To Chelle
and
Dad and Mum
Abstract

A fundamental concern in Psychology is the extent to which we learn to perceive our world and, further, the degree to which perception remains modifiable even in adulthood. Yet despite the significance of these concerns, perceptual learning has been somewhat sporadically studied, and often only at a phenomenal level. This thesis proposes a new theoretical framework for perceptual learning, and argues that a multiplicity of processes have been examined under this single term. The empirical work reported in this thesis examines a range of these different learning processes, and illustrates methods by which the process/processes underlying a particular phenomenon can be revealed. Extended replications of seminal studies on ‘perceptual learning’ demonstrate the non-perceptual learning nature of the processes reported in those studies. Further empirical work presents new evidence for the plasticity of human vision on fundamental dimensions of visual processing. These findings suggest that even adults' perceptual experience is modifiable as a result of changes at an early stage of visual processing. Final empirical work considers the types of learning that may occur in the more complex and naturalistic task of detecting features in X-rays, and this leads on to an examination of visual search learning.

It is concluded that, given the varied nature of the learning processes identified, a unified theory of perceptual learning may be an unrealistic goal. Instead, a detailed understanding of the different mechanisms underlying each of the identified learning processes is likely to prove more useful. Finally, it is argued that all of the identified processes, previously regarded as perceptual learning, could underlie improvements on complex ‘real-world’ discrimination tasks. This is illustrated through the application of the theoretical framework, developed in this thesis, to mammographic film reading. It is argued that by isolating and systematically targeting each of the learning processes involved in a task, more effective training programmes could be designed.
Acknowledgements

There are many people who have contributed to the work reported in this thesis in a significant way, and the few words which I write here could not possibly adequately thank them all.

First of all, I would like to thank my two supervisors, Ian Davies and David Rose: Ian for his open communication, for always being ready to act as a sounding board, for his incisive comments, for pushing me on at just the right level, and for teaching me the ‘art’ of English grammar: David for helping me get to grips with the ‘techno-jargon’ of psychophysics and its methods, and for the loan of his precious equipment - I can only apologise that it no longer all works. I would also like to thank both Ian and David for their careful reading of earlier versions of this thesis.

Second, I would like to thank the two people who in particular have helped with the programming work that arose during the course of this thesis. Ian McQueen for never finishing the first program that I required - if he had I would probably never have learned to program at all! - and, Keith Rosser, my father-in-law, for sharing his expertise and for writing the stimulus display routines used in the experiment reported in chapter five.

Third, I would like to thank the many people who contributed to the development of stimuli and equipment used in some of the experiments reported in this thesis. In particular, Terry Thacker for his electronics expertise and the building of pieces of equipment such as the ‘beeper’ box; Dick Casey for his remarkable precision hole drilling skills, and the staff of the Jarvis Breast Screening Centre for taking X-rays of the resulting stimuli. Whilst acknowledging the support of outside agencies, I would also like to thank the Human Factors section of Flight Systems at DRA Farnborough for the use of their stereoscopic display facilities, and the SERC
(latterly the EPSRC) and DRA Farnborough, in combination, for the CASE award which supported me financially during the course of this PhD.

Finally, and most importantly of all I would like to thank my wife, Chelle, who despite her own considerable work commitments provided me with never ending support and patience. In particular, I would like to thank her for enduring many soliloquies during which I espoused ideas on the nature of perceptual learning; for always being willing to read through my work and for the unique perspective she was able to provide; for being a willing guinea pig ever ready to help with the piloting of all manner of uncomfortable experiments and their set-up, and for seeing me through the many crises of confidence which might otherwise have prevented me from completing this work.
# Table of contents

1 Distinguishing the phenomena of perceptual learning 1  
1.1 The phenomena of perceptual learning 1  
1.2 Historical context: towards a working definition of perceptual learning 2  
1.3 The distinctive features hypothesis 3  
1.4 Distinguishing perceptual learning from discrimination learning 5  
1.5 Towards a taxonomy of learning 6  
1.5.1 Early perceptual learning research 21  
1.5.2 Studies of (previously referred to as) perceptual learning  
1) hypothesis generation, rule testing and concept formation 23  
1.5.3 Studies of (previously referred to as) perceptual learning  
2) perceptual re-organisation 26  
1.5.4 Studies of (previously referred to as) perceptual learning  
3) search learning 29  
1.5.5 Studies of (previously referred to as) perceptual learning  
4) ‘low-level’ perceptual learning 32  
   Grating waveform discrimination 33  
   Learning for visual motion discrimination 33  
   Learning for orientation discrimination 34  
   Learning for vernier acuity 36  
   ‘Fast’ and ‘slow’ phases in low-level learning? 37  
1.6 General issues in perceptual learning 40  
1.6.1 The influence of feedback on perceptual learning 41  
1.6.2 The importance of attention for low-level learning 42  
1.7 Summary 43  
1.8 Outline of remaining chapters 44  

2 When is perceptual learning not perceptual learning? 46  
2.1 A replication and extension of Gibson and Gibson (1955) 46  
2.1.1 Method 51  
2.1.2 Results 51  
2.1.3 Discussion 52  
2.2 Learning to sex day-old chicks 54  
2.2.1 An elaboration of Biederman and Shiffrar (1987) 58  
2.2.2 Method 59  
2.2.3 Results 61  
2.2.4 Discussion 63  
2.3 Summary 64  

3 Investigating low-level perceptual learning 66  
3.1.1 The importance of plasticity in adult vision 66  
3.1.2 Low-level learning 67  
3.2 Learning of stereoscopic depth discrimination 68  
3.2.1 Learning to perceive stereograms 68  
3.2.2 The nature of learning in stereopsis 68
4.4 General discussion of X-ray experiments

5 Visual search learning
5.1 Changes in search process as explanations for the learning of discrimination tasks
5.2 Note on terminology
5.3 Factors influencing the conduct of visual search tasks
5.3.1 Feature integration theory
5.3.2 Attentional engagement theory
5.4 The influence of categorisation on perceptual discrimination
5.5 Possible mechanisms of long-term learning in visual search
5.6 Design of the present experiment
5.7 Method
5.8 Results
5.9 Discussion
5.9.1 Improvements in search speed 1) Specificity of learning
5.9.2 Improvements in search speed 2) General changes in search process
5.9.3 Improvements in search accuracy
5.9.4 Summary and implications of results for discrimination learning

6 General Discussion
6.1 Review of thesis
6.2 Hypothesis generation, rule testing and concept formation
6.3 Perceptual re-organisation
6.4 Low-level learning
6.5 Search learning
6.6 Implications for the taxonomy of studies of perceptual learning
6.7 Implications for a general theory of perceptual learning
6.8 Modelling discrimination learning
6.9 Implications for examining the learning of ‘real world’ discrimination
6.10 Implications for the learning of complex real world discrimination tasks: the importance of task analysis in the design of more effective training programmes for mammographic film reading
6.11 Limitations and suggestions for future work

References

Appendix A Random-dot anaglyph, stereo stimulus
Appendix B Instructions for stereo experiment
Appendix C Raw scores for observers in stereo study
Appendix D Raw scores (contrast threshold %) for observers in contrast sensitivity experiment one
Appendix E Raw scores (contrast threshold %) for observers in contrast sensitivity experiment three
| Figure 1.1 | An example of a standard shape (a), and a transformation created by right-left reversal (b). (after Pick, 1965) | 24 |
| Figure 1.2 | A triangle created by subjective contours | 26 |
| Figure 1.3 | An example of a dot pattern (a), and a transformation of that pattern (b). (after Hock, Webb and Cavedo, 1987) | 27 |
| Figure 2.1 | Related and unrelated scribbles used by Gibson and Gibson (1955) | 48 |
| Figure 2.2 | Photographs of the genital eminences of male and female day-old chicks | 56 |
| Figure 2.3 | The experimental instructions for chick sexing. (after Biederman and Shiffrar, 1987) | 57 |
| Figure 2.4 | An example from four observers of their drawing of one of the photographs, and of how the drawings were scored | 62 |
| Figure 3.1 | Diagram of screen display showing nonius lines, fixation dot, and the four possible locations in which a stereogram could be displayed | 76 |
| Figure 3.2 | Change in mean accuracy (% correct) over time (error bars indicate ± 2 standard errors of the mean) | 80 |
| Figure 3.3 | Change in reaction times over days (again error bars indicate ± 2 standard errors of the mean) | 82 |
| Figure 3.4 | Schematic representation of the fixation dot and the three possible locations in which a stimulus was displayed | 93 |
| Figure 3.5 | Percent correct for each observer for each day | 98 |
| Figure 3.6 | Percent correct per block of 100 trials averaged across all days of training | 99 |
| Figure 3.7 | Mean contrast thresholds for the trained locations, the 2 cpd transfer stimulus, and for the mean of the untrained locations | 100 |
| Figure 3.8 | Mean contrast thresholds for observer MS for trained location, 2 cpd stimulus, and untrained locations up to six months after training | 102 |
| Figure 3.9 | Diagram of masking over the oscilloscope screen | 114 |
| Figure 3.10 | Number of false alarms for each day of training | 118 |
| Figure 3.11 | Percentage of direction changes noticed over each day of training | 118 |
| Figure 3.12 | Mean change in contrast threshold between baseline session and each follow-up for each type of stimulus | 121 |
| Figure 4.1 | Schematic illustration of a grid showing the different possible hole locations | 137 |
| Figure 4.2 | Search times during training for each hole size | 142 |
| Figure 4.3 | Mean search times (secs.) for the dominant and non-dominant (the trained eye) eyes before and after training | 143 |
| Figure 4.4 | Search times across days of training, for each contrast level and for each training group | 155 |
| Figure 5.1 | Search times for target present (TP) and target absent (TA) trials, for each day of training, for the different numbers of distractors | 179 |
Figure 5.2 Change in reaction times over the course of training for target present (TP) and target absent (TA) trials and for the different numbers of distractors 181

Figure 6.1 Diagram to show relationship between processes involved in conducting a visual discrimination task 211
List of tables

Table 1.1 A taxonomy of studies of perceptual learning .......................... 8
Table 2.1 Results from Gibson and Gibson (1955) and an extended replication .......................... 49
Table 2.2 Scores for each participant from the drawing analysis .......................... 63
Table 3.1 Disparity values for each stimulus, and absolute difference .......................... 77
Table 3.2 Mean accuracy scores and standard deviations (in brackets) for each of the three disparity increments .......................... 81
Table 3.3 Mean accuracy scores (% correct) and standard deviations (in brackets) .......................... 82
Table 3.4 Mean reaction times (msec) and standard deviations (in brackets) .......................... 83
Table 3.5 Observers' contrast thresholds (Michelson %) for each orientation pre- and post-training .......................... 110
Table 3.6 Proportion of correctly noticed direction changes to false alarms for each observer collapsed across days of training .......................... 119
Table 4.1 Mean accuracy scores (out of 25) and standard deviations (in brackets) for each of the hole depths .......................... 140
Table 4.2 Mean accuracy scores (out of 25) and standard deviations (in brackets) for each hole depth .......................... 141
Table 4.3 Mean search times (secs.) and standard deviations (in brackets) for each of the hole depths .......................... 141
Table 4.4 Mean search times (secs.) and standard deviations (in brackets) for each of the hole depths .......................... 143
Table 4.5 Mean search times and standard deviations (in brackets) for the different hole depths, before and after training .......................... 144
Table 4.6 Mean accuracy scores (out of 25) and standard deviations (in brackets) for each hole depth and for each training group .......................... 153
Table 4.7 Mean accuracy scores and standard deviations (in brackets) at baseline and transfer for each training group .......................... 154
Table 5.1 The eight different search tasks .......................... 175
Table 5.2 Mean scores for target present and target absent trials with the different numbers of distractors .......................... 178
Table 5.3 Results of analyses of variance: main effect of target present /target absent for each search task .......................... 182
Table 5.4 Results of analyses of variance: main effect of number of distractors for each search task .......................... 183
Table 5.5 Results of polynomial contrasts: main effect of number of distractors .......................... 183
Table 5.6 Results of analyses of variance: interaction between number of distractors and target present (TP)/target absent(TA) for each search task .......................... 184
Table 5.7 Results of correlations: number of distractors with target present and target absent separately .......................... 185
Table 5.8 Results of analyses of variance: main effect of change over time for each search task .......................... 185
Table 5.9 Results of analyses of variance: interaction between search task and time .......................... 187
Table 5.10 Mean scores for interaction effects between dark circles, light ellipses and dark ellipses, with the training search task and target present/target absent 188
Table 5.11 Results of analyses of variance: main effect of target present/target absent for each search task 189
Table 5.12 Results of analyses of variance: main effect of number of distractors for each search task 189
Table 5.13 Results of polynomial contrasts: effect of number of distractors 190
Table 5.14 Results of analyses of variance: interaction between number of distractors and target present(TP)/target absent(TA) for each search task 190
Table 5.15 Results of analyses of variance: main effect of change over time for each search task 192
Table 5.16 Results of analyses of variance: interaction between search task and change over time for each search task paired with the training search task 193
Prologue

Aims and objectives
The main aim of this thesis is to investigate perceptual learning in the context of vision and to develop a more thoroughly integrated theory of the many processes involved in the phenomena historically studied under this 'umbrella' term. The study of perceptual learning cuts across the whole of the study of perception by asking the fundamental question "in what sense do we learn to perceive?" (Gibson and Gibson, 1955). Thus, a vast range of, in many senses, unconnected phenomena are available for study. Consequently, in examining this area, the present thesis does not represent an attempt to provide a series of experiments, which in combination answer a specific empirical question on a particular perceptual learning phenomenon; this would not achieve the goal of integrating the many processes previously studied in this area. Rather this thesis seeks to clarify, through discussion and a series of illustrative experiments, the nature of the different processes involved in tasks previously studied as perceptual learning.

To achieve this aim, experiments were conducted using a variety of methods in order to examine a range of phenomena and to demonstrate a set of 'tools' which can be used to analyse a particular learning task into its component processes. In reporting these experiments a number of specific and widely varying areas of literature are discussed and many avenues for future research in these specific areas are opened. However, ultimately this thesis aims to identify the different processes and to develop a model of how the many processes which are identified, including more specifically defined perceptual learning processes, may combine in everyday discrimination learning tasks. This model will provide a framework within which future work can be contextualised and discussed. Finally, it will be argued that a task analysis approach, using methods such as those demonstrated in this thesis, is the best way of identifying the processes involved in a particular learning
phenomenon and that by using such an analysis, training methods for various real world tasks can be better specified.

Vision as a domain of study

Vision is perhaps the most immediate and the most highly developed sense in humans. As such it has been the most frequently studied sensory process. The present research was also conducted in the domain of vision since a much richer and larger body of research on vision is available from which to formulate ideas and within which to interpret the findings of experiments. However, ultimately, the work of this thesis aims to formulate a model of human discrimination learning which will be applicable to other modalities to the extent that other human sensory processes operate in a broadly equivalent way to vision. This supposition finds some support from studies of cortical architecture which indicate "that the architecture of the visual cortex is very similar to the architecture of other cortical areas, the pattern of cortical connections in it resembles that found elsewhere and many general features of its functional organization are reflected in other cortical areas" (Zeki, 1993). However, assessment of the validity of the model developed in the present research for other sensory domains remains a subject for future research; the validity of any model developed in the context of a specific modality needs to be assessed for other modalities. The work of this thesis will highlight critical methodological issues which should be considered when doing this.

The problem of defining perception

In examining perceptual learning one is immediately drawn into the debate about whether perception is directly sensed (e.g. Gibson, 1966) or constructed (Bruner, 1957; Gregory, 1970) and hence what can be considered perceptual learning varies with one's definition of perception. Resolution of this debate is not the aim of this thesis, nor, given the varied nature of the processes studied as perceptual learning would it seem appropriate to adopt one particular stance. Consequently, in the
present thesis, processes are distinguished at an operational level which is independent of one's definition of perception. Adopting this approach facilitates discussion of perceptual learning phenomena without requiring resolution of the 'nature of perception' debate.

Chapter one of this thesis begins the process of examining learning phenomena at an operational level. Previous research on perceptual learning has generally investigated the learning of some difficult discrimination task, on the assumption that the learning of the discrimination rests on perceptual learning. Chapter one examines this previous research on perceptual learning and argues that in fact a number of different processes have been studied under the guise of a unitary perceptual learning process. As a result of this argument a taxonomy of the learning processes studied is proposed. Some of the processes distinguished within this taxonomy are conceptualised as perceptual learning.
Chapter one
Distinguishing the phenomena of perceptual learning

1.1 The phenomena of perceptual learning

William James (1890; cited in Gibson, 1967, pp. 4-5) eloquently described some of the many anecdotal examples of individuals with highly developed discrimination skills when he wrote:

"that "practice makes perfect" is notorious in the field of motor accomplishments. But motor accomplishments depend in part on sensory discrimination. Billiard-playing, rifle shooting, tightrope-dancing demand the most delicate appreciation of minute disparities of sensation, as well as the power to make accurately graduated muscular responses thereto. In the purely sensorial field we have the well-known virtuosity displayed by the professional buyers and testers of various kinds of goods. One man will distinguish by taste between the upper and lower half of a bottle of old Madeira. Another will recognize, by feeling the flour in a barrel, whether the wheat was grown in Iowa or Tennessee. The blind deaf mute, Laura Bridgman, has so improved her touch as to recognize, after a years interval, the hand of a person who once had shaken hers; and her sister in misfortune, Julia Brace, is said to have been employed in the Hartford Asylum to sort the linen of its multitudinous inmates, after it came from the wash, by her wonderfully educated sense of smell."

Yet we do not need anecdotal examples such as these to consider the extraordinary power of human discriminative ability. Our senses are constantly bombarded with physical stimulation, and by discriminating between elements of this stimulation, we make sense of and perceive the world around us. In doing so, we demonstrate our own highly developed discrimination skills. That many of these skills are a product of early sensory experience, is demonstrated by experiments on the effects of early visual experience on later discrimination behaviour (e.g. Gibson and Walk, 1956;
Tees, 1968; Blakemore and Cooper, 1970), and by neurophysiological studies which have indicated that during the first three months of life there is a critical period for visual development (e.g. Hubel and Wiesel, 1970; Rakic, 1977; Hubel, 1979). For instance, Hubel (1979) reports that when one eye of a monkey is deprived of vision during the critical period, subsequent neurophysiological investigation reveals a paucity of cells responsive to visual input to that eye, the great majority of cells being driven by input to the eye which was left open during the critical period. Thus, during the critical period for visual development, there is considerable plasticity of the brain. However, many complex discrimination skills are learned later in life, and these have often been regarded as examples of adult perceptual learning. Whether these skills, in part, develop as a result of more limited plasticity in the adult brain, is a matter of some debate. It is with adult perceptual learning that this thesis is concerned.

1.2 Historical context: towards a working definition of perceptual learning

Bishop Berkeley (1709) when considering the subject of perceptual learning, stated that elementary sense impressions were welded together by association with images of past impressions to form meaningful perceptions. Yet despite this early statement and the interest of the associationists in learning in general, subsequent associative psychology tended to ignore the issue of perceptual learning entirely. Those associationists who did consider the subject continued to regard perceptual learning as arising from the association of past events with sensory experience. Essentially, the argument ran that we perceive the visual stimulation which defines a dustbin as a dustbin, rather than anything else, because in the past this pattern of stimulation was associated with, for instance, putting rubbish into it. Consequently, it was not until the 1950's, when some of the problems of associative psychology were recognised, that Eleanor Gibson and others began to address the subject of perceptual learning, and an alternative and more frequent dialogue developed. In her review paper (1953) on the subject, she described a large number of examples
of what she regarded as perceptual learning, and set out the proposal that "practice results in a closer approximation of discriminative responses to differential stimulation........perceptions become better differentiated and permit finer discriminations within the dimension." (Gibson, 1953, p. 422). She went on to outline a stimulus driven view of learning, which resulted in a greater specificity of responses to stimuli. This theme was picked up in the paper by Gibson and Gibson (1955). In their paper they proposed a differentiation theory of perceptual learning arguing:

"that percepts change over time by progressive elaboration of qualities, features, and dimensions of variation; that perceptual experience even at the outset consists of a world, not of sensation, and that the world gets more and more properties as the objects in it get more distinctive; finally, that the phenomenal properties and the phenomenal objects correspond to physical properties and physical objects in the environment whenever learning is successful. In this theory perception gets richer in differential responses, not in images. It is progressively in greater correspondence with stimulation not less. Instead of becoming more imaginary it becomes more discriminating. Perceptual learning then consists of responding to variables of physical stimulation not previously responded to."

This definition of perceptual learning provides a working definition from which the present thesis starts to explore the phenomena.

1.3 The distinctive features hypothesis

One of the key debates in early work on perceptual learning was whether distinctive features and the 'acquired distinctiveness of cues' (Miller and Dollard, 1941; Lawrence, 1949) were critical for learning to make a discrimination. Essentially, the proposals of Gibson and Gibson (1955) mapped onto this early debate when they argued, that as a result of the increasing differentiation which occurred following
practice, certain features not previously responded to acquired subjective distinctiveness. These additional distinctive features could then be used to make a discrimination. The importance of distinctive features was investigated in an experiment by Pick (1965). She distinguished between the "distinctive features" hypothesis, as previously described, and the "schema" hypothesis. The latter proposed that discrimination and identification involve matching sensory data, or cues about objects, to prototypes or schemata of the objects, which have been built up through repeated experience with the objects and stored in memory. Following her investigation of these two alternatives, Pick concluded that, whilst there was evidence for both processes in discrimination learning, "when no memory requirement is imposed by the task, schema learning does not occur. In short, detection of features may be the necessary and sufficient condition for improvement of discrimination." However, Pick's (1965) findings were challenged by Caldwell and Hall (1970) who argued that her study had confounded discrimination learning with concept learning. When they conducted an experiment which unconfounded these two learning processes, they found no evidence for the use of distinctive features in discrimination learning over and above prototypes. Further, they went on to argue that the apparent dichotomy of distinctive feature learning vs. prototype learning is fallacious since prototypes are necessarily composed of distinctive features. Although this proposal is obviously true, it does miss the point that it may only require a single distinctive feature to make a discrimination, and that as a result, sufficient features to develop refined prototypes of the to-be-discriminated objects may not need to be extracted. Caldwell and Hall (1970) themselves state that "when enough distinctive features are in storage, then one has a schema or prototype..............and that when many distinctive features have been stored one has a refined schemata." (my emphasis). However, it is clear that, rather than being opposed to each other, the two hypotheses are essentially related, with prototypes developing from the extraction of, and even constituting, distinctive features.
1.4 Distinguishing perceptual learning from discrimination learning

In their paper the Gibsons (1955) briefly discussed what sort of experiment appropriately explores perceptual learning and stated that "true perceptual learning experiments are limited to those concerned with discrimination." Whilst this statement may be true, it, or similar thinking on the part of other researchers, seems to have led to the assumption that any discrimination learning experiment is in fact investigating perceptual learning. It is ironic that so many researchers have fallen into the syllogistic reasoning trap which has been so well documented by Chapman and Chapman (1959) and Ceraso and Provitera (1971).

The experiments discussed in the previous section clearly illustrate the outcome of this reasoning. For instance, Caldwell and Hall (1970) stated that they were investigating "the distinctive features hypothesis vs. the schema hypothesis in perceptual learning" (my emphasis) and then proceeded throughout much of their paper to discuss discrimination learning as though the two were necessarily synonymous. The pervasiveness of this line of thinking is even apparent in contemporary work. For instance, a recent book by Hall (1991) was entitled "Perceptual and Associative Learning" yet nearly every experiment on perceptual learning reviewed in the book might have been better regarded as a discrimination learning experiment in which the learning rested not so much on learning to perceive anything, as on other processes such as associative (as Hall concludes) or conceptual learning.

If one adopts a broad definition of perception then simply attending to features can be considered a part of perception. Consequently, as the definition of perceptual learning taken from Gibson and Gibson (1955) suggests, it can be argued that even if a feature, which an observer learns to use to discriminate between two objects, was previously perceived, the change in behavioural response to that feature constitutes perceptual learning. This supposition is only valid under the Gibsons'
assumed synonymity of perceptual and discrimination learning. Consequently, in the present thesis, to distinguish these processes the definition of perceptual learning is narrowed to exclude any simple change in response at a behavioural level. Perceptual learning, then, is considered to be 'perceiving variables of physical stimulation not previously sensed or perceived'. However, this definition may seem a retrograde step for as pointed out by Fellows (1968, p. 6), in his book on the discrimination process, the perception of a difference is a private event, and so cannot be observed by anyone but the subject, leading to all the ensuing problems of introspectionism. Unless the experimenter observes a change in response he or she cannot know whether any change in discrimination has actually occurred, and a change in discrimination is a necessary but not sufficient indicator of perceptual learning. Thus, although in the present thesis the definition of perceptual learning is narrowed to exclude simple changes in response it is still necessary to observe responses in order to determine whether any learning has occurred. How then are we to distinguish perceptual from discrimination learning? The answer lies in a set of operational methods demonstrated in the work of this thesis which, in part, draw on the psychophysical and neurophysiological evidence of the last 30 years. However, as the majority of the literature purporting to investigate perceptual learning uses discrimination learning paradigms, this literature will be included in a review of previous research on perceptual learning in order to indicate how the learning processes previously investigated can be distinguished.

1.5 Towards a Taxonomy of Learning

So far, a working definition of perceptual learning has been developed, and the point has been made that discrimination learning does not necessarily indicate the operation of perceptual learning processes. The remainder of this chapter goes on to argue that, as a result of the previously under-defined definition of perceptual learning, a range of quite different learning processes have been studied as perceptual learning. The present discussion will examine how these processes can
be distinguished, and argues that there are at least four distinct types of process, which together form a set of processes involved in improvements in discrimination. The first process, which many experiments, ostensibly on 'perceptual learning' illustrate, is that of hypothesis generation, rule testing and concept formation. The second process, which may be revealed in experiments on 'perceptual learning', is the learning of new perceptual groups, through a process of 'perceptual reorganisation'. The third process, is referred to here as 'search learning' and the fourth process as 'low-level' learning. These processes will be defined and discussed in detail shortly.

Table 1.1 presents a summary of many of the published 'perceptual learning' experiments which have been conducted during the last 40 or so years arranged in chronological order (published abstracts are generally not included in this table). In order to understand fully the categorisations made in table 1.1 the reader is advised to read on and to pay special attention to sections 1.5.2 - 1.5.5 before examining the table in detail. The experiments have been taxonomised according to the type of learning process that might underlie the observed phenomenon. Obviously, it has not been possible to include every perceptual learning study conducted, and some that probably should have been included will almost certainly have been overlooked. However, an attempt has been made to represent the major studies in the area conducted in the last forty years. Studies conducted prior to this period are reviewed by Gibson (1953) and a discussion of the studies reviewed in her paper is made shortly.

In the present taxonomy, a fairly complete coverage of studies on 'low-level' learning and perceptual reorganisation has been attempted, as there are relatively few studies in these areas, and yet they represent the types of processes for which the strongest argument can be made that observers really do learn to perceive something new. Studies which have reported solely computational modelling of
<table>
<thead>
<tr>
<th>Author and date</th>
<th>Nature of study</th>
<th>Data, analyses, results &amp; conclusions</th>
<th>Type of learning &amp; comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gibson and Gibson (1955)</td>
<td>Proposed a differentiation theory of perceptual learning and examined the process by which observers learned to identify a particular nonsense squiggle from amongst a set of nonsense squiggles.</td>
<td>Number of trials taken to achieve perfect identification performance. No statistical analysis. Adults and older children learned to make the identification with practice (adults almost immediately). Younger children (6 to 8 years) improved but did not generally reach perfect performance before becoming too fatigued to continue. Observers' descriptions of the squiggles changed from naming responses to descriptions of their properties.</td>
<td>Hypothesis generation, rule testing and concept formation. Although their theory still has credibility the Gibson's experiment provides no real evidence that observers learned to perceive anything new.</td>
</tr>
<tr>
<td>Pick (1965)</td>
<td>Investigated the process by which observers learned to discriminate geometric shapes from transformations of them.</td>
<td>Number of confusion errors. ANOVA. Results suggested distinctive features training led to superior learning compared with prototype training.</td>
<td>Hypothesis generation, rule testing and concept formation.</td>
</tr>
<tr>
<td>Caldwell and Hall (1970)</td>
<td>Examined the process by which observers learned to discriminate geometric shapes from transformations of them.</td>
<td>Number of confusion errors. ANOVA. Results indicated distinctive features vs. prototype training did not lead to a difference in learning.</td>
<td>Hypothesis generation, rule testing and concept formation.</td>
</tr>
<tr>
<td>Ramachandran and Braddick (1973)</td>
<td>Studied specificity of learning for time taken to perceive stereograms comprised of uniformly orientated line elements with vernier breaks between them.</td>
<td>Time to perceive figure in depth. No statistical analysis. Results indicated a reduction in perception times with practice. This reduction transferred to a novel figure when the line elements maintained the practised orientation. However, perception times increased when the line elements were orthogonal to those practised on.</td>
<td>Low-level learning indicated by lack of transfer across orientation of line elements. However, transfer of learning to a novel figure suggests a more general learning process may also have occurred.</td>
</tr>
<tr>
<td>Study Source</td>
<td>Methodology</td>
<td>Results</td>
<td>Conclusion</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Frisby and Clattworthy (1975)</td>
<td>Examined the effect of <em>a priori</em> information on learning to perceive random-dot stereograms.</td>
<td>Time to perceive object in depth (sec's). No statistical analysis. Results indicated that none of the types of <em>a priori</em> information had an effect on perception times. Observers' perception times decreased following practice and this change was partially retained (38%) after three weeks.</td>
<td>Study design does not allow reliable identification of type of learning.</td>
</tr>
<tr>
<td>MacCracken and Hayes (1976)</td>
<td>Studied learning to perceive complex random-dot stereograms (more than two depth planes) both within a single session and over time.</td>
<td>Time to perceive object in depth (sec's). ANOVA. Results revealed a decrease in perception times with repeated exposure during sessions, but an increase between sessions.</td>
<td>Study design does not allow reliable identification of type of learning.</td>
</tr>
<tr>
<td>Ramachandran (1976)</td>
<td>Examined learning to perceive random dot stereograms and assessed specificity of learning to retinal location.</td>
<td>Time to perceive figure in depth. No statistical analysis. Perception times decreased with practice but increased when the retinal location of the stereostimulus was changed. A further change of location back to the starting position also resulted in an increase in perception times to around their starting levels.</td>
<td>Results do not allow reliable identification of type of learning. But, low-level learning could be ruled out (see chapter three).</td>
</tr>
<tr>
<td>De Valois (1977)</td>
<td>Investigated the effect of adapting to high contrast sinusoidal gratings on contrast sensitivity.</td>
<td>Log contrast sensitivity. No statistical analysis. Contrast sensitivity increased across the entire spatial frequency range measured by between 0.35 and 0.8 log units over a 1.5 yr. period.</td>
<td>Study design does not allow reliable identification of type of learning. Low-level learning might be suspected but use of method of adjustment to measure thresholds does not allow ruling out of observer criterion shifts.</td>
</tr>
<tr>
<td>McKee and Westheimer (1978)</td>
<td>Considered the effect of practice on vernier acuity.</td>
<td>Threshold (sec/arc). ANOVA. Results indicated that thresholds decreased over the course of practice.</td>
<td>In the light of more recent research these results are probably attributable to low-level learning. The study design itself does not allow reliable identification of the type of learning.</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Studied/Examined</td>
<td>Methodology</td>
<td>Results</td>
</tr>
<tr>
<td>----------</td>
<td>------------------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>Fiorentini and Berardi (1980, 1981)</td>
<td>Studied discrimination of complex waveforms which differed in either the contrast, relative spatial phase or number of their components.</td>
<td>Percent correct. No statistical analysis. Discrimination performance improved from around 70% to 90% correct within between 30 &amp; 150 temporal 2AFC trials. Learning did not transfer across orientation, spatial frequency or retinal location but did transfer across eyes. Improvement was retained over different days of the experiment.</td>
<td>Low-level learning. These studies may illustrate a fast learning process.</td>
</tr>
<tr>
<td>Ball and Sekuler (1982, 1987)</td>
<td>Considered the effect of training on ability to discriminate between two similar directions of motion.</td>
<td>Hit and false alarm rates from same-different judgements converted to d' ANOVA. Results indicated a steady improvement in discrimination for the trained directions, with or without feedback, but not for untrained directions of motion. The majority of learning (74%) transferred across eyes but not retinal location.</td>
<td>Low-level learning.</td>
</tr>
<tr>
<td>Gellatly (1982)</td>
<td>Examined learning of illusory contours and colour as a result of experience.</td>
<td>Self report from observers. 14/16 learned to see illusory contours and 13/18 learned to see illusory colour.</td>
<td>Perceptual reorganisation. However use of self report makes ruling out of demand or other effects unconvincing.</td>
</tr>
<tr>
<td>Long (1982)</td>
<td>Investigated learning to perceive random-dot stereograms and transfer of learning to transformed stereostimuli.</td>
<td>Time to perceive figure in depth. ANOVA. Results revealed that learning for a control stimulus did not transfer to spatially transformed stimuli (e.g. different spatial frequency). However, a control group who only practised the transformed stimuli were, on average, 10-15 sec's slower to perceive them suggesting some transfer of learning had occurred.</td>
<td>The results are suggestive of a low-level learning process. However, the partial transfer of learning suggests that there was also some more general learning. The significance of this was unfortunately not assessed in the ANOVA.</td>
</tr>
<tr>
<td>Study</td>
<td>Description</td>
<td>Method</td>
<td>Results</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
<td>--------</td>
<td>---------</td>
</tr>
<tr>
<td>Weinman and Cooke (1982)</td>
<td>Examined the effect of previous non-specific experience in perceiving random-dot stereograms on perception of a previously unseen stereogram.</td>
<td>Time to perceive figure in depth. Mann Whitney U test. Results revealed that group with previous experience were faster to perceive a novel stereogram than group without previous experience. However, further examination suggested that experience only helped observers who were initially particularly slow perceivers.</td>
<td>Study design does not allow reliable identification of type of learning.</td>
</tr>
<tr>
<td>Fendick and Westheimer (1983)</td>
<td>Evaluated the effect of practice on peripheral and foveal stereoacuity thresholds.</td>
<td>Threshold as assessed by method of constant stimuli. No statistical analysis. Observers' thresholds decreased 60-80% for peripheral locations but less on the fovea.</td>
<td>Study design does not allow reliable identification of type of learning. However, given the findings of other studies of a similar nature low-level learning might be expected.</td>
</tr>
<tr>
<td>Mayer (1983)</td>
<td>Investigated the effect of practice on contrast sensitivity using sinusoidal gratings and a signal detection paradigm.</td>
<td>Contrast sensitivity. ANOVA. Results indicated improvements in sensitivity for obliquely orientated gratings which did not transfer to horizontal or vertical gratings. No improvement in sensitivity was found following practice on horizontal or vertical gratings.</td>
<td>Low-level learning.</td>
</tr>
<tr>
<td>Study</td>
<td>Methodology</td>
<td>Results</td>
<td></td>
</tr>
<tr>
<td>------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Green, Pisoni and Carrell (1984)</td>
<td>Examined the effects of training on observers' ability to recognise words displayed as speech spectrograms.</td>
<td>Percent correct. No statistical analysis. Observers learned to identify a set of words from their speech spectrograms with 100% accuracy after 24 days of training. This learning transferred to the same words spoken by the same talker on another occasion (91.3% correct) and to the same words spoken by a new male talker (76%) and a new female talker (76%) but less to a synthetic talker (48%). Learning did not transfer to recognition of a novel set of words.</td>
<td></td>
</tr>
<tr>
<td>Vogels and Orban (1985)</td>
<td>Investigated the effects of practice on observers' ability to discriminate the orientation of two lines presented in temporal succession.</td>
<td>d' and JND's calculated from 2AFC data. No statistical analysis. JND's decreased for observers who practised on an oblique orientation and this learning did not transfer to vertical or horizontal stimuli. There was no decrease in JND's for observers who practised on horizontal or vertical stimuli.</td>
<td></td>
</tr>
<tr>
<td>Biederman and Shiffrar (1987)</td>
<td>Examined the sexing of day-old chicks. Specifically they conducted an expert systems analysis of the task and evaluated the effect of the resulting knowledge on the performance of naive observers.</td>
<td>Percent correctly classified chicks. Pearson's correlations. The presentation through simple line drawings of the diagnostic feature and its location improved naive observers performance from around chance levels to 84% correct. Pearson's correlations indicated a high correlation with professionals performance following training.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hypothesis generation, rule testing and concept formation. However, the possibility that learning resulted from some form of perceptual reorganisation following training cannot be ruled out from this study design.</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Research Question</td>
<td>Methodology</td>
<td>Findings</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Bressan and Vallortigara (1987)</td>
<td>Evaluated whether with practice the time taken to perceive the Saturn Illusion (a</td>
<td>Perception time (sec's). ANOVA. Results indicated that with repeated exposure the time taken to perceive the effect reduced.</td>
<td>Study design does not allow reliable identification of the type of learning.</td>
</tr>
<tr>
<td></td>
<td>stereokinetic effect) reduced.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hock, Webb and Cavedo (1987)</td>
<td>Studied whether observers changed the way they grouped dot patterns as a result of</td>
<td>Groups scored for size, shape, orientation and relative location. ANOVA. Results revealed categorisation training led to a change in the way dots were grouped.</td>
<td>Perceptual reorganisation.</td>
</tr>
<tr>
<td></td>
<td>categorisation training.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rabin (1988)</td>
<td>Investigated the effect of experience, either already held or obtained through a</td>
<td>ROC curves were constructed from hit and false alarm rates in a same different task and observers' confidence ratings. ANOVA. Results indicated that following training observers area under the curve (A') had increased. This suggested that experience improved olfactory discrimination.</td>
<td>Study design does not allow reliable identification of type of learning.</td>
</tr>
<tr>
<td></td>
<td>training procedure, on ability to discriminate between a set of odorants.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epstein, Hughes, Schneider and Bachy-Rita (1989)</td>
<td>Examined the effect of practice on learning to identify vibrotactile representations of simple unfamiliar visual patterns.</td>
<td>Probability of a correct response. ANOVA. Results indicated an improvement in identification performance with practice. Learning transferred significantly to identification of novel visual patterns from vibrotactile stimulation.</td>
<td>Study design does not allow reliable identification of type of learning.</td>
</tr>
<tr>
<td>Study</td>
<td>Description</td>
<td>Methodology</td>
<td>Results/Findings</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
<td>-------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Hartley, Higgins, MacLeod and Arnold (1990)</td>
<td>Investigated effect of training programmes on detection of a pseudo-weed amongst cash crop stubble (exp. 1) and on detection of <em>centaurea junica</em> amongst cash crop stubble (exp. 2).</td>
<td>Percentage of pseudo-weeds detected (exp. 1 &amp; 2), no. of false alarms (exp. 1) and reaction times (exp. 1). ANOVA. Results indicated improved hit rate, reduced false alarms and faster RT's following training (exp. 1). No statistical analysis but training increased hit rate (exp. 2).</td>
<td>Study design does not allow reliable identification of type of learning. The complex nature of the task examined is such that all of the types of learning distinguished in the taxonomy could be involved.</td>
</tr>
<tr>
<td>Nazir and O'Regan (1990)</td>
<td>Examined specificity of learning to discriminate simple visual patterns for retinal location.</td>
<td>Proportion of errors for trained and untrained retinal locations. T-tests. Results indicated performance worse at untrained locations although considerable transfer took place.</td>
<td>Low-level learning (but likely to be later than basic visual analysis) and concept formation. Pattern of results somewhat inconsistent. This may have resulted from use of multiple t-tests leading to a number of Type I errors.</td>
</tr>
<tr>
<td>Bennet and Westheimer (1991)</td>
<td>Evaluated the effects of practice on an observer's ability to detect the misalignment of three points and to resolve a six-line grating.</td>
<td>Threshold (arcsec). No statistical analysis. Little change in thresholds for either task was observed over a large number of practice trials (10,500 - three point alignment task; 6000 - grating resolution task).</td>
<td>This study is one of the few published studies which indicate no perceptual learning.</td>
</tr>
<tr>
<td>Karni and Sagi (1991)</td>
<td>Studied learning of a texture discrimination task.</td>
<td>SOA required for 80% correct performance. No statistical analysis. Detection speed improved as a result of practice and was specific for eye, visual field location and background element orientation.</td>
<td>Search learning localised at a low-level of visual processing.</td>
</tr>
<tr>
<td>O'Toole and Kersten (1992)</td>
<td>Examined learning to perceive random dot stereograms.</td>
<td>Reaction times and percent correct scores in 2AFC tasks. ANOVA. Results revealed that performance improved with practice. Learning did not transfer across retinal location. Further experiments suggested this was due to selective visual attention.</td>
<td>Results of study suggest that specificity effects observed are due to selective spatial attention. Further clarification of type of learning occurring is not possible from the data.</td>
</tr>
<tr>
<td>Reference</td>
<td>Methodology</td>
<td>Results</td>
<td>Conclusions</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>Poggio, Fahle and Edelman (1992)</td>
<td>Studied short-term learning of vernier hyperacuity and conducted computer simulations of the learning process based on a HyperBF network.</td>
<td>Percent correct. No statistical analysis. Results for human observers indicated an increase in percent correct with practice. This learning did not transfer across orientation. The HyperBF simulation showed similar results.</td>
<td>Low-level learning. The learning in this study may illustrate a distinct fast learning process.</td>
</tr>
<tr>
<td>Shiu and Pashler (1992)</td>
<td>Investigated learning for line orientation discrimination.</td>
<td>Percent correct. ANOVA. Performance improved over 396 trials with feedback. Improvement without feedback only occurred when training extended over days. Learning was specific to retinal location and orientation. Further, learning did not occur without attention.</td>
<td>Low-level learning. Both fast and slow learning may be illustrated here but if so the pattern of transfer is identical for both.</td>
</tr>
<tr>
<td>Treisman, Vieira and Hayes (1992)</td>
<td>Examined the development of automated performance in visual search tasks.</td>
<td>Search times (ms). No statistical analysis. Examination of search slopes indicated, for conjunction targets, a search process which sped up enormously following practice. However, search did not become parallel and learning did not transfer across even small task changes and did not transfer across spatial locations.</td>
<td>Search learning.</td>
</tr>
<tr>
<td>Ahissar and Hochstein (1993)</td>
<td>Studied learning of a local feature detection task and a global identification task.</td>
<td>Percent correct as a function of SOA and threshold SOA for 81.6% correct performance. No statistical analysis. Performance improved as a result of practice. Local learning was partly specific to element orientation and specific to element size. Global learning was specific to physical array size. Learning required attention to task.</td>
<td>Search learning partly localised at a low-level of processing. However, conclusion that high-level attentional mechanisms are essential in low-level learning may need some modification as some learning occurs without attention.</td>
</tr>
<tr>
<td>Source</td>
<td>Description</td>
<td>Metrics/Analysis</td>
<td>Learning Type</td>
</tr>
<tr>
<td>---------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>Fahle and Edelman (1993)</td>
<td>Studied effects of extensive practice on long term learning of vernier hyperacuity.</td>
<td>Thresholds (sec/arc). Regression analysis. Results indicated a reduction in thresholds over the course of training with or without feedback. This learning did not transfer across orientation but did transfer from a larger to a smaller range of vernier offsets.</td>
<td>Low-level learning.</td>
</tr>
<tr>
<td>Karni and Sagi (1993)</td>
<td>Investigated time course of learning in a texture discrimination task.</td>
<td>Percent correct for a particular SOA. No statistical analysis. Fast learning phase occurred in first session only and was specific to orientation and visual field location but not eye. Slow learning occurred between sessions following an eight hour latent phase. Learning was retained for at least 2-3 years.</td>
<td>Search learning localised at a low-level of visual processing. Authors suggest that slow learning may occur as a result of a neural consolidation mechanism.</td>
</tr>
<tr>
<td>Source</td>
<td>Description</td>
<td>Learning Effect</td>
<td>Notes</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>Weiss, Edelman and Fahle (1993)</td>
<td>Evaluated various computational models for perceptual learning of visual hyperacuity and the degree to which they mapped onto the results of a psychophysical experiment.</td>
<td>Percent correct. No statistical analysis. Some observers improved on a horizontal line vernier task and were subsequently better on a horizontal vernier task which had been simultaneously presented but to which they did not have to respond (other observers showed no improvement on either). An unsupervised exposure-dependent learning rule (EDL rule) for synaptic modification provided the closest approximation to the psychophysical data.</td>
<td>Low-level learning. The findings suggest that low-level perceptual learning in hyperacuity can occur independently of any feedback mechanism perhaps because temporal correlation between pre- and postsynaptic activity leads to amplification of neural responses. The results further suggested that attention to the task may not be a pre-requisite for learning to occur.</td>
</tr>
<tr>
<td>Fahle (1994b)</td>
<td>Reviewed experiments and conducted some further investigation on fast and slow perceptual learning for visual hyperacuity, visual search for vernier breaks and electrophysiological correlates of learning in hyperacuity.</td>
<td>Fast learning for vernier hyperacuity - Percent of error responses in 2AFC. Performance improved with practice but did not transfer across orientation or retinal location. Slow learning for vernier hyperacuity - Performance improved with practice and did not transfer across orientation or eye. Electrophysiological data revealed significant decreases in response latencies and spatial distribution of responses accompanying fast learning of jump displacement detection. Visual search - results suggested that vernier breaks are detected in parallel at least for displays sizes up to eight items.</td>
<td>Low-level learning. Both fast and slow phases of learning are illustrated here. The visual search evidence needs expanding as Pashler (1987) indicates that displays of up to eight items (but not larger) can be searched for conjunction targets (which are not perceptual primitives), in parallel.</td>
</tr>
<tr>
<td>Study</td>
<td>Summary</td>
<td>Methodology</td>
<td>Results/Findings</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Fahle (1994c)</td>
<td>Evaluated whether there were any electrophysiological correlates of learning in a jump displacement task.</td>
<td>Task data - percent correct. Wilcoxon signed-rank tests. Results indicated improvements in performance with practice. This improvement did not transfer across orientation. Electrophysiological data - latency (ms) and amplitude (μV). Wilcoxon signed-rank tests and t-tests. Results indicated a significant reduction in latencies for practised but not unpractised orientations. Further there were significant changes in scalp potential distributions.</td>
<td>Low-level learning. This study probably provides the first direct evidence for cortical changes accompanying low-level learning.</td>
</tr>
<tr>
<td>Goldstone (1994)</td>
<td>Investigated the effect of categorisation training on discriminating between squares of different brightness and size (separable dimensions) and different brightness and saturation (integral dimensions).</td>
<td>d' calculated from percent correct for same and different trials. t-tests. For separable dimensions discrimination ability improved with practice on the practised dimension but could get worse on the unpractised dimension. For integral dimensions discrimination ability improved on the practised dimension and also on the unpractised dimension.</td>
<td>An attentionally mediated low-level learning process can be postulated for these results although the lack of independent transfer tasks does not allow confident exclusion of other possibilities.</td>
</tr>
<tr>
<td>Sireteanu and Rettenbach (1994)</td>
<td>Examined improvements in search speed for visual search tasks.</td>
<td>Reaction times. Search speed increased as a result of practice and for some observers serial search became parallel. Learning transferred to a different visual search task.</td>
<td>Search learning localised at a higher level of processing.</td>
</tr>
<tr>
<td>Vidyasagar and Stuart (1994)</td>
<td>Studied learning to perceive global form from apparent motion detection</td>
<td>Time required to perceive global form. No statistical analysis. Inspection of data indicated learning which transferred to another direction of motion and across orientation of pattern elements.</td>
<td>No reliable conclusion possible as task confounds global vs. local motion detection and results could be explained by observer criterion shifts.</td>
</tr>
</tbody>
</table>
perceptual learning processes and neurophysiological studies of neural plasticity have not been included in the taxonomy. Although both of these areas have strong links with the study of perceptual learning, neither provide direct data on the nature of the perceptual learning phenomena observed at the behavioural level in adults. By taxonomising the studies, it can be seen how they cannot be viewed simply as studies of a unitary process referred to as perceptual learning, but as studies of many and varied processes. Examining table 1.1, a clear structure can be seen in the type of perceptual learning experiments conducted at various points in time. Early experiments (1950's and 1960's) tended to reveal processes which essentially pertained to hypothesis generation, rule testing and concept formation, reflecting contemporary research interests and paradigms at that time, whilst more recent experiments tend to focus on learning localised at early levels of visual processing reflecting more recent thinking that "a person cannot understand the process of perception while ignoring the known physiology of the sensory systems that underlie the perceptual process." (Levine and Shefner, 1991, p. 1). It may also be noted that a number of studies summarised in the taxonomy were designed in such a way that it is not possible to identify reliably the type of learning process which they investigated. A substantial number of these studies were investigations of learning to perceive stereograms. This may constitute a relatively unique learning process and this possibility guided the choice of a study to examine the learning of stereopsis, reported in chapter three of this thesis. Following a brief critique of the studies of perceptual learning reviewed in Gibson (1953), and thus prior to the period included in the taxonomy, subsequent sections provide specific examples of each type of study and learning process in order to give an indication of how the classification has been conducted. It should be noted that all of these learning processes have previously been referred to as perceptual learning, but that under the definition of perceptual learning proposed in this thesis, only 'low-level' learning and perceptual re-organisation are considered to indicate the observer learning to
perceive anything, whilst any or all of the processes could underlie improvements in discrimination performance.

