, J. D., Brewer, J. B., Desmond, J. E., & Glover, G. H. (1997). Separate neural bases of two fundamental memory processes in the human medial temporal lobe. *Science, 276*(5310), 264-266.

Gaffan, D. (1977). Monkey's recognition memory for complex pictures and the effect of fornix transection. *Quarterly Journal of Experimental Psychology, 29*, 505-514.

Gaffan, D. (1992). Amnesia for complex naturalistic scenes and for objects following fornix transection in the rhesus monkey. *European Journal of Neuroscience, 4*(5), 381-388.

Gaffan, D. (1994a). Dissociated effects of perirhinal cortex ablation, fornix transection and amygdalectomy: evidence for multiple memory systems in the primate temporal lobe. *Experimental Brain Research, 99*(3), 411-422.

Gaffan, D. (1994b). Scene-specific memory for objects: a model of episodic memory impairment in monkeys with fornix transection. *Journal of Cognitive Neuroscience, 6*, 305-320.

Gaffan, D., & Harrison, S. (1989). Place memory and scene memory: effects of fornix transection in the monkey. *Experimental Brain Research, 74*(1), 202-212.

Gaffan, D., Harrison, S., & Gaffan, E. A. (1986). Visual identification following inferotemporal ablation in the monkey. *Quarterly Journal of Experimental Psychology B, 38*(1), 5-30.

Gaffan, D., & Murray, E. A. (1992). Monkeys (Macaca fascicularis) with rhinal cortex ablations succeed in object discrimination learning despite 24-hr intertrial intervals and fail at matching to sample despite double sample presentations. *Behavioral Neuroscience, 106*(1), 30-38.

Gaffan, D., & Parker, A. (1996). Interaction of perirhinal cortex with the fornix-fimbria: memory for objects and "object-in-place" memory. *Journal of Neuroscience, 16*(18), 5864-5869.

Gaffan, D., Parker, A., & Easton, A. (2001). Dense amnesia in the monkey after transection of fornix, amygdala and anterior temporal stem. *Neuropsychologia, 39*(1), 51-70.

Gardiner, J. M., & Java, R. I. (1993). Recognizing and remembering. In A. Collins, M. A. Conway, S. E. Gathercole & P. E. Morris (Eds.), *Theories of memory* (pp. 163-188). Hillsdale, NJ: Erlbaum.

Gardiner, J. M., & Parkin, A. J. (1990). Attention and recollective experience in recognition memory. *Memory & Cognition, 18*(6), 579-583.

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (1999). Response deadline and subjective awareness in recognition memory. *Consciousness & Cognition, 8*(4), 484-496.

Gardiner, J. M., & Richardson-Klavehn, A. (2000). Remembering and knowing. In E. Tulving & F. I. M. Craik (Eds.), *Oxford handbook of memory* (pp. 229-244). New York: Oxford University Press.

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

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91*(1), 1-67.

Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory & Cognition, 13*(1), 8-20.

Glanzer, M., & Adams, J. K. (1990). The Mirror Effect in Recognition Memory: Data and Theory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16*, 5-16.

Glanzer, M., Adams, J. K., Iverson, G. J., & Kim, K. (1993). The regularities of recognition memory. *Psychological Review, 100*(3), 546-567.

Glanzer, M., & Bowles, N. (1976). Analysis of the word-frequency effect in recognition memory. *Journal of Experimental Psychology, 61*, 23-29.

Goldstein, A. G., & Chance, J. E. (1970). Visual recognition memory for complex configurations. *Perception & Psychophysics, 9*(2B), 237-241.

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

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

Gratton, G., Corballis, P. M., & Jain, S. (1997). Hemispheric organization of visual memories. *Journal of Cognitive Neuroscience, 9*(1), 92-104.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: John Wiley and Sons.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences, 10*(1), 14-23.

Grill-Spector, K., & Kanwisher, N. (2005). Visual recognition: as soon as you know it is there, you know what it is. *Psychological Science, 16*(2), 152-160.

Gronlund, S. D., & Ratcliff, R. (1989). Time course of item and associative information: implications for global memory models. *Journal of Experimental Psychology: Learning, Memory & Cognition, 15*(5), 846-858.

Gross, C. G., & Mishkin, M. (1977). The neural basis of stimulus equivalence across retinal translation. In S. Haxnad, R. W. Doty, L. Goldstein, J. Jaynes & G. Krauthammer (Eds.), *Lateralization in the nervous system* (pp. 109-122). New York: Academic Press.

Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of Neurophysiology, 35*, 96-111.

Guillem, F., N'Kaoua, B., Rougier, A., & Claverie, B. (1998). Location of the epileptic zone and its physiopathological effects on memory-related activity of the temporal lobe structures: a study with intracranial event-related potentials. *Epilepsia, 39*(9), 928-941.

Haist, F., Shimamura, A. P., & Squire, L. R. (1992). On the relationship between recall and recognition memory. *Journal of Experimental Psychology: Learning, Memory & Cognition, 18*(4), 691-702.

Hankey, G. J., & Stewart-Wynne, E. G. (1988). Amnesia following thalamic hemorrhage. Another stroke syndrome. *Stroke, 19*(6), 776-778.

Hanley, J. R., & Davies, A. D. M. (1997). Impaired recall and preserved cognition. In A. J. Parkin (Ed.), *Case Studies in the Neuropsychology of Memory* (pp. 111-126). Hove: Psychology Press.

Hannay, H. J., & Levin, H. S. (1988). *The Continuous Recognition Memory test: A manual*: [Available from the first author].

Hannay, H. J., & Levin, H. S. (1989). Visual Continuous Recognition Memory in normal and closed-head-injured adolescents. *Journal of Clinical & Experimental Neuropsychology, 11*(4), 444-460.

Hannay, H. J., Levin, H. S., & Grossman, R. G. (1979). Impaired recognition memory after head injury. *Cortex, 15*(2), 269-283.