1.5.1 Early perceptual learning research

Gibson's (1953) review paper on improvements in perceptual judgements with practice provides an excellent illustration of the methodological problems of much early research in this area. She carried out an extensive review of studies conducted in the area, and grouped them into several categories. Of these, two categories of study can be criticised in general, and within other categories methodological problems exist for most of the studies reported. Looking at the two categories which are problematic in general, the first of these concerns experiments on absolute estimation and rating. A study of this type was conducted by Horowitz and Kappauf (1946, cited in Gibson, 1953, p. 409) who required observers to estimate the range in yards of aerial targets which were up to 8000 yards away. Observers showed improvements on this task with practice. The problem with this study is that one cannot distinguish whether the improvement was due to perceptual learning for depth perception, or perhaps more likely, to the observer acquiring a better concept of what the appropriate label in yards for a given depth is (e.g. a more refined idea of what 10 yards or 100 yards looks like). As early as 1901, Thorndike and Woodworth noted that "in the case of unfamiliar standards such as grams or centimetres, the acquisition of the mere idea of what a gram or centimetre is, makes a tremendous difference in all judgements". This idea can be extended such that a more refined concept of what a familiar standard is could similarly improve 'judgements'. The other studies on absolute estimation and rating reviewed by Gibson (1953) suffer from the same type of methodological problem.

The second category of studies she reviews, which also suffer from a common problem, are those on the recognition of patterned stimuli under impoverished conditions of stimulation. For instance, typical studies in this area test observers'
ability to recognise a familiar stimulus at very low illumination levels, with very brief exposure times, or when presented to parts of the sensory surface where sensitivity is lower or less acute (e.g. peripheral vision). Performance on these types of task generally improves with practice. For instance, Seward (1931, cited in Gibson, 1953, p. 411) examined the effects of practice on identification of letters presented dimly on a ground-glass screen for 1.5 sec. Practice resulted in a gradual, continuous improvement for all observers. However, it is not possible to tell in this study whether learning resulted from some improvement in sensitivity localised at an early stage of visual processing, a higher level change such as learning to use features which perhaps were more distinctive under low illumination levels to identify the letters, or even familiarity with the task in general. To distinguish between these possibilities interocular, retinal location and even orientation transfer tasks could be employed (the use of transfer tasks to distinguish different types of learning is discussed in subsequent sections of this chapter). The majority of other studies on recognition of patterned stimuli under impoverished conditions, that Gibson (1953) reviews, suffer from similar problems.

In the remaining categories of perceptual learning study that Gibson (1953) reviews, most have other methodological problems. For instance, studies on the correction of visual anomalies often employed methods which involved exercising or strengthening of eye muscles, thus confounding interpretation of results, whilst studies on the reduction of the effects of spatial illusions frequently did not adequately distinguish between whether the illusion had really been reduced or whether the observer had merely learnt to correct their judgements by some constant factor. For instance, related to this last problem are studies which investigated observers' ability to aim at underwater targets such as Judd's (1908, cited in Gibson, 1953, p. 419) study of throwing darts at such a target or Hendrickson and Shroeders' (1941, cited in Gibson, 1953, p.419) repeat of the experiment with an airgun to minimise the motor control required. The latter study found that
performance on the original task, and on a transfer task, were both improved when observers were taught about the principle of refraction; observers were then able to institute a constant error correction procedure.

In summary, a review of the studies on improvements in perceptual judgements with practice reviewed by Gibson (1953) indicates that in general they suffered from a range of methodological problems which at best makes the type of learning occurring hard to identify, and which at worst means they were not investigating perceptual learning at all. More recent studies, while not necessarily suffering from the same problems have, for the reasons outlined previously, investigated a range of different processes as perceptual learning. The next four sections propose a series of categories into which these studies can be classified.

1.5.2 Studies of (previously referred to as) perceptual learning 1) hypothesis generation, rule testing and concept formation

Many complex discrimination tasks rely not only upon the observer being able to perceive the features which can be used to make a discrimination, but upon them actually knowing which of the perceived features are relevant to making the discrimination. When learning such a discrimination task, the observer is required to form possible discrimination rules, and to test them out until they discover a successful rule (see Fellows, 1968 for a discussion of the different models of this rule testing process). When many features are available on which a discrimination can be based, it is likely that many rules will have to be tried and tested before a successful rule is found. Further, the required discrimination rule may be a complex conjunctive or even disjunctive rule, which may require an extended rule testing process to discover. In these circumstances, the learning of a complex visual discrimination can take a considerable period of time, which has led a number of researchers to assume that some form of perceptual learning is required to successfully accomplish the task because, as LaBerge (1973) comments, "perceptual
learning.....is considered to be a slow process". The conclusions drawn from the spontaneous comments often made by observers have further compounded the problem. For instance, observers have often reported that they learned to make a discrimination because they noticed a discriminative feature which they had not previously noticed; as will be discussed in section 1.5.4, which discusses search learning, other types of learning may underlie these reports.

An example of an experiment that illustrates the above type of reasoning, and which falls under the present category, was carried out by Pick (1965). The experiment was designed to investigate whether discriminations were based on distinctive features or prototypes. In her experiment observers (in this case kindergarteners) learned to discriminate geometric shapes from transformations of those shapes. The transformations were made according to various rules (e.g. change in scale, addition of curvature to shape components, or right-left reversal). Figure 1.1 shows an example of a standard shape and the type of transformation made.

![Figure 1.1 An example of a standard shape (a) and a transformation created by right-left reversal (b). (after Pick, 1965).](image)

The observer's task was to compare a constantly displayed standard, with each card in a pack of 15 cards containing standards and transformations, and to indicate whether the shape on each card was the same as the standard. The observer repeatedly went through the pack until he or she made one perfect run (i.e. he or she picked out all the cards which displayed the standards and only those cards). Pick concluded in favour of the distinctive features hypothesis after comparison of
performance under various experimental conditions. However, examination of the differences between the standards and the transformations, such as the example shown in Figure 1.1, suggests that it is extremely unlikely that the difficulty with this task lay in perceiving the different forms. Instead, it would seem likely that the observer has to acquire a concept of what the experimenter considers the same and different. For instance, the observer has to learn that the transformed shape shown in Figure 1.1, which can be rotated through 180° such that it exactly matches the standard, is not considered the same. Other studies (e.g. Caldwell and Hall, 1970) used similar types of stimuli and methods and can also be classified as studies of hypothesis generation, rule testing and concept formation. In fact, examination of studies such as these, reveals that they have much in common with research on concept formation that was conducted at around the same time (i.e. prior to the 1970's), such as the work by Bruner, Goodnow and Austin (1956). Thus, the point made here is not that studies such as Pick's are flawed per se, but that they had little to do with learning to perceive anything, in other words, that they were not investigating perceptual learning.

However, it could still be argued, if one held an extreme version of the constructionist view of perception, that the type of task classified here as hypothesis generation, rule testing and concept formation, may involve formation of new perceptual units (e.g. visual schemata), representations of which are stored in memory. That this sort of explanation is unlikely, is indicated by the work of, amongst others, Gibson and Gibson (1955), Pick (1965), and Walk and Schwartz (1975, cited in Walk, 1978) all of whom present evidence that individuals use specific features rather than complete schemata to discriminate objects. It is ironic that, only if one took an extreme constructionist view of perception, could the findings of Gibson and Gibson (1955) be considered to illustrate perceptual learning. Further, regardless of one's view on the nature of perception, these studies can be distinguished from others at an operational level. First, they have
generally utilised simple perceptual stimuli which vary along only a few salient dimensions, and second, no improvement in sensory ability is required to perform the discrimination. Further consideration of the types of conclusion drawn from studies of hypothesis generation, rule testing and concept formation is made in chapter two.

1.5.3 Studies of (previously referred to as) perceptual learning 2) perceptual re-organisation

It was the Gestalt psychologists who first concerned themselves with the implications of the proposition that the whole is greater than the sum of the parts. There are many striking visual examples of this proposition such as the triangle formed from subjective contours shown in Figure 1.2.

![Figure 1.2 A triangle created by subjective contours.](image)

Essentially, the Gestalt psychologists proposed that we segregate visual images into figure and ground as a result of grouping image elements together. They put forward a set of principles which in combination determine which elements will be grouped together to form a figure. Examples of these principles are the principle of proximity, which states that we tend to group nearby items together, and the principle of similarity, which states that we tend to group items which are like each other together. Further Gestalt principles concern the concept of Pragnanz, which is the idea that we tend to see things as belonging together if they combine to make a "good" figure (see Garner, 1974, chapter two for a discussion of pattern goodness). More recent research related to the Gestalt psychologists' work has
tended to refer to the process of texture segregation: texture segregation is contingent upon the principles which the Gestalt psychologists proposed. Julesz (1981) suggests that we have a preattentive system which recognises units of a pattern referred to as textons, and Sagi and Julesz (1985) have indicated that detecting where textons are, can be accomplished via parallel search, but that identifying what they are, requires an attentive serial search process. Improvements in the speed with which a pattern can be segregated are discussed in section 1.5.4. This section is concerned with changes in the way in which a pattern is segregated.

Thus, to summarise, as a result of visual processing we segregate images into figures, made by grouping together image elements, and ground, and the former constitute distinct entities which can sometimes be preattentively detected. These entities may often constitute the distinctive features upon which a discrimination can be based. However, on some occasions, it may be that a discrimination can only be accomplished when observers have learned to group together elements to make a new figure (initially it may not have been of good form and so would not be perceived as a distinct entity); this type of learning could be considered perceptual learning, the observer may actually learn to perceive something which initially did not emerge as a distinct figure. An example of a study which illustrates this type of learning was conducted by Hock, Webb and Cavedo (1987). They took a number of dot patterns (an example pattern is shown in Figure 1.3) and constructed eight transformations of each pattern by moving groups of dots around (see Figure 1.3).

Figure 1.3 An example of a dot pattern a), and a transformation of that pattern b). (after Hock, Webb and Cavedo, 1987)
Observers were required to parse the dots and circle those which grouped together. One group of observers carried out this parsing procedure with no prior experience of the dot patterns. Two other groups of observers first took part in extensive training with either a concept-formation or paired associate training procedure in which they learned to categorise the dot patterns. Results indicated that, following these training procedures, observers segmented the dots in the parsing procedure in a different way from those who had received no categorisation training. Thus, Hock, Webb and Cavedo (1987) concluded that perceptual learning had occurred, since observers who took part in the categorisation training detected novel dot groupings (ones which were not naturally the most salient) that were common to members of the same category, and that thus facilitated discrimination. Further consideration of this type of learning in relation to a real world discrimination task (chick-sexing) is made in chapter two.

At this point it is also important to draw a distinction between grouping at the level discussed here, where a new group is actually perceived that consequently facilitates discrimination, and ‘chunking’ as discussed by LaBerge (1973). He proposed that as a result of practice, a task requiring detection of novel letter-like forms became automatic, and suggested that this occurred as a result of the unitization of the feature outputs stimulated by that form, as a result of practice. However, his experiments were not concerned with features which did not actually form a good group, but with improvements in processing speed for features which already grouped together well. Further, Treisman, Vieira and Hayes (1992) have recently re-examined development of automaticity in processing. Following a series of experiments, they suggested that in fact, rather than formation of a new, unitary, preattentively available feature, the improvement in speed on this type of task results from improvement at a later comparison stage. Drawing on the proposals of Logan (1980), they suggested that practice could result in more efficient comparison between features and a representation of the target held in memory. They suggested
that if memory traces of previous trials accumulate and are automatically matched to the currently attended stimulus, the correct responses may be immediately available, with no decision process needed. The better the match, and the more frequently the particular stimulus has appeared, the more rapidly the appropriate decision will be evoked. Tasks such as those of LaBerge (1973), and Treisman, Vieira and Hayes (1992) are viewed here as examples of search learning, which is discussed in the next section.

1.5.4 Studies of (previously referred to as) perceptual learning 3) search learning

There are two areas of research on search learning which are pertinent to the present thesis. The first is concerned with the development of automaticity in search processes, and with the extent to which detection tasks, which initially required a serial attentive search process, could come to be accomplished via a preattentive parallel search process, as a result of practice. This type of learning could have an effect on discrimination learning tasks. As discussed in section 1.3 of this chapter, on some level, any discrimination learning task will involve the identification of discriminative features. These features may not be of a type which 'pop-out' at an observer as a result of being detected by preattentive search mechanisms (e.g. see Treisman and Gormican, 1988) and consequently may not be so perceptually salient as features which do pop-out. This may have an influence on the formation of hypotheses about possible discrimination rules, such that less perceptually salient features are only used to formulate hypotheses after perceptually more salient features (i.e. those which pop-out) have been discarded. Further, there is a possibility that discriminative features which must be detected via an attentive serial search process will be missed, leading to failures of detection and poor discrimination performance. Thus, if as a result of practice, features which previously were only detected via an attentive serial search process come to be detected by parallel search mechanisms, this could lead to an improvement in the
speed with which a discrimination task is learned. In addition, this could also lead to an improvement in the accuracy of discrimination and identification performance based upon improved detection rates for those features.

The second area of research on search learning, which is pertinent to the present thesis, is concerned with improvement in the speed of detection for features, which may already be detected preattentively (Fiorentini (1992) proposes that an improvement in processing speed is indicative of perceptual learning). However, this type of learning may have little impact on the learning of discrimination tasks that are contingent upon detection of features which are already detected by preattentive mechanisms. As will be illustrated when the relevant research is reviewed, this type of learning may occur at an early level of visual processing.

There has been relatively little work conducted which can be categorised as the first type of search learning (i.e. investigation of changes in search process and the development of automaticity). Typical studies include one by Treisman, Vieira and Hayes (1992) who provided evidence which suggested that even with practice that reduces search slopes, conjunction search does not move from a serial to a parallel search process. However, another study in this category has provided evidence for changes in search process. Sireteanu and Rettenbach (1994) investigated detection performance for four search tasks. Their results indicated that for some observers, after a few hundred practice trials, tasks, which initially required serial search, could be conducted in parallel, and that learning was retained for at least four months. Further, they found that learning transferred almost completely to another search task and that there was complete interocular transfer. Their results suggested that the search learning process takes place at a relatively late stage of visual processing. However, this is not to say that all search learning will take place at a high level. There are many conceivable processes by which serial search could
become parallel or by which an improvement in visual search speeds could occur and some of these are discussed in detail in chapter five of this thesis.

There has been relatively more research conducted on the process by which search tasks, which are already accomplished by preattentive mechanisms, could be accomplished with even greater speed as a result of practice. Examples of studies which can be categorised in this way are those conducted by Karni and Sagi (1991, 1993) who investigated learning of a simple preattentive texture discrimination task. Observers were presented with a display of oriented line elements for 10 msec. Within the display there was a small segment (foreground texture) with differently oriented line elements. The global orientation of this foreground texture could be either vertical or horizontal and the observer's task was to report which orientation it had on each trial. A variable time after the stimulus onset a mask was displayed (stimulus-mask-onset asynchrony, SOA) and, thus, results were considered in terms of percent correct as a function of SOA. Results indicated that there was a fast learning period during the first few blocks (50 trials per block) of the first session which was indicated by an increase in percent correct for a particular SOA; this learning was specific to background element orientation and visual field location but not eye. Thereafter, there was no improvement within a practice session (100 trials per session approx.) but improvement occurred between sessions after a latent phase of around eight hours. This was indicated by a reduction in the threshold SOA for 80% correct performance on subsequent sessions. Learning was largely specific to the trained eye (only 18% transfer) and to visual field location (an estimate made from their Fig 2 (1991) indicates around 21% transfer) and was specific to the orientation of the background elements but not the foreground elements. This learning was retained for at least 2-3 years. Previous neurophysiological research has indicated that visual input is processed along various dimensions, such as orientation and spatial frequency (e.g. De Valois and De Valois, 1988). Further, populations of cells tuned to respond to specific values on those dimensions have
been shown to exist in the visual cortex. Thus, the results of Karni and Sagi's (1991, 1993) studies are indicative of learning at an early stage of visual processing (i.e. the visual cortex), since learning did not transfer across basic stimulus parameters such as background element orientation. Karni and Sagi (1991) suggested that the learning may be localised in area V1 where orientation-selective monocular and binocular cells are located.

In addition to the evidence that some search learning can be localised at early stages of visual processing, there has recently been increasing evidence for learning at early stages of visual processing with respect to the basic sensitivity of sensory systems, and this is discussed in the next sections.

1.5.5 Studies of (previously referred to as) perceptual learning 4) 'low-level' perceptual learning

In recent years, an increasing number of studies of perceptual learning have been conducted, that show improvement in sensitivity for some fundamental dimension of visual analysis. Typically, these studies assess transfer of learning from one stimulus parameter to another, with the parameter values chosen such that it can be expected, on the basis of previous psychophysical and neurophysiological work, that they would be analysed by separate visual processing channels. Lack of transfer is thus taken to be indicative of learning specific to early or 'low' levels of visual processing. This type of argument is particularly convincing when the pattern of transfer in an experiment reveals transfer to parameter values within the bandwidth of a typical channel but not to parameter values beyond. The term 'low-level' learning is used to refer to learning of this type in the present thesis. A review of some typical low-level learning studies will help to clarify the essential features of this type of experiment. A more comprehensive range of studies is reviewed here than for the other types of learning distinguished since a substantial proportion of
the experiments reported in chapters three to five are concerned with low-level learning.

**Grating waveform discrimination**

Fiorentini and Berardi (1980, 1981) reported a series of experiments in which observers learned to discriminate complex waveforms which differed either in the relative spatial phase of their components, or the contrast of one of their components, or their number of components. Learning typically reached an asymptotic value (increasing from around 70% to over 90% correct) within between 30 and 150 temporal forced-choice discrimination trials; the number of trials required varied between observers and across tasks. The improvement was retained from day to day of the experiments, but it did not transfer when observers had to complete the same discrimination with the gratings orientated 45° or 90° differently, or when the spatial frequency of the gratings was changed to ± one octave. Neither did learning for one retinal location transfer to a new retinal location. The learning did transfer across eyes, across spatial frequencies ± 0.5 octaves different, and across an orientation difference of 30°. Thus, the results indicated a low-level learning process at a site where cells are selective for orientation and spatial frequency, but respond to input from both eyes.

**Learning for visual motion discrimination**

In a series of experiments, Ball and Sekuler (1982, 1987) examined observers' ability to discriminate the direction of motion of a large number of randomly positioned dots all moving in the same direction. The basic task was a temporal same-different discrimination task, in which observers were shown dots moving along parallel pathways in two 500 msec intervals. The dots could be moving in one of eight different directions (0° to 315° inclusive, in 45° steps). Each observer was assigned one 'standard' direction of motion to practise discriminating from another similar direction of motion (± 3°). The dots in both presentation intervals
were either moving in the same direction (the standard direction), or the dots in one
interval were moving in the standard direction, whilst the others were moving in a
direction $3^\circ$ different from the standard. The observer had to indicate whether the
dots in both intervals were moving in the same direction or different directions.
Results showed that observers' ability to discriminate direction of motion improved
with practice for the standard location with or without feedback, and that this
improvement was maintained for at least 10 weeks. Learning partially transferred
to dots with a direction of motion up to $45^\circ$ different from the standard direction,
but not to dots with a direction of motion more than $45^\circ$ different. The majority of
learning (74%) transferred across eyes, but was specific to retinal location. Thus,
the results indicated a low-level learning process at a neural site where cells were
primarily binocularly driven and directionally selective, but relatively broadly
tuned.

**Learning for orientation discrimination**

Vogels and Orban (1985) conducted two experiments in which they investigated
observers' ability to discriminate the orientation of two lines presented in temporal
succession. One of the lines had a 'standard' orientation whilst the other differed
by a small increment or decrement (adjustable in steps of 0.06°). Some observers
practised discriminating oblique lines, whilst others practised discriminating lines
where the standard was at a cardinal orientation (horizontal or vertical). For
observers who trained on the oblique orientations, just noticeable differences
(JND's) decreased to about half their original value over the course of 4320 practice
trials, and remained stable thereafter. This improvement did not transfer to the
cardinal orientations, and only to a small extent to an unpractised oblique
orientation. For observers who trained on the cardinal orientations, there was no
change in their JND's over the course of training. Vogels and Orban (1985)
attributed their results to an improvement in the sensory mechanisms which compare
the signals of the visual cortical cells (Regan and Beverley, 1985, cited in Vogels
and Orban, 1985), or to a later sensorial processing stage. For instance, optimisation of the rule defining the way information from different cells is combined (Johnson, 1980, cited in Vogels and Orban, 1985).

Shiu and Pashler (1991, 1992) also investigated learning for the discrimination of line orientations. In their experiments, observers were required to judge whether the orientation of two lines shown in temporal succession was the same or different. The lines were always presented to a constant retinal location during the training phase of the experiments. Observers' discrimination performance improved over 396 trials with either trial by trial, or block by block feedback, but not without feedback. However, improvement without feedback occurred when training was extended over a period of days and thus, they concluded that in the absence of more immediate feedback to maintain motivation, fatigue effects masked any improvement within sessions. That improvement occurred within sessions when feedback was given, ruled out the hypothesis that it occurred due to a slow consolidation process of the type described by Karni and Sagi (1991, 1993). The improvement in performance did not transfer across retinal locations or across a 90° change in orientation. Shiu and Pashler (1993) also found that if observers were required to judge whether the lines differed in luminance (the lines simultaneously varied in orientation orthogonally to luminance) then no improvement in orientation discrimination was seen (further consideration of the importance of attention for learning is given in section 1.6.2). Overall, the results indicated that a low-level learning process occurred for line orientation discrimination judgements at a site where retinotopic organisation of the visual input was maintained, and where cells were orientationally tuned. This learning appeared to be dependent upon the cognitive set of the observer.

Finally, Schoups, Vogels and Orban (1993) have also recently reported learning for orientation discrimination which was specific to retinal location.
Learning for vernier acuity

Perhaps the most intensively investigated low-level learning phenomena are those concerned with improvements in vernier acuity. One of the first studies conducted in this area was by McKee and Westheimer (1978) who were interested in reports made by Volkmann (1863) and later on others, that their resolution acuity improved with practice. Their observers completed over 2000 practice trials at detecting the direction of offset of a vernier stimulus for three different orientations (vertical, horizontal and oblique). Thresholds decreased over the course of practice for all observers and for all orientations. However, conclusive interpretation of the source of this learning effect was not possible as they did not employ any form of transfer task to measure the neural specificity of learning. Thus, it could be argued that their results were attributable to some more general learning factor such as improved attention. As it has turned out, later research has suggested otherwise.

A great deal of the more recent work on learning of vernier acuity has been conducted by Fahle and co-workers (Poggio, Edelman and Fahle, 1992; Poggio, Fahle and Edelman, 1992; Fahle, 1993b; Fahle and Edelman, 1993; Fahle, 1994b; Fahle, 1994c; Fahle, Edelman and Poggio, 1994; Fahle, Edelman and Poggio, 1995). For instance, Fahle and Edelman (1993) reported a series of experiments in which they investigated learning of vernier acuity in observers who completed 10,000 trials in which they had to detect the direction of a vernier offset for two types of hyperacuity task (line verniers and three point verniers). Observers' thresholds steadily decreased over the course of training for both line and three point vernier tasks (learning was faster for line verniers). Additional experimentation using line verniers revealed that learning occurred with or without feedback although the rate of learning in the absence of feedback was reduced. The decrease in thresholds transferred from a larger to a smaller set of offset ranges, but did not transfer across stimulus orientation. In fact, there was a tendency for an 'overshoot' when stimulus orientation was changed by 90°, such that observers
were significantly worse on the first block at the new orientation, than they were on the first block of the experiment (one possible explanation for this phenomena is discussed in the next section). Thus overall, the results indicated a learning process which occurred at an early level of visual processing where cells were orientationally tuned. A further interesting observation in this study was that learning was faster over the initial blocks of training. This might suggest a distinct fast phase of learning, a possibility which was assessed in a second series of experiments (Fahle, Edelman and Poggio, 1995) and which is discussed in the next section.

Further detailed reviews of low-level perceptual learning studies examining stereoacuity and contrast sensitivity, which are specifically related to experimental work conducted in this thesis, are made in chapter three.

'Fast' and 'slow' phases in low-level learning?
It has recently been suggested that there may be distinct ‘fast’ and ‘slow’ phases of low-level perceptual learning (Poggio, Fahle and Edelman, 1992; Fahle, 1993b; Karni and Sagi, 1993; Fahle 1994b; Fahle, Edelman and Poggio, 1994). For instance, Poggio, Fahle and Edelman (1992) showed that learning can occur in a few tens of trials, and that it is orientation and stimulus specific. They proposed that underlying the fast learning stage of each task, there may be a process whereby the brain synthesises a task specific module which receives its input from retinotopic cells, and they forwarded a detailed computational model based upon a mathematical technique called HyperBF interpolation (Poggio, Edelman and Fahle, 1992; Poggio, Fahle and Edelman, 1992), although, Poggio, Fahle and Edelman (1992) do acknowledge that other types of network could be used as an alternative. Further to this, Weiss, Edelman and Fahle (1993) have proposed a model which incorporates a hardwired set of templates which may have been further tuned by
experience, and a learning component, which may synthesise new modules when required, based upon a HyperBF network.

The suggestion that the brain synthesises new modules on a task driven basis leads to two testable propositions. The first is that learning should be task specific and should therefore not transfer to different tasks. The second is that there should be a fast learning phase which occurs as a result of modules being synthesised. Kumar and Glaser (1993) tested these propositions by measuring observers' performance for a small number of trials, for each of 34 different hyperacuity tasks. By combining across trials on these different types of task a measure of threshold could be obtained. If fast learning was dependent on synthesis of task-specific modules, then initial threshold estimates should be high with this paradigm, as very few trials were conducted for each type of stimulus, and no transfer across different types of task would be expected. They found performance was in the hyperacuity range even with this paradigm, and thus questioned the proposal that task-specific modules are synthesised. However, their findings can easily be accommodated by the model of Weiss, Edelman and Fahle (1993), as resulting from the operation of hard-wired components which could have been sufficient to perform Kumar and Glaser's various tasks at hyperacuity levels. Interestingly Weiss, Edelman and Fahle's (1993) model could map neatly onto Gerald Edelman's theory of Neural Darwinism (1987) which is discussed briefly in the introduction to chapter three of this thesis.

Fahle (1994a) has recently provided evidence that fast perceptual learning does not transfer across similar hyperacuity tasks. However, Fahle, Edelman and Poggio (1995) have found a more varied pattern of transfer. They found that there was partial transfer of fast learning of hyperacuity across eyes and size of offset but that learning was specific to retinal location. Similarly, Karni and Sagi (1991, 1993) have indicated a fast phase to perceptual learning of texture segregation, which they found was specific for orientation and visual field location, but not for eye, and
Fiorentini and Berardi (1980, 1981) also found transfer of learning across eyes, but not across orientation or spatial frequency; their experiment can be considered to examine fast learning, as observers only contributed around 70 responses each. Shiu and Pashler (1991, 1992), whose observers exhibited learning within 396 practice trials, also found learning that was specific to orientation and visual field location, but did not test transfer across eyes. All of these results suggest a low-level learning process that may occur at a different neural site to the slow learning that has been observed (as described in previous sections); this is implied by the slightly different transfer patterns observed (i.e. the interocular transfer found in fast learning).

However, it is likely that a model of this learning, such as that of Weiss, Edelman and Fahle (1993), does not provide a complete explanation of the fast learning observed in these experiments. For instance, the mixed pattern of transfer observed by Fahle, Edelman and Poggio (1995), could have occurred because of multiple processes operating in the early stages of learning, some of which are related to high level factors such as attentional mechanisms and learning of mechanical or other aspects of the task, whilst others involve the initial stages of low-level learning processes. The high level processes could have led to the partial transfer of learning observed across some tasks. Indeed, it would be rather surprising if there were no such learning involved in the early stages of this type of task. Further, it could be argued that observers develop a set of expectations as a result of practising a task, such that when the task changes, they are at a disadvantage on the new task (this may be indicated in the overshoot effect often reported, e.g. see Fahle and Edelman (1993) Fig. 7). Thus, a greater degree of transfer would be expected across tasks which were more similar. This type of explanation could explain why transfer across eyes is observed in fast learning studies or why there is sometimes partial transfer across other task dimensions. A final issue is that although a number of experiments are suggestive of fast learning these experiments may well
indicate a number of different fast learning processes, for instance, fast search learning (Karni and Sagi, 1991, 1993) as opposed to fast low-level learning (Fiorentini and Berardi, 1980, 1981; Poggio, Fahle and Edelman, 1992). The possibility that this is the case is supported by the fact that the patterns of transfer found are not identical in different fast learning experiments (e.g. some find more interocular transfer than others).

Thus, to summarise, recent research has suggested that there may be grounds to sub-divide low-level perceptual learning into two qualitatively distinct phases, a fast phase and a slow phase, and computational models have been proposed to account for the fast phase of learning. Further, the distinctions made in this thesis between different types of learning process, and the results of others' experiments, suggest that there may be grounds to expect multiple fast learning processes. For the present, sufficient information is not available to allow confident distinction between different types of fast learning or to state confidently that a fast learning process which is qualitatively distinct from slow learning exists; although the broadly consistent pattern of evidence is suggestive, it may be that a unitary learning process defined by an exponential curve underlies the two phases. Consequently, in table 1.1 appropriate comments are added to indicate when a study may indicate a fast learning process although no separate category or categories for fast learning are defined in the taxonomy.

1.6 General issues in perceptual learning

There are two further issues, which are particularly relevant to learning at a low-level of visual processing, that should be considered at this point. These issues concern the role of feedback, and the importance of attention for perceptual learning.
1.6.1 The influence of feedback on perceptual learning

The effects of feedback on learning in general have been extensively researched and documented. However, of primary interest here is whether feedback is important for perceptual learning to take place. Gibson (1969 pp. 136-140) suggests that in the case of perceptual learning, knowledge of results (e.g. right/wrong feedback given by the experimenter) is not necessary for learning to take place since perceptual learning is not contingent upon learning particular responses (i.e. associative learning). The proposal that perceptual learning can occur in the absence of feedback has been examined in a number of experiments (e.g. McKee and Westheimer, 1978; Ball and Sekuler, 1987; Shiu and Pashler, 1993; Fahle and Edelman 1993). In general, the picture which emerges from these experiments is that learning occurs to a greater extent, or more rapidly, when external feedback is provided, but that substantial improvement still occurs even without feedback. For instance, Ball and Sekuler (1987), in their studies on learning for motion discrimination, found that for some tasks, provision of feedback did not lead to any significant improvement in learning, whilst in other tasks, less improvement in performance occurred without feedback, but significant improvements in performance did still occur.

At this point, it is pertinent to discuss what can actually be considered to constitute feedback. Experiments, such as those reviewed in this chapter, that have investigated whether learning can occur without feedback, have generally considered feedback to be provided externally, for instance, by the experimenter. However, in many tasks, such as those used in the experiments that have been reviewed, other feedback may still be available to the observer. An example is what will be referred to here as intrinsic feedback, which can be defined as feedback that the observer receives as the result of his/her own assessment of the results of his/her performance on a task. This is likely to be especially important when a task is easy and the observer will therefore possess a clear indication of the
correctness of their responses. However, even on more difficult tasks, such as the detection of threshold level signals, the observer will receive intrinsic feedback. For instance, the variability of thresholds means that on some trials a given signal strength may be clearly perceptible leading to intrinsic positive feedback as a result of the observer being confident that they have perceived the signal. Weiss, Edelman and Fahle (1993) briefly discuss potential types of feedback in supervised and unsupervised learning models including the possibility of intrinsic feedback in unsupervised learning (in their terms an "internal teacher"). Thus, although experiments have indicated that external feedback is not necessary for perceptual learning to occur, they have not demonstrated that perceptual learning is independent of feedback per se. In fact some results which have been interpreted as indicating that learning does not occur without attention can alternatively be considered to result from the absence of feedback when a task is not attended to, as discussed in the next section.

1.6.2 The importance of attention for low-level learning

An important issue in low-level perceptual learning is whether attention to the particular stimuli to be detected or discriminated is essential for learning to occur. If learning occurs simply as the result of the stimulation of early receptors and analysers, then even when the stimuli are not attended to, learning should still occur, providing the stimuli still impinge on the receptor cells. However, if learning is dependent upon some form of top down modulation, then learning would not be expected when the stimuli are not attended to. This issue has been investigated by a few researchers. For instance, as described earlier, Shiu and Pashler (1992) found that observers' ability to discriminate line orientation did not improve when they attended to the luminance of the lines rather than their orientation. Ahissar and Hochstein (1993) conducted a study in which observers practised one of two possible visual tasks both of which used exactly the same visual stimuli, but which depended on different stimulus attributes. They found that for
both tasks observers learned (i.e. performance improved) as a result of practice, but that this learning did not transfer to the unpractised task. They interpreted their results as indicating that high level attentional mechanisms, controlling changes at early visual processing levels, are essential in low-level perceptual learning. However, Weiss, Edelman and Fahle (1993) conducted a study in which observers were simultaneously presented with horizontal and vertical line verniers in a cross formation. Following baseline measurement for each orientation observers completed a practice phase during which they only had to judge the direction of offset for verniers of one orientation. Observers who improved following practice on the horizontal task also showed an improvement on the vertical task, whilst observers who practised the vertical task showed improvement on neither. Thus, these results suggested that learning, which is independent of attention, can occur for vernier acuity.

Further, the conclusions of the studies which suggest attention is necessary for learning can be questioned, if one assumes that some form of intrinsic feedback is required for low-level perceptual learning to take place. Thus, in tasks where the observer does not attend to the relevant task dimensions they consequently receive no intrinsic feedback and it could be this absence of feedback which results in a failure to learn. It should, however, be noted, that Weiss, Edelman and Fahle (1993) have provided some evidence to indicate that learning in which there is intrinsic feedback does not map so neatly onto psychophysical results as a use-dependent learning rule (EDL algorithm).

1.7 Summary
This chapter has examined previous research on perceptual learning. It has been argued that, since perception itself is a private event, in order to infer an improvement in perceptual abilities one must observe discrimination learning. This fact has had the consequence of leading many researchers to assume that an
improvement in discrimination necessarily indicates perceptual learning. However, it has been argued in the present chapter that, by examining previous research at an operational level, it can be seen that many different learning processes have been studied. Consequently, a taxonomy of studies previously referred to as perceptual learning has been proposed. Only some of the processes distinguished in this taxonomy can still be considered to be perceptual learning under the more stringent definition derived in this chapter.

However, the proposed taxonomy is only tentative and it is likely that subsequent research may suggest modifications to the classificatory scheme. The examples used to indicate the method by which the categorisation has been conducted were chosen as particularly clear illustrations of the distinctions made. In many cases previous research cannot necessarily be categorised so clearly, and this is particularly true for examinations of complex, ‘real world’ discrimination tasks. Thus, there is a need to establish clear methods for distinguishing between when the learning of a discrimination task requires perceptual learning and when it does not. Further, although a new theoretical framework for the study of perceptual learning has been proposed, at a specific level there are many issues relating to the nature of individual learning phenomena which are unresolved. For instance, search learning is relatively poorly defined and further investigation is required to understand more thoroughly the nature of improvements in visual search. As a result of establishing the new theoretical framework it is now possible to design studies to investigate these phenomena in a more meaningful manner such that the implications of their results for a wider context can be understood.

1.8 Outline of remaining chapters

The remaining chapters of this thesis address the issues raised above, seeking to develop methods to allow confident identification of different learning processes where such methods are lacking, and to examine the nature of some of the learning
processes distinguished in greater detail. Chapter two develops methods for revealing the non-perceptual learning nature of the processes observed in some previous research by conducting replications and extensions of two seminal studies (Gibson and Gibson, 1955; Biederman and Shiffrar, 1987). Chapter three demonstrates a set of experimental methods for examining phenomena referred to here as ‘low-level’ perceptual learning, and through a series of four experiments examines such learning for two aspects of visual analysis: detection of contrast and stereo depth discrimination. Chapter four examines, through two further experiments, the types of learning which may actually occur in a naturalistic learning task, and a final experiment, reported in chapter five, goes on to consider visual search learning issues arising from this naturalistic task. Chapter six discusses the implications of the findings from the empirical work described in this thesis, in combination with results reported previously by other researchers, for the theoretical framework developed in the present chapter. It is proposed that whilst not all of the processes examined can be regarded as perceptual learning, they could all still be integrated as processes which can underlie the learning of discrimination tasks. A model of discrimination learning is outlined to indicate one way in which these processes could interrelate. Finally, the application of this model and the theoretical framework established in this thesis to real world discrimination tasks, and the training of individuals to make complex visual discriminations, is discussed.
Chapter two

When is perceptual learning not perceptual learning?

In chapter one, it was argued that many previous experiments which have used discrimination learning paradigms in an attempt to explore perceptual learning, may in fact have revealed learning which was based on other processes. This chapter provides a more detailed illustration of this, by reporting the results of a replication and extension of the study used by Gibson and Gibson (1955) to demonstrate their theory of perceptual learning. By conducting this replication, a first way in which it can be revealed whether perceptual learning really underlies an improvement in discrimination, will be demonstrated.

Following the replication of Gibson and Gibson (1955), an elaboration of a study conducted by Biederman and Shiffrar (1987) on the learning of chick-sexing will be reported. This study was conducted in order to further confirm, by ruling out perceptual re-organisation as the main source of the learning observed, that the learning of a complex discrimination task which had been characterised as perceptual learning could in fact be demonstrated to rest mainly upon hypothesis generation, rule testing and concept formation. In conducting this elaboration, a method for distinguishing whether perceptual re-organisation may underlie an improvement in discrimination performance is demonstrated.

Together, these two elaborations on previous investigations of perceptual learning demonstrate ways in which perceptual learning may be ruled out as the process underlying the successful learning of a new discrimination task.

2.1 A replication and extension of Gibson and Gibson (1955)

In their paper on perceptual learning, Gibson and Gibson (1955) proposed their 'specificity' theory of perceptual learning. To recap briefly, they suggested that
simply through repeated exposure to the stimulus array, an individual learns to differentiate finer and finer properties of the physical stimulation, such that perception is increasingly in correspondence with the physical world: this is what they considered to be perceptual learning. They suggested that as a consequence of this learning, features could come to acquire distinctiveness, thus, facilitating discrimination between objects. The Gibsons presented a demonstration experiment to illustrate their 'specificity' theory of perceptual learning. They devised a set of 30 drawings. Eighteen of these drawings were a set of related 'scribbles' which were made up of one 'standard' scribble, and seventeen others that varied either in their number of coils, or their degree of horizontal compression, or their orientation (right-left reversal). These items were designed to be relatively indistinguishable at the outset of the experiment. The remaining twelve drawings were unrelated 'nonsense' items which were designed to be easily distinguishable from each other, and from the set of eighteen related scribbles. All of the items except for the left-right reversed version of the standard scribble are illustrated in Figure 2.1 (copied from Gibson and Gibson, 1955). The unrelated scribbles are shown in the top half of the figure and the related scribbles are shown in the bottom half. The drawing in the centre of the related scribbles was the standard scribble. The items were printed on 2 x 4 inch cards and made up in a pack which included four copies of the standard scribble.

The experimental task proceeded as follows. On each run the observers were first shown the standard scribble for five seconds. After this they were shown each of the cards in the pack one at a time, in a random order. Each card was presented for three seconds. After presentation of each card the observer had to state whether it was the same as the standard scribble.
Figure 2.1 Related and unrelated scribbles used by Gibson and Gibson (1955).
Observers were not given any feedback regarding the correctness of their responses. At the end of the run if the observer had made any mistakes (i.e. identifying a non-standard as the standard) then he/she was again shown the standard for five seconds and another run through the pack in a new random order was initiated. This procedure continued until the observer made no mistakes during an entire run.

The results of the study for two groups of observers the Gibsons' used (adults and older children aged 8½ to 11 years), as well as the results from an elaborative replication to be described shortly are shown in table 2.1

<table>
<thead>
<tr>
<th></th>
<th>Adults (N = 12)</th>
<th>Older Children (N = 10)</th>
<th>Replication Group (N = 10)</th>
<th>Simultaneous Group (N = 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Av. no. of runs to achieve criterion</td>
<td>3.10</td>
<td>4.70</td>
<td>3.10</td>
<td>1.10</td>
</tr>
<tr>
<td>Av. no. of errors on first run</td>
<td>3.00</td>
<td>7.90</td>
<td>1.40</td>
<td>0.10</td>
</tr>
<tr>
<td>Percent of errors for items differing by one quality</td>
<td>17.00</td>
<td>27.00</td>
<td>9.14</td>
<td>0.53</td>
</tr>
<tr>
<td>Percent of errors for items differing by two qualities</td>
<td>2.00</td>
<td>7.00</td>
<td>2.02</td>
<td>0.00</td>
</tr>
<tr>
<td>Percent of errors for items differing by three qualities</td>
<td>0.70</td>
<td>2.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

As can be seen from table 2.1, both of the Gibsons' groups took a number of runs through the pack to achieve perfect performance, with the older children taking a little longer to learn the task, and making more errors on average, on the first run through (the Gibsons did not include errors where the observer failed to identify a copy of the standard item as such in their error rates). On the basis of these results the Gibsons suggested that learning (implicitly in their terms perceptual learning) was required to achieve criterion performance on the task. Analysis of the older childrens' spontaneous verbal comments indicated that during the course of the experiment the proportion of these comments which described relational properties, such as "too thin" increased. The Gibsons interpreted these comments as being
correlated with the discovery of further discriminative features in the drawings, which they suggested supported the notion that learning consists of responding to ever more units of physical stimulation as a result of an increase in differentiation. Further evidence, which they suggested supported this notion, came from analysis of the items which were wrongly identified as a copy of the standard. As can be seen from table 2.1 the greatest percentage of errors (i.e. calculated as a percentage of the total number of times they were presented) were made, by both groups of observers, on the items which differed from the standard on only one quality, and a greater percentage of errors were made on items that differed from the standard by two qualities, than on those that differed by three qualities.

Thus, on the basis of their experiment the Gibsons stated that "the results show clearly that the kind of perceptual learning hypothesized has occurred" (p38). However, it is questionable whether they really demonstrated that the observers learned to perceive any difference which was imperceptible at the outset. For instance, the comments recorded from the older children, which the Gibsons suggested were correlated with an increase in differentiation, could merely show a change in strategy, from the use of item labels to the use of specific features, rather than an improvement in perception. This possibility is further supported by the fact that even on first inspection, the supposedly indistinguishable shapes appear to be quite distinctive when presented together (e.g. as in Figure 2.1) and indeed Gibson and Gibson (1955) make a comment to this effect. A direct test of this proposition could be made by requiring observers to perform the Gibsons' original task, but with the standard constantly displayed for them to compare each card with. If they still made errors then perhaps it could justifiably be argued that some form of perceptual learning was required in order for them to learn to make the discrimination. If, on the other hand, they made no errors, then this would suggest that observers did not need to learn to perceive any additional features in order to be able to complete adequately the task, and that the learning observed in the Gibsons'
original experiment resulted from non-perceptual learning processes. An extension and replication of the Gibsons' original experiment was conducted to investigate this possibility.

2.1.1 Method
Two groups of ten adult observers were recruited. One group of observers (replication group) completed the task exactly as it was originally conducted by Gibson and Gibson (1955). The second group of observers (simultaneous group) completed the task with the standard constantly displayed. Thus, they could make a simultaneous comparison between each card in the pack and the standard.

2.1.2 Results
A set of results for each group, scored in the same way as Gibson and Gibson (1955) can be seen in table 2.1. Like the group of adult observers in the original study the replication group took a mean of 3.1 runs to achieve perfect performance but they made fewer errors, on average, on the first run through. However, the distribution of errors was in general similar to the Gibsons' results, with the greatest percentage of errors made on items that differed from the standard on only one quality, and a greater percentage of errors made on items that differed from the standard by two qualities than by three qualities. Across all runs and all observers, a total of 34 errors were made. Of these, 11 were errors where the standard was called a non-standard (misses), whilst the remaining were errors where a non-standard was identified as the standard (false alarms). Of the 23 false alarms (equivalent to the errors Gibson and Gibson include in their error counts and, thus, the errors shown in table 2.1) 17 were made with items that differed from the standard by only one quality, but of these, 15 were confusions with just one particular stimulus, the right-left reversed version of the standard. Unfortunately, the Gibsons did not break their error responses down in this level of detail, but given the overall similarity of the results, it might be expected that a similar
proportion of the errors they observed were with this one stimulus. This finding alone suggests that, as with Pick's (1965) study, perhaps most of the difficulty in this task lay in understanding the experimenters concept of same and different: if an individual sees a familiar face in a mirror (i.e. right-left reversed) they would probably still consider it to be the same face. However, some additional false alarms were made and they may have arisen because the scribbles concerned were initially indistinguishable from the standard. Examination of the results from the simultaneous group allows assessment of this possibility. All but one of the ten observers in the simultaneous group performed the task perfectly on the first run through. The remaining observer made one error on the first run, identifying the right-left reversed version of the standard as a copy of the standard. When questioned afterwards she indicated that she had not realised this would be considered different.

2.1.3 Discussion

The results of the present extension and replication of Gibson and Gibson (1955) would seem to indicate that the discrimination learning which they observed, did not result from observers learning to perceive anything. When observers were able to compare simultaneously the stimuli with the standard they were readily able to distinguish them. Thus, it would seem more feasible that the discrimination learning which the Gibsons observed related to the memory load imposed by the paradigm they used. In order to perform the original version of the task there were many strategies that an observer could attempt to use. For instance, they could try to develop and remember a discrimination rule (or a set of rules) to be tried on a particular run through (e.g. the target has four loops and is a clockwise spiral). Alternatively, they might try to hold a visual representation of the standard in memory in order to conduct a template matching operation. Observers might of course try combinations of strategies such as these. In fact, under the alternative interpretation of the older childrens' spontaneous comments, made earlier, it could
be argued that they moved from the latter strategy to the former. What all the strategies would have in common is that an observer's performance would be limited by memory. For instance, if an observer can only extract and remember a limited number of possible discrimination rules to apply on a particular run, then it may take several runs before they have had sufficient time to discover a successful discrimination rule. Alternatively, if an observer attempts a template matching strategy then it may take a number of runs before they develop a visual representation which is detailed enough to facilitate discrimination.

In summary, it would seem highly unlikely, given the ease with which the task could be accomplished when the standard was constantly displayed, that the Gibsons' experiment demonstrated perceptual learning. The current evidence suggests that it is more likely that their task required learning because the paradigm imposed memory restrictions that limited the speed with which observers could formulate a suitable discrimination rule. Thus, although there may be nothing intrinsically wrong with the Gibsons' theory (observers may well learn the discrimination on the basis of identifying the discriminative features) their experiment does not really provide any good evidence that the observer is perceiving variables of physical stimulation not previously sensed or perceived, and as such does not illustrate perceptual learning. The Gibsons have almost certainly fallen into the trap of thinking that because, "true perceptual learning experiments are limited to those concerned with discrimination" (p. 35), that any discrimination learning experiment is necessarily a perceptual learning one.

Given that the Gibsons' task did not require perceptual learning, then by default the learning can be described as resulting from a process of hypothesis generation, rule testing and concept formation; this process is necessary for any discrimination task to be successfully learned. A second learning phenomena which had been thought to rest upon perceptual learning but which more recent research (Biederman and
Shiffrar, 1987) has indicated may also arise from the difficulty of discovering a successful discrimination rule, is the task of learning to sex chickens. However, in this case there was a further possibility to be ruled out before such a conclusion could be confidently reached. The next section describes the study conducted by Biederman and Shiffrar (1987) which suggested that the task was not a perceptual learning one, and an elaboration of their study which adds weight to this conclusion, by ruling out the additional possibility as a main source of the learning observed.

2.2 Learning to sex day-old chicks

Egg producers have long been interested in ways of separating female from male chickens at as early a stage as possible. The latter represent a considerable feed cost and can deny the egg producing females easy access to food and water. In the early part of the present century a method for telling male and female day-old chicks apart on a perceptual basis was discovered by the Japanese, and this was communicated to American ‘poultrymen’ in 1934. As a result of this, schools to teach chick sexing were set up in various American cities (see Lunn (1948) for a history of chick-sexing and a description of the skill). The skill of chick-sexing has generally been regarded as an extraordinarily difficult perceptual task requiring years of extensive practice to achieve high levels of performance. Biederman and Shiffrar (1987) reported that professional sexers estimated that an average of 2.4 months was required to achieve 95% accuracy, which, considering that sexers can examine up to 1000 chicks per hour is a considerable amount of practice. Professionals estimated that it took between two and six years of experience to achieve maximum performance levels of between 99.3% and 99.5% accuracy at an average rate of 960 birds per hour. Given the difficulty of the task and the lengthy time required to learn the skill, mastery of chick sexing has been considered to be an example of perceptual learning.
Biederman and Shiffrar (1987) conducted an expert systems analysis of the task of chick-sexing. They took a set of 18 photographs (shown in Figure 2.2) of the genital eminences of day-old chicks, that were first published in Canfield (1941), and asked an expert sexer to circle the area in each that was examined to indicate the sex of a chick.

They then examined this area and were able to extract a simple difference between males and females. In males the genital eminence was "round" whilst in females it was "flat" or "pointy". The expert sexer agreed as to the diagnosticity of this distinction. Next, Biederman and Shiffrar (1987) took a group of 36 naive observers and asked them to classify the 18 pictures as male or female on the basis of their own intuition (they were told that there were equal numbers of males and females). Observers were not given any error feedback regarding their classification performance. Following this first classification, half of the observers were shown a set of diagnostic diagrams and instructions (reproduced in Figure 2.3) on how to perform the sexing task, which were derived from the earlier analysis of the pictures. It took about a minute to read through these instructions. All of the observers were then required to perform a second classification of the same 18 pictures they had previously classified. A group of five professional sexers were also asked to classify the pictures. On their first classification the naive observers averaged 59.8% correct. On their second classification, the group of 18 observers who received instruction on how to perform the sexing achieved 84% accuracy, whilst the remaining 18 observers, who had not had any instruction, achieved 54.1% accuracy. The professional sexers achieved 72% correct choices. As the photographs were primarily of specially selected rare types of eminence, a weighting of accuracy scores according to the relative frequency of the different types was calculated. Following this, professional sexers performance was 84.1% correct and the performance of observers who had received instruction was 89.8% correct.
Figure 2.2 Photographs of the genital eminences of male and female day-old chicks.
Now you will be asked to sex-type another set of chickens. You should use the following rules to discriminate between the male and female genitals. The first part of your task should be to locate the informative region. In order to find this point, you should look for the two large cylindrical side lobes near the bottom of each picture. The genitals are located either between the ends of these two lobes (1) or slightly above the ends (2).

After you have found this area, you should study the swellings present there. It is at this point that males may be differentiated from females. Male chicken genitals tend to look round and fullish like a ball or watermelon. Here are two examples:

Female chicken genitals can take on two different appearances. They can look pointed, like an upside down pine tree, or flatish. Here are two examples:

Usually, but not always, male genitalia are larger. Sometimes either sex will appear to have double genitalia. You should differentiate the following 18 pictures based on this set of rules. Again, nine of the pictures are male and nine female in random order. You should begin the task by studying all of the pictures carefully. If you decide that a pictured set of chicken genitals belongs to a male, circle the "M" after the corresponding number on your answer sheet. If you decide that the pictured genitals belong to a female chicken then circle the "F" after the corresponding number on your answer sheet. Please work carefully. Once you understand these rules, you may turn this paper over and begin. Any questions?

Figure 2.3 The experimental instructions for chick sexing (after Biederman and Shiffrar, 1987).
Further, correlational analysis indicated that, following instruction, the performance of the naive observers more closely resembled that of the experts ($r=0.84$) than it did their own pre-training performance ($r=0.63$). Prior to instruction the naive observers performance only correlated slightly ($r=0.21$) with the expert sexers.

Thus, in general, these results suggested that a large part of the experts skill rested on the examination of a particular diagnostic contour, and that once naive observers were aware of the location of this contour and what to look for, their performance closely mimicked that of the experts. So, under the taxonomy established in this thesis, these results indicate that perhaps instead of requiring perceptual learning, this complex discrimination task required a procedure of hypothesis generation and rule testing that continued until a rule which led to a satisfactory level of discrimination performance was discovered. However, there does remain a possible role for one type of perceptual learning in this task which Biederman and Shiffrars' (1987) experiment did not rule out. It could be that the instructions given to the naive observers encouraged some form of perceptual re-organisation in the way that they viewed the eminences which, in the normal course of learning the task, only occurred following extensive sexing experience. In other words, prior to instruction or experience it could be that in some instances the critical feature does not have 'good form' and, thus, is not available to be used as a discriminative feature (see section 1.5.3 for a more detailed discussion of perceptual re-organisation). In order to test this possibility an elaboration of Biederman and Shiffrars' (1987) experiment was designed.

2.2.1 An elaboration of Biederman and Shiffrar (1987)

Parsing techniques where observers are required to circle critical regions in an image have previously been used to elicit expert knowledge (e.g. Biederman and Shiffrar, 1987), to examine the development of expertise (e.g. Lesgold, Glaser, Rubinson, Klopfer, Feltovich, and Wang, 1988) and to study image segmentation
(e.g. Hock, Webb and Cavedo, 1987). In addition, neuropsychological tests such as the Rey Complex Figure Copy Task (Rey, 1941 cited in Lezak, 1983, pp. 395-402) have been devised for use in neuropsychological assessment. They examine individuals' copies of a complex figure and the order in which components of the figure are drawn; individuals with a variety of brain injuries show abnormal copying patterns. The present study was designed to utilise a parsing paradigm, which drew upon the methods used in previous research and upon neuropsychological testing procedures, in order to study individuals' initial perceptual segmentation of the photographs of chick genitals shown in Figure 2.2. By examining the way in which individuals segmented the images, as revealed through their drawings, it was possible to discover whether the critical discriminative feature was perceived as a distinct figure, or, whether perceptual re-organisation of the stimulus input would be required prior to the accomplishment of successful discrimination between males and females. Further, the relative distinctiveness of the features was examined by requiring participants to draw the most distinctive features first. It might be expected that the more distinct features would be utilised in the formulation of initial discrimination rules. Thus, the distinctiveness of the critical feature might be one predictor of the length of time required to learn a discrimination.

2.2.2 Method

Ten adult participants were required to produce a drawing of each of the eighteen photographs shown in figure 2.2 in a carefully structured manner. The drawings were made using coloured felt tip pens. The exact instructions given to participants provide an account of the procedure and were as follows:

"In this task you are going to have to produce a drawing of each of a set of eighteen photographs of chicken genitals. However, we are going to do this in an unusual way."
On the response sheets which I have given you there are a series of blank boxes. You will have to draw one photo in each box. I need to see how the drawing develops step by step. What I want to end up with is a picture that shows how you break each photo down into its separate components.

We will proceed in the following way. I will hand you a coloured pen. Then I will ask you to draw the feature in the photo which stands out the most to you. It is very important that you really do draw the feature which stands out the most and don't just do the drawing as you normally would. For instance, many people would normally start with the big features and work towards the finer details. When you have drawn the first feature you will give me the coloured pen back. I will then give you a different coloured pen. I then want you to draw the next feature in the photo that stands out the most to you. Again when you have completed drawing that feature you will hand the pen back to me. I will then hand you another pen. We will continue to proceed in this way until you are happy that you have drawn all the features in the photo. Thus, I should end up with a drawing that shows how you segment the photo down into its separate features and the relative distinctiveness of each feature.

Now I want you to do your first drawing. Put your name at the top of the response sheet. Put number" (number of photo given to them by experimenter) "next to the box.

Here is the first coloured pen. Draw the feature that stands out the most to you and when you have done that hand the pen back to me.

Here is the second colour pen. Draw the next feature that stands out the most to you".

The latter instruction was repeated (with the pen number changed) until the participant was satisfied that he/she had drawn all the features in a particular photo. In all there were 30 different coloured pens. The pens were handed out in exactly
the same order each time. The photographs were passed to participants in a random order.

2.2.3 Results
Each of the drawings a participant produced was scored by two independent judges. They scored each drawing in three ways. First, whether the critical feature was actually clearly present as a distinct feature. Second, the position of the critical feature in terms of drawing order was determined. This could be calculated since the coloured pens were always handed out in the same order (e.g. black was always first, dark brown second, a light brown third etc.). Third, the total number of features drawn was counted. Having completed the scoring individually, the two judges compared their results for each participant. Any disagreements were discussed until a resolution was agreed upon. There were some disagreements regarding the total number of colours which occurred as a result of one or other judge miscounting. There was a 10% disagreement rate regarding the presence or absence of the critical discriminating feature as a distinct entity. The majority of these disagreements were easily resolved upon discussion as one judge was over-emphasising absolute accuracy of representation; although the critical feature needed to be recognisable as such (i.e. in the correct location and of approximately the same form), placing too much weight on drawing accuracy might have led to scoring based upon drawing ability per se. Figure 2.4 presents an original photograph, four observers' drawings of that photograph and details of how each drawing was scored. For instance, participant ST's drawing did not distinguish the critical feature. Instead, she considered the dark region below the critical feature to emerge as a distinct entity. Thus, it is likely that she may have missed the diagnostic flat contour indicating the fact that this chick was a female. In total this participant distinguished four distinct structures in her drawing as indicated by the use of four colours.
Figure 2.4 An example from four observers of their drawing of one of the photographs and how the drawings were scored.
To the right of figure 2.4 is a key to indicate the order in which the first ten colours were handed out (none of the drawings illustrated distinguish more than ten features although this was not always the case). Table 2.2 illustrates each participant’s score for the three ways in which each drawing was scored.

### Table 2.2 Scores for each participant from the drawing analysis.

<table>
<thead>
<tr>
<th>Participant</th>
<th>No. of drawings critical feature present (out of 18)</th>
<th>Average position of critical feature</th>
<th>Average no. of features for drawings.</th>
</tr>
</thead>
<tbody>
<tr>
<td>EB</td>
<td>16.0</td>
<td>1.6</td>
<td>6.3</td>
</tr>
<tr>
<td>AB</td>
<td>15.0</td>
<td>1.3</td>
<td>4.7</td>
</tr>
<tr>
<td>LG</td>
<td>9.0</td>
<td>1.2</td>
<td>4.0</td>
</tr>
<tr>
<td>MP</td>
<td>18.0</td>
<td>1.2</td>
<td>4.3</td>
</tr>
<tr>
<td>IR</td>
<td>14.0</td>
<td>2.4</td>
<td>5.2</td>
</tr>
<tr>
<td>ST</td>
<td>9.0</td>
<td>1.1</td>
<td>3.8</td>
</tr>
<tr>
<td>CF</td>
<td>10.0</td>
<td>2.4</td>
<td>6.2</td>
</tr>
<tr>
<td>JH</td>
<td>14.0</td>
<td>1.4</td>
<td>4.1</td>
</tr>
<tr>
<td>MI</td>
<td>15.0</td>
<td>1.3</td>
<td>5.0</td>
</tr>
<tr>
<td>PA</td>
<td>17.0</td>
<td>1.1</td>
<td>5.2</td>
</tr>
<tr>
<td>Mean</td>
<td>13.7</td>
<td>1.5</td>
<td>4.9</td>
</tr>
</tbody>
</table>

As can be seen from table 2.2, participants did tend to perceive the critical feature as a distinct feature a greater proportion of the time (on average 13.7 out of 18 drawings). They also tended to perceive the critical feature as relatively distinct (average position in drawing order = 1.5) compared with the total number of features they actually perceived as distinct figures (average number of features in drawings with critical feature present = 4.88).

### 2.2.4 Discussion

The results from this parsing study of observers' perceptual segmentation of photographs of chick genitals suggest that the critical feature which must be examined to perform the task of chick-sexing is generally perceived as a distinct entity, and further, that it is perceived as one of the most distinctive features present in the images. Thus, the results would seem to suggest that in general the task of chick-sexing does not require a perceptual re-organisation process. Instead the main
difficulty (in addition to those associated with the manual aspects of chick-sexing) with the task probably lies in selecting an appropriate rule from the many possible rules that could be generated, as a result of the variety of features present in a single chick and the wide variation in genital structure between chicks. The latter source of variation increases the difficulty of discovering a discrimination rule by preventing it from being based on a single dimension. For instance, the rule discovered by Biederman and Shiffrar (1987) which distinguishes a female is a disjunctive rule, "either flat or pointy". However, the fact that in some cases the critical feature was not segmented as a distinct entity (e.g. participant ST's drawing in Figure 2.4) suggests that in some instances perceptual learning in the form of perceptual re-organisation (such as that reported by Hock, Webb and Cavedo, 1987) may be required before discrimination can be accomplished.

2.3 Summary

In this chapter, two learning phenomena have been re-examined, and it has been shown that rather than improvements in performance arising from perceptual learning, the majority of learning is likely to occur in these tasks as a result of the operation of other processes. It has further been suggested that in the case of the studies examined here, these processes are likely to centre on hypothesis generation, rule testing and concept formation. That the learning observed did not involve perceptual learning, was established through the use of two methods. One involved an operational test of the need for perceptual learning through the use of a simultaneous comparison procedure. The other method involved a parsing procedure which can be used to examine observers' perceptual segmentation of an image. Both of these methods can be used to rule out when perceptual learning is not required, or does not underlie the successful learning of a discrimination task. The next chapter turns to paradigms which can provide a positive indication of the occurrence of perceptual learning, and more specifically the occurrence of low-level learning. It is this type of learning for which the strongest argument can be made
that an observer really learns to perceive something new, and, consequently, it is this type of learning which was of the greatest interest, and on which there is the greatest emphasis, in this thesis.
3.1.1 The importance of plasticity in adult vision

That human vision should retain a degree of plasticity in adulthood would seem necessary for adaptive responses to the world around us, for as Zeki (1993, p. 207) writes:

"A totally hard-wired nervous system might make sense in a rigidly coded world, where everything can be identified by a unique label, where nothing changes and where every condition that an organism is likely to encounter, and every possible reaction to it, is predetermined and known beforehand. This, of course, is far from true."

In fact a visual system which was totally hard-wired, or which became fixed after early development, would seem counter to the biological principle of survival set out in Darwinian theory. The application of Darwin's principles to neural development has been conducted by Edelman (1987). He proposes that "the brain is dynamically organized into cellular populations containing individually variant networks, the structure and function of which are selected by different means during development and behaviour." Essentially, he proposes that patterns of connection among groups of cells (hundreds of thousands of strongly interconnected neurones) are initially established as a result of the developmental action of mechanochemical events; he describes this as establishment of the primary repertoires. Following this, sensory input leads to competition between groups for selection leading to the strengthening or weakening of synaptic connections; this he calls formation of secondary repertoires. Finally, Edelman (1987) proposes that through connection between local maps serving different modalities, the spatiotemporal continuity required for perceptual categorisation is provided. Thus, his theory describes a set
of processes which allow perceptual categorisation to occur without assuming a prearranged world. Although his theory is primarily concerned with the earlier development of the nervous system it may be that the types of processes he proposes, particularly regarding the formation of the secondary repertoire, underpin the adult perceptual learning phenomena examined in this chapter. In line with the sort of processes Edelman (1987) proposes, Frégnac, Shulz, Thorpe and Bienenstock (1988) have found evidence, in kittens and adult cats, to support the role of temporal correlation between pre- and post-synaptic activity during learning, in the induction of long-lasting modifications of synaptic transmission.

3.1.2 Low-level learning

The present chapter reports the results of four psychophysical studies conducted to examine evidence for low-level changes in perception as a result of experience, for two different aspects of visual analysis, the discrimination of stereoscopic depth, and contrast sensitivity. The first presents an interesting topic for examination since previous research on stereoacuity, as described in the taxonomy of learning presented in chapter one, suggests that it may constitute a relatively unique form of perceptual learning. The second is of fundamental importance for human visual ability, and represents an investigation of a perceptual learning process which could occur at an earlier level of visual processing than those previously studied.

In reporting these experiments this chapter demonstrates the use of methods which can positively indicate the occurrence of perceptual learning in contrast to the methods used in chapter two which can be used to rule out its occurrence.
Learning of stereoscopic depth discrimination

3.2.1 Learning to perceive stereograms

It has often been found that with repeated presentation of the same Julesz random-dot stereogram, the time required to see the global figure in depth, reduces from minutes to almost instantaneous perception. (e.g. Ramachandran, 1976). Early researchers often suggested that changes in the efficiency of vergence eye movements led to the learning effect observed. For instance, Julesz (1971) suggested that the long perception times found with random-dot stereograms arose because the observer had to make a series of vergence eye movements in order to bring corresponding dots into Panum's fusional area. Thus, over time, learning could occur through observers learning to make a more efficient series of vergence eye movements, perhaps avoiding divergent movements when convergent movements were required, and vice versa. However, more recently the role of eye movements has been questioned; Bradshaw, De Bruyn and Rogers (1992) state that although eye movements are an important limiting factor in perceiving complex stereograms they do not believe that they are the critical variable underpinning stereoscopic learning. They showed that complex random-dot stereograms with large disparities did not necessarily take longer to perceive than random-dot stereograms with small disparities. Other researchers have also suggested eye movements are not responsible for improved stereogram perception times (Ramachandran, 1976; Fendick and Westheimer, 1982) as learning is maintained even when stimuli are presented for 150 msec or less.

3.2.2 The nature of learning in stereopsis

More recently, there has been some debate over the exact nature of the learning process underlying improvements in stereopsis. Some researchers have reported findings that the learning process appears to be specific to certain dimensions such as retinal location (Ramachandran, 1976; O'Toole and Kersten, 1992) and
orientation of the stereogram elements (Ramachandran and Braddick, 1973) whilst others have found evidence for a more general learning process (e.g. Weinman and Cooke, 1982).

3.2.3 Specific learning in stereopsis

An example of specific learning was reported by Ramachandran (1976) who observed learning specific to retinal location. In his experiment, observers first learned to perceive a particular random-dot stereogram in one retinal location. Once perception times reached an asymptotic value the position of the stimulus was shifted to a new retinal location. Perception times immediately returned back to near their original starting levels and then proceeded to decrease with practice. For some observers, once perception times had reached asymptotic values in the new location, the stimulus was shifted back to the original position, and once again perception times were found to increase immediately. Ramachandran (1976) suggested this effect might indicate a learning process specific to anatomical areas of the brain. However, this type of explanation is implausible given that, for observers who took part in the second stimulus shift back to the original starting position, perception times increased to around their starting levels; ‘low-level’ learning is generally found to be much more enduring (e.g. Ball and Sekuler, 1987). A more likely explanation for this type of result was proposed by O'Toole and Kersten (1992) who also reported location specific learning effects for stereopsis. They suggested that the learning appeared specific because of selective spatial attention to the probable location of stimulus appearance (this may be particularly likely if trials are in close temporal succession). If this selection were to occur at a pre-cognitive level, then the ensuing expectancy could cause the activation of a series of cortical shifts3.1 for the previous rather than the current

3.1 Ramachandran (1976) suggests that as observers become more skilled vergence eye movements are replaced by ‘cortical shifts’ (see Anderson and Van Essen (1987) for ideas about neurophysiology of cortical shifting) leading to the perception of global stereopsis. It is likely that a sequence of cortical shifts is automatically activated for the expected spatial location of target appearance when subsequent trials are in close enough temporal succession.
spatial location; it has been shown that other aspects of stereo-perception are not amenable to high-level cognitive intervention so that just being told where to expect a stimulus to appear may not be enough to counter automatically activated processes. For instance, Frisby and Clattworthy (1975) examined the effect on perception times of presenting observers with a variety of different types of information about the nature of a 'hidden' cyclopean object in a random-dot stereogram. They reported that none of the types of prior information, ranging from a verbal description of the cyclopean object through to a full scale replica model, had an impact on perception times as compared with a no prior-information control group. This finding is also consistent with the idea of a 'low-level' perceptual learning process in learning to perceive random-dot stereograms where some kind of neural or other tuning is required which is not amenable to change through top-down control.

Ramachandran (1976) reported that a similar pattern of results to those he found within a single session, occurred when practice was distributed over a period of days. However, this could be explained by the finding that overnight increases in perception times are likely to occur in general as reported by MacCracken and Hayes (1976) and MacCracken, Bourne and Hayes (1977).

In a rather different paradigm Long (1982) reported specificity of learning to spatial dimensions of stereo-stimuli such as spatial frequency of the dots. In his study observers allocated to an experimental group first took part in training on a standard random-dot stereo-stimulus before transfer of learning to spatially transformed stereo-stimuli was tested (the stimuli were spatial frequency filtered). Observers in a control group just took part in training on the transformed stimuli. Long reports no transfer of learning, for observers in the experimental group, between the standard stimulus and the transformed stimuli; observers showed a reduction in perception times over the course of training with the standard stimulus.
but showed a large increase in perception times when they were switched to the transformed stimuli. However, close examination of the results suggests that considerable transfer of learning did take place. Observers in the control groups who only trained on the transformed stimuli were considerably slower to perceive them (by 10 - 15 seconds on average). This strongly suggests that in the experimental group the much quicker perception times for these same stimuli were due to considerable transfer of learning from their prior training on the standard stimuli. In fact O'Toole and Kersten (1992) suggest that stereopsis is not specifically tuned to spatial dimensions of the target quoting work by Mayhew and Frisby (1978, 1981) which indicates that contrary to the findings of Ramachandran and Braddick (1973) stereopsis is not tuned to the spatial dimension of orientation.

3.2.4 General learning effects

In contrast to the reported specific learning effects many researchers have made the informal observation that previous experience with random-dot stereograms leads to a general facilitation of perception for other random-dot stereograms on subsequent exposures. Weinman and Cooke (1982) formally assessed this possibility and reported that previous experience with random-dot stereograms led to significantly faster perception times for a test stereogram as compared with a naive control group.

3.2.5 Improvements in stereoacuity

Evidence for fine tuning of stereoscopic depth perception mechanisms comes from Fendick and Westheimer (1983) who investigated the degree to which stereoacuity at the fovea and various retinal eccentricities could be improved through extensive practice (6000 trials). Their task required observers to discriminate the relative depth of squares which were defined by an outline of dots. As such it is unclear whether this task was examining local or global stereopsis (or both). They found considerable improvement in thresholds at peripheral locations over the first 3000-
4000 trials. However, improvement at the fovea was only observed in some observers. The process by which this improvement occurred was not speculated upon in their paper, but it is conceivable that the kind of fine tuning of stereo perception mechanisms observed, resulted from 'low-level' perceptual learning. Evidence which supports this supposition has been reported by Fahle (1993b). Using a stereoscopic display system he presented observers with two dots which were at different depths. The observer's task was to indicate which of the dots was the closer. The dots were either displayed one above the other (vertically) or they were displayed side by side (horizontally). Observers completed the first half of training on one dot arrangement and then transferred to the other arrangement. The results indicated that observers' sensitivity to the stereoscopic depth differences improved as a function of practice but did not transfer when the arrangement of the dots was changed (e.g. from vertical to horizontal or vice versa). However, it is not possible to determine from this study whether the specificity of learning observed resulted from the operation of selective spatial attention as discussed in section 3.2.3. The present study further investigates this possibility.

3.2.6 Learning processes in stereopsis

As the above review indicates, there appear to be multiple processes implicated in stereo learning. First, there is some suggestion that initial learning about correct vergence eye movements may later be superseded by cortical shifts. This type of learning may result in remembered 'phase-sequences', which can be automatically activated in anticipation of the likely spatial location of a target, leading to observations of apparently location specific learning that may have little to do with change in the tuning of visual perception mechanisms at a neural level. In order to control for learning about eye movements, the present study utilises presentation times which are faster than the time taken to initiate a voluntary eye movement. Second, there are more general learning processes which appear to result in a general improvement in perception times for stereograms. Third, there is a much
longer term fine-tuning process leading to improved stereoacuity. This last process is the most likely candidate for 'low-level' learning and the present study further investigates this possibility through the use of a transfer of training design which investigates the degree of neural specificity of learning.

3.2.7 Outline of the present study

The present study re-examined the specificity of learning of stereoscopic depth discrimination. Previous research has shown that some binocular cells respond only when their preferred features appear at the same depth plane as the point of fixation, whilst others only respond when their preferred features fall on a different depth plane to the point of fixation. These cells have been called disparity selective cells. Further, whilst some of these cells are responsive only to features falling behind the point of fixation (uncrossed disparities), others are responsive only to features falling in front of the point of fixation (crossed disparities) (e.g. Poggio and Fischer, 1978; Ferster, 1981). Further, research has indicated that when making judgements about depth based on disparity detection, there are some binocular cells that respond to specific areas of the visual field and to specific sizes and types of disparity (Barlow, Blakemore and Pettigrew, 1967). In this context it is hard to conceive of disparity specific learning in the primary visual cortex without retinal location specific learning, or vice versa. The present study utilised these findings to design the transfer of training tasks which investigated the neural specificity of learning. Observers first took part in a pre-training measurement phase during which their ability to discriminate the relative depth of two stereograms with various disparities, presented in either the upper or lower visual field and for crossed and uncrossed disparities was measured. All possible combinations of these stimulus parameters were measured. Following the baseline measurement phase, observers took part in a training phase, distributed over a number of days, during which they practised discrimination of the relative depth of two stereograms presented either only at crossed (half the observers) or uncrossed disparities (the other half of
observers) for any one retinal location. Because during training, a particular location of presentation was always combined with only one direction of disparity (crossed or uncrossed), the neural specificity of learning could be assessed following training, through testing transfer from trained locations and disparities, to untrained locations and disparities. Consequently, following the training phase, observers' ability to discriminate the relative depth of the two stereograms for all the types of trial that were measured in the baseline phase was re-measured in a post-training measurement phase. By presenting stimuli at each location equally often during training, any specific learning phenomenon revealed in the post-training transfer testing which was due to the operation of selective spatial attention mechanisms could be ruled out.