Henson, R. N., Rugg, M. D., Shallice, T., Josephs, O., & Dolan, R. J. (1999). Recollection and familiarity in recognition memory: an event-related functional magnetic resonance imaging study. *Journal of Neuroscience, 19*(10), 3962-3972.

Hines, D. (1975). Immediate and delayed recognition of sequentially presented random shapes. *Journal of Experimental Psychology: Human Learning & Memory, 1*(5), 634-639.

Hintzman, D. L. (1969). Recognition time: Effects of recency, frequency, and the spacing of repetitions. *Journal of Experimental Psychology, 79*(1), 192-194.

Hintzman, D. L. (1984). Episodic versus semantic memory: a distinction whose time has come - and gone? *Behavioral & Brain Sciences, 7*, 240-241.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review, 93*, 411-428.

Hintzman, D. L. (1988). Judgements of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95*, 528-551.

Hintzman, D. L., Caulton, D. A., & Levitin, D. J. (1998). Retrieval dynamics in recognition and list discrimination: further evidence of separate processes of familiarity and recall. *Memory & Cognition, 26*(3), 449-462.

Hirshman, E., & Master, S. (1997). Modeling the conscious correlates of recognition memory: reflections on the remember-know paradigm. *Memory & Cognition, 25*(3), 345-351.

Hirst, W., Johnson, M. K., Kim, J. K., Phelps, E. A., Risse, G., & Volpe, B. T. (1986). Recognition and recall in amnesics. *Journal of Experimental Psychology: Learning, Memory & Cognition, 12*(3), 445-451.

Hockley, W. E. (1982). Retrieval processes in continuous recognition. *Journal of Experimental Psychology: Learning, Memory & Cognition, 8*(6), 497-512.

Hockley, W. E. (1992). Item versus associative information: Further comparisons of forgetting rates. *Journal of Experimental Psychology: Learning, Memory & Cognition, 18*(6), 1321-1330.

Hockley, W. E., & Murdock, B. B. (1987). A decision model for accuracy and response latency in recognition memory. *Psychological Review, 94*, 351-358.

Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia. Progressive fluent aphasia with temporal lobe atrophy. *Brain, 115*(6), 1783-1806.

Holdstock, J. S., Mayes, A. R., Cezayirli, E., Isaac, C. L., Aggleton, J. P., & Roberts, N. (2000). A comparison of egocentric and allocentric spatial memory in a patient with selective hippocampal damage. *Neuropsychologia, 38*(4), 410-425.

Holt, E. B. (1903). Eye movements and central anaesthesia. *Psychological Review, 4*, 3-5.

Hornak, J., Duncan, J., & Gaffan, D. (2002). The role of the vertical meridian in visual memory for objects. *Neuropsychologia, 40*, 1873-1880.

Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: evidence for a long-term memory contribution to short-term memory span. *Journal of Memory & Language, 30*, 685-701.

Humphreys, M. S. (1978). Item and relational information: a case for context independent retrieval. *Journal of Verbal Learning &Verbal Behavior, 17*, 179-187.

Humphreys, M. S., & Bain, J. D. (1983). Recognition memory: a cue and information analysis. *Memory & Cognition, 11*(6), 583-600.

Hunt, R. R. (1995). The subtlety of distinctiveness: What von Restorff really did. *Psychonomic Bulletin & Review, 2*, 105-112.

Huxlin, K. R., Saunders, R. C., Marchionini, D., Pham, H. A., & Merigan, W. H. (2000). Perceptual deficits after lesions of inferotemporal cortex in macaques. *Cerebral Cortex, 10*(7), 671-683.

Inoue, C., & Bellezza, F. S. (1998). The detection model of recognition using know and remember judgments. *Memory & Cognition, 26*(2), 299-308.

Insausti, R., Amaral, D. G., & Cowan, W. M. (1987). The entorhinal cortex of the monkey: II. Cortical afferents. *Journal of Comparative Neurology, 264*(3), 356-395.

Ito, M., Tamura, H., Fujita, I., & Tanaka, K. (1995). Size and position invariance of neuronal responses in monkey inferotemporal cortex. *Journal of Neurophysiology, 73*(1), 218-226.

Iwai, E., & Mishkin, M. (1969). Further evidence on the locus of the visual area in the temporal lobe of the monkey. *Experimental Neurology, 25*(4), 585-594.

Iwai, E., & Yukie, M. (1987). Amygdalofugal and amygdalopetal connections with modality-specific visual cortical areas in macaques (Macaca fuscata, M. mulatta, and M. fascicularis). *Journal of Comparative Neurology, 261*(3), 362-387.

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

Janowsky, J. S., Shimamura, A. P., Kritchevsky, M., & Squire, L. R. (1989). Cognitive impairment following frontal lobe damage and its relevance to human amnesia. *Behavioral Neuroscience, 103*(3), 548-560.

Jessen, F., Flacke, S., Granath, D. O., Manka, C., Scheef, L., Papassotiropoulos, A., et al. (2001). Encoding and retrieval related cerebral activation in continuous verbal recognition. *Brain Research: Cognitive Brain Research, 12*(2), 199-206.

Jiang, Y., Haxby, J. V., Martin, A., Ungerleider, L. G., & Parasuraman, R. (2000). Complementary neural mechanisms for tracking items in human working memory. *Science, 287*(5453), 643-646.

Johnson, C. J., Paivio, A., & Clark, J. M. (1996). Cognitive components of picture naming. *Psychological Bulletin, 120*(1), 113-139.

Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory & Cognition, 13*(4), 289-303.

Jolicoeur, P., Gluck, M. A., & Kosslyn, S. (1984). Pictures and names: Making the connection. *Cognitive Psychology, 16*, 243-275.

Jones, T. C., & Atchley, P. (2002). Conjunction error rates on a continuous recognition memory test: little evidence for recollection. *Journal of Experimental Psychology: Learning, Memory & Cognition, 28*(2), 374-379.

Joray, S., Herrmann, F., Mulligan, R., & Schnider, A. (2004). Mechanism of disorientation in Alzheimer's disease. *European Neurology, 52*(4), 193-197.