Throughout the experiment stimuli were presented for shorter times than that required to initiate a voluntary eye movement (Carpenter, 1977, pp. 56-77) leading to presentation to a constant retinal location and thus excluding the learning of vergence eye movements as an explanation for any learning observed.

3.2.8 Method

Observers

Sixteen observers were recruited to take part in the study, five women and eleven men. Observers' ages ranged between 20 and 48 years, and they had a mean age of 29.8 years. All of the observers had either no experience or very limited experience of the type of stereoscopic task to be used in this experiment. The observers were all screened for stereo blindness using random dot anaglyph stereo stimuli (see Appendix A for stereo-blindness test) before being included in the study. All of the observers had normal or corrected to normal vision.
**Apparatus and stimuli**

Stimuli for the task were generated by a Silicon Graphics Iris 4D Series Computer. They were displayed on a Tektronix SGS625 17" stereoscopic display system at a resolution of 1280 by 1024 pixels. Observers viewed the display from a distance of 115 cm. At this distance the screen subtended a visual angle of 16.72 degrees by 13.51 degrees.

A single colour gun was used to generate the stimuli because the red, green and blue guns of the monitor address pixels that are slightly displaced from each other causing the location of an edge of mixed colour to be poorly defined. Green was chosen because it is closest to the peak of the human spectral sensitivity function.

In the centre of the screen a small green fixation dot was displayed. This subtended a visual angle of 3 min arc. Above and below the dot vertical nonius lines appeared. The nearest end of the lines was 24 min arc away from the edge of the fixation dot. The lines subtended a visual angle of 3 min arc by 5.29 degrees. The top nonius line was displayed to the left eye and the bottom nonius line was displayed to the right eye. The stereograms used in the experiment appeared as green squares of the same brightness and hue as the fixation dot and nonius lines (C.I.E. co-ordinates, \( Y = 25.1 \text{ cd/m}^2, x = 0.281, y = 0.591 \): C.I.E. co-ordinates when viewed through circularly polarising spectacles, \( Y = 9.47 \text{ cd/m}^2, x = 0.279, y = 0.599 \)). Two laterally separated stereograms were displayed at a time against a blank background (\( Y = 0.005 \text{ cd/m}^2 \)) either above the fixation dot or below the fixation dot. Each square had side lengths that subtended a visual angle of 2.69 degrees. The nearest corner of the squares was 1.9 degrees away from the fixation dot in the zero disparity depth plane. A diagram of the screen layout can be seen in Figure 1.1
Each stereogram was generated from two squares which had been laterally separated by 2' 25", 3' 13", 4' 1" or 4' 49" degrees of visual angle (experimental stimuli) or by 4' 49", 9' 41" or 14' 31" degrees of visual angle (familiarisation stimuli); the size of this disparity determined the apparent depth of each stereogram. The disparities of all the stimuli were such that fusion would be possible even when they were displayed for only 117 msec (Ogle, 1952; Schor and Tyler, 1980; Woo, 1974). The fixation dot was defined as the zero disparity depth plane. One half of each stereogram pair was presented to each eye through the use of a time-multiplexed display system (see Wickens, 1990 for a detailed description), essentially, by a combination of the display of each eye’s view on alternate frames (monitor refresh rate 60 Hz) together with the use of light polarisation and polarised filters.

**Design**

There were twelve different types of experimental stimuli, resulting from all the possible combinations of three different factors. First, the stereograms in each pair were always presented at a different depth from each other. Table 3.1 shows the different combinations of disparity which pairs of stimuli could take and the
absolute difference in disparity between each pair (note that there are three possible absolute disparities).

Table 3.1 Disparity values for each stimulus and absolute difference.

<table>
<thead>
<tr>
<th>Stimulus 1</th>
<th>Stimulus 2</th>
<th>Absolute difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>2' 25&quot;</td>
<td>3' 13&quot;</td>
<td>0' 48&quot;</td>
</tr>
<tr>
<td>2' 25&quot;</td>
<td>4' 01&quot;</td>
<td>1' 36&quot;</td>
</tr>
<tr>
<td>2' 25&quot;</td>
<td>4' 49&quot;</td>
<td>2' 24&quot;</td>
</tr>
<tr>
<td>3' 13&quot;</td>
<td>4' 01&quot;</td>
<td>0' 48&quot;</td>
</tr>
<tr>
<td>3' 13&quot;</td>
<td>4' 49&quot;</td>
<td>1' 36&quot;</td>
</tr>
<tr>
<td>4' 01&quot;</td>
<td>4' 49&quot;</td>
<td>0' 48&quot;</td>
</tr>
</tbody>
</table>

The resulting differences in the relative depth of the squares should have been at around observers' stereoacuity threshold (Fendick and Westheimer, 1983; Woo, 1974). Second, the pair of stereograms was either presented at crossed disparities (in front of fixation) or uncrossed disparities (behind fixation). Finally, the pair of stereograms could either be presented above the fixation dot or below the fixation dot.

The experiment used a test-training-retest design spread over five days. On the first day observers completed a number of practice trials using the familiarisation stimuli followed by a baseline measurement of their performance for each of the twelve types of experimental stimulus. On the next three days observers took part in a training procedure using a subset consisting of half of the experimental stimuli; the subset used varied between observers (see below). On the final day the baseline measures were repeated. Observers were allocated to one of two training groups to counterbalance the design. For one group, all stimuli that were presented during training above the fixation dot were presented at crossed disparities, and all stimuli presented below the fixation dot were presented at uncrossed disparities. For the second group, all of the stimuli that were presented during training above the fixation dot were presented at uncrossed disparities, and all stimuli that were presented below the fixation dot were presented at crossed disparities. Presentation
of the stereograms to specific retinal locations was achieved through 'faster than eye movements presentation' (in this case stimuli were presented for 117 msec).

Procedure

On the first day observers were given a standard instruction sheet (See Appendix B). Once observers had read this, they were given the opportunity to ask for further explanation about anything they were unsure of. Once observers were sure that they understood the experimental procedure, the room lights were turned out and a ten minute dark adaptation period was started. Following this, observers began the practice trials. The observers' task was to indicate which of the two stereograms appeared to be closest to them. Observers completed 50 practice trials on the familiarisation stimulus set. This allowed them to accustom themselves to the task procedure and also allowed time for variability in reaction times to be reduced. After finishing the practice trials, observers completed the baseline measurement phase consisting of 120 trials in total, ten with each of the twelve types of experimental stimulus. The trials were presented in random order. On the next three days, following a ten minute dark adaptation period, observers completed 200 training trials per day with the appropriate subset of experimental stimuli. On half the trials stimuli were presented above the fixation dot, and on the other half below the fixation dot. Again presentation order was randomised. On each trial one of the three possible disparity differences between the two stereograms was pseudo-randomly chosen (this ensured that approximately equal numbers of trials were presented for each absolute difference in disparity). On approximately half the trials the left stereogram would appear to be closer when seen in depth and on the other half of the trials the right stereogram would appear to be closer. On the final day, following the dark adaptation period, the 120 baseline trials were repeated. On all days the presentation of the stereograms was observer paced. Once observers were sure they were correctly fixated they initiated a trial by pressing the spacebar. The stimuli were then displayed for 117 msec. A tone sounded simultaneously with
the presentation of a stimulus to aid observers' temporal attention. Observers indicated which of the two stereograms appeared closer to them by pressing a mouse button. They pressed the left mouse button if the left stereogram appeared closer and the right mouse button if the right stereogram appeared closer. Observers were given no external feedback about the accuracy of their choices. The computer recorded reaction times, from appearance of the stimulus to the pressing of a mouse button, in milliseconds. Each experimental session lasted approximately 20 minutes.

3.2.9 Results

The data from one observer, who had extensive practice on the experimental task during its piloting and development, were excluded from the following analysis as he showed a high level of performance from the first day of the actual experiment (raw scores for the remaining fifteen observers can be seen in Appendix C).

Note on effects that can be tested

As described above disparity and location are separately considered in terms of the number of types of experimental stimuli. However, as detailed in the introduction to this experiment, for each observer a particular location of presentation was always linked with a particular type of disparity during the training sessions. Thus, separate effects of location of presentation and type of disparity cannot be tested. Consequently, data from the trials which presented the types of experimental stimuli that an observer would have seen during training, have been combined to provide a measure of performance on training trials. Data from trials on which the remaining types of experimental stimuli were presented, those which an observer would not have seen during training, were combined to provide a measure of performance for untrained types of trial (transfer trials). This allows testing of a single factor, training vs. transfer trials.
Change over time in observers' performance.

The first question to be assessed was whether, as expected, observers did learn over the course of the experiment. Mean accuracy scores (percent correct) and reaction times were calculated for each observer for training trials for pre- and post-training days, and for each of the three training-days. These means were further subdivided into one mean for each of the three disparities. Data for the transfer trials were not included in the pre- and post-training data, as this could have led to any trends over time being confounded with whether transfer of learning did or did not occur following training.

Accuracy data

A two way analysis of variance (time of measurement (5) - pre-training, training-day 1, training-day 2, training-day 3, post-training; disparity (3) - 0.01°, 0.03°, 0.04°) with repeated measures on both factors revealed a significant change over time in observers' accuracy ($F(4,56)=2.79$ $p=0.035$). Figure 3.2 illustrates this change in mean accuracy over time.

![Figure 3.2](image_url) Change in mean accuracy (% correct) over time (error bars indicate ± 2 standard errors of the mean).
Polynomial contrasts revealed a significant quadratic trend \((t=3.01 \ p=0.009)\). Observers improved over the first four days of the experiment but there was a slight drop off in performance on the post-training session. Longer term follow-ups would be required to assess whether this drop-off continued, or whether it was simply a function of other factors operating in the post training session; for instance, the fact that transfer trials were randomly interleaved with training trials during this session (this was obviously not the case during the training sessions).

As expected, the analysis of variance also revealed a significant difference in mean accuracy between the three different levels of disparity \((F(2,28)=11.70 \ p=0.001)\). Table 3.2 shows the mean accuracy scores and standard deviations for each disparity level.

Table 3.2 Mean accuracy scores and standard deviations (in brackets) for each of the three disparity increments.

<table>
<thead>
<tr>
<th>Disparity (degrees)</th>
<th>Mean percent correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01°</td>
<td>56.65 (6.86)</td>
</tr>
<tr>
<td>0.03°</td>
<td>59.84 (9.33)</td>
</tr>
<tr>
<td>0.04°</td>
<td>65.20 (11.57)</td>
</tr>
</tbody>
</table>

Polynomial contrasts revealed a significant linear trend \((t=4.00 \ p=0.001)\). Observers became more accurate as disparity increased. There was no significant interaction between disparity and time of measurement.

**Reaction times**

A two way analysis of variance with repeated measures on time of testing and size of disparity, indicated a significant change in observers' reaction times over the course of training \((F(4,56)=19.03 \ p<0.0001)\). Figure 3.3 illustrates this change.
Figure 3.3 Change in reaction times over days (again error bars show ± 2 standard errors of the mean).

Polynomial contrasts revealed a significant linear trend for observers' reaction times to decrease over the experimental sessions ($t=5.59 \ p<0.0005$). The main effect of disparity and the interaction between disparity and time of measurement were not significant.

**Comparison of pre- and post-training data**

**Accuracy data**

The mean scores and standard deviations for observers' accuracy on trained and untrained types of trial over time can be seen in table 3.3.

**Table 3.3 Mean accuracy scores (% correct) and standard deviations (in brackets).**

<table>
<thead>
<tr>
<th>Disparity</th>
<th>Pre-training</th>
<th>Post-training</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training trials</td>
<td>Transfer trials</td>
</tr>
<tr>
<td>0.01°</td>
<td>54.03 (9.67)</td>
<td>50.07 (10.98)</td>
</tr>
<tr>
<td>0.03°</td>
<td>52.59 (10.85)</td>
<td>56.18 (15.31)</td>
</tr>
<tr>
<td>0.04°</td>
<td>59.17 (14.66)</td>
<td>58.39 (17.40)</td>
</tr>
<tr>
<td>Mean</td>
<td>54.26</td>
<td>54.88</td>
</tr>
</tbody>
</table>

In line with the previous results a three way analysis of variance (time of measurement (2) - pre-training, post-training; disparity (3) - 0.01°, 0.03°, 0.04°;
type of trial (2) - training, transfer) with repeated measures on time of measurement, type of trial and size of disparity, indicated a significant change in observers' accuracy from pre- to post-training ($F(1,14)=4.89, p=0.044$). Examination of the mean scores revealed that observers were more accurate following training (pre-training mean = 54.57%, post-training mean = 63.49%). However, the prediction that learning would be specific to types of trial on which observers trained was not supported, as indicated by the absence of a significant time by type of trial interaction ($F(1,14)=1.91, p=0.189$). In fact examination of the mean scores tends to suggest that there is more improvement on transfer trials than on the training trials.

The analysis of variance also indicated a significant effect of the size of the disparity ($F(2,28)=4.03 p=0.029$). Polynomial contrasts revealed a significant linear trend ($t=2.36 p=0.03$). Observers were more accurate as the size of the disparity increased ($0.01^\circ$ disparity mean = 55.66%, $0.03^\circ$ disparity mean = 59.14%, $0.04^\circ$ disparity mean = 63.04%). There were no other significant main effects, or interactions.

**Reaction time data**

The mean scores and standard deviations for observers' reaction times on training and transfer trials over time can be seen in table 3.4.

<table>
<thead>
<tr>
<th>Disparity types</th>
<th>Pre-training</th>
<th>Post-training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trained trial types</td>
<td>Untrained trial types</td>
<td>Trained trial types</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>$0.01^\circ$</td>
<td>866 (261)</td>
<td>864 (284)</td>
</tr>
<tr>
<td>$0.03^\circ$</td>
<td>888 (282)</td>
<td>888 (292)</td>
</tr>
<tr>
<td>$0.04^\circ$</td>
<td>859 (250)</td>
<td>837 (252)</td>
</tr>
<tr>
<td>Mean</td>
<td>871</td>
<td>863</td>
</tr>
</tbody>
</table>

Table 3.4 Mean reaction times (msec) and standard deviations (in brackets).
Analysis of variance indicated a significant difference in reaction times between pre- and post-training (\(F(1,14)=18.64, \ p=0.001\)). Examination of the mean scores reveals that observers were faster following training (pre-training mean = 867 msec, post-training mean = 656 msec). Again the prediction that learning would be specific to the types of trial on which observers trained was not supported (\(F(1,14)=0.24 \ p=0.632\)). There were no other significant main effects or interaction effects.

### 3.2.10 Discussion

This study sought to re-examine the evidence for 'low-level' learning in stereopsis. Observers' ability to discriminate which of two stereo stimuli was closer, and the speed with which they made these discriminations, improved as a result of practice. However, no evidence was found to support the hypothesis that this learning occurred at a level which was specific to the cells stimulated during training. The improvement in discrimination accuracy and in reaction times transferred fully across either type of disparity, or retinal location, or both.

### 3.2.11 Specificity of learning in stereopsis

Previous research has indicated that some of the learning in stereopsis is specific to retinal location (e.g. Ramachandran, 1976; O'Toole and Kersten, 1992). These results can be explained in terms of selective visual attention, based on the signal probability for spatial locations, of the kind demonstrated by Bashinski and Bacharach (1980). The present study sought to examine whether in addition to this type of specific learning, a specific 'low-level' form of learning occurred. This was examined through a very specific measure of neural specificity of learning, and a paradigm in which stimuli were presented in a random sequence such that an observer would not develop expectations about where the next pair of stimuli would be presented, or about what type of disparity they would be presented with. Given these controls, a specific learning effect would have been strongly suggestive of
learning occurring at a ‘low-level’ of visual processing. However, the absence of such an effect indicates that there is no reason to reject the suggestion that the specific learning effects observed by other researchers can be attributed to selective visual attention.

3.2.12 Improvements in stereoacuity

Given that no evidence for a ‘low-level’ learning process was obtained in the present study, this leaves open the question of what is the mechanism by which the improvements in stereoacuity reported by Fendick and Westheimer (1983) and Fahle (1993b) occur, and where is the likely location of such learning? The present study, consistent with Fendick and Westheimer (1983) and Fahle (1993b), provided evidence that observers could improve their ability to make judgements about small differences in relative depth on the basis of stereo depth information. However, the results of the present experiment would suggest that this learning is not specific to the cells stimulated during training. Thus, in that respect the mechanism by which this learning occurs is likely to be different to that seen with learning in vernier hyperacuity tasks (e.g. Fahle and Edelman, 1992) where specificity of learning to location of presentation, and to a number of stimulus dimensions, such as stimulus orientation, is usually observed.

3.2.13 Monocular cues for depth perception

The present study used non random-dot stereograms as stimuli. One of the traditionally claimed advantages of random-dot stereograms over non random-dot stereograms, was that random-dot stereograms were thought to be devoid of monocular cues to the nature of the stereo form; this claim has been called into question by O’Toole and Kersten (1992). Given that there can be monocular cues in non random-dot stereograms, it might be proposed that observers in the present study could have made depth judgements based upon the monocular cue arising from the horizontal separation of the half stereograms: the size of this separation
would have determined the relative depth of the stereograms. The difference in the size of the separation for the two stereograms shown on any one trial would have been well within naive observers hyperacuity thresholds (Kumar and Glaser, 1993). However, this explanation can almost certainly be ruled out, as during the pre- and post-testing sessions, trials were pseudo-randomly presented, such that prior to each trial, observers would not have known whether to expect a crossed or uncrossed disparity trial. On crossed disparity trials, the stereogram with the larger separation would have appeared closer, whilst for uncrossed disparity trials, it would have been the further away. Since observers performed at above chance levels in the post-training session they must have at least been perceiving the stereo images in depth in order to know how to interpret the monocular separation information. Given this fact it would seem likely that it would be the stereo information rather than the monocular information which observers used to make the depth discrimination.

3.2.14 Summary
The present study has demonstrated the use of a paradigm which allows determination of whether a learning phenomenon occurs at an early level of visual processing, by using knowledge regarding the structure of the brain obtained from previous psychophysical and neurophysiological research. Through the use of these methods, it was indicated that the locus of the perceptual learning observed was likely to be at later levels of visual processing, or in higher level mechanisms. That the learning can be called perceptual learning, is justified in that, even with simultaneous stimulus presentation, observers' depth discrimination improved as a result of practice. Thus, they learned to "perceive variables of physical stimulation not previously sensed or perceived". However, the exact mechanism of this learning process is left open, as its determination is beyond the scope of this thesis. The next three experiments develop further the use of transfer measures based on knowledge about the neural structure of visual processing mechanisms in an
investigation of learning for contrast sensitivity. In doing so it is demonstrated that with a sufficient range of transfer measures a fairly detailed localisation of learning can be obtained.
Changes in contrast sensitivity with experience

3.3.1 The importance of contrast sensitivity for vision

One of the fundamental steps in perceiving our visual environment is to segment it using contrast boundaries which define texture and objects (Marr and Hildreth, 1980; Marr 1982). Indeed, in addition to the limits imposed by visual acuity, our ability to perceive these boundaries places a finite limit on the detail level of the texture we perceive in the world around us. Frequently, the boundaries in our visual environment are of a high enough contrast to be easily perceived. However, in some circumstances boundaries are of lower contrast, for instance, as distance increases contrast boundaries become less distinct. Further, some tasks, such as examining X-rays, can be critically dependent on the detection of low contrast boundaries and may depend on heightening contrast sensitivity (Davies, Sowden, Hammond and Ansell, 1994). An improvement in our sensitivity to low contrast levels could consequently have an impact on our ability to perceive the visual world in circumstances such as these.

3.3.2 The use of sinusoidal gratings to measure contrast sensitivity

Sinusoidal gratings appear as a pattern of light and dark bars. At their simplest the variation in the luminance of these bars is described by a single sine wave. More complex patterns can be described by a combination of sinusoidal components, and indeed the Fourier Theorem demonstrates that a periodic waveform of any complexity can be broken down into the linear sum of harmonically related sine and cosine waves of specified frequencies, amplitudes and phases. Sinusoidal gratings are good stimuli for the systems analysis approach generally used to examine vision since they can be passed through a linear system without alteration in their basic shape but merely in amplitude and phase. The Michelson/Rayleigh contrast of a sinusoidal grating can be calculated as follows:
An observer's contrast sensitivity can be calculated using sinusoidal stimuli and an appropriate psychophysical method to measure contrast threshold.

3.3.3 Evidence for long term changes in contrast sensitivity as a result of experience

Two studies reported in the literature, provide some evidence for improvements in contrast sensitivity as a result of extensive and focused exposure to sinusoidal gratings. De Valois (1977) conducted an experiment in which observers' contrast thresholds were measured for a wide range of spatial frequencies using sinusoidal gratings as stimuli. Following this, they adapted to a high contrast grating (training stimulus: contrast = 95%) for a five minute period. Immediately after this adaptation period, all of the observers' thresholds were again measured. De Valois found that observers' contrast sensitivity decreased for stimuli with spatial frequencies within ± one octave of the training stimulus. For stimuli with spatial frequencies within ± one-two octaves of the training stimulus there was no reliable change in contrast sensitivity. As the spatial frequency of the test stimulus increased to ± two octaves and beyond that of the training stimulus observers showed a significant increase in their contrast sensitivity. This finding has been subsequently confirmed by the research of others (e.g. Tolhurst and Barfield, 1978).

An interesting 'side-effect' of this study was that after taking part in tasks such as these over a one year period observers showed an increase in contrast sensitivity. At spatial frequencies up to 3 cpd this increase was about 0.8 of a log unit. At spatial frequencies above 3 cpd the increase was progressively reduced as the spatial frequency increased such that at the maximum spatial frequency measured (22.63 cpd) the increase was about 0.35 of a log unit. Thus, this study provided almost incidental evidence that long term changes in contrast sensitivity could occur as a result of extensive experience.
A second study which also provided some evidence that contrast sensitivity could be increased through experience was conducted by Mayer (1983). She investigated whether the usual pattern of anisotropy\(^3\) (variation in sensitivity for contours at different orientations) could be influenced through experience. In her experiment observers' contrast sensitivities were measured for gratings presented at four different orientations: vertical, horizontal, left diagonal and right diagonal. Following this, observers completed 3000 signal detection trials with feedback, for gratings presented at the orientation to which they were least sensitive; for five out of six observers that orientation was a diagonal reflecting the usual pattern of anisotropy observed in Caucasian adults. Following the training period, observers' thresholds were re-measured for all four orientations. The results indicated that sensitivity increased for the trained orientation and also for the orthogonal orientation, whilst there was no change in sensitivity for the horizontally or vertically oriented stimulus (or in the case of the observer trained on a cardinal orientation no change had occurred for the obliquely orientated stimuli). One observer then went on to practice detecting gratings presented either vertically or horizontally. No change in contrast sensitivity occurred following this practice. The results thus indicated that an increase in contrast sensitivity was possible, but only for stimuli presented at orientations for which a hypothetical peak sensitivity had not been reached. Unfortunately the durability of learning was not assessed in this study.

The present series of studies seek to study systematically and in greater detail the degree to which contrast sensitivity can be changed in adults through extensive practice. By using various transfer of learning tasks it will be possible to give an indication of the likely neural locus of any learning observed. In addition, the studies use various training paradigms in an attempt to determine the factors

---

\(^3\) Previous evidence has indicated that in general eyes without optical error nevertheless vary in sensitivity for contours at different orientations (e.g. Appelle, 1972), this is referred to as anisotropic sensitivity.
influencing the size of any change in sensitivity observed. Finally the retention of learning will be assessed in order to provide further information regarding the nature of any increase in sensitivity which occurs.

Contrast sensitivity experiment one

This experiment examined whether observers could improve their ability to detect low-contrast sine-wave gratings. Further, the study sought to determine the likely neural locus of any learning which was observed, through transfer of learning tasks. Three dimensions to assess transfer across were selected based upon the findings of previous psychophysical and neurophysiological research. First, it is now well established that one of the fundamental dimensions along which visual input is analysed is spatial frequency (for a review of evidence see Shapley and Lennie, 1985); spatial frequency loosely corresponds to size. Essentially, the spatial frequency which will cause the peak response in a neuron varies between neurons, and further, the size of the range of spatial frequencies to which neurons will respond (their bandwidth) varies between neurons located in different areas of the brain. Second, previous research has established that a topographical representation of the visual world is maintained in the visual processing areas of the brain (e.g. see Zeki, 1993, chapter three for a general review). In other words points, which are next to each other in the real world, are analysed by neurons whose receptive fields are located next to each other, thus, maintaining the spatial layout of the world. Third, it has been found that whilst cells in the LGN are monocular (i.e. they respond only to the input from one eye, e.g. Hubel and Wiesel, 1961), cells in the visual cortex can be either monocular or binocular, and further, in the case of binocular cells a preference may be exhibited for input from one eye (e.g. Hubel and Wiesel, 1965, 1968). Consequently, in the present experiment observers' baseline thresholds were measured for a range of stimuli, prior to them starting a ten day training period over which they completed 10,000 temporal forced-choice
trials for one type of stimulus. These stimuli had various spatial frequencies, could be presented in different spatial locations, and were presented separately to each eye. There was a weekend break in between the first five days and the second five days of training. On each trial, observers had to detect in which of two intervals a briefly presented sinusoidally modulated luminance grating appeared. The grating was presented near an observer's contrast threshold (<1 dB above). Following training, observers' baseline thresholds were re-measured to assess whether any learning had occurred, and if so whether it had transferred across eyes, retinal locations or spatial frequency. These measurements were repeated 19, 50, 80 and 195 days after the start of the experiment. Specifically, transfer of learning to a grating with a spatial frequency one octave below the training stimulus and transfer of learning to other retinal locations were examined, including inter- and intra-ocular transfer and inter- and intra-hemispheric transfer. In combination, these tasks could be used to give an indication of the neural substrate underpinning any learning that occurred.

3.4.1 Method

Observers

Four male observers volunteered to take part in the experiment. All had normal or corrected to normal vision. The observers were aged 24, 25, 46 and 53. None of the observers had participated in any experiments involving contrast grating detection before.

Apparatus and Stimuli

The stimuli were produced by an Innisfree Picasso waveform generator which was controlled by a Cromemco System 3 computer via a D+7A, 8-bit analogue and digital output card, and were displayed on a Tektronix 608 oscilloscope with a P31 phosphor. Observers were seated 100 cm away from the display. Viewing distance
was maintained through the use of a chin and forehead rest. From this distance the display subtended a visual angle of 6.84° by 5.71°. The monitor was placed on a high table such that it was at around observers' eye level when they were seated. The table was covered with black felt to prevent screen reflections from the table surface. Behind the monitor was an illuminated background; a piece of hardboard extending 17.4° above the monitor and 10.4° to each side. The mean luminance of this background was 9.41 cd/m². The rough side of the hardboard faced the observer to diffuse reflection of the light source, which was located underneath the table on which the monitor was placed, facing the hardboard, such that the observer would not have received any direct illumination.

The stimuli were circular patches of sine wave gratings with a spatial frequency of 4 cpd (or in the case of one stimulus 2 cpd), and subtended a visual angle of 2.0°. A fixation dot subtending a visual angle of 6' was placed in the centre of the oscilloscope screen. The nearest edge of the grating was 1.37° away from this fixation point. The gratings were displayed in one of three locations, either in the top left, top right or bottom left quadrants of the screen. Figure 3.4 shows a schematic diagram of the screen set-up and the three possible locations in which a grating could be displayed.

![Figure 3.4 Schematic representation of the fixation dot and the three possible locations in which a stimulus was displayed.](image)
The contrast of the gratings was varied in 0.05 log unit steps using a purpose built dB attenuator. The gratings were displayed against a uniform luminance field which matched the spatially averaged luminance of the gratings (Y=9.5 cd/m^2).

**Design**

The overall design of the study was that observers initially had their contrast thresholds measured for three retinal locations in the left eye, and two retinal locations in the right eye, using the 4 cpd stimulus. Specifically, in the left eye these were sections of the lower nasal retina, the lower temporal retina and the upper nasal retina, and in the right eye they were in the lower nasal and lower temporal retina. In addition, one threshold was measured for the lower nasal retina in the left eye using the 2 cpd stimulus. Thresholds were measured using two interleaved staircases (Cornsweet, 1962). Each staircase consisted of a run of 30 trials with an average of 6.16 turnarounds in each staircase. The threshold was calculated by taking the mean of the contrast values from all of the turnarounds within a staircase and then averaging across the two staircase means. For runs where the staircases clearly did not even come close to converging (average separation > 6 dB), the threshold measurement was repeated and an average over the four staircases was taken. Presentation of the gratings to a constant retinal location was achieved through exposure times which were quicker than the speed with which a voluntary eye movement can be initiated (Carpenter, 1977); in this case exposure times were 117 msec.

Following the baseline measurements taken on the first day, observers completed ten days training. On each day observers completed 1000 temporal two-alternative forced-choice trials, for one retinal location in the left eye (lower nasal retina), with a grating which had a spatial frequency of 4 cpd.
After the ten days training observers thresholds were once again measured, on the twelfth day, for all of the baseline locations and spatial frequencies. These measurements were repeated 19, 50, 80 and in the case of one observer 195 days after the first day of training. The measurements taken 19 and 50 days after training were for the training location only.

Procedure

On each day the equipment was switched on at least half an hour prior to the beginning of the experiment in order to allow time for it to stabilise. Throughout experimental sessions observers were seated on a comfortable chair in a dark room. Each experimental session began with a ten minute dark adaptation period during which time observers adapted to the uniform luminance field displayed on the oscilloscope.

Threshold measurement procedure

A computer program was written in Cromemco BASIC which allowed semi-automatic measurement of thresholds. Once the dark adaptation period was over observers were asked if they were ready to begin. Thresholds were measured for the following: the left eye lower nasal retina with a 4 cpd stimulus; the left eye lower nasal retina with a 2 cpd stimulus; the left eye upper nasal retina with a 4 cpd stimulus; the left eye lower temporal retina with a 4 cpd stimulus; the right eye lower nasal retina with a 4 cpd stimulus, and finally the right eye lower temporal retina with a 4 cpd stimulus. The order of measurement was varied across observers. Monocular presentation was achieved through observers wearing an eye patch over the eye on which measurements were not currently being conducted. There was a five minute time interval between measurements on the left eye and measurements on the right eye, during which time observers re-adapted their right eye to the oscilloscope screen. Observers were shown, prior to starting each run, where the stimulus for that run would be presented. One staircase began with the
attenuator set to 30 dB attenuation (the resulting Michelson contrast of the grating was 1.3%) whilst the other began with it set to 20 dB attenuation (the resulting Michelson contrast of the grating was 4.0%). The experimenter randomly selected which staircase to begin the measurement procedure with. Once observers were happy that they were fixating correctly, they initiated a trial by pressing a key, a grating was then presented for 117 msec. A single tone sounded concurrently with grating presentation to signal to observers the exact time interval during which a grating was presented. Following the presentation interval, observers indicated whether they had seen the grating or not. In accordance with standard staircase procedure, if they said that they had seen the grating the attenuation level was increased for the next trial in that staircase, whereas, if they said that they had not seen the grating, the attenuation level was decreased for the next trial in that staircase. Until the first turnaround in each staircase the attenuation level was adjusted in 2 dB steps, thereafter it was adjusted in 1 dB steps. As there were two interleaved staircases, alternate trials came from alternate staircases. The experimenter recorded the attenuation level for each trial. After a run of 30 trials in each staircase, the measurement of that particular threshold was stopped. If the two staircases did not converge, the threshold measurement was repeated. This was necessary on five occasions over the course of the whole experiment. The experimenter calculated the thresholds after each experimental session.

Training sessions procedure

A computer program was written in Cromemco BASIC which fully automated the training sessions procedure. Following the ten minute dark adaptation period observers pressed a key on the computer keyboard to start the experiment. Once the observer was happy that he/she was fixating correctly he/she pressed a key to initiate the first trial. The task was a temporal two-alternative forced-choice. After a 500 msec pause, the first presentation interval occurred signalled by a concurrent 'beep'. Following this, there was a 500 msec inter-stimulus-interval during which a
uniform luminance field of the same luminance as the spatial average of the grating was displayed, and then the second presentation interval occurred. The presentation intervals were 117 msec long each. The observer's task was to indicate in which of the two intervals a grating was presented. He/she responded by entering either a '1' or a '2' on the computer keyboard. Once he/she was satisfied that he/she was fixating correctly, the observer initiated the next trial by pressing the ENTER key. Observers completed 1000 trials on every training day (10,000 training trials in total). After each block of 100 trials there was a pause so that observers could have a brief rest if required and so that the attenuation level could be increased one dB step if they had scored more than 80% correct on the previous block. The stimulus was always presented to one retinal location: a section of the lower nasal retina of the left eye. The starting attenuation level on day one of training, and hence the resulting contrast of the grating, was set at just above the observer's threshold for the training location (<1 dB above) as measured during the first threshold measurement session.

3.4.2 Results

Training data

Figure 3.5 shows the percentage of correct decisions for each observer over each of the ten training days. An increase in attenuation level resulting from an observer scoring over 80% correct is indicated by a break in the line between the day of training on which the increase occurred and the previous day. Blocks on which observers got over 80% of their choices correct were always the first block of a training day.
Figure 3.5 Percent correct for each observer for each day.

The data points in Figure 3.5 for training days on which an increase in attenuation level occurred are plotted as percent correct for blocks at the higher attenuation level. They thus indicate percent correct for 900 trials rather than the 1000 trials which all other data points represent.

As can be seen from Figure 3.5 only two observers actually got more than 80% correct in a block of trials, thus, necessitating an increase in attenuation level. Observer MS had the attenuation level increased on the second and third days of training. Observer PS had the attenuation level increased on the sixth and ninth days of training. Observers GH and ID had the same attenuation level throughout training. From these data it would appear that there were no major improvements in performance across the training sessions.
Figure 3.6 Percent correct per block of 100 trials averaged across all days of training.

Figure 3.6 shows percent correct per block of 100 trials averaged across all training sessions, for each observer. Thus, each data point represents percent correct for 1000 trials. First blocks on which observers scored over 80% correct have been excluded from calculation of the scores, as the attenuation on the next block would have been higher, and this would have led to an unrepresentative decrease in performance on the second block of trials. A one way analysis of variance (training block (10)) with repeated measures on percent correct for each block revealed that there was no trend for improvement within training sessions \(F(9,27)=1.49\), \(p=0.202\).

**Threshold measurement sessions data**

**Calculation of contrast levels from attenuation data.**

The raw data were in the form of the attenuation level at which each turnaround occurred. In order to calculate contrast thresholds this data had to be transformed to calculate the actual contrast of the grating for each attenuation level. The Michelson contrast of the gratings with no attenuation was 0.403 and a 1 log unit change was equivalent to 20, 1 dB steps. Thus the attenuated contrast level is given by the following formula:
Attenuated contrast = \frac{0.403}{10^{20}} \quad \text{Where } x \text{ is the attenuation level in dB}

Initial examination of the results suggested that there was little difference in thresholds between the various untrained retinal locations, and that they all varied over time in the same way. A two way analysis of variance (stimulus type (4) - left eye upper nasal retina, left eye lower temporal retina, right eye lower nasal retina and right eye lower temporal retina; time of measurement (3) - baseline, transfer and 80 day follow-up) with repeated measures on time of measurement, was carried out to compare variation in contrast thresholds over time for each of the various transfer retinal locations. There was no significant difference in thresholds between the locations \(F(3,9)=0.33 \ p=0.822\), or in the way thresholds for the different locations varied over time \(F(6,18)=0.82 \ p=0.566\). As there were no significant differences between the locations a mean threshold was calculated for each time interval across the four transfer retinal locations.

Figure 3.7 shows mean threshold across all four observers for the trained location, the 2 cpd stimulus and the untrained retinal locations at each of the time intervals they were measured (raw scores for each observer can be seen in Appendix D).

![Figure 3.7](image_url)
As can be seen from figure 3.7 there is a trend for observers' thresholds to decrease for the trained location, but this decrease in threshold does not appear to transfer across spatial frequency or across retinal location (i.e. there is no change in thresholds for these latter stimuli). A two way analysis of variance (stimulus type (3) - training stimulus, 2 cpd stimulus, average of other retinal locations stimuli; time of measurement (3) - pre-training, post-training, 80 day follow-up) with repeated measures on time of measurement, was conducted to assess formally this change for each type of stimulus. There was a significant interaction between time of measurement and the contrast threshold measured ($F(4,12)=3.64$ $p=0.037$). Protected t-tests revealed that there was no significant change in contrast thresholds between the pre-training and post-training measurement sessions for the trained location, the transfer retinal locations, or the 2 cpd stimulus. However, for the trained location there was a significant decrease in contrast thresholds between pre-training and the 19 day follow-up ($t(3)=15.12$ $p=0.0005$), a significant decrease between the 19 day follow-up and the 50 day follow-up ($t(3)=2.57$ $p=0.041$) and no significant change between the 50 day and 80 day follow-up. There were no significant differences between pre- or post-training and the 80 day follow-up for the other retinal locations, or the 2cpd stimulus. It is particularly interesting to note that for the trained location there is a trend for thresholds to reduce after completion of training such that they have reached a minimum value by the 50 day time interval. From the current data it is not possible to say exactly when this minimum value is reached, but it is after the 19 day measurement interval and has occurred by the 50 day measurement interval. This decrease in thresholds appears to be maintained at the 80 day measurement interval. For one observer, all the thresholds were again measured 195 days after the first day of training. Figure 3.8 shows the threshold data for this observer across all the time intervals. It is apparent from figure 3.8 that the decrease in threshold for the trained location, for this observer, was maintained even after 195 days. Figure 3.8 also illustrates that...
for this observer there was a decrease in threshold for the 2 cpd stimulus following training.

Figure 3.8 Mean contrast thresholds for observer MS for trained location, 2 cpd stimulus, and untrained locations up to six months after training.

This is not entirely unexpected since the 2 cpd stimulus had a spatial frequency which was only an octave below the training stimulus. This is within some estimates of bandwidth for spatial frequency processing channels in the visual cortex (De Valois and De Valois, 1988, p 203-205) and the LGN (Kaplan and Shapley, 1982; Hicks, Lee and Vidyasagar, 1983; Derington and Lennie, 1984). The exact tuning of these channels is likely to vary between individuals.

Contrast sensitivity experiment two

Contrast sensitivity experiment one provided some evidence for an improvement in contrast sensitivity which was specific to three fundamental visual processing parameters: retinal location, spatial frequency, and eye. In addition, previous research has established that there are orientation specific visual processing channels, and neural correlates of these channels have been widely found in the
form of orientation specific cells in the visual cortex (e.g. Hubel and Wiesel, 1962). Whilst the majority of cells in the striate cortex have a relatively narrow orientation bandwidth (15° to 30°), cells at a pre-cortical level (e.g. LGN cells) are much more broadly tuned (De Valois and De Valois, 1988, pp. 264-267). Hence, orientation specific learning would imply change at a cortical level, whilst non-orientation specific learning could imply change at a pre-cortical level. Contrast sensitivity experiment two specifically sought to examine specificity of learning for orientation in order to further determine whether learning occurred at a cortical or a pre-cortical level.

In addition, two possible methodological problems from contrast sensitivity experiment one were controlled for in contrast sensitivity experiment two. First, contrast sensitivity experiment one used fast presentations, and a forced choice training paradigm. It is possible that learning occurs through some kind of tuning for stimuli presented in very short temporal intervals (<150 msec). In fact, a review of previous research on 'low-level' perceptual learning shows that a large number of such studies used similar fast presentation times (e.g. Ramachandran, 1976; McKee and Westheimer, 1978; Fiorentini and Berardi, 1980, 1981; Karni and Sagi, 1991; O'Toole and Kersten, 1992; Shiu and Pashler, 1992; Fahle and Edelman, 1993; Fahle, Edelman and Poggio, 1994; Rapf and Wehrman, 1994 - this is not an exhaustive list). The present study addresses this issue by using a training paradigm in which observers adapt to a constantly present drifting grating. Any learning which occurs could thus be attributed to tuning which is not specific to channels processing rapidly presented stimuli (e.g. transient channels; Tolhurst 1975a, 1975b; Graham, 1989, p 502). Further, if the key factor in learning is the total time for which observers are exposed to the low contrast gratings, then in the current study by presenting gratings for ten minutes a day, observers' total exposure time will be five and a half times that which they received in the previous study. Thus, we might expect to see a larger learning effect.
Second, contrast sensitivity experiment one used a threshold measurement procedure which may be susceptible to shifts in an observer's criterion for responding. To control for this possibility, contrast sensitivity experiment two uses a 'criterion free', forced-choice threshold measurement procedure.