Juola, J. F., Fischler, I., Wood, C. T., & Atkinson, R. C. (1971). Recognition time for information stored in long-term memory. *Perception and Psychophysics, 10*, 8-14.

Kahana, M. J., & Bennett, P. J. (1994). Classification and Perceived Similarity of Compound Gratings that Differ in Relative Spatial Phase. *Perception & Psychophysics, 55*, 642-656.

Kahana, M. J., & Sekuler, R. (2002). Recognizing spatial patterns: A noisy exemplar approach. *Vision Research, 42*, 2177-2192.

Kahn, J. I., & Foster, D. H. (1981). Visual comparison of rotated and reflected random-dot patterns as a function of their positional symmetry and separation in the field. *Quarterly Journal of Experimental Psychology, 33A*, 155-166.

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.

Kawato, M., Hayakawa, H., & Inui, T. (1993). A forward-inverse optics model of reciprocal connections between visual cortical areas. *Network, 4*, 415-422.

Kelley, W. M., Miezin, F. M., McDermott, K. B., Buckner, R. L., Raichle, M. E., Cohen, N. J., et al. (1998). Hemispheric specialization in human dorsal frontal cortex and medial temporal lobe for verbal and nonverbal memory encoding. *Neuron, 20*, 927-936.

Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning & Verbal Behavior, 1*, 153-161.

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

Kim, M. H., Hong, S. B., & Roh, J. K. (1994). Amnesia following left thalamic infarction. *Journal of Korean Medical Science, 9*, 427-431.

Kim, M. S., Kim, Y. Y., Kim, E. N., Lee, K. J., Ha, T. H., & Kwon, J. S. (2005). Implicit and explicit memory in patients with obsessive-compulsive disorder: An event-related potential study. *Journal of Psychiatric Research*.

Kim, M. S., Kwon, J. S., Kang, S. S., Youn, T., & Kang, K. W. (2004). Impairment of recognition memory in schizophrenia: event-related potential study using a continuous recognition task. *Psychiatry & Clinical Neurosciences, 58*(5), 465-472.

Kimura, D. (1963). Right temporal-lobe damage. Perception of unfamiliar stimuli after damage. *Archives of Neurology, 8*, 264-271.

Kintsch, W. (1967). Memory and decision aspects of recognition learning. *Psychological Review, 74*(6), 496-504.

Kintsch, W. (1970). Models for Free Recall and Recognition. In D. A. Norman (Ed.), *Models of Human Memory*. New York: Academic Press.

Knowlton, B. J., & Squire, L. R. (1995). Remembering and knowing: two different expressions of declarative memory. *Journal of Experimental Psychology: Learning, Memory & Cognition, 21*(3), 699-710.

Kobatake, E., & Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology, 71*(3), 856-867.

Kobatake, E., Tanaka, K., Wang, G., & Tamori, Y. (1993). Effects of adult learning on the stimulus selectivity of cells in the inferotemporal cortex. *Society of Neuroscience Abstracts, 19*, 975.

Komatsu, H., & Ideura, Y. (1993). Relationships between color, shape, and pattern selectivities of neurons in the inferior temporal cortex of the monkey. *Journal of Neurophysiology, 70*(2), 677-694.

Kopelman, M. D. (1995). The Korsakoff syndrome. *British Journal of Psychiatry, 166*(2), 154-173.

Korsnes, M. S., Magnussen, S., & Reinvang, I. (1996). Serial position effects in visual short-term memory for words and abstract spatial patterns. *Scandinavian Journal of Psychology, 37*, 62-73.

Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J. P., Thompson, W. L., et al. (1999). The Role of Area 17 in Visual Imagery: Convergent Evidence from PET and rTMS. *Science, 284*, 167-170.

Larrabee, G. J., Trahan, D. E., & Curtiss, G. (1992). Construct validity of the Continuous Visual Memory test. *Archives of Clinical Neuropsychology, 7*(5), 395-405.

Larsen, A., & Bundesen, C. (1998). Effects of spatial separation in visual pattern matching: evidence on the role of mental translation. *Journal of Experimental Psychology: Human Perception & Performance, 24*(3), 719-731.

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

Lee, A. C., Barense, M. D., & Graham, K. S. (2005). The contribution of the human medial temporal lobe to perception: bridging the gap between animal and human studies. *Quarterly Journal of Experimental Psychology B, 58*(3-4), 300-325.

Lee, A. C. H., Robbins, T. W., Pickard, J. D., & Owen, A. M. (2000). Asymmetric frontal activation during episodic memory: The effects of stimulus type on encoding and retrieval. *Neuropsychologia, 38*, 677-692.

Lehmann, S., & Murray, M. M. (2005). The role of multisensory memories in unisensory object discrimination. *Brain Research: Cognitive Brain Research, 24*(2), 326-334.

Lewine, J. D. (1989). The Temporal Dynamics of Event Memory: A Stage Analysis of Mnemonic Processing by Man and Macaque. *Journal of Cognitive Neuroscience, 1*, 356-371.

Lewis, M. B., & Johnston, R. A. (1998). Understanding caricatures of faces. *Quarterly Journal of Experimental Psychology A, 51*(2), 321-346.

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

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

Lockhead, G. R. (1970). Identification and the form of multidimensional discrimination space. *Journal of Experimental Psychology, 85*(1), 1-10.

Logothetis, N. K., & Pauls, J. (1995). Psychophysical and physiological evidence for viewer-centered object representations in the primate. *Cerebral Cortex, 5*(3), 270-288.

Logothetis, N. K., Pauls, J., Buelthoff, H. H., & Poggio, T. (1994). View-dependent object recognition by monkeys. *Current Biology, 4*, 401-414.

Logothetis, N. K., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. *Current Biology, 5*(5), 552-563.

Logothetis, N. K., & Sheinberg, D. L. (1996). Visual object recognition. *Annual Review of Neuroscience, 19*, 577-621.

Lueschow, A., Miller, E. K., & Desimone, R. (1994). Inferior temporal mechanisms for invariant object recognition. *Cerebral Cortex, 4*(5), 523-531.

Magnussen, S. (2000). Low-Level Memory Processes in Vision. *Trends in Neuroscience, 23*, 247-251.

Mandler, G. (1980). Recognizing: The judgement of previous occurrence. *Psychological Review, 87*(3), 252-271.

Manning, S. K., & Schreier, H. (1988). Recency and suffix effects in pictures as a function of recall method. *American Journal of Psycholgy, 101*(1), 97-109.

Mayes, A. R., Van Eijk, R., Gooding, P. A., Isaac, C. L., & Holdstock, J. S. (1999). What are the functional deficits produced by hippocampal and perirhinal lesions? *Behavioral & Brain Sciences, 22*, 36-37.

McBride, D. A., & Dosher, B. A. (1997). A comparison of forgetting in an implicit and explicit memory task. *Journal of Experimental Psychology: General, 126*(4), 371-392.

McDermott, K. B., Buckner, R. L., Peterson, 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*, 631-640.

McElree, B., Dolan, P. O., & Jacoby, L. L. (1999). Isolating the contributions of familiarity and source information to item recognition: a time course analysis. *Journal of Experimental Psychology: Learning, Memory & Cognition, 25*(3), 563-582.

McFarland, D. J., & Cacace, A. T. (1995). Comparisons of memory for nonverbal auditory and visual sequential stimuli. *Psychological Research, 57*(2), 80-87.

McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological Review, 39*, 352-370.

McMackin, D., Cockburn, J., Anslow, P., & Gaffan, D. (1995). Correlation of fornix damage with memory impairment in six cases of colloid cyst removal. *Acta Neurochir (Wien), 135*(1-2), 12-18.

Menon, V., White, C. D., Eliez, S., Glover, G. H., & Reiss, A. L. (2000). Analysis of a distributed neural system involved in spatial information, novelty, and memory processing. *Human Brain Mapping, 11*(2), 117-129.

Meunier, M., Bachevalier, J., Mishkin, M., & Murray, E. A. (1993). Effects on visual recognition of combined and separate ablations of the entorhinal and perirhinal cortex in rhesus monkeys. *Journal of Neuroscience, 13*(12), 5418-5432.

Meunier, M., Hadfield, W., Bachevalier, J., & Murray, E. A. (1996). Effects of rhinal cortex lesions combined with hippocampectomy on visual recognition memory in rhesus monkeys. *Journal of Neurophysiology, 75*(3), 1190-1205.

Miller, E., & Lewis, P. (1977). Recognition memory in elderly patients with depression and dementia: a signal detection analysis. *Journal of Abnormal Psychology, 86*(1), 84-86.

Miller, E. K. (2000). Organization through experience. *Nature Neuroscience, 3*(11), 1066-1068.

Miller, E. K., & Desimone, R. (1993). Scopolamine affects short-term memory but not inferior temporal neurons. *Neuroreport, 4*(1), 81-84.

Miller, E. K., & Desimone, R. (1994). Parallel neuronal mechanisms for short-term memory. *Science, 263*(5146), 520-522.

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

Miller, E. K., Li, L., & Desimone, R. (1993). Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *Journal of Neuroscience, 13*(4), 1460-1478.

Mishkin, M. (1966). Visual mechanisms beyond the striate cortex. In R. Russell (Ed.), *Frontiers of Physiological Psychology* (pp. 93-119): Academic Press.

Mishkin, M. (1982). A memory system in the monkey. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 298*(1089), 83-95.

Miyashita, Y. (1988). Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nature, 335*(6193), 817-820.

Miyashita, Y. (2000). Visual associative long-term memory: encoding and retrieval in inferotemporal cortex of the primate. In M. Gazzaniga (Ed.), *The New Cognitive Neurosciences* (pp. 379-392). Cambridge, MA: MIT Press.

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.

Miyashita, Y., Higuchi, S.-I., Sakai, K., & Masui, N. (1991). Generation of fractal patterns for probing the visual memory. *Neuroscience Research, 12*, 307-311.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science, 229*(4715), 782-784.

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

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 Psychology, 9*(5), 555-604.

Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review, 89*(6), 609-626.

Murdock, B. B. J. (1960). The distinctiveness of stimuli. *Psychological Review, 67*, 16-31.

Murdock, B. B. J. (1961). The retention of individual items. *Journal of Experimental Psychology, 62*(6), 618-625.

Murdock, B. B. J., & Dufty, P. O. (1972). Strength theory and recognition memory. *Journal of Experimental Psychology, 94*, 284-290.

Murdock, B. B. J., & Hockley, W. E. (1989). Short-term memory for associations. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 24, pp. 71-108). San Diego, CA: Academic Press.

Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology, 78*, 679-684.

Murray, E. A., & Bussey, T. J. (1999). Perceptual-mnemonic functions of the perirhinal cortex. *Trends in Cognitive Sciences, 3*(4), 142-151.

Murray, E. A., Davidson, M., Gaffan, D., Olton, D. S., & Suomi, S. (1989). Effects of fornix transection and cingulate cortical ablation on spatial memory in rhesus monkeys. *Experimental Brain Research, 74*(1), 173-186.

Murray, E. A., & Mishkin, M. (1998). Object recognition and location memory in monkeys with excitotoxic lesions of the amygdala and hippocampus. *Journal of Neuroscience, 18*(16), 6568-6582.

Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology, 53*, 53-81.

Nazir, T. A., & O'Regan, J. K. (1990). Some results on tranlsation invariance in the human visual system. *Spatial Vision, 5*(2), 81-100.

Neath, I. (1993). Distinctiveness and serial position effects in recognition. *Memory & Cognition, 21*(5), 689-698.

Nelson, D. L., Reed, V. S., & Walling, J. R. (1976). Pictorial superiority effect. *Journal of Experimental Psychology: Human Learning & Memory, 2*(5), 523-528.

Ng, K. T., Gibbs, M. E., Crow, S. F., Sedman, G. L., Hua, F., Zhao, W., et al. (1991). Molecular Mechanisms of Memory Formation. *Molecular Neurobiology, 5*, 333-350.