In brief, in contrast sensitivity experiment two, observers' contrast thresholds for gratings presented at eight different orientations were measured using a criterion free measurement procedure. Following this, they completed ten days training on detecting change in the direction of motion of a drifting grating presented at their least sensitive orientation, and at their threshold contrast for that orientation; as described previously Mayer (1983) suggests that improvement in contrast sensitivity can occur, but only for orientations to which observers have not reached a theoretical peak sensitivity. Following training, observers' baseline thresholds were re-measured in order to assess whether learning had occurred and if so whether it had transferred to the untrained orientations.

3.5.1 Method

Observers
Two observers participated in this experiment, one male and one female. Both were in their mid twenties and had normal vision. Neither of the observers had taken part in an experiment involving contrast grating detection before.

Apparatus and stimuli
Stimuli were generated, controlled and displayed by the same equipment as was used in contrast sensitivity experiment one. In addition, the temporal waveform of the gratings presented during the threshold measuring procedure was controlled by a Feedback TWG 500 temporal waveform generator. Observers were seated 100 cm away from the display. The room conditions with respect to illumination levels,
spatial location of the equipment etc. were identical to those in contrast sensitivity experiment one.

The stimuli were sine wave gratings with a spatial frequency of 4 cpd. A black mask was placed over the oscilloscope screen such that a circular patch of grating subtending a visual angle of 2° was visible through an aperture cut in the centre of the mask. In the centre of the screen a fixation dot was placed subtending a visual angle of 6'. The orientation of the grating could be changed in 20' steps through hardware control.

**Training procedure stimuli**
The gratings used in the training procedure were drifting at a rate of 0.5° per second (2 cycles per second). Drifting gratings were used in order to prevent the build up of conventional after images over time (Arend and Skavenski, 1979). The drift rate was chosen such that it would be slow enough to activate the same velocity channels as a stationary grating (Graham, 1989, p 462-465); i.e. channels which are not selective for direction of motion. In addition, it would be expected that the drifting gratings would also excite some direction of motion specific channels which were tuned to low velocities. The contrast of the grating was set to just above an observer's threshold (<1 dB above) for a grating of that orientation.

**Threshold measurement procedure stimuli**
The gratings used in the threshold measurement stages were stationary rather than drifting. However, it was important to ensure that they were equivalent in certain respects to those used in the training procedure, if they were to provide relevant threshold measurements. First, they were displayed for 250 msec: this is equivalent to the time it took for the drifting grating to move half a cycle (e.g. from mean, to peak, to mean luminance). Second, in order to ensure that the gratings had the same temporal form as those used in the training procedure, the TWG 500 temporal
waveform generator was used to increase and decrease their intensity in a sinusoidal manner (mean to peak to mean). Thus in a 250 msec presentation, a spatial channel centred on a light bar in the grating would observe the stimulus go from mean, to peak, to mean luminance in a sinusoidal manner equivalent to that seen with the drifting gratings. Due to these equivalencies it could be expected that the training and test gratings would, to a large extent, stimulate the same processing channels.

However, the situation is somewhat complicated by the fact that channels which were not centred on a bar in the threshold measurement stages would be stimulated by a stimulus with lower amplitude than that observed with the drifting gratings. Thus, in the threshold measurement procedure, there would be less stimulation for some contrast analysers than during training, or, different analysers to those excited during training could be excited. This potential problem is at least partially offset by the involuntary movements of the eyes (Fuchs, 1971). These effectively create phase shifts for the gratings thus ensuring that channels would be centred on different parts of the grating bars on different trials during the threshold measurement procedure, and, in addition, that they would not always be exposed to a full amplitude stimulus during the training procedure. Thus, it could be expected that, in both the threshold measurement and training procedures, channels would sometimes be exposed to full amplitude stimuli, and sometimes to stimuli which were less than full amplitude. From the point of view of interpreting the experimental data these differences in the stimuli are likely to make it harder to perceive any change in thresholds following the training procedure rather than easier; in other words a conservative estimate of change might be expected.

In summary, the gratings used in the training and threshold measurement procedures were approximately equivalent in the following respects. They would both excite the same spatial frequency channels, the same velocity channels, the same contrast analysers, and they would both excite sustained and transient processing channels.
On this last point Tolhurst (1975a, 1975b) indicates that gratings with spatial frequencies between 2 and 7.6 cpd, and with presentation times of less than 800 msec will excite both sustained and transient processing channels.

**Design**

On the first day of the study, observers had their contrast thresholds measured for a grating with a spatial frequency of 4 cpd presented at eight different orientations ranging from 0° (vertical) to 157.5°, in 22.5° steps. Thresholds were measured using a two-up, one-down, three-alternative temporal forced-choice, staircase procedure; this gives an efficient estimate of 71% detection threshold with minimal bias, in the fewest trials (Rose, 1987). Each staircase consisted of a run of fifteen turnarounds. Attenuation levels were changed in steps of three dB until the first two turnarounds had occurred, then in steps of two dB until the next two turnarounds had occurred and thereafter in steps of one dB. The threshold was calculated as the average of the last ten turnarounds. On the next ten days observers took part in a training procedure lasting 20 minutes per day.

On these training days, observers adapted to a drifting grating with an orientation set to that for which they had the highest threshold (i.e. the orientation for which they were least sensitive\(^{3,3}\) as determined by the first threshold measurement session. Following this training period, on the twelfth day observers' thresholds were re-measured. Measurement of thresholds was planned to occur at various time intervals after the end of training. However, due to breakdown of the Cromemco computer this was not possible. In both the threshold measurement and training stages, gratings were always presented to the fovea of an observer's non-dominant eye.

---

\(^{3,3}\) The results from previous psychophysical studies investigating perceptual learning, such as Mayer's experiment (1983), suggest that there is a limit on the degree of improvement in sensitivity possible. This may reflect finite limits on human visual sensitivity. By training individuals on stimuli at the orientation to which they are least sensitive we might expect that there will be more 'room for improvement' before this finite limit is reached. The results of Mayer (1983) reported earlier, were consistent with this supposition.
eye (sighting eye dominance was measured using the Porta test (Porta, 1593; Crovitz and Zener, 1962; Gronwall and Sampson, 1971).

**Procedure**

As in contrast sensitivity experiment one, equipment was switched on at least half an hour before the start of an experimental session to allow time for stabilisation. Observers were seated on a chair in a dark room. Prior to the start of each experimental session, observers spent ten minutes adapting to the uniform luminance field displayed on the oscilloscope screen.

**Threshold measurement procedure**

A computer program was written in Cromemco BASIC which allowed fully automatic measurement of thresholds. Once the dark adaptation period was over, observers were asked if they were ready to begin. For both observers the order of threshold measurement was the same. Starting with 0°, a threshold was measured sequentially for every 22.5° change in orientation, until the final measurement for a grating with an orientation of 157.5° (eight threshold measurements in all). Gratings were presented monocularly to the observer's non-dominant eye. Monocular presentation was achieved through observers' wearing spectacles with one lens removed and with the other translucent such that it only allowed very diffuse light through (this lens had a refractive power of zero diopters prior to being made translucent). It was hoped that these spectacles would be an improvement over the eye patch used in contrast sensitivity experiment one, which, on occasion, had led some observers to complain of binocular rivalry (or similar) problems. Prior to the start of each run observers were shown the orientation at which the grating would be displayed. Following this the attenuator was set to 20 dB. The observer initiated the first and subsequent trials by pressing the ENTER key once they were fixating correctly. Three 250 msec presentation intervals were then signalled, in sequence, by a concurrent computer generated tone. In between each
presentation interval there was a 500 msec inter-stimulus interval during which a uniform luminance field of the same luminance as the spatial average of the grating was displayed. The observer's task was to indicate which of the three intervals the grating appeared in. They did this by pressing a key on the computer keyboard's numeric keypad ('1' for the first interval, '2' for the second interval and '3' for the third interval). If an observer got two responses correct in a row the attenuation was increased for the next trial in a step size according to the schedule previously described. If an observer got a response wrong the attenuation level was decreased for the next trial. The observer then initiated the next trial by pressing the ENTER key. At the end of each run the computer program automatically calculated the observer's threshold and saved the information to floppy disk.

**Training sessions procedure**

A computer program was written in Cromemco BASIC which fully automated the training sessions procedure. After the ten minute dark adaptation period the computer signalled that training was about to begin by giving ten short beeps. Following this the drifting grating appeared. The grating drifted in a constant direction for a period of time (between four and ten seconds) randomly selected by the computer program, and then changed to the opposite direction of drift. The observer's task was to press the ENTER key every time they thought the direction of drift changed; this task was introduced to ensure observers' attention to the training stimulus. The contrast of the grating was set to just above the observer's threshold (\(<1\) dB above) for a grating of the orientation on which they were training. The observers completed the task monocularly with their non-dominant eye.

3.5.2 Results

Table 3.5 shows observers' contrast thresholds for each orientation measured before and after training.
Table 3.5 Observers' contrast thresholds (Michelson %) for each orientation pre- and post-training.

<table>
<thead>
<tr>
<th>Orientation</th>
<th>Baseline</th>
<th>Transfer</th>
<th>Baseline</th>
<th>Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0°</td>
<td>1.71</td>
<td>1.60</td>
<td>3.71</td>
<td>4.24</td>
</tr>
<tr>
<td>22.5°</td>
<td>1.76</td>
<td>2.08</td>
<td>5.02</td>
<td>6.31</td>
</tr>
<tr>
<td>45.0°</td>
<td>2.24</td>
<td>3.04</td>
<td>3.45</td>
<td>3.74</td>
</tr>
<tr>
<td>67.5°</td>
<td>2.57</td>
<td>4.38</td>
<td>5.46</td>
<td>3.83</td>
</tr>
<tr>
<td>90.0°</td>
<td>2.68</td>
<td>4.57</td>
<td>3.04</td>
<td>4.71</td>
</tr>
<tr>
<td>112.5° ⋆</td>
<td>3.41</td>
<td>4.43</td>
<td>4.87</td>
<td>8.84</td>
</tr>
<tr>
<td>135.0°</td>
<td>2.68</td>
<td>3.91</td>
<td>3.41</td>
<td>5.81</td>
</tr>
<tr>
<td>157.5° ⋆</td>
<td>1.82</td>
<td>2.24</td>
<td>6.60</td>
<td>6.26</td>
</tr>
<tr>
<td>Mean</td>
<td>2.36</td>
<td>3.28</td>
<td>4.45</td>
<td>5.47</td>
</tr>
</tbody>
</table>

* Training orientation for observer MS
** Training orientation for observer JR

As can be seen from table 3.5 neither observer shows a large improvement immediately after training for the trained orientation. In fact somewhat unexpectedly both observers show a trend for their contrast thresholds to increase for almost all orientations. This change was assessed using a matched pairs t-test for each observer which treated each pair of pre- and post-training threshold measurements as a single case; a significant increase in thresholds following training was indicated for observer MS (t(7)=3.68 p=0.008) but not for observer JR (t(7)=1.68 p=0.134). However, closer examination of JR's data suggested that this result might be due to the effects of one strong outlying change - that for stimuli presented at an orientation of 67.5° - which showed a large decrease in contrast thresholds. When this pair of scores were excluded, observer JR also showed a significant increase in contrast thresholds following training (t(6)=2.55 p=0.044). It should be noted that there is no clear reason for the marked difference in change for the 67.5° measurement for this observer.

Contrast sensitivity experiment three

Contrast sensitivity experiments one and two appear to give quite different results. Whilst experiment one shows improvements in contrast sensitivity, experiment two
shows a decrement in performance following training. However, the improvements in contrast sensitivity experiment one occurred some time after training whilst it was not possible to collect data at similar time intervals in contrast sensitivity experiment two. If it had been possible to do so a similar consolidation effect to that observed in contrast sensitivity experiment one might have been observed; the changes observed by De Valois (1977) over an 18 month period are consistent with this possibility. Another possibility is that the different results were observed because of the different threshold measurement and training paradigms used in contrast sensitivity experiment two. It might be that frequent adaptation to a constantly presented, contrast threshold level grating, over a period of time, causes a reduction in sensitivity as a result of the repeated fatiguing (e.g. Tolhurst and Barfield, 1978) of the stimulated analysers; at this point it may be relevant to consider that persistence of simple motion aftereffects for between one and two weeks has been reported by Favreau (1979), and that extremely long term (at least three months) persistence of orientation-contingent colour aftereffects has been reported by Jones and Holding (1975). The duration and size of the aftereffect is related to the length of the adaptation period - the longer the period the greater the aftereffect and the longer it endures. However, the results of contrast sensitivity experiment two are even more surprising when one considers that De Valois (1977) reported an increase in contrast sensitivity following completion of adaptation experiments over an 18 month period. However, her observers were also taking part in other (unspecified) related experiments and it may have been these which led to the increase in contrast sensitivity.

If frequent prolonged exposure to a threshold level grating results in a decrease in contrast sensitivity for the ‘trained’ channels, then one might expect to see enhancement in sensitivity for channels which have a mutually inhibitory relationship with those trained, especially in the spatial frequency domain (e.g. De Valois, 1977); results from psychophysical studies investigating independence of
and inhibitory relationships between orientation analysers have been rather less conclusive (see Graham, 1989, p121) although the neurophysiological evidence is very conclusive. Further, we might postulate that the results from contrast sensitivity experiment two suggest that the general increase in contrast thresholds occurred due to changes at a pre-cortical level where analysers are not orientationally tuned or in cortical cells which are not orientationally tuned (e.g. layer 4 cells - Hubel and Wiesel, 1977).

Contrast sensitivity experiment three investigated the above possibilities in more detail. It utilised the same threshold measurement procedures and training procedures as were used in contrast sensitivity experiment two. However, in the present experiment baseline thresholds were measured for a range of orientations and for a range of spatial frequencies. Observers then took part in the same training procedure as was used in contrast sensitivity experiment two. Following this all the baseline thresholds were re-measured. These measurements were repeated a number of times after the end of training (up to four months later) as in contrast sensitivity experiment one. Thus, contrast sensitivity experiment three allows further investigation of the effects observed in contrast sensitivity experiments one and two. By following up over a long time period it allows investigation of whether a consolidation process, similar to that observed in contrast sensitivity experiment one, occurs after training with the adaptation paradigm. By including orientation and spatial frequency transfer tasks it allows for possible replication of the results of contrast sensitivity experiment two; a decrease in sensitivity for stimuli presented at any orientation (trained or untrained). In addition, if the hypothesis regarding the effects of repeated fatiguing is correct, then given the well established inhibitory links between spatial frequency processing channels, we might expect to see a reduction in sensitivity for the trained channels and an increase in sensitivity for channels more than one bandwidth different in spatial frequency from the trained
spatial frequency (e.g. for striate cortex cells ≥ 1.5 octaves, De Valois, 1988, p203-206).

3.6.1 Method

Observers

Four observers were recruited to take part in the study. Two were male aged 25 and 36 and the other two were female aged 20 and 28. All of the observers had normal or corrected to normal vision. Two of the observers (JH and IR) had never taken part in any experiments involving contrast grating detection before. Observer PS had previously participated in contrast sensitivity experiment one, and observer JR had previously participated in contrast sensitivity experiment two.

Apparatus and stimuli

Stimuli were produced by a Picasso waveform generator, controlled by an IBM compatible PC/AT computer with a digital I/O card and a D-A/A-D card fitted. The stimuli were displayed on a Tektronix 608 oscilloscope. The precise contrast of the gratings was varied in 0.05 log unit steps using the purpose built dB attenuator. The temporal waveform of the gratings presented during the threshold measuring procedure was again controlled by a TWG 500 temporal waveform generator. The auditory signals were generated by a purpose built 'beeper' with a dial to control precisely their duration (msec resolution). The beeper was triggered by a TTL type input.

The seating position of observers and location of the oscilloscope were unchanged from contrast sensitivity experiments one and two. Consequently the viewing distance was again 100 cm.
As before a mask was placed over the oscilloscope screen such that a circular patch of grating subtending a visual angle of $2^\circ$ was visible. However, in order to reduce complications resulting from the presence of a high contrast edge, a ring of translucent perspex with a diameter subtending a visual angle of $1^\circ$ surrounded the viewing aperture (see Figure 3.9 for an illustration).

![Figure 3.9 Diagram of masking over the oscilloscope screen.](image)

The stimuli were sine wave gratings with various spatial frequencies (1, 2, 3, 4, 6 and 8 cpd). The gratings were displayed at orientations of $45^\circ$, $90^\circ$ and $135^\circ$, with $90^\circ$ being horizontal. A fixation dot subtending a visual angle of $6'$ was placed in the centre of the screen.

**Training procedure stimuli**

The gratings used in the training procedure had a spatial frequency of 4 cpd and were drifting at a rate of $1/4^\circ$ per second (1 cycle per second). A slower drift rate was chosen than that used in contrast sensitivity experiment two to make sure that the drift rate would be well within the velocity limits of a channel responding to stationary gratings as opposed to approaching some estimates of the boundary (Graham, 1989, pp. 462-465).
Threshold measurement procedure stimuli

As before the gratings used in this procedure were stationary rather than drifting. However, they were again equivalent in a number of respects to those used in the training procedure. This time the gratings were displayed for 500 msec as this was now the time in which a drifting grating would move a half cycle. The TWG 500 temporal waveform generator was used to generate an equivalent half cycle of a sinusoidal temporal waveform of 500 msec duration.

Design

The design of the experiment was the same as that for contrast sensitivity experiment two in all respects other than those detailed below. On the first day of the study observers had their contrast threshold measured for a grating with a spatial frequency of 4 cpd at an orientation of 45° in their dominant eye. They also had their contrast thresholds measured for a grating with a spatial frequency of 4 cpd at orientations of 45°, 90° and 135° in their non-dominant eye and also their thresholds were measured for gratings with a spatial frequency of 1, 2, 3, 6 and 8 cpd oriented at 45° again in their non-dominant eye. Thresholds were measured using a two-up one-down, three-alternative temporal forced-choice staircase procedure identical to that used in contrast sensitivity experiment two. On the eight consecutive days after the baseline threshold measurement session observers took part in a training procedure identical to that in contrast sensitivity experiment two. However, in this case all observers adapted to the same stimulus; a grating with a spatial frequency of 4 cpd and an orientation of 45° presented only to their non-dominant eye. Following the training period, on day 10 all of the observers' baseline thresholds were re-measured. These measurements were repeated 19, 50 and 130 days from the start of training for all observers.

In the present experiment, as thresholds were measured for only a limited number of orientations, it was not possible to determine with any reasonable accuracy the
orientation to which individuals were least sensitive. However, Appelle (1972) indicates that in Caucasian adults (e.g. such as observers in the present experiments) the typical pattern of anisotropy is for greater sensitivity to horizontal and vertical orientations than to diagonals. Thus in the present study it was decided to train observers on stimuli with an orientation of 45°.

**Procedure**

The equipment was switched on at least half an hour before the beginning of an experimental session to allow time for stabilisation. Observers were seated on a comfortable chair in a dark room. As in the previous experiments they spent ten minutes adapting to the uniform luminance field displayed on the oscilloscope screen prior to the start of an experimental session.

**Threshold measurement procedure**

A computer program was written in QuickBasic, compiled and linked to timing routines written in Assembly language in order to produce an executable program to replace that lost with the Cromemco system. After the dark adaptation period was over the threshold measurement procedure began. The procedure was identical to that for contrast sensitivity experiment two except that the presentation intervals were 500 rather than 250 msec.

**Training sessions procedure**

A computer program was written in QuickBasic, compiled and linked to the Assembly Language timing routines in order to produce an executable program to replace that lost with the Cromemco system. Following the ten minute dark adaptation period the training period began. The training procedure was identical to

---

3,4 The timing routines are based upon those developed by Graves and Bradley (1987, 1988, 1991) and return sub millisecond resolution timing information directly from the PC hardware timer (8253 timing chip). By adding this information to the low resolution BIOS time counts, timing with better than millisecond accuracy and no upper limit on the interval that can be timed is possible.
that used in contrast sensitivity experiment two except for the following details. In this experiment all observers adapted to a grating with a spatial frequency of 4 cpd and an orientation of 45° using only their non-dominant eye (the grating was displayed with a contrast level slightly above an observer's threshold (<1 dB above), as previously measured). The observers were required to press the SPACEBAR every time they noticed a change of drift direction rather than the ENTER key. The computer program recorded the number of times that an observer correctly noticed the grating change direction of drift (if they pushed the spacebar within three seconds of a direction change it was counted as having been noticed), the total number of times the grating changed direction of drift, and the number of times an observer pushed the spacebar when the grating had not changed direction of drift within the last three seconds (false alarms).

3.6.2 Results

Training data

The percentage of direction changes that an observer noticed for each day was calculated. Figures 3.10 and 3.11 show the number of false alarms and the percentage of noticed direction changes for each observer over the eight days of training (on average the grating changed direction of drift between 80 and 90 times during the ten minute adaptation interval).
From figures 3.10 and 3.11 it is apparent that there is no trend over the course of training for observers to notice either an increased percentage of direction of drift changes or to reduce the number of false alarms they make. Also there is quite a high degree of variability apparent with scores varying quite widely between observers and over time. Two observers performed at close to ceiling levels (PTS and JR) whilst the other two observers (IR and JH) noticed a much smaller percentage of direction changes despite making a greater number of false alarms. For these latter two observers, if they were not noticing any direction changes then a random response pattern would be expected which would lead to more false alarms than correctly noticed direction changes. This can be understood as follows:
the grating changed direction of drift between four and ten seconds after the previous change of direction, thus, on average, it would change direction of drift every seven seconds. If an observer pushed the button within three seconds of a direction change they were scored as noticing the change of direction. If an observer pushed the button after the initial three second period they were scored as making a false alarm. Thus, given the average interval of seven seconds for drift in a constant direction, the proportion of correctly noticed direction changes to false alarms for a random response pattern would be $3/7:4/7 = 0.43:0.57$. Inspection of the data for each observer indicated that for each day all observers were correctly noticing more direction changes than they were making false alarms, which would suggest a similar non-random response pattern on each day of training. The number and proportion of correctly noticed direction changes to false alarms, collapsed across the eight days of training, for each observer is shown in table 3.6.

Table 3.6 Proportion of correctly noticed direction changes to false alarms for each observer collapsed across days of training.

<table>
<thead>
<tr>
<th></th>
<th>JR</th>
<th>PTS</th>
<th>IR</th>
<th>JH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportions</td>
<td>0.96:0.04</td>
<td>0.93:0.07</td>
<td>0.66:0.34</td>
<td>0.62:0.38</td>
</tr>
<tr>
<td>Numbers</td>
<td>588:23</td>
<td>623:47</td>
<td>396:204</td>
<td>364:220</td>
</tr>
</tbody>
</table>

From table 3.6 it would seem that none of the observers were following a random response pattern. Overall the results imply that the measure of threshold derived using the staircase procedure gave quite a good estimate of threshold for the drifting grating task for two of the observers (IR and JH) but was not so well matched for the other two.

Analysis of variance was used to assess formally changes over time in false alarms and percentage of direction changes noticed. A one way analysis of variance (day of training (8)) with repeated measures on the number of false alarms for each day revealed no significant change over time in the number of false alarms ($F(7,21)=1.69$ $p=0.166$). A second one way analysis (day of training (8)) of
variance with repeated measures on the percentage of correctly noticed direction changes for each day revealed no significant change over time in the percentage of correctly noticed direction changes \( F(7,21) = 1.24 \ p = 0.326 \).

**Threshold measurement sessions data**

*Calculation of contrast levels from attenuation data*

With the new equipment the voltages supplied to the waveform generator and hence the resulting unattenuated contrast of the grating was different to that in contrast sensitivity experiments one and two. The new unattenuated Michelson contrast was 0.248 and hence the formula to calculate contrast level is now given by:

\[
\text{Attenuated contrast} = \frac{0.248}{10^x} \quad \text{Where} \ x \ \text{is the attenuation level}
\]

Figure 3.12 shows the mean change in contrast thresholds between the baseline threshold measurement session and each follow-up measurement session for each of the different types of stimuli (raw scores for each observer can be seen in Appendix E). The means are calculated across all four observers for all of the time intervals other than for the difference in thresholds between the baseline and transfer threshold measurement sessions. For this difference the data from observer IR has been excluded. This was because routine checking of the voltages output from the TWG 500 temporal waveform generator, which was carried out following every session, revealed that for this observer's transfer session, they were far higher than they should have been \( (> +4 \ V \ \text{rather than} +2.55 \ V) \). This meant that the resulting grating contrasts were higher than they should have been leading to an apparent overall increase in sensitivity, thus, confounding these results. The data from subject JR has been excluded from calculations of the threshold for her dominant eye as she suffered damage to her eye following an accident resulting in an inability to detect reliably an unattenuated grating (Michelson contrast 24.8%).
Figure 3.12 Mean change in contrast threshold between baseline session and each follow-up session for each type of stimulus.

An additional female observer aged 26 was recruited, and her thresholds for all the stimuli were measured at the same time intervals as the measurements taken for observers in the present study. However, this observer did not take part in any training procedure. From the threshold measurements made on this observer the average standard error for re-measuring contrast thresholds at successive time intervals was calculated. Following the method of De Valois (1977) two regions are indicated in Figure 3.12: an inner region lying between zero ± 1.96 standard errors of the mean and an outer region lying between zero ± 2.58 standard errors of the mean; changes which exceed the boundaries of these regions are significant at the 5% level and 1% levels respectively.

Using the 5% level of significance, it can be seen from Figure 3.12 that, following training, the observers' contrast thresholds increased, in at least one session, for all types of stimulus other than the 1 cpd stimulus and the interocular transfer stimulus. For the training stimulus this increase is significant for the 50 and 130 day follow-

121
up sessions. For the 2 cpd stimulus the increase is significant for one of the four follow-up sessions. For the 90° orientation stimulus, the 3 cpd stimulus and the 6 cpd stimulus the increase is significant for three of the four follow-up sessions. For the 135° stimulus and the 8 cpd stimulus the increase is significant for all four of the follow-up sessions.

The widespread increase in contrast thresholds could be attributed to some higher level change in observers' strategy when completing the task. However, this is probably not the case since, interestingly, for two out of the four follow-up sessions there is a significant decrease in contrast thresholds for the 1 cpd stimulus, and for one out of the four follow-up sessions there is a significant decrease in contrast thresholds for the interocular transfer stimulus. For the remaining follow-up sessions for both these stimuli there was no significant change in thresholds.

**General Discussion of contrast sensitivity experiments**

The results from the contrast sensitivity experiments reveal an extremely interesting pattern of findings. Contrast sensitivity experiment one indicates that a stimulus specific, long term, and enduring decrease in contrast thresholds can occur following extensive training. Further, this decrease in thresholds appears to result from a consolidation process, which occurs following the end of training. In contrast to this, contrast sensitivity experiment two, using an adaptation training paradigm, indicates an increase in contrast thresholds following training, and further, that this increase transfers to stimuli at a variety of different orientations. Contrast sensitivity experiment three confirmed and extended the findings of contrast sensitivity experiment two. An enduring increase in contrast thresholds was observed which transferred to other orientations and to other spatial frequencies within the bandwidth of a typical cortical spatial frequency processing channel. A decrease in contrast thresholds was observed for a stimulus probably processed by a
spatial frequency processing channel which had a mutually inhibitory relationship with the channels processing the training stimulus (the 1 cpd stimulus). A decrease in thresholds was also observed for the interocular transfer stimulus which is likely to have been processed by channels independent from those which processed the training stimulus. Taken together, the results from contrast sensitivity experiments one, two and three are suggestive of limited, bi-directional plasticity in human contrast sensitivity either in layer 4 of the visual cortex or at a pre-cortical level.

3.7.1 Long term change of contrast sensitivity

The present series of experiments produced enduring changes in contrast thresholds (both increases and decreases) the nature of which varied across experiments. The most obvious difference between contrast sensitivity experiment one, and contrast sensitivity experiments two and three (other than the different transfer tasks) is the training regimes used. Contrast sensitivity experiment one used a temporal forced choice task with gratings presented for 117 msec, whilst contrast sensitivity experiments two and three used a direction of drift detection task during which observers constantly adapted to a grating presented at their threshold contrast. Whilst participation in contrast sensitivity experiment one resulted in a specific decrease in contrast thresholds for the training stimulus, participation in contrast sensitivity experiments two and three, resulted in an increase in thresholds for the training stimulus.

One explanation for the results of contrast sensitivity experiment one is that they may reflect the use of a threshold measurement procedure which could be susceptible to observer criterion shifts, and given that contrast sensitivity experiments two and three gave a different, although not inconsistent, pattern of results, there is still a need to replicate contrast sensitivity experiment one using a criterion free method. However, this explanation of the results would seem unlikely for a number of reasons. First, enduring changes in contrast sensitivity were found
in contrast sensitivity experiments two and three, albeit in the opposite direction, and these experiments used a criterion free method of threshold estimation. Second, research typically shows a stable relationship between thresholds measured using the same psychophysical procedure (e.g. Kelly and Savoie, 1973): hence, observers generally do not change their response pattern when the same threshold is measured, using the same technique, on a number of occasions. Third, the consistency in the pattern of change across subjects would argue against a criterion shift, since it is unlikely that all observers would show such a shift. Fourth, the fact that a decrease in thresholds was seen only for the trained location and eye also argues against a criterion shift, since such a shift might lead to a decrease in thresholds for all stimuli and locations. Finally, the fact that the decrease in learning occurred after a consolidation period rather than immediately after training, and that this was uniform across observers, also suggests that a criterion shift is an unlikely explanation.

A second possible explanation for the results of contrast sensitivity experiment one is that, as previously discussed, they are due to the short exposure times for which observers viewed the stimuli. It may be that in such experiments learning results from improved temporal tuning of the channels stimulated by these fast presentations. The results of contrast sensitivity experiment one are consistent with this possibility, and contrast sensitivity experiments two and three provide no direct evidence counter to this suggestion. However, long term change did occur in contrast sensitivity experiments two and three, and this does provide evidence for plasticity in contrast sensitivity.

The change observed in contrast sensitivity experiments two and three may have resulted from adaptive processes. Previous research has suggested that contrast adaptation may play a functional role since it can serve to increase the gain of the system, although the absolute signal is smaller at all contrast levels (Blakemore,
In addition, in much the same way as for adaptation to luminance, the dynamic ability to shift the range of contrasts to which a cell responds can prevent saturation at low contrast levels. Georgeson (1985) has shown that following adaptation to a grating, response amplitudes are smaller for a lower contrast grating, but are unaffected for higher contrast gratings. Bonds (1993) has confirmed and extended these findings in the visual cortex of cats using single cell recordings. He found a reduction in response even for very low adapting contrasts, for example, a reduction in response to a grating with 1% contrast following adaptation to a grating with 3% contrast. He also found that very short adaptation periods for a grating with 28% contrast (as low as 50 msec) are sufficient to lead to a reduction in response amplitude to a grating with 14% contrast at the end of a two second period. The durability of these reductions was not assessed. Further, Bonds (1993) goes on to indicate that whilst these reductions in response amplitude following adaptation are observed for cortical cells they are not observed in LGN cells implying a cortical basis for these mechanisms. The long term change observed in contrast sensitivity experiments two and three may have resulted from the repeated adaptation of contrast analysers such that over a period of time during and after training there was a shift in the preferred contrast level of the cells exposed to the training stimulus.

3.7.2 Multiple contrast analysers and analysers near threshold?
Contrast sensitivity experiments two and three used training stimuli with contrasts a little above observers' thresholds. Bond's (1993) work confirms that such contrasts are sufficient to induce an adaptation effect. This adaptation effect is likely to occur either in cells which respond to a dynamic range of contrasts or to cells specifically tuned to low contrast levels; although psychophysical evidence has tended to suggest that there are probably no contrast analysers near threshold (Graham, 1989) physiological evidence has indicated that the tuning of cells along the contrast axis is subject to considerable variation. For instance, Albrecht and Hamilton (1982)
found that whilst some cells begin responding at 1% contrast and saturate by 10% other cells do not begin responding until 10% contrast. It may be that the results of the present study occur as a result of change in cells specifically tuned to low contrast levels.

3.7.3 The likely locus of plasticity for contrast sensitivity

The present studies have provided considerable evidence of low-level plasticity for contrast sensitivity. The results of contrast sensitivity experiment one indicated learning which was specific to retinal location and eye of origin. This would suggest change in monocularly driven cells; these occur primarily at a pre-cortical level and in layer 4 of the cortex (Hubel and Wiesel, 1968). However, there was a mixed pattern of transfer across spatial frequency in this experiment. For three observers there was no transfer to a stimulus with a spatial frequency one octave lower than the training stimulus, whilst for the final observer there was transfer of learning. Estimates of the full bandwidth at half amplitude for striate cortex cells typically indicate a bandwidth of 1 to 1.4 octaves with stimuli about 1 octave away being almost totally ineffective (De Valois and De Valois, 1988). Thus, in general, the specificity of learning to spatial frequency observed in experiment one would suggest change at a cortical level; pre-cortical cells (e.g. LGN cells) are generally more broadly tuned for spatial frequency and, thus, if learning were occurring at this level some transfer would have been expected for most observers with stimuli one octave different from the training stimulus.

The results of contrast sensitivity experiment three indicated a slightly different pattern of change with respect to spatial frequency. The increase in thresholds in this experiment transferred to stimuli up to and including one octave different in spatial frequency from the training stimulus. This suggests change in relatively broadly tuned spatial frequency channels most likely at a pre-cortical level or in layer 4 of the visual cortex where spatial frequency tuning tends to be broader. The
lack of transfer to a stimulus two octaves away from the training stimulus indicates
the extent of the tuning breadth of these channels and also argues against the
possibility that the decrease in thresholds was due to some more general factor.
Similarly, the specificity of the increase in thresholds to the eye of origin argues
against this possibility and agrees with the results of contrast sensitivity experiment
one, again indicating change in monocularly driven cells either at a pre-cortical
level or in layer 4 of the visual cortex. The wide transfer of change across stimuli
of different orientations observed in contrast sensitivity experiments two and three
further supports this possibility since cortical cells other than those in layer 4 are
orientationally selective whilst pre-cortical cells and many layer 4 cortical cells are
not (Hubel and Wiesel, 1977).

It might be noted, when examining transfer to stimuli with different orientations
that, in experiment three, there is an asymmetry in the transfer patterns. Observers'
thresholds increase a greater amount for the 135° transfer stimulus than for the 90°
transfer stimulus. Examination of the raw scores for individual observers (shown in
Appendix E) suggests that this asymmetry largely results from observer JR, who
shows a very pronounced increase in contrast thresholds for the 135° transfer
stimulus as compared with the 90° transfer stimulus: for the remaining observers a
pronounced asymmetry does not appear to exist. Further research is required to
assess whether this effect was indeed specific to this one observer rather than
indicating the workings of some more widespread, and systematic process.

Finally, as previously discussed, the results of contrast sensitivity experiments two
and three may reflect change in contrast gain control mechanisms. Such
mechanisms have been indicated as operating at cortical, LGN and retinal levels (De
Valois and De Valois, 1988, p163; Bonds, 1993) depending on the precise
definition and time constant of the gain control being considered. Further, as has
been discussed, the results of contrast sensitivity experiment one could be attributed
to some form of temporal tuning for brief duration stimuli; sustained and transient processing channels operate both at the level of the LGN and the primary visual cortex.

So, to summarise, taken together the results from contrast sensitivity experiments one, two and three build a picture of cortical plasticity in the early stages of visual processing either at a pre-cortical level or in layer 4 of the visual cortex or, perhaps more likely, at all these levels; it is possible that the results from these experiments arise from plasticity in more than one mechanism. The process revealed by contrast sensitivity experiment one has a longer time course than that observed in contrast sensitivity experiments two and three, and may be localised in the visual cortex, whilst the process observed in contrast sensitivity experiments two and three has a more immediate effect and most probably results from a change in gain control mechanisms in layer 4 of the visual cortex or at pre-cortical levels. It is clear that further research is needed to clarify these possibilities and suggestions for such research will be discussed in a subsequent section. The neurophysiological mechanisms underlying these plasticity processes are left open by these experiments, however, recent research has provided some clues to the type of change which could be occurring.

3.7.4 The possible neurophysiological basis underlying change and consolidation

Neurophysiological evidence for cortical plasticity has been accumulating in recent years. For example, in the somatosensory system Merzenich, Nelson and Stryker (1984) showed that following amputation of a digit, part of the cortical territory occupied by that digit before amputation was occupied by other digits following amputation. Pons, Garraghty, Ommaya, Kaas, Taub and Mishkin (1991) found that twelve years after the severance of the sensory nerves coming from the forelimbs in a monkey, the representation of the lower face area had expanded by a distance of
12-14 mm into the cortical territory that had belonged to the arm and hand. This latter study illustrates the long time scale over which such changes can occur. In the visual system, Wurtz, Yamasaki, Duffy and Roy (1991) found that following lesions in area V5 of the monkey, the receptive fields of cells at the borders of the lesion expanded very substantially in all directions, as if the input to them had been modified as a consequence of the lesion. Similarly in area V1, Pettet and Gilbert (1992) and Gilbert (1994) have shown that following a focal retinal lesion (thus removing input to a restricted region of the visual cortex) the silenced cortical area regains visually driven activity over a period of months (some of the change occurs within minutes), with the receptive fields shifting to new positions outside the lesioned part of the retina. Further, Volchan and Gilbert (1995) have indicated a cortical basis for the receptive field re-organisation that occurs following the introduction of an artificial scotoma. Andersen, Trommald, Jensen and Paulsen (1994) have shown that over a period of one to three weeks deafferented neurones can sprout new spines to form synapses and further have found the development of new long filamentous protrusions (up to 30μm).

So, overall there is considerable neurophysiological evidence for neural plasticity, and some of this operates over the same kind of time course as the consolidation process observed in contrast sensitivity experiment one. It may be that processes similar to those reported here account for the consolidation observed. In addition, there was some evidence for an overnight consolidation process as indicated by the fact that in contrast sensitivity experiment one, any improvement in detection performance during the training sessions always occurred on the first block of the next day’s training session. Sagi (1994) suggests that some low-level learning requires an overnight consolidation process that may be dependent on REM sleep; observers in contrast sensitivity experiment one had been to sleep in between sessions which were always held on separate days. However, an equally plausible explanation for this result is that any improvement within sessions was masked by
observer fatigue building up during the session such that the improvement was only apparent at the beginning of the next session.

3.7.5 Limitations and suggestions for further research

The present studies have indicated an interesting and broadly consistent pattern of results. However, there are a number of inconsistencies and potential sources of error in the data. First, there was considerable observer variability both in terms of the specificity and the magnitude of change. Other researchers investigating low-level perceptual change have reported similar variability (e.g. Fahle, 1994); further investigation is required to identify the potential sources of such variability.

Second, as was discussed previously there is a need to replicate contrast sensitivity experiment one using a criterion free threshold measurement procedure. It would also be of interest to replicate contrast sensitivity experiment one with the addition of an orientation transfer task in order to assess whether the learning observed in this experiment is occurring at the same levels of visual processing as in contrast sensitivity experiments two and three, or whether in fact learning at multiple levels is possible. Third, the data from the transfer to other spatial frequencies in contrast sensitivity experiment three is somewhat surprising since it reveals an asymmetric transfer pattern. There is a greater increase in thresholds for the higher spatial frequencies than occurs for the training stimulus but this is not true for the stimuli with spatial frequencies lower than the training stimulus. Typically, one would expect a decreasing effect as the spatial frequency of the transfer stimulus becomes increasingly different from that of the training stimulus. There is thus a need to examine the replicability of this effect. Fourth, it would be of interest to examine whether training with higher contrast gratings in contrast sensitivity experiments two and three would lead to larger long term increases in observers' thresholds, since it is possible, that the size of any effect which could be observed was limited by the fact that the adapting stimulus was close to threshold. Previous results (Georgeson, 1985; Bonds, 1993) indicate that only stimuli with a contrast below the adapting
stimulus contrast are affected and, hence, with a near contrast adaptation stimulus, the potential size of any change is greatly restricted. Fifth, the results of Bonds (1993) indicate that even very short stimulus presentations (50 msec) are sufficient to induce an adaptation effect. Given this finding, it leaves open the question of why the 117 msec stimulus presentations used in contrast sensitivity experiment one did not also produce an adaptation effect similar to that seen in contrast sensitivity experiments two and three. It may be that with such short presentation intervals the very short duration of the adaptation effect is insufficient to induce long term change of the type observed in contrast sensitivity experiments two and three. Finally, there is a need to examine in more precise detail the time course of the consolidation effect observed in contrast sensitivity experiment one.

Thus, to summarise, there is scope for a great deal of further investigation of the phenomena observed in the present studies. However, overall a broadly consistent pattern of findings which indicate bi-directional plasticity for contrast sensitivity, localised in layer 4 of the striate cortex and/or at pre-cortical levels, has been observed.