Nickerson, R. S. (1965). Short-Term Memory for Complex Meaningful Visual Configurations: A Demonstration of Capacity. *Review of Canadian Psychology, 19*(2), 155-160.

Nickerson, R. S. (1968). A note on long-term recognition memory for pictorial material. *Psychonomic Sciences, 11*(2), 58.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115*(1), 39-61.

Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory & Cognition, 14*, 700-708.

Nosofsky, R. M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Human Perception & Performance, 17*(1), 3-27.

Nosofsky, R. M. (1992). Exemplar-based approach to relating categorization, identification, and recognition. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition*. Hillsdale, NJ: Lawrence Erlbaum.

Nosofsky, R. M., & Zaki, S. R. (2003). A hybrid-similarity exemplar model for predicting distinctiveness effects in perceptual old-new recognition. *Journal of Experimental Psychology: Learning, Memory & Cognition, 29*(6), 1194-1209.

Okada, R. (1971). Decision latencies in short-term recognition memory. *Journal of Experimental Psychology, 90*(1), 27-32.

Owen, A. M., Sahakian, B. J., Semple, J., Polkey, C. E., & Robbins, T. W. (1995). Visuo-spatial short-term recognition memory and learning after temporal lobe excisions, frontal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia, 33*(1), 1-24.

Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.

Paivio, A., Clark, J. M., Digdon, N., & Bons, T. (1989). Referential processing: Reciprocity and correlates of naming and imaging. *Memory & Cognition, 17*(2), 163-174.

Papez, J. W. (1995). A proposed mechanism of emotion. 1937. *Journal of Neuropsychiatry and Clinical Neuroscience, 7*(1), 103-112.

Parker, A., & Gaffan, D. (1997a). The effect of anterior thalamic and cingulate cortex lesions on object-in-place memory in monkeys. *Neuropsychologia, 35*(8), 1093-1102.

Parker, A., & Gaffan, D. (1997b). Mamillary body lesions in monkeys impair object-in-place memory: functional unity of the fornix-mamillary system. *Journal of Cognitive Neuroscience, 9*, 512-521.

Parker, A., & Gaffan, D. (1998a). Interaction of frontal and perirhinal cortices in visual object recognition memory in monkeys. *European Journal of Neuroscience, 10*, 3044-3057.

Parker, A., & Gaffan, D. (1998b). Memory systems in primates: Episodic, semantic, and perceptual learning. In A. D. Milner (Ed.), *Comparative Neuropsychology* (pp. 109-126). Oxford: Oxford University Press.

Parkin, A. J., Dunn, J. C., Lee, C., O'Hara, P. F., & Nussbaum, L. (1993). Neuropsychological sequelae of Wernicke's encephalopathy in a 20-year-old woman: selective impairment of a frontal memory system. *Brain & Cognition, 21*(1), 1-19.

Parkin, A. J., Rees, J. E., Hunkin, N. M., & Rose, P. E. (1994). Impairment of memory following discrete thalamic infarction. *Neuropsychologia, 32*(1), 39-51.

Parks, T. E. (1966). Signal-detectability theory of recognition-memory performance. *Psychological Review, 73*(1), 44-58.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58*, 193-198.

Phillips, W. A., & Christie, D. F. (1977). Interference with visualization. *Quarterly Journal of Experimental Psychology, 29*(4), 637-650.

Phillips, W. A., & Christie, D. F. M. (1977). Components of visual memory. *Quarterly Journal of Experimental Psychology, 29*, 117-133.

Pike, R. (1984). Comparison of convolution and matrix distributed memory systems for associative recall and recognition. *Psychological Review, 91*(3), 281-293.

Pike, R., Dalgleish, L., & Wright, L. (1977). A multiple-observations model for response latency and the latencies of correct and incorrect responses in recognition memory. *Memory & Cognition, 5*, 580-589.

Potter, M. C., & Levy, E. I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology, 81*(1), 10-15.

Proverbio, A. M., Burco, F., del Zotto, M., & Zani, A. (2004). Blue piglets? Electrophysiological evidence for the primacy of shape over color in object recognition. *Cognitive Brain Research, 18*, 288-300.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88*, 93-134.

Rainer, G., & Miller, E. K. (2000). Effects of visual experience on the representation of objects in the prefrontal cortex. *Neuron, 27*(1), 179-189.

Rapaczynski, W., & Ehrlichman, H. (1979). Opposite visual hemifield superiorities in face recognition as a function of cognitive style. *Neuropsychologia, 17*(6), 645-652.

Raser, G. A. (1972). False recognition as a function of encoding dimension and lag. *Journal of Experimental Psychology, 93*, 333-337.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*, 59-108.

Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember-know judgements in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory & Cognition, 26*(2), 294-320.

Reynolds, T. J., & Medin, D. L. (1981). Stimulus interaction and between-trials proactive interference in monkeys. *Journal of Experimental Psychology: Animal Behavior Processes, 7*(4), 334-347.

Rhodes, G. (1985). Lateralized processes in face recognition. *British Journal of Psychology, 76 (Pt 2)*, 249-271.

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

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

Rhodes, G., & McLean, I. G. (1990). Distinctiveness and expertise effects with homogeneous stimuli: towards a model of configural coding. *Perception, 19*(6), 773-794.

Riches, I. P., Wilson, F. A., & Brown, M. W. (1991). The effects of visual stimulation and memory on neurons of the hippocampal formation and the neighboring parahippocampal gyrus and inferior temporal cortex of the primate. *Journal of Neuroscience, 11*(6), 1763-1779.

Richmond, B. J., & Sato, T. (1987). Enhancement of inferior temporal neurons during visual discrimination. *Journal of Neurophysiology, 58*(6), 1292-1306.

Riege, W. (1977). Inconstant nonverbal recognition memory in Korsakoff patients and controls. *Neuropsychologia, 15*(2), 269-276.

Riege, W. H., Klane, L. T., Metter, E. J., & Hanson, W. R. (1982). Decision speed and bias after unilateral stroke. *Cortex, 18*(3), 345-355.