Summary of chapter three

The present chapter has examined perceptual learning phenomena for two aspects of visual analysis. In both cases perceptual learning as defined in this thesis was observed. However, as has been discussed different perceptual learning processes such as perceptual re-organisation and low-level learning can be distinguished even within the broad process defined as perceptual learning. The type of learning which underlies a given learning phenomenon can be better specified through the use of transfer of learning tasks. In the present studies, transfer of learning tasks were designed on the basis of the findings of previous psychophysical and neurophysiological research on the aspects of visual analysis under investigation. In
the case of the stereo task, these transfer tasks indicated that a form of perceptual learning occurred which was not localised at a very early level of processing. That perceptual learning did occur is indicated by the improvement in observers' discrimination performance on the task, which involved simultaneous comparison. However, further transfer of learning tasks would be required to define more precisely the exact nature of the learning process observed. It may be that the learning of stereo depth discrimination constitutes a unique perceptual learning process which would warrant a separate taxonomic category. This is obviously a question for future research. For the present, a more detailed set of transfer tasks were employed in the series of experiments investigating learning for contrast sensitivity. These tasks not only indicated that the perceptual learning which was observed was low-level in nature but gave a good indication of the likely neural locus of that learning. A more detailed indication of the likely locus of learning could be expected for the stereo learning with the use of a similarly comprehensive range of transfer tasks. Thus, in the present chapter a general methodological approach to indicating positively the occurrence of perceptual learning and, further, to identifying when that learning is low-level in nature, has been demonstrated.

However, it remains open to question whether the types of low-level learning phenomena shown in the present chapter, and in the work of previous researchers, are likely to occur in real world, non-laboratory, discrimination tasks. In other words are the sorts of improvements in discrimination observed in everyday tasks such as the examination of X-rays likely to rest in part upon low-level learning, or is it more likely that other types of learning, such as the processes revealed in chapter two, underlie the improvements in discrimination often observed? The next

3.5 There may be grounds for many separate taxonomic categories which this thesis does not distinguish. However, the aim of this thesis was not to generate a complete list but to make the fundamental point that there are different types of learning processes underlying phenomena previously considered to rest upon a unitary perceptual learning process and to define a set of useful broad categories into which learning phenomena can be placed (this topic is discussed further in the final chapter of this thesis).
chapter makes a first attempt to address this issue, and reports the results of two studies which examine learning for the detection of features in X-rays under normal observation conditions.
Chapter four
Examining learning in a ‘naturalistic’ task

4.1.1 The importance of contrast sensitivity for detecting features in X-rays

The results of the three experiments reported in chapter three investigating the plasticity of contrast sensitivity, and the experiments of De Valois (1977) and Mayer (1983) have indicated that a limited amount of plasticity exists for contrast sensitivity. Some ‘real world’ tasks, such as the preliminary stages of examining X-rays for abnormality, may be critically dependent on contrast sensitivity; the detection of abnormality in radiographs often requires distinction of low contrast targets from fuzzy backgrounds (Revesz, Kundel, and Graber, 1974; Kundel and Revesz, 1980). Pauli (1993) showed that a simple measure of contrast sensitivity (The Cambridge Low Contrast Gratings, Wilkins and Robson, 1987; Wilkins, Della Sala, Somazzi, Nimmo-Smith, 1988) predicted mammographic film reading performance following training. Further, there is some evidence to suggest that, in part, the skill of the expert radiologist is acquired through perceptual learning some of which may result in heightened contrast sensitivity (Davies, Sowden, Hammond and Ansell, 1994). This latter series of studies included an experiment in which the performance of naive observers on the detection of targets in radiographs was compared to that of radiologists. The stimuli were X-rays of perspex blocks with small holes drilled into them which produced low contrast dots in the X-ray image. Thus, although the task was one of target detection in X-rays, the experts' knowledge in the form of visual schema specific to radiology, effective search patterns, probable combinations of features in targets, anatomical structure and so on, would have been either irrelevant, or may even have presented a handicap. The results showed that the experts were significantly better at the target detection task, suggesting that a component of their skill may have resulted from low-level perceptual learning, possibly for contrast sensitivity.
4.1.2 Perceptual learning and radiological expertise

Most models of radiological expertise include stimulus driven and top-down components (e.g. Swensson, 1980; Gale, 1992; Pauli, 1993). Models of the stimulus driven components typically propose a pre-attentive global search process which guides visual attention towards image features which may represent an abnormality. However, in certain circumstances a feature indicative of abnormality may be below some detection threshold. In this situation neither the pre-attentive nor the attentive search processes which follow will detect the feature. In circumstances such as these, in order for the observer to be able to detect the abnormal features low-level learning, which results in improved sensitivity to the relevant stimulus dimension, may be required. The present studies examine the type of learning which may occur, in naive observers, for the detection of features (dots) in X-rays under 'standard' viewing conditions\(^4\) rather than under the sorts of conditions used to study low-level learning, as reported in chapter three. Standard viewing conditions are defined here as the sort of viewing conditions under which X-rays are viewed in normal medical, or other, practice (i.e. a well lit room, with the X-rays constantly displayed on a light box). Through the use of transfer tasks the present studies aim to determine whether any improvement occurs in low-level visual processes, in higher level processes or in a combination of both. The task of dot-detection in X-ray images can be expected to simulate some of the main aspects of the target detection processes involved in examining medical X-ray images whilst requiring none of the experts' conceptual medical knowledge.

X-ray experiment one

This experiment examined observers' ability to detect high spatial frequency, low contrast dots in X-rays, and whether this changed as a result of extended practice.

---
\(^4\) Certain deviations from the sorts of viewing conditions used in real medical, or other, practise (see below) were made in the current studies in the interests of designing transfer tasks which would help to identify the type of process underlying any learning observed.
The stimuli were X-rays of perspex blocks with holes of constant diameter and varying depth drilled into them. The holes showed up as dots in the resultant X-ray images and the depth of the hole determined the contrast of the dot. Following the results of Pauli (1993), observers' contrast sensitivity was measured using the Cambridge Low Contrast Gratings (Wilkins and Robson, 1987; Wilkins, Della Sala, Somazzi, Nimmo-Smith, 1988) in order to assess whether contrast sensitivity, as measured by the gratings, predicted detection performance or the amount of learning. Following completion of the gratings task, observers' baseline performance was assessed on the dot detection task, for each eye separately\(^{4,2}\). Observers then completed eight days training. On each day their task was to complete 200 dot detection trials using only their non-dominant eye. Immediately after training the baseline measures were repeated for each eye and they were repeated one more time 50 days after the end of training. Given previous research on learning to examine X-rays it was expected that observer detection performance would improve as a result of practice. By assessing the extent to which any learning transferred to the untrained, dominant eye an indication of the level at which any learning occurred could be obtained. Transfer across eyes would indicate learning at a binocular level or in some higher cognitive process (e.g. attentive and/or search processes, concept formation etc.), whilst incomplete or no transfer across eyes would indicate a monocular low-level learning process.

4.2.1 Method

Observers

Fourteen observers participated in the experiment, 13 of whom were women and one of whom was a man. Twelve of the observers were aged between 19 and 23 years whilst the other two were around forty years of age. All of the observers had

\(^{4,2}\) The use of a monocular viewing procedure in the current studies represents the main deviation from the sorts of viewing conditions used in real medical, or other, practice.
normal or corrected to normal vision. Observers received a token payment in return for their participation in the study.

**Apparatus and stimuli**

The stimuli were X-rays of perspex blocks which measured 125 mm by 125 mm. Each block comprised a grid of twenty five squares arranged five by five. The sides of the squares measured 25 mm and there was a hole drilled in each square in one of five possible locations. The locations were the centre of the square or exactly midway along one of the four hemi-diagonals. Figure 4.1 shows a schematic illustration of a grid.

![Figure 4.1 Schematic illustration of a grid showing the different possible hole locations.](image)

The holes all had a diameter of 0.35 mm but they could be one of four possible depths: 0.10 mm, 0.25 mm, 0.50 mm or 0.75 mm; the depth of the hole varied the contrast of the resultant dot in the X-ray image, with deeper holes having greater contrast. All of the squares in a particular grid had holes of the same depth drilled into them. There were eight grids in total, two grids for each of the four hole depths. The location of the hole in each square in a grid was randomly selected, within the constraint that there were an equal number of squares with a hole in each location.
A number of trial X-rays were taken of each grid with various thicknesses of perspex overlaying the blocks and with various X-ray beam densities. The final set of images was chosen on the basis of pilot work which indicated that for this set, the holes could only be detected with scrutiny and that performance ranged from just above chance to just below a ceiling level. All of the final images were produced with the same thicknesses of overlay perspex and with the same beam density. The final X-ray images of the grids measured 132 mm by 132 mm and the diameter of the holes in the X-rays was 0.37 mm. A photograph of an actual X-ray can be seen in Appendix F. The stimuli were displayed on a light box with a luminance of 2610 cd/m² (illumination = 2000 Lux). The area surrounding a stimulus was masked out with black cardboard.

Design

The study followed a test-training-test-retest design. On the first day of the experiment observers' contrast sensitivity was measured, for each eye separately, using the Cambridge Low Contrast Gratings (Wilkins and Robson, 1987; Wilkins, Della Sala, Somazzi, Nimmo-Smith, 1988) and their eye dominance was assessed using the Porta test (Porta, 1593; Crovitz and Zener, 1962; Gronwall and Sampson, 1971). Following this, they completed one grid at each hole depth, with each eye separately (different grids were used for the two eyes). Monocular presentation was achieved through observers wearing a translucent lens over the non-observing eye. On the next eight training days observers completed eight grids per day, two at each of the four hole depths (200 decisions per day), using only their non-dominant eye. There was a weekend break in between the first four and the last four days of training. The specific order in which observers completed the grids was counterbalanced across observers and across days. In addition, to prevent observers memorising the location of the targets in a specific grid, the grids were presented at one of four possible orientations: either normal orientation (e.g. vertical, 0°) or rotated through 90°, 180° or 270°. For each four occasions that a grid was
examined it was presented at a different orientation. After the four training days observers' performance was reassessed for each hole depth, for each eye separately, using the same grids for each eye and in the same order, as was used to make the initial assessment. These baseline measures were completed again, by ten of the subjects, 50 days later.

**Procedure**

On the first day of the experiment, observers were given verbal instructions about the task. Following this, they were shown some examples of a target hole in a grid. Once observers understood the experimental procedure the first session began.

In all sessions observers were seated in an artificially illuminated room facing the light box. In order to maintain viewing conditions similar to those used in normal X-ray examination, no restraints to maintain a constant viewing distance were employed. Thus, as in normal X-ray viewing, observers could vary viewing distance in order to help discriminate image features. Observers examined each square in a grid for the presence of a target hole and indicated which of the five possible locations was the most likely to contain a target hole on an answer sheet; the task was thus a five-alternative spatial forced-choice. The answer sheet was laid out as a grid of five by five squares, with a circle in each of the five possible hole locations, in each square; observers crossed through the circle corresponding to the location in which they thought a hole was present for the respective square. Observers were free to search the squares in each grid in any order they wished. The time taken to complete each grid was recorded in seconds using a stop clock.

**4.2.2 Results**

Initial analysis was conducted to assess whether there were any significant differences in accuracy or search times between the pairs of grids with the same hole depth. Matched pairs t-tests revealed that there were no significant differences
between any of the four pairs of grids with the same hole depth (p > 0.05). Consequently, in the analysis of the data from the training sessions the data have been averaged across each pair of grids with the same hole depth.

**Analysis of accuracy data 1) Training sessions**

A two way analysis of variance (day of training (8); hole depth (4) - 0.10 mm, 0.25 mm, 0.50 mm, 0.75 mm) with repeated measures on all factors was conducted to investigate whether observers' accuracy changed over the course of training. As expected there was a main effect of hole depth (F(3,11)=9.10 p=0.003). Polynomial contrasts indicated a significant linear trend (t=5.42 p=0.0001), observers were more accurate on the deeper hole depths; mean scores and standard deviations are shown in table 4.1.

<table>
<thead>
<tr>
<th>Hole Depth (mm)</th>
<th>Mean Score (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>5.06 (0.78)</td>
</tr>
<tr>
<td>0.25</td>
<td>5.23 (0.80)</td>
</tr>
<tr>
<td>0.50</td>
<td>6.56 (1.82)</td>
</tr>
<tr>
<td>0.75</td>
<td>10.80 (5.13)</td>
</tr>
</tbody>
</table>

There was no main effect of day of training or interaction between day of training and hole depth (p > 0.05) indicating that there was no trend for observers' accuracy to change over the course of training.

**Analysis of accuracy data 2) Baseline and transfer sessions**

A three way analysis of variance (time (2) - baseline, transfer; eye (2) - dominant, non-dominant; hole depth (4) - 0.10 mm, 0.25 mm, 0.50 mm, 0.75 mm) with repeated measures on all factors was conducted to examine whether there was any change in observers' accuracy between the baseline and transfer sessions, for each eye. There was no significant difference in performance between the eyes, but, as expected, there was a significant difference in performance on the different hole depths (F(3,11)=18.04 p<0.0005). Polynomial contrasts revealed a significant
linear trend ($t=6.78 \ p=0.00001$), observers were more accurate on the deeper holes (mean scores and standard deviations are shown in table 4.2).

**Table 4.2** Mean accuracy scores (out of 25) and standard deviations (in brackets) for each hole depth across time and eye.

<table>
<thead>
<tr>
<th>Depth (mm)</th>
<th>0.10 mm</th>
<th>0.25 mm</th>
<th>0.50 mm</th>
<th>0.75 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean score</td>
<td>5.25 (0.70)</td>
<td>5.32 (0.99)</td>
<td>6.61 (1.49)</td>
<td>11.13 (3.17)</td>
</tr>
</tbody>
</table>

There was no significant main effect of time ($p > 0.05$) indicating that there was no overall change in observers' accuracy between the baseline and transfer sessions, and there were no significant interaction effects. Thus, overall the results indicate that there was no change in accuracy over time in general, and that there was no change specific to the trained (non-dominant) eye.

**Analysis of search times 1) Training sessions**

A two way analysis of variance (day of training (8); hole depth (4) - 0.10 mm, 0.25 mm, 0.50 mm, 0.75 mm) with repeated measures on all factors was conducted to investigate whether observers' search times (secs.) changed over the course of training. There was a significant main effect of hole depth ($F(3,39)=3.31 \ p=0.030$). Polynomial contrasts indicated a linear trend of borderline significance ($t=2.08 \ p=0.058$), suggesting that observers were slightly quicker on the deeper hole depths (mean scores and standard deviations can be seen in table 4.3).

**Table 4.3** Mean search times (secs.) and standard deviations (in brackets) for each of the hole depths across day of training.

<table>
<thead>
<tr>
<th>Depth (mm)</th>
<th>0.10 mm</th>
<th>0.25 mm</th>
<th>0.50 mm</th>
<th>0.75 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean score</td>
<td>166 (35)</td>
<td>164 (35)</td>
<td>164 (33)</td>
<td>156 (32)</td>
</tr>
</tbody>
</table>

There was a main effect of day of training ($F(7,91)=14.37 \ p<0.0005$) and there was no interaction between day of training and hole depth ($p > 0.05$) indicating that the change in search times occurred equally for the different hole depths.
Polynomial contrasts revealed a significant linear trend for observers to become quicker over the course of training ($t=9.64 \ p<0.0005$). Figure 4.2 shows mean search times for each day of training for each hole depth.

![Search times during training for each hole size](image)

Figure 4.2 Search times during training for each hole size

The linear reduction in search times for each of the hole sizes can be clearly seen in Figure 4.2. Further, it is of interest to note that there is a slight increase in search times between day four and day five of training (day five followed a weekend break). Post hoc comparisons revealed that the increase in search times following the weekend break was significant ($t=3.34 \ p=0.005$).

**Analysis of search times 2) Baseline and transfer sessions**

A three way analysis of variance (time (2) - baseline, transfer; eye (2) - dominant, non-dominant; hole depth (4) - 0.10 mm, 0.25 mm, 0.50 mm, 0.75 mm) with repeated measures on all factors was conducted to examine whether there was any change in observers' search times (secs.) between the baseline and transfer sessions for each eye. There was a significant main effect of eye ($F(1,13)=5.58 \ p=0.034$). Observers were faster at the task with their dominant eye (mean for dominant eye =
187, mean for non-dominant eye = 203). There was also a main effect of hole depth (F(3,39)=3.84 p=0.017). Polynomial contrasts revealed a significant linear trend (t=2.27 p=0.04), observers were faster on the deeper holes (the mean scores and standard deviations are shown in table 4.4).

Table 4.4 Mean search times (secs.) and standard deviations (in brackets) for each of the hole depths across time and eye.

<table>
<thead>
<tr>
<th>Hole Depth</th>
<th>Mean Score</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10 mm</td>
<td>206 (42)</td>
<td></td>
</tr>
<tr>
<td>0.25 mm</td>
<td>192 (49)</td>
<td></td>
</tr>
<tr>
<td>0.50 mm</td>
<td>195 (45)</td>
<td></td>
</tr>
<tr>
<td>0.75 mm</td>
<td>192 (44)</td>
<td></td>
</tr>
</tbody>
</table>

There was also a main effect of time (F(1,13)=153.47 p<0.005). Observers were faster following training by over 100 seconds per grid (baseline mean = 263 secs., transfer mean = 143 secs.); this is a reduction in search time per hole of nearly five seconds. Further, there was a significant interaction between time and eye (F(1,13)=4.49 p=0.05). Figure 4.3 shows search times averaged across the four hole depths for the baseline, transfer and follow-up sessions, for each eye separately. Further, the data are shown averaged across 14 observers who participated in the baseline and transfer sessions, and separately averaged across the subset of 10 observers who also participated in the follow-up sessions.

Figure 4.3 Mean search times (secs.) for the dominant and non-dominant (the trained eye) eyes before and after training.
As can be seen from figure 4.3 there is little difference between the average search times for all 14 observers as compared with the subset of 10 observers at baseline, and there is no difference at transfer. In general, figure 4.3 indicates that observers became quicker over time with both eyes, but that there was greatest improvement for the trained (non-dominant) eye.

Finally, there was an interaction between time and grid ($F(3,39)=4.89 \ p=0.006$), observers showed a greater decrease in search times following training for the shallowest hole depth (0.10 mm) such that following training, search times for the different hole depths had converged to around the same value (mean scores and standard deviations are shown in table 4.5).

Table 4.5 Mean search times and standard deviations (in brackets) for the different hole depths, before and after training, across eye.

<table>
<thead>
<tr>
<th>Hole Depth</th>
<th>Baseline Mean</th>
<th>Transfer Mean</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10 mm</td>
<td>270 (56)</td>
<td>142 (41)</td>
<td>128</td>
</tr>
<tr>
<td>0.25 mm</td>
<td>235 (65)</td>
<td>142 (41)</td>
<td>93</td>
</tr>
<tr>
<td>0.50 mm</td>
<td>238 (54)</td>
<td>149 (41)</td>
<td>89</td>
</tr>
<tr>
<td>0.75 mm</td>
<td>245 (56)</td>
<td>139 (37)</td>
<td>106</td>
</tr>
</tbody>
</table>

Analysis of change, and maintenance of learning after 50 days

Two additional three way analyses of variance (time (3) - baseline, transfer, follow-up; eye (2) - dominant, non-dominant; hole depth (4) - 0.10 mm, 0.25 mm, 0.50 mm, 0.75 mm) with repeated measures on all factors were conducted, using the data from the ten observers who had participated in the 50 day follow-up, to assess whether there was any change in accuracy by the follow-up session, and to assess whether the decrease in search times observed following training was maintained. There was no main effect of time ($p>0.05$) for observers' accuracy. There was a main effect of time for observers' search times ($F(2,18)=80.10 \ p<0.0005$) and, further, there was an even more pronounced interaction between time and eye ($F(2,18)=9.66 \ p=0.001$) than was observed following comparison of just the baseline and transfer sessions. As can be seen from figure 4.3 the decrease in
search times which occurred following training was maintained on the follow-up session; post hoc comparisons indicated an almost significant further decrease in search times on the follow-up session as compared to the transfer session (t=2.08, p=0.06).

**Does contrast sensitivity predict performance?**

An average contrast sensitivity score was calculated for each observer by taking the mean of their separate scores for each eye as measured using the Cambridge Low Contrast Gratings (Wilkins and Robson, 1987; Wilkins, Della Sala, Somazzi, Nimmo-Smith, 1988). Observers' mean performance in the baseline and transfer sessions was calculated by averaging across the eight grids, for search times and accuracy separately. A mean change score was calculated for search times and accuracy by subtracting observers' performance at baseline from their performance at transfer. Four regression analyses were conducted to assess whether observers' baseline performance or change scores for search times and accuracy were predicted by their contrast sensitivity. None of the equations significantly predicted observer performance (p > 0.05).

**4.2.3 Discussion**

X-ray experiment one examined learning in a naturalistic X-ray examination task. Results indicated that observers' performance on the task was not predicted by a measure of contrast sensitivity. Further, there was no change in observers detection accuracy following the training period, and there was no change in accuracy 50 days after the end of the training period. However, following the training period, there was a decrease in the time which observers required to search the grids for target holes. Most of this decrease in search times transferred to the untrained (dominant) eye but a proportion of the decrease observed was specific to the trained eye. The decrease in search times was maintained 50 days after the end of the training period.
Learning in X-ray feature detection

The lack of improvement in observers' detection accuracy was surprising. One possibility is that observers did not have sufficient attentive experience with the target hole for an increase in sensitivity to occur (cf. Shiu and Pashler, 1993). This possibility is discussed in more detail and addressed in X-ray experiment two. However, observers in the present experiment did show improvement in the time taken to search the grids. Most of the decrease in search times observed for the trained eye transferred to the untrained eye (72.5%) transfer. This would imply that the majority of the learning which occurred was localised either at a binocular level of visual processing or in some higher cognitive process. Further, the increase in search times observed following the weekend break during the training period, also suggests a high level learning process; low-level learning is generally found to be more enduring, or even to show consolidation (e.g. contrast sensitivity experiments one and three; Karni and Sagi, 1991, 1993). However, the incomplete interocular transfer pattern suggests that some of the learning occurred at an early monocular level of visual processing. This learning may have resulted from some form of neural tuning perhaps for contrast sensitivity. The results from contrast sensitivity experiments one, two and three, and those of De Valois (1977) and Mayer (1983) lend some plausibility to this possibility. However, the fact that a test of contrast sensitivity did not predict performance is somewhat inconsistent with this possibility (this is discussed in a subsequent section).

Two alternative explanations for the partial transfer of learning for search times are suggested by the data since it is noticeable that following training search times for the trained and untrained eyes are equal. The first explanation is that observers search times decreased to some theoretical minimum due to physical and other constraints of the task (e.g. the time taken to record each decision on the answer sheet etc.). This explanation would seem unlikely since it is noticeable that there has been an almost significant, further decrease in observers search times by the
follow-up session. The second explanation is that the untrained (dominant eye) can only improve to the same level as the trained eye because learning is due to some general or binocular process. This explanation is more difficult to discount on the basis of the current data.

Eye dominance and detection speed

Prior to starting the training procedure observers were significantly faster at detecting the target holes with their dominant eye whilst there was no difference in the accuracy of detection between the two eyes. This result may reflect a difference in processing speed for visual information as a function of eye dominance. For instance Coren and Porac (1982) reported results that indicated that information from the sighting dominant eye was processed faster than information from the sighting non-dominant eye. However, the sort of difference in processing speed they observed was only very small (14 msec), whilst in the present study observers were on average 33 seconds faster (over four grids) with their dominant eye. It may be that, in addition to speed of processing differences, factors such as more efficient guidance of eye movements for the dominant eye were responsible for the bulk of the difference in search times between the two eyes.

Contrast sensitivity and detection of features in X-ray images

The main perceptual demands of the current task were on visual acuity and contrast sensitivity since the most pertinent defining characteristics of the dots was their small size and low contrast level relative to the background. Further, variation between the stimuli was purely in terms of their contrast and this had the expected effect on observer detection accuracy and to a slight extent detection speed. Given the importance of contrast sensitivity for performance on the task it was somewhat surprising that a measure of contrast sensitivity predicted neither initial performance nor the amount of learning. In this respect the present results do not agree with those of Pauli (1993) who found that the Cambridge Low Contrast Gratings were
the only measure in a cognitive test battery which significantly predicted performance for learning to detect abnormal features in mammograms. The present result is further surprising when one considers that the spatial frequency of the gratings (4 cpd), when viewed from the standard testing distance of six metres, is only slightly different from the fundamental spatial frequency of a target hole viewed from 20 cm (5 cpd); the latter viewing distance is an estimate of the average distance from which observers scrutinised the X-rays. Thus, the result cannot be explained as due to the gratings measuring contrast sensitivity for a spatial frequency which was very different to that for the target holes. It is possible that the learning observed in the present study resulted from some process other than a change in contrast sensitivity and that this explains why in contrast to the results of Pauli (1993) the test did not predict learning. However, the question as to why the test did not predict initial performance remains unanswered especially given that the difference in performance for the different hole depths clearly indicates that performance was dependent on the contrast of the stimuli. Clearly further research is required to identify an as yet unspecified covariate which may have influenced detection performance in the present task.

X-ray experiment two

X-ray experiment one indicated an improvement in speed, but not in accuracy, for detection of high spatial frequency, low contrast targets. Some of this improvement may have resulted from low-level learning processes, but given that such learning usually involves an improvement in sensitivity (e.g. De Valois, 1977; Fendick and Westheimer, 1983; Vogels and Orban, 1985; Ball and Sekuler, 1987; Fahle and Edelman, 1993; Goldstone, 1994) the lack of improvement in accuracy of detection was surprising. Shiu and Pashler (1992) have shown that in order for low-level perceptual learning to occur the observer must have sufficient attentive experience with the relevant stimuli to be discriminated. The detection task in X-ray
experiment one was a spatially uncertain one and many of the targets were at or below observers' detection thresholds. This combination of factors could have led to observers lacking attentive experience with the target, which may have been the reason that they did not show any improvement in accuracy. A further problem, resulting from this combination of factors, was that observers may have been unable to develop an adequate visual schema for the target, effectively resulting in them 'forgetting' what they were looking for. X-ray experiment two assesses this possibility by replicating X-ray experiment one but using a match-to-sample paradigm. Thus, as in X-ray experiment one observers were required to search X-rays of grids of twenty-five squares for target dots. However, there was always a sample hole in the centre of a square and the observers task was to detect another hole, which matched the sample hole, in one of the four remaining possible locations. By referring to the sample hole observers were able to have sufficient attentive experience with the target and also to develop an adequate visual schema of the target. Consequently, if there was still no improvement in detection, this could not be due to lack of attentive experience with the target or to forgetting what the target looked like.

As in X-ray experiment one transfer across eyes was measured to provide some indication of the level at which any learning occurred. In addition, a second transfer task testing transfer across contrast levels was incorporated into the training phase of the experiment. Half the observers were trained on 'low-contrast' targets for the first half of training followed by a switch to higher contrast targets in the second half of training, whilst the other half of the observers did the reverse. Thus, transfer across contrast levels was assessed halfway through training. Albrecht and Hamilton (1982) have provided neurophysiological evidence in the monkey and cat that there is extensive variation between cortical cells in the range of contrasts to which they will respond suggesting that independent contrast channels may exist. If learning occurred through changes in contrast sensitivity, then to the extent that this
was specific to particular contrast channels, and to the extent that the stimuli used in the two halves of training excited independent contrast channels, no transfer would be expected across stimuli between the first and second half of training. A second possibility, which would also lead to no transfer across stimuli, would be if improvement in detection performance resulted from observers developing a specific visual schema for each type of target for which they had had extensive detection practice. On the other hand, improvement in detection performance which resulted from a more general learning process such as learning about the task in general or learning of improved visual search patterns would result in transfer across stimuli.

4.3.1 Method

Observers

Fourteen observers were recruited to take part in the study of whom eight completed the full experiment, three men and five women. Observers' ages ranged between 20 and 35 years and all had normal or corrected to normal vision. Observers received a token payment in return for their participation in the study.

Apparatus and stimuli

The stimuli were constructed in the same way as those in X-ray experiment one with respect to hole depths and widths, size, number and type of grids. However, there were two important differences. First, there was a hole drilled in the centre of every square. A second hole of exactly the same dimensions was drilled exactly midway along one of the four hemi-diagonals, there were thus two holes in every square one of which was always in the centre to provide a sample for observers to refer to. Second, the set-up of the X-ray machines used to produce the images had been changed as a result of routine servicing. Consequently, using the same beam density settings and perspex overlay thicknesses as were used in X-ray experiment one produced images in which the dots were not visible. As a result of this, the
images finally selected for X-ray experiment two were taken with different beam densities and thicknesses to those used in X-ray experiment one.

As in X-ray experiment one, the stimuli were displayed on light boxes with a luminance of 2610 cd/m² (illumination = 2000 Lux) and with the area surrounding a stimulus masked out using black cardboard.

**Design**

The basic design of the study was the same as for X-ray experiment one. However, there were two important differences. First, an additional transfer task was embedded in the training phase of the experiment. The observers were split into two groups one group completed only the grids with the deeper hole depths (0.50 mm and 0.75 mm) on the first four days of training and the grids with the shallower hole depths (0.10 mm and 0.25 mm) on the second four days of training, the other group did the reverse of this. Thus, there was a test for transfer across stimuli with different contrast levels between the two halves of training. As before observers completed eight grids (200 decisions) per day but in the present experiment this resulted from them repeating each grid twice in each session (two grids at each of two hole depths, repeated twice = eight grids). The second important difference from X-ray experiment one was that there was no repeat of the baseline measurements 50 days after the end of the experiment.

As in X-ray experiment one the order in which observers completed grids was counterbalanced across observers and across days, and the grids were presented at varying orientations to prevent observers from memorising the location of holes in any particular grid.
Procedure

Before the beginning of the first session observers were given verbal instructions about the task. These instructions emphasised that there was a sample target in the centre of every square and that observers were to refer to this when deciding upon the likely spatial location of the target hole. The task was thus a match-to-sample, four-alternative spatial forced-choice. As before, following the instructions, observers were shown examples of the target hole and, additionally, their attention was also drawn to the presence of a sample hole in the centre of every square.

On each day the experimental procedure was the same as for X-ray experiment one except that observers indicated in which of four possible locations a target was located. As before observers recorded their decisions on a response sheet and the time to complete each grid was recorded in seconds using a stop clock.

4.3.2 Results

As the present experiment incorporated a transfer task during the training phase of the experiment from lower to higher contrast holes or vice versa, performance was averaged across all grids for each day of training in order to assess this transfer. Thus, for the lower contrast training phase performance is averaged across the 0.10 mm and 0.25 mm hole depths for each day of training whilst, in the higher contrast training phase, performance is averaged across the 0.50 mm and 0.75 mm hole depths.

Accuracy data 1) Training sessions

A three way analysis of variance (training group (2) - lower or higher contrast holes first; contrast level (2) - lower contrast holes, higher contrast holes; day of training (4) - days one to four for each training block) with repeated measures on contrast level and day of training was conducted in order to examine observers' performance during the training phase of the experiment. There were no significant main or
interaction effects. As was expected the effect of contrast level tended towards significance (F(1,6)=4.91 p=0.06); the mean for the lower contrast holes was 6.39 correct and the mean for the higher contrast holes was 7.90 correct. The absence of a main effect of day of training indicates that there was no change in observers' mean accuracy over the course of training.

Accuracy data 2) Baseline/transfer sessions

A four way analysis of variance (training group (2) - lower or higher contrast holes first; time (2) - baseline, transfer; eye (2) - dominant, non-dominant; hole depth (4) - 0.10 mm, 0.25 mm, 0.50 mm, 0.75 mm) was conducted with repeated measures on time, eye and hole depth. There were no significant main effects for time, eye or training group. However, as expected there was a main effect of hole depth (F(3,18)=4.24 p=0.020). Polynomial contrasts revealed a significant linear trend (t=2.44 p=0.05). In addition, there was a significant interaction between hole depth and training group (F(3,18)=9.815 p=0.026). The mean scores and standard deviations for each hole depth are shown in table 4.6 collapsed across groups, and separately for each group.

Table 4.6 Mean accuracy scores (out of 25) and standard deviations (in brackets) for each hole depth and for each training group across time and eye.

<table>
<thead>
<tr>
<th>Group</th>
<th>0.10 mm</th>
<th>0.25 mm</th>
<th>0.50 mm</th>
<th>0.75 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higher contrast first</td>
<td>6.33 (2.75)</td>
<td>5.83 (0.63)</td>
<td>7.42 (0.76)</td>
<td>9.75 (1.50)</td>
</tr>
<tr>
<td>Lower contrast first</td>
<td>6.55 (1.16)</td>
<td>7.85 (1.32)</td>
<td>7.00 (1.51)</td>
<td>8.50 (3.07)</td>
</tr>
<tr>
<td>Mean across groups</td>
<td>6.44 (1.72)</td>
<td>6.84 (1.48)</td>
<td>7.21 (1.23)</td>
<td>9.13 (2.54)</td>
</tr>
</tbody>
</table>

Examination of the mean scores in table 4.6 indicates that as hole depth (and therefore resultant contrast) increases so accuracy also increases. Further, whilst the group who trained on the higher contrast grids first were the more accurate for the higher contrast grids, the group who trained on the lower contrast grids first were the more accurate for the lower contrast grids. It is worth noting at this point, that for the smaller hole depths performance was little better than chance (chance
performance = 6.25), and that overall, task difficulty is obviously quite high, even for the deeper holes as indicated by the relatively low scores for these depths.

There was a significant interaction between training group and time (F(1,6)=6.51 p=0.043). The mean scores and standard deviations collapsed across eye and hole depth are shown in Table 4.7.

<table>
<thead>
<tr>
<th>Table 4.7 Mean accuracy scores and standard deviations (in brackets) at baseline and transfer for each training group across eye and hole depth.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Group</td>
</tr>
<tr>
<td>Higher contrast first</td>
</tr>
<tr>
<td>Lower contrast first</td>
</tr>
</tbody>
</table>

The mean scores in Table 4.7 indicate that whilst the group that trained for the first week on the higher contrast stimuli and for the second week on the lower contrast stimuli showed no change in mean accuracy over time, the group who completed the opposite training procedure showed a decline in overall accuracy following training. There were no other significant interaction effects including no interaction between eye and day indicating that there was no change in accuracy over the course of time specific to the trained (non-dominant) eye.

**Search times 1) Training sessions**

A three way analysis of variance (training group (2) - lower or higher contrast holes first; contrast level (2) - lower contrast holes, higher contrast holes; day of training (4) - days one to four for each training block) with repeated measures on contrast level and day of training was conducted to examine observers' performance during the training phase of the experiment. There were no significant main effects of training group, contrast level or day of training. There was a significant interaction between training group and contrast level (F(1,6)=9.79 p=0.020). The mean scores are plotted in Figure 4.4.
Examination of Figure 4.4 suggests that the group who trained on the grids with the higher contrast holes first had faster detection speeds for the grids with the lower contrast holes than the group who started off training on the grids with the lower contrast holes. In other words, training on the grids with the higher contrast holes first may actually have conferred an advantage, in terms of search speed, when it came to searching the grids with the lower contrast holes, as compared with having no prior experience. Training on the grids with the lower contrast holes first did not lead to faster search speeds on the grids with the higher contrast holes than for the group who started off training on the grids with the higher contrast holes.

However, to assess more adequately this asymmetric transfer pattern a better comparison is to compare for each group, their performance on the first day of training for their respective second training task, with their performance for that task prior to the start of training using their non-dominant (trained) eye. In other words, for the group who trained on the lower contrast grids first, their
performance at baseline on the higher contrast grids is compared with their performance on the first training day for the higher contrast grids (training day 5 in Figure 4.4). For the group who trained on the higher contrast grids first, the equivalent comparison is made between their performance on the lower contrast grids at baseline and their performance on these grids on the first day of training with them (i.e. training day 5). Matched pairs t-tests were conducted to make these comparisons. For the group who started training on the lower contrast holes, there was a significant reduction in their search times on the higher contrast holes, as compared with their search times at baseline (t(4)=4.10 p=0.015; baseline mean search time per grid = 243 secs., training day mean search time per grid = 161 secs.). For the group who trained on the higher contrast holes first there was no significant change in their search times on the lower contrast holes as compared with their search times at baseline. Whilst these results also indicate an asymmetric transfer pattern it is the opposite to that which was suggested by examination of the training data in isolation. Thus, these results indicate that whilst training on the lower contrast grids first is advantageous for search time performance on the higher contrast grids, the reverse is not true for training on the higher contrast grids first.

There was a significant interaction between training group and day of training (F(3,10)=4.18 p=0.021). The mean scores (plotted in Figure 4.4) indicate that the group who started training on the grids with the lower contrast holes showed a tendency to become quicker at inspecting the grids, especially over the first half of training, whilst the group who started training on the grids with the higher contrast holes showed no such tendency. There were no other significant interaction effects.

**Search times 2) Baseline/transfer sessions**

A four way analysis of variance (training group (2) - lower or higher contrast holes first; time (2) - baseline, transfer; eye (2) - dominant, non-dominant; hole depth (4) - 0.10 mm, 0.25 mm, 0.50 mm, 0.75 mm) was conducted with repeated measures
on day, eye and hole depth. There were no significant main effects of training group, eye or hole depth but there was a significant main effect of time \( (F(1,6)=9.78 \ p=0.02) \). Observers were faster to search a grid following training (baseline mean time to complete a grid = 257 secs., transfer mean time to complete a grid = 153 secs.). There were no significant interaction effects including no effect of time by eye \((p>0.05)\) indicating that the decrease in reaction times occurred equally for the trained and the untrained eyes.

4.3.3 Discussion

X-ray experiment two further examined learning in a naturalistic X-ray examination task using a match-to-sample paradigm. The results indicated that as in X-ray experiment one, for a group who started training on grids with higher contrast holes, there was no change in detection accuracy during or after a training period. However, for a second group of observers, who started training on grids with lower contrast holes, there was a decrease in their detection accuracy at transfer relative to baseline measurements. Examination of observers' search times indicated that they decreased following the training period and that this decrease transferred to the untrained eye. Examination of the pattern of transfer for the improvements in search times across contrast levels indicated an asymmetric transfer pattern. Observers who started training on the grids with lower contrast holes, showed a decrease in search times especially over the first half of training, and this decrease transferred to search times for the grids with higher contrast holes such that they were faster than at baseline. However, for observers who started training on the grids with the higher contrast holes there was no beneficial effect on their search times for the grids with the lower contrast holes.

**Learning to detect features in X-rays**

The results of X-ray experiment two essentially replicated those of X-ray experiment one; observers' detection speeds for the holes improved but, even with
the provision of a sample target in a known location, there was no improvement in their detection accuracy. The latter finding suggests that the lack of improvement in accuracy was not due to insufficient attentive experience with the target or to observers 'forgetting' what the target looked like.

**Pattern of transfer**

The transfer tasks yielded a mixed pattern of results. The complete transfer of the improvement in search speeds across eyes indicated a binocular or high level learning process. This is different to the partial interocular transfer observed in X-ray experiment one which suggested a low-level component to the learning. The present results in combination with those from X-ray experiment one suggest that perhaps the alternative binocular or high level learning explanation proposed for the results of X-ray experiment one is a more valid interpretation (i.e. that the untrained (dominant) eye could not improve to a level of performance beyond that achieved by the trained (non-dominant) eye).

The asymmetric pattern of transfer across contrast levels was unexpected. Observers who started training on the grids with lower contrast holes showed a decrease in search times over the course of training on these holes which transferred to the grids with higher contrast holes in the second half of training. Observers who started training on the grids with higher contrast holes showed faster performance on the lower contrast holes, at baseline, than the group who started training on those holes. Further, their performance at baseline on the lower contrast holes was as fast as their performance on the grids with the higher contrast holes on the first day of training. Thus, the asymmetric transfer pattern most likely arose because of the difference between the groups in baseline performance on the grids with the lower contrast holes. In addition, the results indicated that the rate at which learning proceeds varies as a function of task difficulty. This is indicated by the relatively fast performance, from the first day of training, of the group who started training
on the grids with the higher contrast holes, on those grids (i.e. most learning occurred immediately after the baseline session), as compared with the slower performance and the decrease in search times observed on the grids with the lower contrast holes for the group who started training on those grids.

**Effect of order of training on accuracy**

In agreement with the results of X-ray experiment one observers, in the present experiment, who started training on the grids with the higher contrast holes, showed no change in accuracy over the course of the experiment. However, the group who started training on the grids with the lower contrast holes were less accurate following training. This may have resulted from some type of motivational factor since performance on the lower contrast holes was clearly only slightly better than chance. Thus, when these observers came to searching the grids with the higher contrast holes they may have already have 'given-up' engaging in a detailed search. An alternative possibility is that they may have developed a visual schema for the low-contrast targets which when applied to the higher contrast targets led them to systematically ignore these targets leading to the decrease in accuracy observed.

**General discussion of X-ray experiments**

The two experiments reported in this chapter have clearly indicated that observers can learn to detect small, low-contrast features in X-rays with greatly increased speed following extensive practice. The pattern of transfer for this learning suggests that it probably results from a binocular or a high level learning process. This finding raises questions regarding the type of learning which might be expected in real world tasks such as X-ray examination. Although previous experiments (e.g. Fiorentini and Berardi, 1980, 1981) have indicated that low-level learning processes can occur under carefully controlled laboratory conditions, the present experiment suggests that in a more naturalistic task and setting, where conditions are not so
precisely controlled, such learning may not be so readily facilitated (this issue is
discussed further in the final chapter of this thesis).