Ringo, J. L. (1996). Stimulus specific adaptation in inferior temporal and medial temporal cortex of the monkey. *Behavioural Brain Research, 76*(1-2), 191-197.

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

Roberson, D., & Davidoff, J. (2000). The categorical perception of colors and facial expressions: The effect of verbal interference. *Memory & Cognition, 28*(6), 977-986.

Rock, I. (1973). *Orientation and form*. New York: Academic Press.

Rockland, K. S., Saleem, K. S., & Tanaka, K. (1994). Divergent feedback connections from areas V4 and TEO in the macaque. *Visual Neuroscience, 11*(3), 579-600.

Rockland, K. S., & Van Hoesen, G. W. (1994). Direct temporal-occipital feedback connections to striate cortex (V1) in the macaque monkey. *Cerebral Cortex, 4*(3), 300-313.

Roediger, H. L., 3rd, Weldon, M. S., Stadler, M. L., & Riegler, G. L. (1992). Direct comparison of two implicit memory tests: word fragment and word stem completion. *Journal of Experimental Psychology: Learning, Memory & Cognition, 18*(6), 1251-1269.

Rolls, E. T. (1984). Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces. *Human Neurobiology, 3*(4), 209-222.

Rolls, E. T. (1992). Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. *Philosophical*

Transactions of the Royal Society of London. Series B: Biological Sciences, 335(1273), 11-20; discussion 20-11.

Rolls, E. T., Miyashita, Y., Cahusac, P. M., Kesner, R. P., Niki, H., Feigenbaum, J. D., et al. (1989). Hippocampal neurons in the monkey with activity related to the place in which a stimulus is shown. *Journal of Neuroscience, 9*(6), 1835-1845.

Roodenrys, S., Hulme, C., Alban, J., Ellis, A. W., & Brown, G. D. (1994). Effects of word frequency and age of acquisition on short-term memory span. *Memory & Cognition, 22*, 695-701.

Rosenzweig, M. R., Bennett, E. L., Colombo, P. J., Lee, D. W., & Serrano, P. A. (1993). Short-Term, Intermediate-Term, and Long-Term Memories. *Behavioral Brain Research, 57*, 193-198.

Ross, J., Morrone, M. C., & Burr, D. C. (1997). Compression of visual space before saccades. *Nature, 386*(6625), 598-601.

Rossion, B., Dricot, L., Devolder, A., Bodart, J. M., Crommelinck, M., De Gelder, B., et al. (2000). Hemispheric asymmetries for whole-based and part-based face processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience, 12*(5), 793-802.

Rubin, D. C., Hinton, S., & Wenzel, A. (1999). The precise time course of retention. *Journal of Experimental Psychology: Learning, Memory & Cognition, 25*(5), 1161-1176.

Rubin, D. C., & Wenzel, A. (1996). One Hundred Years of Forgetting: A Quantitative Description of Retention. *Psychological Review, 103*(4), 734-760.

Rugg, M. D., Walla, P., Schloerscheidt, A. M., Fletcher, P. C., Frith, C. D., & Dolan, R. J. (1998). Neural correlates of depth of processing effects on recollection: evidence from brain potentials and positron emission tomography. *Experimental Brain Research, 123*(1-2), 18-23.

Sary, G., Vogels, R., & Orban, G. A. (1993). Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science, 260*(5110), 995-997.

Saslow, M. G. (1967). Latency for saccadic eye movement. *Journal of the Optical Society of America, 57*, 1030-1036.

Schiano, D. J., & Watkins, M. J. (1981). Speech-like coding of pictures in short-term memory. *Memory & Cognition, 9*, 110-114.

Schwartz, E. L., Desimone, R., Albright, T. D., & Gross, C. G. (1983). Shape recognition and inferior temporal neurons. *Proceedings of the National Academy of Sciences of the USA, 80*(18), 5776-5778.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery & Psychiatry, 20*(1), 11-21.

Seacord, L., Gross, C. G., & Mishkin, M. (1979). Role of inferior temporal cortex in interhemispheric transfer. *Brain Research, 167*(2), 259-272.

Seeck, M., Michel, C. M., Mainwaring, N., Cosgrove, R., Blume, H., Ives, J., et al. (1997). Evidence for rapid face recognition from human scalp and intracranial electrodes. *Neuroreport, 8*(12), 2749-2754.

Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Learning and Verbal Behavior, 6*, 156-163.

Shepard, R. N., & Teghtsoonian, M. (1961). Retention of information under conditions approaching a steady state. *Journal of Experimental Psychology, 62*(3), 302-309.

Shiffrin, R. M. (1999). 30 years of memory. In C. Izawa (Ed.), *On Human Memory: Evolution Progress and Reflections of the 30th Anniversary of the Atkinson-Shiffrin Model.* (pp. 17-33). Mahwah, NJ: Erlbaum.

Sigala, N., & Logothetis, N. K. (2002). Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature, 415*(6869), 318-320.

Simion, F., Bagnara, S., Roncato, S., & Umilta, C. (1982). Transformation processes upon the visual code. *Perception & Psychophysics, 31*(1), 13-25.

Simons, J. S., Graham, K. S., Owen, A. M., Patterson, K., & Hodges, J. R. (2001). Perceptual and semantic components of memory for objects and faces: A PET study. *Journal of Cognitive Neuroscience, 13*(4), 430-443.

Singer, W. (1993). Synchronization of cortical activity and its putative role in information processing and learning. *Annual Review of Physiology, 55*, 349-374.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*(2), 174-215.

Sobotka, S., & Ringo, J. L. (1993). Investigation of long-term recognition and association memory in unit responses from inferotemporal cortex. *Experimental Brain Research, 96*(1), 28-38.

Spitzer, H., Desimone, R., & Moran, J. (1988). Increased attention enhances both behavioral and neuronal performance. *Science, 240*(4850), 338-340.

Squire, L. R., & Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science, 253*(5026), 1380-1386.

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences of the USA, 93*(24), 13515-13522.