The lack of improvement in accuracy observed in the present experiments could
have occurred for a number of reasons. The results of X-ray experiment two
suggested that the lack of improvement was not due to insufficient attentive
experience with the target. A second possibility was that observers decided to pass
any gains in their performance onto speed of detection rather than accuracy. There
are two factors which would appear to argue against this possibility. The first is that
the instructions for both experiments carefully emphasised that accuracy and speed
were equally important to completion of the task, and the second is that it would
seem unlikely that observers would all consciously choose to pass performance gains
onto accuracy rather than speed. That is unless there was some extraneous factor
operating to influence observers in such a uniform way. In fact there was indeed
one extraneous factor intrinsic to the task which may have led to allocation of any
performance gains to detection speed rather than accuracy. Observers would have
been aware of the approximate time it took them to complete each grid and certainly
of the time it took them to complete each experimental session. Thus, they would
effectively have received feedback on their detection times, and this may have
encouraged them to concentrate on improving the speed of their performance rather
than the accuracy. An experiment recently conducted by Roling (1995) using
exactly the same X-ray stimuli as were used in X-ray experiment one supports this
argument. She gave observers feedback after completion of every grid on the
accuracy of their performance and further, observers were paid a small sum of
money for every hole they detected above chance performance levels. Her results
indicated an improvement in both the speed and the accuracy of observers' performance.

160
The present experiments have indicated through the use of transfer tasks that the improvements in performance which occurred in a naturalistic discrimination task were either not attributable to a low-level learning process or occurred at a binocular level of processing. These findings raise questions regarding the type of learning which might be expected in real world situations even on tasks which ostensibly might be expected to induce low-level learning processes. However, the study did demonstrate learning which may have occurred as a result of higher level processes. One category of learning which would seem likely to underlie the improvements in detection speed observed in the present experiments is search learning. It is possible that as a result of such learning the rate at which observers were able to gather information increased. The next chapter examines search learning in more detail and through the use of transfer tasks assesses the likely type and locus of any learning which occurs following extensive practice on a basic visual search.
Chapter five
Visual search learning

It has frequently been observed that the rate of searching in visual search increases with practice. The increase in the speed with which observers were able to detect targets in the X-ray experiments, reported in chapter four of this thesis, may reflect this type of improvement. However, there has been little investigation of the nature of improvements in visual search other than to investigate whether the changes occur at a low-level (e.g. Karni and Sagi, 1991, 1993), and whether they can occur as a result of serial search becoming parallel (Sireteanu and Rettenbach, 1994). The present chapter reports a more extensive investigation of the nature of learning in basic visual search tasks: understanding the nature of learning in visual search could have important implications for the interpretation of visual discrimination tasks as discussed in the next section.

5.1 Changes in search process as explanations for the learning of discrimination tasks.

Visual discrimination tasks almost invariably involve a visual search for the discriminative features. Initial discrimination rules may be formulated to use highly salient visual features such as those which 'pop-out' in a first-glimpse of the visual scene. Thus, if a discrimination task can only be accomplished through the use of less salient features, which are detected via an attentive serial search process, it is likely to be more difficult to learn (Sowden, Davies and Chivers, 1992). Further, a discrimination task which requires an attentive serial search will be prone to errors resulting from the observer sometimes missing the presence of the discriminative feature. In this context, learning may result from a change in the search process for the discriminative feature. For instance, a change from serial to parallel search could lead to the more rapid learning of a discrimination task, and to the observer missing the presence of the discriminative feature on fewer occasions as a result of
it beginning to pop-out (i.e. an increase in discrimination accuracy). Improvements in discrimination performance which result from search learning could thus mimic perceptual learning.

Evidence for search learning and its effect on subsequent discrimination performance has been shown in the field of radiology. For instance, it has been reported that part of an expert radiologists skill arises from the fact that they can detect abnormalities via a pre-attentive search (Kundel and Nodine, 1975). Hendee (1993, pp. 140-141) suggests that this ability may develop as a result of practice, and general support for this notion was provided by Sireteanu and Rettenbach (1994) who found that serial search could become parallel with practice (for the reader who is unfamiliar with these terms, serial and parallel search are explained in section 5.3.1). The results from the two X-ray experiments reported in chapter four showed an improvement in the rate at which observers were able to detect high spatial frequency, low-contrast dots, from amongst a noisy background. This may also have resulted from an improvement in the way that observers searched for the target dots. It might be, for instance, that instead of using serial search, they learned to detect the target dot via a pre-attentive parallel search process operating over the whole, or at least part, of the visual display.

So, to summarise, there is some evidence to suggest that visual search processes can change as a result of experience, and further, there is some evidence to suggest that such changes can lead to improvements in discrimination accuracy. Thus, the aims of this chapter are twofold. The first aim is to further investigate whether changes in search process could lead to improvements in discrimination accuracy. The second aim is to better understand the nature of the learning processes observed in previous visual search experiments and in the X-ray experiments reported in chapter four. In order to address this second aim, it is important to examine some of the factors that may be involved in search learning, starting with a basic understanding
of some of the factors that influence how a visual search task is initially conducted. These factors are discussed after the following brief definition of the terminology used throughout this chapter.

5.2 Note on terminology
Visual search experiments can be somewhat problematic to describe clearly. Consequently, in order to make the meanings of the terms used throughout this chapter clear, a definition of each is made at this point. In the context of a visual search task observers are presented with ‘displays’ of various ‘elements’ (e.g. red circles and green squares). Each element can have various ‘features’, such as its colour or its shape, which are drawn from ‘dimensions’ of variation. Consequently, each feature represents a ‘value’ on a specific dimension. Some theorists consider that specialised ‘modules’ exist to process each dimension and suggest that these modules form a spatial map for each possible ‘feature-value’. These maps are referred to here as ‘feature maps’.

5.3 Factors influencing the conduct of visual search tasks
The factors that can be considered to influence the way a visual search task is conducted will vary as a function of the theory of visual search one examines. To date there are two main theories of visual search that are of interest for the present discussion. The first theory of interest has been named ‘Feature integration theory’ and was initially proposed as the result of work by Treisman and Gelade (1980). The second theory of interest is referred to as ‘attentional engagement theory’, and was proposed by Duncan and Humphreys (1992).

5.3.1 Feature integration theory
Examining feature integration theory, there are two related points that it is important to understand when considering the possible nature of learning in visual search. The first point concerns the distinction drawn in this theory between
spatially serial and spatially parallel search. The second point concerns the types of search task that will be conducted using parallel search, versus the types of search task that will be conducted using serial search.

Looking at the first point, a search is called serial when in order to detect a target an observer is forced to examine each element in a display one after another (i.e. in series). Thus, in serial search the search time is a linear function of the number of elements in a display. A search is called parallel when an observer is able to detect a target by simultaneously examining all of the elements in a display. In parallel search the search time is independent of the number of elements in a display. From the above it can be seen that in serial self-terminating search (i.e. the observer stops searching when a target is found, or the whole display has been searched), the ratio of the search slopes for trials on which there is no target present compared with trials on which there is a target present will be approximately two to one, whilst in parallel search the ratio will be one to one.

Considering the second point, according to feature integration theory, feature maps are formed for each possible value on dimensions, such as colour and orientation. Thus, detection of a target which is defined by a unique feature can be accomplished by inspecting the map for that feature-value for the presence of activity. Tasks requiring discrimination of a target from distractors on a single dimension have been termed ‘feature search’ (e.g. Treisman and Gormican, 1988). When target-distractor discriminability is high, such tasks can be accomplished using spatially parallel search mechanisms, but when target-distractor discriminability is low, feature search requires focal attention and spatially serial search (Treisman and Gormican, 1988).

5.1 In fact this is a somewhat simplistic model of search: if one assumes that the time taken to check each element is, say, normally distributed, then it can still be expected that search times will always increase with the number of elements present in a display. Thus, even in parallel search, search times will increase with display size, although the rate of increase will be the same for target present and target absent trials.
Tasks requiring discrimination of a target that is defined by a conjunction of features, from distractors which share some features that overlap with those of the target, have been termed 'conjunction search' (e.g. Treisman and Gormican, 1988). Treisman (1988) proposes that in these tasks a 'spotlight of attention' is focused on each location in a master map of locations. When attention is focused on a particular location in the master map, the features which are active in the corresponding location in various feature maps, can be automatically retrieved. Once retrieved the features can be conjoined to see whether the location under scrutiny contains a conjunction of features that define a target. It is proposed that this process continues until either a target has been found, or all possible spatial locations in the master map of locations have been searched. Thus, conjunction search tasks also generally require focal attention and spatially serial search (there are some exceptions e.g. Treisman and Sato, 1990).

Within the context of feature integration theory there are a number of possible ways in which search learning might occur. In particular these relate to general changes in the search process (i.e. from serial to parallel), and the formation of new detectors (e.g. conjunction detectors). These possibilities are discussed in more detail in section 5.5.

5.3.2 Attentional engagement theory

Turning to attentional engagement theory, this theory dispenses with the serial vs. parallel search distinction. Instead it is proposed that following perceptual segmentation and analysis, chosen information is entered into visual short term memory (VSTM) where it becomes available to guide subsequent action. Two influences are proposed to guide entry to VSTM. First, each element in a display is assigned a weight according to how well it matches an attentional template of the to-be-looked-for information. Second, for elements which group together perceptually, any change in the weight for one is distributed to the others; thus, the
more strongly parts of a display are linked, the more likely they are to be selected or rejected together. In essence, the most strongly weighted element is then entered into VSTM, which corresponds to the directing of attention to that particular part of the visual input.

On the basis of this theory, the easiest search tasks will be when the distractors are homogenous and have no features in common with the target, leading to only one element matching the input-template, and to the activation of spreading suppression over the distractors. The most difficult search tasks will be when the target shares more features with the distractors than they share with each other. In this case, the distractors would also match the input-template to a certain degree, whilst there would be little spreading suppression across the distractors, resulting in greatly increased competition for entry to VSTM.

This view of visual search suggests a number of additional ways in which search learning could occur particularly as a consequence of changes in target-distractor discriminability, and hence the weighting and grouping of elements. These possibilities are described further in section 5.5. Evidence that suggests target-distractor discriminability may change as the result of extensive experience of conducting a visual search task is reported in the next section.

5.4 The influence of categorisation on perceptual discrimination

When observers take part in a visual search task they are making a series of categorisations. When an observer decides that a target is present, he/she is assigning it to a target category defined by a certain set of features. Similarly, when an observer decides that no target is present he/she is assigning the distractors to a non-target category defined by a different set of features. In a categorisation

5.2 Duncan and Humphreys refer to the process by which reduction in weight for any non-target is distributed to other stimuli with which it is perceptually grouped as spreading suppression.
experiment which did not involve visual search, Aha and Goldstone (1992) found that as a result of making categorisations (in their experiment these were based on the size of a rectangle and the position of a line within the rectangle) observers' ability to discriminate between categories improved. Goldstone (1994) confirmed and extended these findings when he showed that there was no transfer of learning to discriminations on unpractised separable dimensions (see Garner, 1974, for a discussion on integral and separable features).

5.5 Possible mechanisms of long-term learning in visual search

Bearing in mind the previous sections, a number of possible explanations arise for the improvements in performance seen for visual search tasks. One possibility applicable to conjunction search tasks is that as a result of practice observers develop specialised conjunction detectors for the target features, or that some kind of direct access is gained to already existing conjunction detectors; evidence suggests that many cells in the visual cortex are selective along more than one dimension (e.g. spatial frequency and orientation (De Valois and De Valois, 1980)). The formation of new conjunction detectors could cause features previously detected by serial search to pop-out. This type of learning would be specific to the search tasks that observers had practised on, as the newly formed conjunction detectors would be specific to the practised feature conjunctions.

A second possibility, proposed by Treisman, Vieira and Hayes (1992), is that learning is based on the formation of new and very specific associations between features, their locations and their required responses, such that once the features have been conjoined by serial processes the target is automatically recognised. This type of learning would also, by definition, be specific to the practised search task.

A third possibility, applicable to both feature and conjunction search, is that as a result of carrying out a large number of perceptual categorisations, an increase in
target-distractor discriminability develops by processes similar to those active in the experiments described by Goldstone (1994). In terms of feature integration theory this might cause a target, which was previously detected through serial search, to pop-out (i.e. become detected through parallel search). In terms of attentional engagement theory (see section 5.3.2), this might result in distractors matching the input-template less strongly leading to decreased competition for attention. Following the findings of Goldstone (1994), it could be expected that this type of learning would again be specific to the dimensions on which observers had practised.

Since there are a number of types of learning that would be specific to the practised search task, the present experiment measures transfer of learning to a novel search task, following extensive practice on a training search task, in order to provide a preliminary indication of whether any of these types of learning occur.

In contrast to stimulus specific learning in visual search, another possible outcome is that observers learn about search tasks in a general way; for instance, through acquiring knowledge about the likely frequency and distribution of targets. A general search learning process would be indicated by the transfer of learning to the novel search task. Transfer of learning to this task would also be expected if a more general change in search process occurred.

For instance, Sireteanu and Rettenbach (1994) found that serial search could become parallel with practice, and further, they found that this change in search process transferred to novel stimulus sets. This type of change in search process implies a change in the way in which attention is allocated. Pashler (1987) suggested that observers searched clumps of stimuli in parallel, but moved between clumps in a serial search pattern. Given his suggestion, and the findings of Sireteanu and Rettenbach (1994) it might be that following practice, larger clumps
could be searched using a parallel search process. In addition to measuring transfer to novel search tasks, a more specific assessment of whether there is a change in search process of this type will be made in the present experiment through examination of search slopes before, during and after training.

5.6 Design of the present experiment

In the present experiment observers were required to take part in extended training on one search task after which transfer of learning to various other search tasks was measured. The amount of any transfer to these tasks can help to distinguish between the alternatives described in section 5.5 as potential sources of learning. By examining search learning the possibility that it may sometimes underlie improvements in discrimination previously attributed to perceptual learning is investigated.

In the context of perceptual learning, the type of search tasks that are of interest in the present experiment, are those that require slow and effortful search; discriminations based on the pre-attentive detection of features are unlikely to involve an extended learning process and hence to have been mistaken for perceptual learning in previous research.

It has been well established that many conjunction search tasks require slow and effortful search (e.g. Treisman and Gelade, 1980). In addition, Treisman and Gormican (1988) conducted a series of experiments to investigate detection of targets distinguished from the distractors on a single dimension. They found that targets which were defined by larger values on quantitative dimensions, such as length, number and contrast, by line curvature, by misaligned orientation, and by values that deviated from a standard or prototypical colour or shape, were detected easily and often in parallel. Conversely, targets defined by smaller values on quantitative dimensions, by straightness, by frame aligned orientation, and by
prototypical colours or shapes, required slow serial search. They suggested that whilst larger or non-prototypical values on a dimension result in an increase in activity in feature detectors leading to rapid detection, smaller or prototypical values result in a reduction in activity which is too small in proportion to the overall level of activity to be noticed.

The basic task used in the present experiment required observers to search a display with varying numbers of elements, superimposed on a random-dot background, for a possible target. Whilst describing the experiment each display will be referred to here as a ‘stimulus’. The search tasks were designed such that slow and effortful search was required, at least initially. Two groups of stimuli were designed: one group used elements which were made up from a combination of two possible values on the dimensions of shape (‘circle’ or ‘ellipse’) and luminance (‘light’ or ‘dark’) whilst the other group (‘novel dimensions stimuli’) used elements which were made up from a combination of values on the dimensions of line length (‘short’ or ‘long’) and orientation (‘left’ or ‘right’). The search tasks were either conjunction search tasks requiring consideration of two dimensions (e.g. luminance and shape, or line length and orientation) or slow feature search tasks requiring search for a target defined by a prototypical feature amongst distractors defined by non-prototypical features (e.g. light circle amongst light ellipses). All the search tasks used are listed in table 5.1 and photographic examples of the stimuli can be seen in Appendix G.

Prior to the start of the training procedure, observers' search accuracy and speed using their non-dominant eye\textsuperscript{5.3} was assessed for a number of search tasks using different combinations of the elements described above. On all but one of these

\textsuperscript{5.3} As described in chapter three, observers' performance was measured for each eye separately in order to allow later assessment of interocular transfer. This provides an indication of whether learning is low-level in nature. Further, the non-dominant eye was used during training since, similar to the effects of orientation anisotropy described in chapter three, there may be greater potential for learning with the non-dominant eye.
tasks, the target was a light circle, but the number and heterogeneity of the distractors was varied. On the remaining task, observers were required to perform a conjunction search task using the novel dimensions stimuli. Observers' search speed and accuracy was also assessed using their dominant eye, for one circle search task (the same elements as used in the training task). Following this initial performance assessment, observers completed three days training in which they searched for a conjunctive target (a light circle) embedded in distractors (light ellipses and dark circles) that shared features with the target. After training, performance was assessed again for all the tasks ('transfer tasks') measured at baseline. Observers were followed up two weeks after training to provide some measure of the duration of learning: whilst learning based upon the development of high level strategies is often relatively unenduring other types of learning such as that localised at a low-level generally endure for long periods of time (e.g. Karni and Sagi, 1993).

Design of transfer tasks
To restate briefly, there were two main aims to this experiment. The first aim was to determine whether search learning could lead to an improvement in discrimination accuracy, and could thus mimic perceptual learning. This can be easily determined in the present experiment by examining whether the number of correct target detections observers make increases as a result of practice. The second aim was to provide some indication of the nature of any learning which leads to an improvement in detection accuracy or detection speed, and section 5.5 outlined a number of ways that such search learning could occur. To distinguish adequately between each of these alternatives would require a series of experiments rather than the single experiment presented here. However, some preliminary indications about the nature of any learning observed can be derived.
The present experiment measured transfer of any learning observed to a variety of search tasks to provide information about the nature of search learning. First, transfer to the novel dimensions stimuli was measured to examine whether learning was specific or general in nature. Second, search slope ratios were examined before, during, and after training to determine whether there was a change in search process: a change in search process for all the stimuli (training and transfer) would indicate a general change in the way that attention was allocated, whereas a change in search process specific to the training stimuli would indicate some type of specific learning, such as learning through the processes discussed in section 5.5. Third, interocular transfer was measured to provide some indication of whether any learning observed was localised at a low-level of visual processing (cf. the findings of Karni and Sagi, 1991, 1993).

In addition to the above, general transfer measures, some more specific information regarding the nature of any change in search process could be obtained by measuring the degree of transfer to combinations of the distractor elements other than the combination used in training. Transfer was measured to new, more homogenous backgrounds where the discrimination could be made on the basis of one of the dimensions (luminance or shape); transfer to these feature search conditions would suggest learning was based on increased target-distractor discriminability rather than increased spreading suppression, since, learning based on an improved ability to group and hence reject heterogeneous distractors together, would not lead to improvement on conditions with homogenous distractors which would group together anyway. The pattern of transfer to an even more heterogeneous distractor set, which contained the same distractors as the training stimuli plus an additional type of distractor, allowed further assessment of this latter possibility, as only partial transfer of learning to this set would be expected if learning were based on the improved operation of spreading suppression. Finally, assessment of transfer to backgrounds with the same degree of heterogeneity as used
in training, but consisting of different combinations of features allowed assessment of the degree to which learning resulted from a change in the interaction between the target and the training distractor set rather than in the interaction with more complex backgrounds in general.

5.7 Method

Observers
Eight observers were recruited to take part in the study, seven females and one male. Observers' ages ranged from 20 to 33 years old with a mean age of 26 years. All of the observers had normal or corrected to normal vision.

Apparatus and stimuli
Stimuli were displayed on a 14" colour monitor at a resolution of 640 x 480 pixels. Each stimulus was composed of various elements on a rectangular random dot background which subtended a visual angle of 5.71 degrees by 7.84 degrees. Each element was made up from a combination of two feature-values, taken from two bi-valued dimensions. For most stimuli the dimensions were shape (circle or ellipse) and luminance (light or dark). For the novel dimensions stimulus set the dimensions were line length (short or long) and line orientation (left or right). Half the stimuli contained a target, the spatial position of which was varied at random. Stimuli varied in the number of distractors which were present. All of the stimuli, except those presented during either the familiarisation phase or the training phase, had either 6, 12, 18 or 24 distractors. Those presented in the familiarisation or training phases had either 6, 12 or 18 distractors. Each number of distractors occurred equally often in a particular search task.

The dots comprising the random dot backgrounds were set to one of three possible luminance levels (either 21.7 cd/m², 29.8 cd/m² or 39.7 cd/m²), and the space-
averaged, mean luminance of the backgrounds was 30.4 cd/m². The remainder of the screen surrounding a stimulus was of a uniform luminance (59.8 cd/m²). For the stimuli with elements varying along the dimensions of shape and luminance the target was always a light circle, whilst the distractors took on the other possible combinations of these feature-values. The measurements for the feature-values were as follows: light = 65.5 cd/m²; dark = 54.1 cd/m²; the ellipse subtended 34' x 26' (area = 12.37 mm²) with the long axis aligned vertically; the circle subtended 30' (area = 12.57 mm²). In the novel dimensions stimulus set, the target was always a short line leaning right. The measurements for the line lengths were as follows: short = 5' x 30', long = 5' x 42'. The luminance for all the lines was 9.07 cd/m². The possible combinations of features formed eight different search tasks and these are detailed in table 5.1 below.

<table>
<thead>
<tr>
<th>Distractors</th>
<th>Type of search task (dimensions discrimination made on)</th>
</tr>
</thead>
<tbody>
<tr>
<td>dark circles</td>
<td>feature search (luminance)</td>
</tr>
<tr>
<td>light ellipses</td>
<td>feature search (shape)</td>
</tr>
<tr>
<td>dark ellipses</td>
<td>feature search (luminance or shape)</td>
</tr>
<tr>
<td>dark circles &amp; dark ellipses</td>
<td>feature search (luminance) but with heterogeneous distractors</td>
</tr>
<tr>
<td>light ellipses &amp; dark ellipses</td>
<td>feature search (shape) but with heterogeneous distractors</td>
</tr>
<tr>
<td>dark circles &amp; light ellipses</td>
<td>conjunction search (luminance and shape)</td>
</tr>
<tr>
<td>dark circles &amp; light ellipses &amp; dark ellipses</td>
<td>conjunction search (luminance and shape)</td>
</tr>
<tr>
<td>long lines leaning right &amp; short lines leaning left</td>
<td>conjunction search (length and orientation)</td>
</tr>
</tbody>
</table>

* Training task

For stimuli with multiple types of distractors there were equal numbers of each distractor type. In addition to the experimental stimuli, there was a set of familiarisation stimuli in which the targets were red squares (luminance = 17.1 cd/m²) and the distractors were blue squares (luminance = 13.7 cd/m²). The squares side length was 30'. Photographs of some stimuli are shown in Appendix G.
In the experiment observers responded by pressing a button on a Pro\textsuperscript{TM} game-pad\textsuperscript{5.4} (essentially a joystick where the directional control is via a pad, which you press in the appropriate direction to indicate desired direction of movement, rather than an upright stick) as soon as their decision was made. Recording of the actual choice was done using the computer keyboard for input.

**Design**

On the first day of the experiment ('baseline session'), observers' sighting eye dominance was determined through the use of the Porta test (Porta, 1593; Crovitz and Zener, 1962; Gronwall and Sampson, 1971). Following this, they completed a search task consisting of 30 trials using the familiarisation stimuli, in order to accustom them to visual search tasks in general, and to the specific experimental procedure. Having completed these trials, observers had their performance, using their non-dominant eye, measured for each of the search tasks detailed in table 5.1. In addition, performance with their dominant eye was measured for the training search task. The order of measurement for the tasks was counterbalanced across observers. Monocular presentation of the stimuli was achieved through observers wearing spectacles with one translucent lens (as described for the contrast sensitivity studies).

In total observers completed 80 trials for each search task; 20 trials for each of the four possible numbers of distractors, and with a target present on half the trials. On the second, third and fourth days of the experiment (the 'training sessions'), observers took part in a training procedure in which they completed 480 trials a day on the training search task; 160 trials for each of the three possible numbers of distractors with a target present on half the trials. The stimuli used in training had

\footnote{A gamepad was used as the primary input device as this avoids problems in recording precise reaction times associated with the buffering of keyboard input, and standard hardware interrupt schedules. Graves and Bradley (1987) report that joystick input yields an average error of 0.55 msec as compared with an unpredictable average error ranging from 18.4 to 36.7 msec (varies across different keyboards) with keyboard input.}
6, 12 or 18 distractors so that the generality of any change in search slopes, following training, to larger distractor sets could be measured (stimuli used in the baseline search tasks had 6, 12, 18 or 24 distractors). On the fifth day of the experiment (the ‘transfer session’), observers had their performance measured for each of the search tasks that were measured on the first day of the experiment. This allowed assessment of whether any improvement that had occurred on the training search task, transferred to other types of stimuli. Two weeks later, observers again had their performance measured for each of the different search tasks (the ‘follow-up session’).

**Procedure**

The presentation of stimuli and recording of data was controlled by purpose written computer software. For all sessions the observers were seated in front of the monitor, in an artificially lit room. On each trial, the observer’s task was to examine each stimulus that was presented and to decide whether there was a target present or not (the target was pre-defined before every block of trials by the experimenter and through the presentation of five example stimuli each of which had a box around the target). Each trial was started by the observer pressing the ENTER key on the computer keyboard. Following this, a stimulus was immediately displayed in the centre of the screen. The stimulus remained displayed until the observer pressed a button on the game-pad, after which a prompt was displayed asking the observer to input a number to indicate whether they thought there had been a target present. Observers entered ‘1’ for target present and ‘3’ for target absent. They then started the next trial by pressing the ENTER key again. The computer program recorded the observer’s reaction time from when the stimulus was displayed until they pushed a game-pad button. The program also recorded what the stimulus was and the observer’s response.
5.8 Results

**Note on the effects tested**

Whilst this experiment generated a large amount of data which are of relevance to the visual search literature in general, for the purposes of this thesis only effects which might indicate learning (i.e. main effects and interactions indicating change over time) and which allowed assessment of its nature and transferability were tested. Thus, main effects examining absolute differences in performance between the different types of search task are not reported here.

**Training sessions: accuracy data**

A three way analysis of variance was conducted (time (3) - training day 1, training day 2, training day 3; number of distractors (3) - 6, 12 or 18; target present/absent (2)) with repeated measures on all factors. Observers were more accurate on target absent than target present trials ($F(1,7)=35 \; p=0.001$) and on trials with fewer distractors ($F(2,14)=19.04 \; p<0.0005$). An interaction between target present/target absent and number of distractors ($F(2,14)=9.82 \; p=0.0002$) indicated that for target present trials there was a stronger trend for observers to become less accurate as the number of distractors increased than for target absent trials. The relevant means can be viewed in table 5.2.

<table>
<thead>
<tr>
<th></th>
<th>6 Distractors</th>
<th>12 Distractors</th>
<th>18 Distractors</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Target present</strong></td>
<td>88.60%</td>
<td>79.05%</td>
<td>75.30%</td>
<td>80.98%</td>
</tr>
<tr>
<td><strong>Target absent</strong></td>
<td>97.75%</td>
<td>97.55%</td>
<td>96.10%</td>
<td>97.13%</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>93.18%</td>
<td>88.30%</td>
<td>85.70%</td>
<td>89.06%</td>
</tr>
</tbody>
</table>

Observers showed no significant change in accuracy over time ($p>0.05$), nor was there any significant interaction between time and the number of distractors ($p>0.05$), or target present/target absent ($p>0.05$). Thus, no learning effect was
apparent in the accuracy with which observers completed the task over the three days of training.

**Training sessions: reaction time data**

A three way analysis of variance was conducted (time (3) - training day 1, training day 2, training day 3; number of distractors (3) - 6, 12 or 18; target present/absent (2)) with repeated measures on all factors. Observers were significantly faster on target present trials (F(1,7)=8.9 p=0.02) and on trials with fewer distractors (F(2,6)=10 p=0.012). However, there was no two way interaction between target present/target absent and number of distractors (p>0.05), or three way interaction between target present/target absent, number of distractors and time (p>0.05). Thus, there is no evidence to indicate that search slopes are steeper for negatives than positives on any day of training; the usual finding with a serial search process is a ratio between the search slopes of 2:1, target absent:target present. Figure 5.1 illustrates the search slopes for target present and target absent trials for each day of training.

![Search times for target present (TP) and target absent trials (TA) for each day of training (D1, D2, D3) for the different numbers of distractors.](image)

**Figure 5.1** Search times for target present (TP) and target absent trials (TA) for each day of training (D1, D2, D3) for the different numbers of distractors.
In order to test more specifically the hypothesis that the ratio of the search slopes was greater than or equal to 2:1 (absent:present), individual least square slope estimates were obtained for each observer, for target present and target absent trials, and for each day of training. Next, for each observer, for each day of training, the ratio of the target absent slope to the target present slope was calculated. The mean of this statistic for day 1 was 1.53, for day 2 was 1.69 and for day 3 was 1.85. A t-test was conducted to test the hypothesis that the slope ratio for each day was equal to or greater than 2. For the first day of training this hypothesis was rejected ($t(7)=2.77 \ p<0.05$) whilst for the second and third days of training this hypothesis was accepted (second day, $t(7)=1.27 \ p>0.05$; third day, $t(7)=0.76 \ p>0.05$). These results suggest that the search process moved closer towards a serial search process over time.

Observers showed a significant decrease in reaction times over the course of the three days of training ($F(2,14)=7.32 \ p=0.007$). There was no significant interaction between time and number of distractors ($p>0.05$) or between time and target present/target absent ($p>0.05$) indicating that the decrease in reaction times occurred equally for stimuli with different numbers of distractors and for target present and target absent trials. Figure 5.2 illustrates the change in reaction times over the course of training for each type of trial.
Figure 5.2 Change in reaction times over the course of training for target present (TP) and target absent (TA) trials and for the different numbers of distractors.

**Baseline, transfer and follow-up sessions data**

*Note on the terminology used in the description of the results*

On each session observers completed nine search tasks; seven with circles and ellipses as elements, one with lines as elements, and a repeat of one of the tasks with circles and ellipses as elements (the same task as the training task), but with their dominant eye. For the purposes of explaining the results, these tasks will be described on the basis of their distractor sets. Hence, for example, the task involving search for a *light circle* amongst *dark circles* and *light ellipses* will be referred to simply as *dark circles & light ellipses*. Further, the task completed using the dominant eye will be referred to as *dominant eye* and the task involving search for a short line sloping right amongst short lines sloping left and long lines sloping right will be referred to as *lines*.

**Baseline, transfer and follow-up sessions: accuracy data**

A three way analysis of variance (time (3) - baseline, transfer, follow-up; target present/target absent (2); number of distractors (4) - 6, 12, 18, 24) with repeated
measures on all factors was conducted for each of the different search tasks. For eight out of the nine search tasks observers were significantly more accurate on target absent than target present trials. The test statistics and mean scores are shown in table 5.3. Whilst for target absent trials performance approached a ceiling level this was not the case for target present trials.

Table 5.3 Results of analyses of variance: main effect of target present/target absent for each search task (means are collapsed across time and number of distractors).

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>D.F.</th>
<th>p</th>
<th>Target present mean (%)</th>
<th>Target absent mean (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>dark circles</td>
<td>22.48</td>
<td>1,7</td>
<td>0.002</td>
<td>88.65</td>
<td>95.73</td>
</tr>
<tr>
<td>light ellipses</td>
<td>26.62</td>
<td>1,7</td>
<td>0.001</td>
<td>79.79</td>
<td>96.34</td>
</tr>
<tr>
<td>dark ellipses</td>
<td>17.05</td>
<td>1,7</td>
<td>0.004</td>
<td>92.08</td>
<td>98.85</td>
</tr>
<tr>
<td>dark circles &amp; dark ellipses</td>
<td>1.88</td>
<td>1,7</td>
<td>0.355</td>
<td>92.19</td>
<td>94.17</td>
</tr>
<tr>
<td>light ellipses &amp; dark ellipses</td>
<td>9.52</td>
<td>1,7</td>
<td>0.018</td>
<td>80.63</td>
<td>95.21</td>
</tr>
<tr>
<td>dark circles &amp; light ellipses*</td>
<td>13.48</td>
<td>1,7</td>
<td>0.008</td>
<td>83.44</td>
<td>95.83</td>
</tr>
<tr>
<td>dark circles &amp; light ellipses &amp; dark ellipses</td>
<td>6.99</td>
<td>1,7</td>
<td>0.033</td>
<td>85.10</td>
<td>94.06</td>
</tr>
<tr>
<td>dominant eye</td>
<td>19.00</td>
<td>1,7</td>
<td>0.003</td>
<td>81.15</td>
<td>94.38</td>
</tr>
<tr>
<td>lines</td>
<td>36.90</td>
<td>1,7</td>
<td>0.001</td>
<td>74.27</td>
<td>96.46</td>
</tr>
</tbody>
</table>

* Training search task.

For seven of the nine search tasks observers accuracy varied as the number of distractors varied. The test statistics and mean scores are shown in table 5.4.
Table 5.4 Results of analyses of variance: main effect of number of distractors for each search task (means are collapsed across time and target present/target absent).

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>D.F.</th>
<th>p≤</th>
<th>6</th>
<th>12</th>
<th>18</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>dark circles</td>
<td>3.05</td>
<td>3,21</td>
<td>0.0510</td>
<td>94.79%</td>
<td>93.13%</td>
<td>91.25%</td>
<td>89.58%</td>
</tr>
<tr>
<td>light ellipses</td>
<td>6.77</td>
<td>3,21</td>
<td>0.0020</td>
<td>94.17%</td>
<td>89.38%</td>
<td>85.42%</td>
<td>83.33%</td>
</tr>
<tr>
<td>dark ellipses</td>
<td>10.89</td>
<td>3,21</td>
<td>0.0005</td>
<td>98.96%</td>
<td>95.63%</td>
<td>93.13%</td>
<td>94.17%</td>
</tr>
<tr>
<td>dark circles &amp; dark ellipses</td>
<td>2.84</td>
<td>3,5</td>
<td>0.1450</td>
<td>93.08%</td>
<td>96.25%</td>
<td>93.13%</td>
<td>90.21%</td>
</tr>
<tr>
<td>light ellipses &amp; dark ellipses</td>
<td>7.95</td>
<td>3,21</td>
<td>0.0010</td>
<td>91.67%</td>
<td>91.25%</td>
<td>85.42%</td>
<td>83.33%</td>
</tr>
<tr>
<td>dark circles &amp; light ellipses</td>
<td>12.52</td>
<td>3,21</td>
<td>0.0005</td>
<td>93.75%</td>
<td>91.67%</td>
<td>86.46%</td>
<td>86.67%</td>
</tr>
<tr>
<td>dark circles &amp; light ellipses &amp; dark ellipses</td>
<td>4.71</td>
<td>3,21</td>
<td>0.0110</td>
<td>93.54%</td>
<td>91.04%</td>
<td>88.96%</td>
<td>84.79%</td>
</tr>
<tr>
<td>dominant eye</td>
<td>5.52</td>
<td>3,21</td>
<td>0.0060</td>
<td>91.88%</td>
<td>90.00%</td>
<td>84.58%</td>
<td>84.58%</td>
</tr>
<tr>
<td>lines</td>
<td>20.69</td>
<td>3,21</td>
<td>0.0005</td>
<td>90.63%</td>
<td>88.75%</td>
<td>82.29%</td>
<td>80.21%</td>
</tr>
</tbody>
</table>

* Training search task.

Polynomial contrasts revealed that for these seven search tasks there was a significant linear trend such that as the number of distractors increased accuracy decreased. The test statistics are shown in table 5.5.

Table 5.5 Results of polynomial contrasts: main effect of number of distractors.

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>p≤</th>
</tr>
</thead>
<tbody>
<tr>
<td>light ellipses</td>
<td>4.09</td>
<td>0.0050</td>
</tr>
<tr>
<td>dark ellipses</td>
<td>5.35</td>
<td>0.0010</td>
</tr>
<tr>
<td>light ellipses &amp; dark ellipses</td>
<td>3.82</td>
<td></td>
</tr>
<tr>
<td>dark circles &amp; light ellipses *</td>
<td>4.71</td>
<td></td>
</tr>
<tr>
<td>dark circles &amp; light ellipses &amp; dark ellipses</td>
<td>3.50</td>
<td></td>
</tr>
<tr>
<td>dominant eye</td>
<td>3.01</td>
<td>0.0200</td>
</tr>
<tr>
<td>lines</td>
<td>7.28</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

* Training search task

There was an interaction between target present/target absent and the number of distractors for five out of the nine search tasks. Table 5.6 shows the test statistics and mean scores for each search task.
Table 5.6 Results of analyses of variance: Interaction between number of distractors and target present (TP)/target absent (TA) for each search task (means are collapsed across time).

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>D.F</th>
<th>p≤</th>
<th>6</th>
<th>12</th>
<th>18</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>dark circles</td>
<td>10.95</td>
<td>3,21</td>
<td>0.0005</td>
<td>TA</td>
<td>92.08%</td>
<td>97.50%</td>
<td>95.00%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TP</td>
<td>97.50%</td>
<td>91.25%</td>
<td>85.83%</td>
</tr>
<tr>
<td>light ellipses</td>
<td>1.84</td>
<td>3,21</td>
<td>0.1710</td>
<td>TA</td>
<td>99.17%</td>
<td>97.08%</td>
<td>94.17%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TP</td>
<td>89.17%</td>
<td>81.67%</td>
<td>76.67%</td>
</tr>
<tr>
<td>dark ellipses</td>
<td>4.51</td>
<td>3,21</td>
<td>0.0140</td>
<td>TA</td>
<td>100%</td>
<td>98.33%</td>
<td>98.33%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TP</td>
<td>97.92%</td>
<td>92.92%</td>
<td>87.92%</td>
</tr>
<tr>
<td>dark circles &amp; light ellipses</td>
<td>0.78</td>
<td>3,21</td>
<td>0.5200</td>
<td>TA</td>
<td>93.33%</td>
<td>96.25%</td>
<td>93.75%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TP</td>
<td>92.92%</td>
<td>92.5%</td>
<td>92.5%</td>
</tr>
<tr>
<td>light ellipses &amp; dark ellipses</td>
<td>2.25</td>
<td>3,21</td>
<td>0.1120</td>
<td>TA</td>
<td>96.67%</td>
<td>97.08%</td>
<td>93.75%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TP</td>
<td>86.67%</td>
<td>85.42%</td>
<td>77.08%</td>
</tr>
<tr>
<td>dark circles &amp; light ellipses*</td>
<td>5.46</td>
<td>3,21</td>
<td>0.0060</td>
<td>TA</td>
<td>97.08%</td>
<td>97.08%</td>
<td>94.58%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TP</td>
<td>90.42%</td>
<td>86.25%</td>
<td>78.33%</td>
</tr>
<tr>
<td>dark circles &amp; light ellipses &amp; dark ellipses</td>
<td>2.79</td>
<td>3,5</td>
<td>0.1490</td>
<td>TA</td>
<td>93.75%</td>
<td>93.33%</td>
<td>94.17%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TP</td>
<td>93.33%</td>
<td>88.75%</td>
<td>83.75%</td>
</tr>
<tr>
<td>dominant eye</td>
<td>12.70</td>
<td>3,21</td>
<td>0.0005</td>
<td>TA</td>
<td>94.58%</td>
<td>92.92%</td>
<td>96.67%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TP</td>
<td>89.17%</td>
<td>87.08%</td>
<td>72.50%</td>
</tr>
<tr>
<td>lines</td>
<td>8.40</td>
<td>3,21</td>
<td>0.0010</td>
<td>TA</td>
<td>97.08%</td>
<td>96.25%</td>
<td>96.67%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TP</td>
<td>84.17%</td>
<td>81.25%</td>
<td>67.08%</td>
</tr>
</tbody>
</table>

* Training search task

Pearson's product moment correlations were calculated, for each of the significant interactions, between the percentage of correct decisions and the number of distractors, for target present and target absent trials separately, in order to assess whether there was a significant linear trend. These are displayed in table 5.7. As can be seen from table 5.7, for target absent trials there tends to be no significant linear relationship between accuracy and the number of distractors except in the case of dark circles where there is a significant improvement in accuracy for target absent trials as the number of distractors increases. In contrast to these findings, for target present trials there tends to be a significant linear decrease in accuracy as the number of distractors increases.
Table 5.7 Results of correlations: number of distractors with target present and target absent separately.

<table>
<thead>
<tr>
<th></th>
<th>r</th>
<th>p≤</th>
</tr>
</thead>
<tbody>
<tr>
<td>dark circles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TA</td>
<td>.995</td>
<td>.0020</td>
</tr>
<tr>
<td>TP</td>
<td>-.999</td>
<td>.0005</td>
</tr>
<tr>
<td>dark ellipses</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TA</td>
<td>-.613</td>
<td>.1940</td>
</tr>
<tr>
<td>TP</td>
<td>-.879</td>
<td>.0610</td>
</tr>
<tr>
<td>dark circles &amp; light ellipses *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TA</td>
<td>-.894</td>
<td>.0530</td>
</tr>
<tr>
<td>TP</td>
<td>-.938</td>
<td>.0310</td>
</tr>
<tr>
<td>dominant eye</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TA</td>
<td>.000</td>
<td>.5000</td>
</tr>
<tr>
<td>TP</td>
<td>-.858</td>
<td>.0710</td>
</tr>
<tr>
<td>lines</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TA</td>
<td>-.799</td>
<td>.1010</td>
</tr>
<tr>
<td>TP</td>
<td>-.954</td>
<td>.0230</td>
</tr>
</tbody>
</table>

* Training search task

A significant change in accuracy occurred over time for three search tasks (light ellipses, light ellipses & dark ellipses and dominant eye); test statistics and mean scores are shown in table 5.8.