Squire, L. R., & Zola, S. M. (1998). Episodic memory, semantic memory, and amnesia. *Hippocampus, 8*(3), 205-211.

Standing, L. (1973). Learning 10,000 Pictures. *Quarterly Journal of Experimental Psychology, 25*, 207-222.

Stark, C. E., & Squire, L. R. (2000). fMRI activity in the medial temporal lobe during recognition memory as a function of study-test interval. *Hippocampus, 10*(3), 329-337.

Stern, C. E., Corkin, S., Gonzalez, R. G., Guimaraes, A. R., Baker, J. R., Jennings, P. J., et al. (1996). The hippocampal formation participates in novel picture encoding: evidence from functional magnetic resonance imaging. *Proceedings of the National Academy of Sciences of the USA, 93*(16), 8660-8665.

Sternberg, S. (1966). High-Speed Scanning in Human Memory. *Science, 153*, 652-654.

Strong, E. K., Jr. (1912). The effect of length of series upon recognition memory. *Psychological Review, 19*, 447-462.

Suzuki, W., Saleem, K. S., & Tanaka, K. (2000). Divergent backward projections from the anterior part of the inferotemporal cortex (area TE) in the macaque. *Journal of Comparative Neurology, 422*(2), 206-228.

Suzuki, W. A., Miller, E. K., & Desimone, R. (1997). Object and place memory in the macaque entorhinal cortex. *Journal of Neurophysiology, 78*(2), 1062-1081.

Suzuki, W. A., Zola-Morgan, S., Squire, L. R., & Amaral, D. G. (1993). Lesions of the perirhinal and parahippocampal cortices in the monkey produce long-lasting memory impairment in the visual and tactual modalities. *Journal of Neuroscience, 13*(6), 2430-2451.

Swets, J. A. (1964). *Signal detection and recognition by human observers*. New York: Wiley.

Tanaka, K. (1996). Representation of visual features of objects in inferotemporal cortex. *Neural Networks, 9*(8), 1459-1475.

Tarr, M. J., & Gauthier, I. (1998). Do viewpoint-dependent mechanisms generalize across members of a class? *Cognition, 67*(1-2), 73-110.

Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology, 21*(2), 233-282.

Thapar, A., & Greene, R. L. (1994). Effects of level of processing on implicit and explicit tasks. *Journal of Experimental Psychology Learning, Memory & Cognition, 20*(3), 671-679.

Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York: Teachers College, Columbia University, Bureau of Publications.

Tovee, M. J., & Rolls, E. T. (1995). Information encoding in short firing rate epochs by single neurons in the primate temporal visual cortex. *Visual Cognition, 2*, 35-58.

Tovee, M. J., Rolls, E. T., & Azzopardi, P. (1994). Translation invariance in the responses to faces of single neurons in the temporal visual cortical areas of the alert macaque. *Journal of Neurophysiology, 72*, 1049-1060.

Trahan, D. E., & Larrabee, G. J. (1988). *Continuous Visual Memory test*. Odessa, FL: Psychological Assessment Resources.

Tulving, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.

Tulving, E. (1985). Memory and Consciousness. *Canadian Psychologist, 26*, 1-22.

Tulving, E., Markowitsch, H. J., Kapur, S., Habib, R., & Houle, S. (1994). Novelty encoding networks in the human brain: positron emission tomography data. *Neuroreport, 5*(18), 2525-2528.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale & R. J. W. Mansfield (Eds.), *Analsysis of Visual Behavior* (pp. 549-586). Cambridge, Mass: MIT Press.

Valenstein, E., Bowers, D., Verfaellie, M., Heilman, K. M., Day, A., & Watson, R. T. (1987). Retrosplenial amnesia. *Brain, 110 (Pt 6)*, 1631-1646.

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

Valentine, T., & Ferrara, A. (1991). Typicality in categorisation, recognition, and identification: Evidence from face recognition. *British Journal of Psychology, 82*, 87-102.

Van Hoesen, G. W. (1982). The parahippocampal gyrus: new observations regarding is cortical connections in the monkey. *Trends in Neurosciences, 52*, 345-350.

Van Strien, J. W., Hagenbeek, R. E., Stam, C. J., Rombouts, S. A., & Barkhof, F. (2005). Changes in brain electrical activity during extended continuous word recognition. *Neuroimage, 26*(3), 952-959.

van Turennout, M., Bielamowicz, L., & Martin, A. (2003). Modulation of neural activity during object naming: effects of time and practice. *Cerebral Cortex, 13*(4), 381-391.

Vandenberghe, R., Dupont, P., Bormans, G., Mortelmans, L., & Orban, G. (1995). Blood flow in human anterior temporal cortex decreases with stimulus familiarity. *NeuroImage, 2*, 306-313.

Vargha-Khadem, F., Gadian, D. G., Watkins, K. E., Connelly, A., Van Paesschen, W., & Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. *Science, 277*(5324), 376-380.

Vogels, R. (1999). Categorization of complex visual images by rhesus monkeys. Part 2: single-cell study. *European Journal of Neuroscience, 11*(4), 1239-1255.

Vogels, R., Biederman, I., Bar, M., & Lorincz, A. (2001). Inferior temporal neurons show greater sensitivity to nonaccidental than to metric shape differences. *Journal of Cognitive Neuroscience, 13*(4), 444-453.

Vokey, J. R., & Read, J. D. (1992). Familiarity, memorability, and the effect of typicality on the recognition of faces. *Memory & Cognition, 20*(3), 291-302.

Volkmann, F. (1962). Vision during voluntary saccadic eye movements. *Journal of the Optical Society of America, 52*, 571-578.

Von Bonin, G., & Bailey, P. (1947). *The neocortex of Macaca mulatta*. Urbana, IL: University of Illinois.

Von Bonin, G., & Bailey, P. (1950). *The neocortex of the chimpanzee* (4th ed.). Urbana, IL: University of Illinois Press.

Wagner, A. D., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1998). Prefrontal cortex and recognition memory: Functional-MRI evidence for context-dependent retrieval processes. *Brain, 121*(1985-2002).

Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract: evidence for a contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory & Cognition, 25*, 1256-1271.

Walker, P., Hitch, G. J., & Duroe, S. (1993). The effect of visual similarity on short-term memory for spatial location: implications for the capacity of visual short-term memory. *Acta Psychologica, 83*, 203-224.

Wallis, G., & Rolls, E. T. (1997). Invariant face and object recognition in the visual system. *Progress in Neurobiology, 51*(2), 167-194.

Wan, H., Aggleton, J. P., & Brown, M. W. (1999). Different contributions of the hippocampus and perirhinal cortex to recognition memory. *Journal of Neuroscience, 19*(3), 1142-1148.

Wang, G., Tanaka, K., & Tanifuji, M. (1996). Optical imaging of functional organization in the monkey inferotemporal cortex. *Science, 272*(5268), 1665-1668.

Ward, G., Avons, S. E., & Melling, L. (2005). Serial position curves in short-term memory: functional equivalence across modalities. *Memory, 13*(3-4), 308-317.

Washburn, D. A., & Astur, R. A. (1998). Nonverbal working memory of humans and monkeys: Rehearsal in the sketchpad? *Memory & Cognition, 26*(2), 277-286.

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review, 72*(2), 89-104.

Weaver, G. E., & Stanny, C. J. (1978). Short-term retention of pictorial stimuli as assessed by a probe recognition technique. *Journal of Experimental Psychology: Human Learning & Memory, 4*, 55-65.

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory: the frontal lobes and autonoetic consciousness. *Psychological Bulletin, 121*(3), 331-354.

Wickelgren, W. A. (1965). Acoustic similarity and intrusion errors in short-term memory. *Journal of Experimental Psychology, 70*, 102-108.

Wickens, D. D., Born, D. G., & Allen, C. K. (1963). Proactive inhibition and item similarity in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 2*, 440-445.

Witter, M. P., & Amaral, D. G. (1991). Entorhinal cortex of the monkey: V. Projections to the dentate gyrus, hippocampus, and subicular complex. *Journal of Comparative Neurology, 307*(3), 437-459.

Witter, M. P., Groenewegen, H. J., Lopes da Silva, F. H., & Lohman, A. H. (1989). Functional organization of the extrinsic and intrinsic circuitry of the parahippocampal region. *Progress in Neurobiology, 33*(3), 161-253.

Wixted, J. T. (2004). The Psychology and Neuroscience of Forgetting. *Annual Review of Psychology, 55*, 235-269.

Wright, A. A., Cook, R. G., Rivera, J. J., Shyan, M. R., Neiworth, J. J., & Jitsumori, M. (1990). Naming, rehearsal, and interstimulus interval effects in memory processing. *Journal of Experimental Psychology Learning, Memory & Cognition, 16*(6), 1043-1059.

Wright, A. A., Santiago, H. C., & Sands, S. F. (1984). Monkey memory: Same/different concept learning, serial probe acquisition, and probe delay effects. *Journal of Experimental Psychology: Animal Behavior Processes, 10*(4), 513-529.

Wright, A. A., Santiago, H. C., Sands, S. F., Kendrick, D. F., & Cook, R. G. (1985). Memory processing of serial lists by pigeons, monkeys, and people. *Science, 229*(4710), 287-289.

Xiang, J.-Z., & Brown, M. W. (1997). Neuronal encoding of the prior occurrence of visual stimuli in rhinal cortex and area TE of the monkey. *Brain Research Abstracts, 14*, 42.

Xiang, J.-Z., & Brown, M. W. (1998). Differential neuronal encoding of novelty, familiarity and recency in regions of the anterior temporal lobe. *Neuropharmacology, 37*, 657-676.

Yago, E., & Ishai, A. (2006). Recognition memory is modulated by visual similarity. *Neuroimage*.

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

Yin, R. K. (1970). Face recognition by brain-injured patients: a dissociable ability? *Neuropsychologia, 8*(4), 395-402.

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: evidence for a dual-process model. *Journal of Experimental Psychology Learning, Memory & Cognition, 20*(6), 1341-1354.

Yonelinas, A. P., Kroll, N. E., Dobbins, I., Lazzara, M., & Knight, R. T. (1998). Recollection and familiarity deficits in amnesia: convergence of remember-know, process dissociation, and receiver operating characteristic data. *Neuropsychology, 12*(3), 323-339.

Yonelinas, A. P., Otten, L. J., Shaw, K. N., & Rugg, M. D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *Journal of Neuroscience, 25*(11), 3002-3008.

Young, A. W., & Bion, P. J. (1980). Absence of any developmental trend in right hemisphere superiority for face recognition. *Cortex, 16*(2), 213-221.

Yukie, M., & Iwai, E. (1988). Direct projections from the ventral TE area of the inferotemporal cortex to hippocampal field CA1 in the monkey. *Neuroscience Letters, 88*(1), 6-10.

Zhou, F., Kahana, M. J., & Sekuler, R. (2004). Short-term episodic memory for visual textures: A roving probe gathers some memory. *Psychological Science, 15*(2), 112-118.

Zhu, X. O., Brown, M. W., McCabe, B. J., & Aggleton, J. P. (1995). Effects of novelty or familiarity of visual stimuli on the expression of the immediate early gene c-fos in rat brain. *Neuroscience, 69*, 821-829.

Zohary, E., Celebrini, S., Britten, K. H., & Newsome, W. T. (1994). Neuronal plasticity that underlies improvement in perceptual performance. *Science, 263*(5151), 1289-1292.

Zola-Morgan, S., Squire, L. R., Rempel, N. L., Clower, R. P., & Amaral, D. G. (1992). Enduring memory impairment in monkeys after ischemic damage to the hippocampus. *Journal of Neuroscience, 12*(7), 2582-2596.

Zola, S. M., Squire, L. R., Teng, E., Stefanicci, L., Buffalo, E. A., & Clark, R. E. (2000). Impaired recognition memory in monkeys after damage limited to the hippocampal region. *Journal of Neuroscience, 20*(1), 451-463.