Table 5.8 Results of analyses of variance: main effect of change over time for each search task (means are collapsed across target present/target absent and number of distractors).

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>D.F.</th>
<th>p≤</th>
<th>Baseline mean (%)</th>
<th>Transfer mean (%)</th>
<th>Follow-up mean (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>dark circles</td>
<td>1.37</td>
<td>2,14</td>
<td>0.286</td>
<td>92.03</td>
<td>90.94</td>
<td>93.59</td>
</tr>
<tr>
<td>light ellipses</td>
<td>12.41</td>
<td>2,14</td>
<td>0.007</td>
<td>83.59</td>
<td>91.25</td>
<td>89.38</td>
</tr>
<tr>
<td>dark ellipses</td>
<td>2.69</td>
<td>2,14</td>
<td>0.103</td>
<td>93.75</td>
<td>96.41</td>
<td>96.25</td>
</tr>
<tr>
<td>dark circles &amp; dark ellipses</td>
<td>2.03</td>
<td>2,6</td>
<td>0.212</td>
<td>92.50</td>
<td>93.13</td>
<td>93.91</td>
</tr>
<tr>
<td>light ellipses &amp; dark ellipses</td>
<td>4.18</td>
<td>2,14</td>
<td>0.038</td>
<td>84.06</td>
<td>90.94</td>
<td>88.75</td>
</tr>
<tr>
<td>dark circles &amp; light ellipses &amp; dark ellipses *</td>
<td>0.18</td>
<td>2,14</td>
<td>0.834</td>
<td>89.22</td>
<td>89.06</td>
<td>90.63</td>
</tr>
<tr>
<td>dominant eye</td>
<td>0.75</td>
<td>2,14</td>
<td>0.491</td>
<td>88.75</td>
<td>90.75</td>
<td>89.22</td>
</tr>
<tr>
<td>lines</td>
<td>4.23</td>
<td>2,14</td>
<td>0.037</td>
<td>86.09</td>
<td>86.25</td>
<td>90.94</td>
</tr>
<tr>
<td></td>
<td>0.17</td>
<td>2,14</td>
<td>0.846</td>
<td>85.94</td>
<td>85.63</td>
<td>84.53</td>
</tr>
</tbody>
</table>

* Training search task.

Looking at the mean scores in table 5.8 it would seem that for light ellipses and light ellipses & dark ellipses there was an increase in accuracy immediately after training but this was less pronounced by the follow-up session. Polynomial contrasts revealed a significant quadratic trend for these two types of trial (light ellipses, t=4.31 p=0.004; light ellipses & dark ellipses, t=2.36 p=0.05). For the
search task conducted with the dominant eye there appeared to be an increase in accuracy by the follow-up session and polynomial contrasts revealed a significant linear trend (t=3.05 p=0.019). Examination of the interaction effects revealed that for all search tasks except dark circles there was no interaction between time and target present/target absent or between time and the number of distractors. For the search task with dark circles as distractors there was an improvement over time for trials with a target present but not for trials with a target absent (F(2,6)=27.88 p=0.0001; target present means (% correct): baseline = 86.25, transfer = 86.56, follow-up = 93.13; target absent means (% correct): baseline = 97.81, transfer = 95.31, follow-up = 94.06).

It can be concluded from these results that the improvement in accuracy seen is not due to observers adopting a more lax criterion for reporting the presence of a target as there is either no decrease, or an improvement in accuracy for target absent trials; a decrease in accuracy on these trials would result from observers reporting the presence of a target when there was none (i.e. a false positive).

Four way analyses of variance (search task (2) - dark circles & light ellipses (training task), one of the other types of search task; time (3) - baseline, transfer, follow-up; target present/target absent (2); number of distractors (4) - 6, 12, 18, 24) with repeated measures on all factors were conducted to compare variation in accuracy over time for each search task with the training search task (dark circles & light ellipses). Examination of the interaction effects between time and search task revealed significant interactions for light ellipses, light ellipses & dark ellipses and dominant eye compared with the training search task. Table 5.9 shows F values, significance levels and mean scores (% correct) for each search task apart from the training search task. Each row of the table shows the F values, degrees of freedom and significance levels for the interaction effect involving the search task shown at
the start of the row and the training search task. The means represent the individual means for each search task.

Table 5.9 Results of analyses of variance: interaction between search task and time (means are collapsed across target present/target absent and number of distractors).

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>D.F.</th>
<th>p&lt;</th>
<th>Baseline mean</th>
<th>Transfer mean</th>
<th>Follow-up mean</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>dark circles</em></td>
<td>2.14</td>
<td>2,14</td>
<td>0.1550</td>
<td>92.03</td>
<td>90.94</td>
<td>93.59</td>
</tr>
<tr>
<td><em>light ellipses</em></td>
<td>17.87</td>
<td>2,14</td>
<td>0.0005</td>
<td>83.59</td>
<td>91.25</td>
<td>89.38</td>
</tr>
<tr>
<td><em>dark ellipses</em></td>
<td>1.79</td>
<td>2,14</td>
<td>0.2040</td>
<td>93.75</td>
<td>94.41</td>
<td>96.25</td>
</tr>
<tr>
<td><em>light ellipses &amp; dark ellipses</em></td>
<td>0.56</td>
<td>2,14</td>
<td>0.5860</td>
<td>92.50</td>
<td>93.13</td>
<td>93.91</td>
</tr>
<tr>
<td><em>dark circles &amp; light ellipses</em></td>
<td>5.17</td>
<td>2,14</td>
<td>0.0210</td>
<td>84.06</td>
<td>90.94</td>
<td>88.75</td>
</tr>
<tr>
<td><em>light ellipses &amp; dark ellipses</em></td>
<td>10.30</td>
<td>2,14</td>
<td>0.0020</td>
<td>86.09</td>
<td>86.25</td>
<td>90.94</td>
</tr>
<tr>
<td><em>lines</em></td>
<td>0.81</td>
<td>2,14</td>
<td>0.4670</td>
<td>88.75</td>
<td>90.75</td>
<td>89.22</td>
</tr>
</tbody>
</table>

* Training search task

It can be seen from table 5.9 that for the search tasks with light ellipses or light ellipses & dark ellipses as distractors, and for the search task completed with the dominant eye, observers started off performing at a lower level than they did on the training search task, but that by the transfer session in the case of light ellipses and light ellipses & dark ellipses, and by the follow-up session in the case of trials conducted with the dominant eye there was no longer any difference. In addition, there were three way interactions between time, search task and target present/target absent for three comparisons of the training search task with another type of search task. These were the training search task with dark circles (F(2,14)=11.68 p=0.001), the training search task with light ellipses (F(2,6)=6.14 p=0.035) and the training search task with dark ellipses (F(2,14)=5.74 p=0.015). The relevant means can be seen in table 5.10.
Table 5.10 Means scores for interaction effects between dark circles, light ellipses and dark ellipses, with the training search task, time and target present/target absent (means are collapsed across number of distractors).

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Transfer</th>
<th>Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>dark circles</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Target Present</td>
<td>86.25%</td>
<td>86.56%</td>
<td>93.13%</td>
</tr>
<tr>
<td>Target Absent</td>
<td>97.81%</td>
<td>95.31%</td>
<td>94.06%</td>
</tr>
<tr>
<td><strong>light ellipses</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Target Present</td>
<td>75.31%</td>
<td>84.38%</td>
<td>79.69%</td>
</tr>
<tr>
<td>Target Absent</td>
<td>91.88%</td>
<td>98.13%</td>
<td>99.06%</td>
</tr>
<tr>
<td><strong>dark ellipses</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Target Present</td>
<td>89.38%</td>
<td>93.13%</td>
<td>93.75%</td>
</tr>
<tr>
<td>Target Absent</td>
<td>98.13%</td>
<td>99.69%</td>
<td>98.75%</td>
</tr>
<tr>
<td><strong>dark circles &amp; light ellipses</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Target Present</td>
<td>84.06%</td>
<td>80.63%</td>
<td>85.63%</td>
</tr>
<tr>
<td>Target Absent</td>
<td>94.38%</td>
<td>96.25%</td>
<td>95.63%</td>
</tr>
</tbody>
</table>

* Training search task

The mean scores in table 5.10 suggest that whilst, for the training search task, there is little change over time for target absent trials and a decrease in accuracy immediately after training for target present trials, for the three different types of transfer search task there is an improvement on target present trials. It is interesting to note that the improvement occurs exclusively on those search tasks with a homogeneous distractor set.

**Baseline, transfer and follow-up sessions: reaction time data**

A three way analysis of variance (time (3) - baseline, transfer, follow-up; target present/target absent (2); number of distractors (4) - 6, 12, 18, 24) with repeated measures on all factors was conducted for each of the different search tasks. For all nine search tasks observers were significantly faster on target present trials than they were on target absent trials. The test statistics and mean scores are shown in table 5.11.
Table 5.11 Results of analyses of variance: main effect of target present/target absent for each search task (means are collapsed across time and number of distractors).

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>D.F.</th>
<th>p</th>
<th>Target present mean (msec)</th>
<th>Target absent mean (msec)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>dark circles</strong></td>
<td>7.33</td>
<td>1,7</td>
<td>0.001</td>
<td>909</td>
<td>1130</td>
</tr>
<tr>
<td><strong>light ellipses</strong></td>
<td>13.27</td>
<td>1,7</td>
<td>0.008</td>
<td>1666</td>
<td>2161</td>
</tr>
<tr>
<td><strong>dark ellipses</strong></td>
<td>6.21</td>
<td>1,7</td>
<td>0.042</td>
<td>857</td>
<td>1057</td>
</tr>
<tr>
<td><strong>dark circles &amp; dark ellipses</strong></td>
<td>9.48</td>
<td>1,7</td>
<td>0.018</td>
<td>920</td>
<td>1312</td>
</tr>
<tr>
<td><strong>light ellipses &amp; dark ellipses</strong></td>
<td>11.68</td>
<td>1,7</td>
<td>0.011</td>
<td>1418</td>
<td>1910</td>
</tr>
<tr>
<td><strong>dark circles &amp; light ellipses</strong></td>
<td>9.65</td>
<td>1,7</td>
<td>0.017</td>
<td>1530</td>
<td>2394</td>
</tr>
<tr>
<td><strong>dark circles &amp; light ellipses &amp; dark ellipses</strong></td>
<td>10.97</td>
<td>1,7</td>
<td>0.013</td>
<td>1334</td>
<td>1988</td>
</tr>
<tr>
<td><strong>dominant eye</strong></td>
<td>13.15</td>
<td>1,7</td>
<td>0.008</td>
<td>1532</td>
<td>1441</td>
</tr>
<tr>
<td><strong>lines</strong></td>
<td>15.01</td>
<td>1,7</td>
<td>0.006</td>
<td>2454</td>
<td>3234</td>
</tr>
</tbody>
</table>

* Training search task

Also for all nine search tasks observers reaction times varied significantly with the number of distractors. The means scores and tests statistics are shown in table 5.12.

Table 5.12 Results of analyses of variance: main effect of number of distractors for each search task (means are collapsed across time and target task).

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>D.F.</th>
<th>p</th>
<th>6</th>
<th>12</th>
<th>18</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>dark circles</strong></td>
<td>12.00</td>
<td>3,21</td>
<td>0.0005</td>
<td>909</td>
<td>1013</td>
<td>1048</td>
<td>1110</td>
</tr>
<tr>
<td><strong>light ellipses</strong></td>
<td>6.94</td>
<td>3,5</td>
<td>0.0310</td>
<td>1318</td>
<td>1805</td>
<td>2157</td>
<td>2375</td>
</tr>
<tr>
<td><strong>dark ellipses</strong></td>
<td>10.86</td>
<td>3,21</td>
<td>0.0005</td>
<td>809</td>
<td>932</td>
<td>1029</td>
<td>1075</td>
</tr>
<tr>
<td><strong>dark circles &amp; dark ellipses</strong></td>
<td>12.31</td>
<td>3,21</td>
<td>0.0005</td>
<td>965</td>
<td>1106</td>
<td>1129</td>
<td>1264</td>
</tr>
<tr>
<td><strong>light ellipses &amp; dark ellipses</strong></td>
<td>12.72</td>
<td>3,5</td>
<td>0.0090</td>
<td>1146</td>
<td>1553</td>
<td>1872</td>
<td>2085</td>
</tr>
<tr>
<td><strong>dark circles &amp; light ellipses</strong></td>
<td>6.74</td>
<td>3,5</td>
<td>0.0330</td>
<td>1299</td>
<td>1883</td>
<td>2170</td>
<td>2497</td>
</tr>
<tr>
<td><strong>dark circles &amp; light ellipses &amp; dark ellipses</strong></td>
<td>10.89</td>
<td>3,5</td>
<td>0.0120</td>
<td>1077</td>
<td>1523</td>
<td>1894</td>
<td>2150</td>
</tr>
<tr>
<td><strong>dominant eye</strong></td>
<td>8.45</td>
<td>3,5</td>
<td>0.0210</td>
<td>1298</td>
<td>1810</td>
<td>2195</td>
<td>2564</td>
</tr>
<tr>
<td><strong>lines</strong></td>
<td>12.00</td>
<td>3,5</td>
<td>0.0100</td>
<td>1794</td>
<td>2584</td>
<td>3228</td>
<td>3771</td>
</tr>
</tbody>
</table>

* Training search task

Polynomial contrasts revealed that for all nine types of search task there was a significant linear trend, such that, as the number of distractors increased so did observers' reaction times. The test statistics are shown in table 5.13.
Table 5.13 Results of polynomial contrasts: effect of number of distractors.

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>p≤</th>
</tr>
</thead>
<tbody>
<tr>
<td>dark circles</td>
<td>8.17</td>
<td>0.0001</td>
</tr>
<tr>
<td>light ellipses</td>
<td>5.01</td>
<td>0.0015</td>
</tr>
<tr>
<td>dark circles</td>
<td>4.44</td>
<td>0.0030</td>
</tr>
<tr>
<td>dark circles &amp; dark ellipses</td>
<td>4.75</td>
<td>0.0020</td>
</tr>
<tr>
<td>light ellipses &amp; dark ellipses</td>
<td>5.22</td>
<td>0.0010</td>
</tr>
<tr>
<td>dark circles &amp; light ellipses*</td>
<td>4.88</td>
<td>0.0020</td>
</tr>
<tr>
<td>dark circles &amp; light ellipses &amp; dark ellipses</td>
<td>6.05</td>
<td>0.0005</td>
</tr>
<tr>
<td>dominant eye</td>
<td>5.53</td>
<td>0.0010</td>
</tr>
<tr>
<td>lines</td>
<td>6.57</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

* Training search task

For four of the nine search tasks (light ellipses, light ellipses & dark ellipses, dark circles & light ellipses & dark ellipses and dominant eye) there was a significant two way interaction between the number of distractors and whether there was a target present or not. Test statistics and mean scores are shown in table 5.14.

Table 5.14 Results of analyses of variance: Interaction between number of distractors and target present(TP)/target absent(TA) for each search task (means are collapsed across time).

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>D.F.</th>
<th>p≤</th>
<th>6</th>
<th>12</th>
<th>18</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>dark circles</td>
<td>2.70</td>
<td>3,21</td>
<td>0.071</td>
<td>TA</td>
<td>1062</td>
<td>1143</td>
<td>1121</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TP</td>
<td>755</td>
<td>882</td>
<td>975</td>
</tr>
<tr>
<td>light ellipses</td>
<td>3.16</td>
<td>3,21</td>
<td>0.046</td>
<td>TA</td>
<td>1456</td>
<td>2080</td>
<td>2416</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TP</td>
<td>1179</td>
<td>1531</td>
<td>1897</td>
</tr>
<tr>
<td>dark ellipses</td>
<td>1.79</td>
<td>3,21</td>
<td>0.180</td>
<td>TA</td>
<td>896</td>
<td>1008</td>
<td>1177</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TP</td>
<td>723</td>
<td>856</td>
<td>881</td>
</tr>
<tr>
<td>dark circles &amp; dark ellipses</td>
<td>1.60</td>
<td>3,21</td>
<td>0.218</td>
<td>TA</td>
<td>1132</td>
<td>1344</td>
<td>1327</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TP</td>
<td>798</td>
<td>867</td>
<td>931</td>
</tr>
<tr>
<td>light ellipses &amp; dark ellipses</td>
<td>5.71</td>
<td>3,21</td>
<td>0.005</td>
<td>TA</td>
<td>1233</td>
<td>1856</td>
<td>2186</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TP</td>
<td>1059</td>
<td>1250</td>
<td>1558</td>
</tr>
<tr>
<td>dark circles &amp; light ellipses*</td>
<td>2.54</td>
<td>3,5</td>
<td>0.170</td>
<td>TA</td>
<td>1571</td>
<td>2287</td>
<td>2700</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TP</td>
<td>1027</td>
<td>1478</td>
<td>1639</td>
</tr>
<tr>
<td>dark circles &amp; light ellipses &amp; dark ellipses</td>
<td>4.14</td>
<td>3,21</td>
<td>0.019</td>
<td>TA</td>
<td>1242</td>
<td>1882</td>
<td>2269</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TP</td>
<td>911</td>
<td>1165</td>
<td>1519</td>
</tr>
<tr>
<td>dominant eye</td>
<td>3.83</td>
<td>3,21</td>
<td>0.025</td>
<td>TA</td>
<td>1530</td>
<td>2303</td>
<td>2669</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TP</td>
<td>1066</td>
<td>1317</td>
<td>1721</td>
</tr>
<tr>
<td>lines</td>
<td>1.83</td>
<td>3,21</td>
<td>0.173</td>
<td>TA</td>
<td>2062</td>
<td>3034</td>
<td>3653</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TP</td>
<td>1526</td>
<td>2133</td>
<td>2802</td>
</tr>
</tbody>
</table>

* Training search task
The mean scores in table 5.14 indicate that, for these four search tasks, as the number of distractors increased search times increased by a larger amount for target absent trials than they did for target present trials. This would suggest that observers examined these stimuli using a serial search process. The pattern of mean scores is also similar for all the other search tasks although obviously falling short of a significant interaction effect. If observers were examining the stimuli using a serial search process, as these data suggest and as the training data also suggested, then one would expect a ratio between the search slopes for target absent and target present trials of approximately 2:1. In order to examine this possibility in more detail across all types of trial, individual least square slope estimates were obtained for each observer for target present and target absent trials for the baseline, transfer and follow-up sessions. The slopes were calculated on values averaged across the different types of search task in order to provide a general indication of the search process adopted by observers, as the previous interaction effects and mean scores would indicate a similar process for all the search tasks. Following calculation of the slopes for each observer, for each session the ratio of the target absent to the target present slope was calculated. The mean of this value was 1.25 for the baseline session, 1.29 for the transfer session and 1.32 for the follow-up session. A t-test was conducted to test the hypothesis that the slope ratio for each day was equal to or greater than 2. This hypothesis was rejected for all three sessions (baseline session, t(7)=10.10 p<0.005; transfer session, t(7)=6.13 p<0.005; follow-up session t(7)=6.20 p<0.005). These results indicate that overall although search slopes are generally steeper for target absent trials observers are not exhibiting a true serial search pattern.

Observers' reaction times varied significantly over time for all the search tasks. F values, degrees of freedom, p values and mean scores are shown in table 5.15.
Table 5.15 Results of analyses of variance: main effect of change over time for each search task (means are collapsed across target present/target absent and number of distractors).

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>D.F.</th>
<th>( p \leq )</th>
<th>Baseline mean</th>
<th>Transfer mean</th>
<th>Follow-up mean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>dark circles</strong></td>
<td>11.71</td>
<td>2,14</td>
<td>0.0010</td>
<td>1251</td>
<td>886</td>
<td>922</td>
</tr>
<tr>
<td><strong>light ellipses</strong></td>
<td>8.57</td>
<td>2,14</td>
<td>0.0040</td>
<td>2380</td>
<td>1629</td>
<td>1732</td>
</tr>
<tr>
<td><strong>dark ellipses</strong></td>
<td>7.09</td>
<td>2,6</td>
<td>0.0260</td>
<td>1160</td>
<td>854</td>
<td>858</td>
</tr>
<tr>
<td><strong>dark circles &amp; dark ellipses</strong></td>
<td>7.04</td>
<td>2,6</td>
<td>0.0270</td>
<td>1418</td>
<td>965</td>
<td>964</td>
</tr>
<tr>
<td><strong>light ellipses &amp; dark ellipses</strong></td>
<td>13.77</td>
<td>2,14</td>
<td>0.0005</td>
<td>2075</td>
<td>1458</td>
<td>1459</td>
</tr>
<tr>
<td><strong>dark circles &amp; light ellipses</strong></td>
<td>9.06</td>
<td>2,14</td>
<td>0.0030</td>
<td>2470</td>
<td>1715</td>
<td>1700</td>
</tr>
<tr>
<td><strong>dark circles &amp; light ellipses &amp; dark ellipses</strong></td>
<td>7.79</td>
<td>2,6</td>
<td>0.0220</td>
<td>2120</td>
<td>1449</td>
<td>1414</td>
</tr>
<tr>
<td><strong>dominant eye</strong></td>
<td>12.35</td>
<td>2,14</td>
<td>0.0010</td>
<td>2418</td>
<td>1757</td>
<td>1726</td>
</tr>
<tr>
<td><strong>lines</strong></td>
<td>12.28</td>
<td>2,14</td>
<td>0.0010</td>
<td>3261</td>
<td>2692</td>
<td>2580</td>
</tr>
</tbody>
</table>

* Training search task

Post hoc testing using Scheffe's test revealed that for all of the search tasks there was a significant decrease in reaction times between the baseline and transfer sessions (\( p < 0.01 \)) but that there was no change in reaction times between the transfer and the follow-up sessions (\( p > 0.05 \)).

Four way analyses of variance (search task (2) - **dark circles & light ellipses** (training task), one of the other types of search task; time (3) - baseline, transfer, follow-up; target present/target absent (2); number of distractors (4) - 6, 12, 18, 24) with repeated measures on all factors were conducted to compare variation in accuracy over time for each search task with the training search task (**dark circles & light ellipses**). No significant interactions were found between change over time and search task and there were no significant higher order interactions. The test statistics and mean scores are shown in table 5.16. The search task shown in the first cell of each row in the table represents the search task which was paired with the training search task.
Table 5.16  Results of analyses of variance: Interaction between search task and change over time for each search task paired with training search task (means are collapsed across target present/target absent and number of distractors).

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>D.F.</th>
<th>( p )</th>
<th>Baseline mean</th>
<th>Transfer mean</th>
<th>Follow-up mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>dark circles</td>
<td>1.39</td>
<td>2,6</td>
<td>0.319</td>
<td>1251</td>
<td>886</td>
<td>922</td>
</tr>
<tr>
<td>light ellipses</td>
<td>1.83</td>
<td>2,14</td>
<td>0.197</td>
<td>2380</td>
<td>1629</td>
<td>1732</td>
</tr>
<tr>
<td>dark ellipses</td>
<td>2.26</td>
<td>2,6</td>
<td>0.186</td>
<td>1160</td>
<td>854</td>
<td>858</td>
</tr>
<tr>
<td>dark circles &amp; dark ellipses</td>
<td>0.87</td>
<td>2,6</td>
<td>0.466</td>
<td>1418</td>
<td>965</td>
<td>964</td>
</tr>
<tr>
<td>light ellipses &amp; dark ellipses</td>
<td>1.18</td>
<td>2,14</td>
<td>0.335</td>
<td>2075</td>
<td>1458</td>
<td>1459</td>
</tr>
<tr>
<td>dark circles &amp; light ellipses</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2470</td>
<td>1715</td>
<td>1700</td>
</tr>
<tr>
<td>dark circles &amp; light ellipses &amp; dark ellipses</td>
<td>0.41</td>
<td>2,6</td>
<td>0.682</td>
<td>2120</td>
<td>1449</td>
<td>1414</td>
</tr>
<tr>
<td>dominant eye</td>
<td>0.05</td>
<td>2,6</td>
<td>0.954</td>
<td>2418</td>
<td>1757</td>
<td>1726</td>
</tr>
<tr>
<td>lines</td>
<td>0.94</td>
<td>2,6</td>
<td>0.440</td>
<td>3261</td>
<td>2692</td>
<td>2580</td>
</tr>
</tbody>
</table>

* Training search task

The results shown in table 5.16 indicate that the magnitude of the decrease in reaction times did not differ significantly across the different types of search task.

5.9 Discussion

The present experiment was designed to examine in more detail the nature of learning in visual search and to investigate the possibility that such learning could sometimes underlie improvements in discrimination performance that had been attributed to perceptual learning. The results indicated that with extensive practice observers became faster at conducting a search task for a known stimulus with no resulting loss of accuracy. This improvement in search speed transferred to a novel dimensions search task suggesting a general learning process. Close examination of the ratio of search slopes for target present to target absent trials suggested that in contrast to results reported previously (e.g. Sireteanu and Rettenbach, 1994), the search process did not move from a spatially serial to a spatially parallel search process and in fact some data from the training portion of the study suggested that search became increasingly serial with practice. In addition, for some types of
search task which had the same target as the training search task there was an improvement in the accuracy of target detection.

5.9.1 Improvements in search speed. 1) Specificity of learning

The improvement in search speed observed, in the current experiment, for the training search task, transferred fully to all the other search tasks. Transfer of learning to the novel dimensions stimuli (line stimuli) indicated a learning effect which was not specific to the stimulus features with which observers trained. This result is consistent with that reported by Sireteanu and Rettenbach (1994). It suggests that stimulus specific effects such as the formation of new pre-attentive detectors, or an increase in target-distractor discriminability through processes such as those observed by Goldstone (1994), or strengthening of associations such as suggested by Treisman, Vieira and Hayes (1992), do not provide an explanation for the non-specific decrease in reaction times observed in the current study. Further, the transfer of learning to the dominant eye is consistent with the idea that the improvement in performance was not contingent upon some form of stimulus specific learning and was not localised at a very low-level of visual processing.

Before accepting that there was a non-specific learning effect there are two alternative explanations for the wide transfer of learning observed. First of all, it could be argued that the experience observers had with the search tasks during the baseline and transfer measurement sessions constituted a certain amount of training and that this led to the general decrease in reaction times seen. However, this explanation can be ruled out since the reduction in search times for these 'untrained' search tasks was as large as for the training search task, whilst examination of the results suggests that the size of the reduction in search times was contingent upon amount of practice. That this is the case can be deduced from two findings. First, examination of the decrease in search times for the training search task showed a consistent trend for a decrease following each training session with no sign of search
times reaching an asymptotic value by the end of training. Second, for all types of search task there was virtually no change in search times between the transfer and follow-up sessions during which time observers had no further experience with any of the stimuli used in the experiments. The second alternative explanation for the general reduction in search times is that it arose as a result of observers improving on mechanical aspects of the task (e.g. faster button pushing). This explanation would seem unlikely as observers showed a consistent trend for search times to decrease with practice; improvements in mechanical ability usually occur early on in the learning of a new task and show a rapid tendency towards asymptote (Blank, 1934 cited in Woodworth and Schlosberg, 1938 p. 35), no such tendency was observed in the present study.

5.9.2 Improvements in search speed. 2) General changes in search process
Sireteanu and Rettenbach (1994) reported a study in which observers learned, with practice, to search for a known target in parallel. This learning transferred to other stimuli with different elements. The present study using a similar paradigm has not completely replicated this finding since although observers' search times became quicker and this improvement transferred to other search tasks this improvement did not occur as the result of a change from a serial to a parallel search process. Instead the data tend to suggest that observers adopted a consistent search process during the course of the experiment which reflected neither the pattern usually seen with a parallel nor with a serial search process. This finding is in accord with the theory outlined by Duncan and Humphreys (1992) which dispensed with the serial vs. parallel search dichotomy.

One possible explanation for the search process observed in the present study derives from proposals made by Bundesen (1990) whose mathematical formulation of visual search processes maps onto the theory of Duncan and Humphreys (1992). He suggested that in a conjunction or in a feature search task, observers can select
and test elements in the display until either a target has been found or all the
elements have been selected and tested. For an \( N \) element display this can be
accomplished in \( k \) views where \( 1 \leq k \leq N \). Hence, a view is defined as the selection
and sampling of one or more elements (no allowance is made in this theory for
sequentially overlapping views). With a multi-view model of the search process
such as this observers must choose the number of views taken to search a display on
the basis of a trade-off between the time taken to shift attention or make an eye
movement from one group of elements to the next, and the improvement in
sampling rate due to improved spatial integration of feature-values once a shift has
been made. The total time taken to search a display is consequently a function of
these two factors and an increase in reaction times with increasing display sizes
would be expected. The data of the present study reflected such a pattern. The
general improvement in reaction times seen in the present study can be understood
within the framework of this model as resulting from optimisation of this trade-off.
Further, it may be that with practice observers were able to search a larger number
of distractors with no loss of spatial resolution. This would mean that fewer views
were required to search a display of the same size resulting in faster reaction times.
A general learning process such as this provides the most plausible explanation for
the wide transfer of learning observed in the present study.

5.9.3 Improvements in search accuracy

The present study provided some evidence for improvements in accuracy following
training. These improvements in accuracy appeared to be specific to the feature
search conditions and in some cases to target present trials. No improvement was
seen in the accuracy of search on the novel dimensions search task. Two factors
can be proposed to account for these results.

The training task involved a discrimination which could only be made after
consideration of two dimensions. This practice may have led to increased
discriminability at each of the category boundaries (light vs. dark and circle vs. ellipse) in the same way as was reported by Goldstone (1994). However, in the present task any improvement in discriminability between, say, dark and light may have led to stronger perceptual grouping of the light features. Thus, whilst there would have been an increase in target-distractor discriminability making elements match an input template less well, there would also be an increase in the perceptual grouping of light elements leading to an increase in the effectiveness of spreading suppression across these elements and, thus, a greater chance of a target being wrongly rejected with distractors. Such a mechanism may result in little change in accuracy of search for conjunction search tasks but an improvement in accuracy for feature search tasks where a target can be identified on the basis of one dimension. The present results are generally consistent with this explanation. However, in a number of cases improvement was also specific to target present trials. Additional factors must be considered to account for this finding.

In order to understand why improvement specific to target present trials was observed it is first of interest to consider some of the more general findings on the effect on performance of the number of distractors, and of target present vs. target absent trials.

Looking first at the increased accuracy observed on trials with fewer distractors, if we return to the model of search proposed earlier and developed from Bundesen (1990), he views visual search as a processing race between possible perceptual categorisations towards visual short-term memory (VSTM) where attention is then engaged. The efficiency of this race varies as a function of factors such as element visibility and target-distractor discriminability. Let us suppose that for a target there is a certain probability $p_t$ of an observer failing to detect that target (a 'miss'). This probability will be a function of factors such as target visibility and target-distractor discriminability. Further, for each distractor let us suppose that there is a
certain probability $q_1$ that it will be wrongly classified as a target ('false alarm') which is a function of distractor visibility and target-distractor discriminability. Then for a given search task the probability of missing a target is given by $tp_1$ where $t$ is the number of targets and the probability of making a false alarm is $(n - t)q_1$ where $n$ is the total number of elements in the display. Thus, the probability of making an error overall is $tp_1 + (n - t)q_1$. From this it follows that the greater the number of elements the greater the chance that an error will be made. The results from the present study support this possibility as observers were indeed less accurate on trials with a greater number of distractors.

Next considering the finding that observers made fewer errors on target absent trials, this can be understood as resulting from two factors. First, if an observer completes the search of a display and has not seen a target then a negative response will be made by default. Second, in the present experiment targets were deliberately selected to be less visible than distractors (i.e. they took on 'standard' values; cf. Treisman and Gormican, 1988) and consequently $p_1 > q_1$. If observers double check their categorisations then the probability of making a false alarm on both occasions is $(q_1)^2$, whilst the probability of mis-categorising a target twice is $(p_1)^2$ and since $p_1 > q_1$ then the size of the ratio between the two probabilities increases. Thus, as the number of checks made increases the relative ratio of misses to false alarms will also increase although of course the absolute number of both will decrease.

In addition to better search accuracy on target absent trials and trials with fewer distractors, an interaction between these two factors was observed. This can be understood as follows. If we assume that target visibility is partially determined by the strength of the sensory evidence that a target is present as suggested by Treisman and Gormican (1987), then this will vary as a function of the ratio of the number of distractors to the number of targets. If the number of distractors
increases whilst the number of targets remains constant, then the sensory response to these distractors will increase, leading to the relative proportions of the total sensory response to target and distractors changing in favour of the distractors, thus, making the target less visible overall. From this it follows, that as the number of distractors increases, the probability of a miss also increases, leading to an increase in errors on target present trials, but not on target absent trials. The interaction between target present/target absent and the number of distractors agrees with this explanation.

Finally, returning to the question of why improvement specific to target present trials was observed, it can be seen that as a result of the efficiency of double checking for false alarms, and the effect of default responding, few errors would be expected on target absent trials, and thus, as observed, performance would be expected to approach a ceiling level. For target present trials, it can be seen that if target visibility improves as a result of training, then the probability of missing a target will decrease, thus, leading to an improvement in accuracy on target present trials such as that observed in the present study.

5.9.4 Summary and implications of results for discrimination learning

The results of the present study provide a limited amount of evidence that a change in search strategy involving more efficient allocation of visual attention, rather than a change in search process (e.g. from serial to parallel), may lead to improvements in visual search speed, such as was observed in the present experiment and in X-ray experiments one and two. Improvements in accuracy in visual search tasks, such as were observed in the present experiment, are more likely to result from a change in target-distractor discriminability; this suggestion is mathematically tenable and supported by the finding that there was no change in accuracy for the novel dimensions search task. An improvement in target-distractor discriminability is unlikely to provide a full explanation for the improvement in search speeds.
observed, since there was complete transfer of this learning to a novel dimensions search task (line stimuli). It may be that any specific learning effects for search speeds were masked by the more general learning process which occurred.

The fact that the present study has demonstrated that, in addition to improvements in search speed, improvements in search accuracy can be made, supports the proposition that such learning may then have an impact on discrimination performance. It is likely that in a discrimination task, as a consequence of the features becoming more readily detected, fewer discrimination failures resulting from missed features would be made, and also that the features would become more perceptually salient leading to the more ready learning of novel discrimination tasks. This proposition is supported by the results of Sowden, Davies and Chivers (1992) who found that whilst discrimination tasks based upon highly salient features, such as features which popped out, could be easily learned, those based on the discovery of less salient, prototypical features were not learned at all during the course of a short practice procedure.
Chapter six
General Discussion

The previous chapters have suggested a new theoretical framework under which perceptual learning can be studied, and have examined the range of learning processes distinguished within that framework. The present chapter briefly summarises and reviews this framework, and the findings of the empirical studies which have been conducted, before going on to propose a re-integration of the processes which have been distinguished, in a model of discrimination learning. Following the development of this model, the importance of distinguishing the various processes which can underlie improvements in discrimination at a practical level is discussed using the real world task of mammographic screening as an illustration.

6.1 Review of thesis

This thesis has examined the phenomena previously referred to as perceptual learning. Until now this term had been used to refer to a variety of very different learning processes. In the first chapter of this thesis, a redefinition of perceptual learning was proposed such that it was considered to be 'perceiving variables of physical stimulation not previously sensed or perceived'. An examination of previous research on perceptual learning indicated that part of the reason for the over generalisation in the use of this term was the oft made assumption that an improvement in performance on a discrimination task necessarily indicated perceptual learning. Further examination of previous research on perceptual learning led to the development of a taxonomy of studies previously referred to as perceptual learning. It was suggested that whilst all of the processes identified in the taxonomy could underlie improvements in discrimination, only some of them could still be regarded as perceptual learning. The empirical work of this thesis went on to examine the learning processes distinguished in
the taxonomy in order to examine further their nature, to investigate ways in which their operation could be indicated, and to evaluate evidence for the distinctions made. This empirical work revealed a number of important points about the operation of these learning processes. First, it was illustrated that the results of some previous studies, which had been proposed to show perceptual learning could, under the newly established theoretical framework, be better described as learning resulting from the operation of hypothesis generation, rule testing and concept formation. Second, a method that could reveal the necessity of perceptual re-organisation for a discrimination task, or that could indicate when perceptual re-organisation had occurred, was developed. Third, examination of learning for stereo depth discrimination and contrast sensitivity indicated ways in which the operation of low-level learning could be shown, and revealed new evidence for plasticity of the visual system on the dimension of contrast detection. Fourth, investigation of learning in a naturalistic task requiring detection of features in X-rays, indicated that, whilst improvement in performance occurred, this was probably not localised at a low-level. Instead it was proposed that some type of search learning had occurred. Consequently, the final experiment of this thesis made a preliminary investigation of the type of learning which occurred in a standard visual search task. Although the empirical findings of this thesis have been specifically discussed in their respective chapters, a brief consideration of the implications of their results will be made here.

6.2 Hypothesis generation, rule testing and concept formation

A great deal of early research on perceptual learning required observers to learn simple visual discrimination tasks. A much referenced example of this type of work was conducted by Gibson and Gibson (1955). However, their task placed a substantial memory load on observers which was likely to limit performance. A replication and extension of the Gibsons' experiment indicated that when this memory load was
substantially reduced, observers could perform the task without difficulty. On the basis of this finding, it would seem difficult to argue that the Gibsons had demonstrated a perceptual learning process (i.e. that observers were *perceiving variables of physical stimulation not previously sensed or perceived*). However, what was almost certainly indicated by the Gibsons' task was the operation of a fundamental discrimination learning process, that of hypothesis generation, rule testing and concept formation; the learning of any discrimination task will require the formation of a discrimination rule which can be derived through a process of hypothesis generation and rule testing. The operation of this process may lead to the formation of new conceptual knowledge. This knowledge could be of two kinds, either visual schemata, or propositional knowledge, including knowledge about features and knowledge of discrimination rules (an observer may of course form both kinds of conceptual knowledge when learning a new discrimination task). Since deriving a successful rule may take a number of attempts, particularly in a task such as the Gibsons' where an observer only has a brief chance to examine the target stimulus before each run through, these tasks may mimic some aspects of perceptual learning. However, as has been demonstrated, careful examination of these tasks reveals that the learning they require is of a very different nature. Thus, the important point made is that when examining improvement on discrimination tasks, learning due to the discovery of discrimination rules based upon already perceptible features should be separated out. This should be the first step towards discovering whether the task also requires other learning processes. An example of a case in which such further examination was required was the task of chick sexing.

6.3 Perceptual re-organisation

Biederman and Shiffrar (1987) examined the difficult skill of chick-sexing. The learning of this skill had previously been considered an example of perceptual learning.
Following an expert systems analysis, they were able to extract a simple discrimination rule upon which performance of the task could be based. When this rule was communicated to naive observers their performance changed such that it closely mimicked that of experts. This finding suggested that the main difficulty of learning the task was not that it required perceptual learning but that observers had to hit upon the correct discrimination rule from amongst the large number of potential discrimination rules which could be tested. However, Biederman and Shiffrar (1987) did not investigate an alternative possibility, which was that their instructions could also have encouraged observers to carry out some sort of perceptual re-organisation of the visual input. Such re-organisation could have caused observers to perceive new structures in the visual input composed of features that had previously been segmented as parts of other perceptual groups; perceptual re-organisation can be considered to be one form of perceptual learning under the definition used in this thesis. The degree to which perceptual re-organisation was required for the learning of the task was investigated in an elaboration of Biederman and Shiffrars' (1987) experiment through the use of a parsing technique. Observers were required to examine and draw the photographs used in the original experiment feature by feature, starting with the most distinctive feature first. The decision to use a drawing method to examine observers' perceptual groupings was based upon the findings of research on knowledge elicitation which has indicated that, in the case of eliciting knowledge regarding performance on visual tasks, language based techniques such as verbal protocol analysis do not adequately tap into underlying perceptual processes (Ericsson and Simon, 1984). The specific technique used was developed from the work of others (e.g. Hock, Webb and Cavedo, 1987; Rey, 1941 cited in Lezak, 1983, pp. 395-402). The results of the elaboration indicated that, in the case of chick sexing, perceptual re-organisation was not generally required, although it was in some instances.
The elaboration of Biederman and Shiffrar (1987) demonstrated an important second step in separating out the processes involved in the learning of any complex visual discrimination task. That is, preliminary determination of whether learning the task may require perceptual re-organisation, or even low-level perceptual learning; in addition to examining observers' perceptual grouping, this method can indicate whether observers detect the discriminative feature at all. Even when the discriminative feature is not known to the experimenter, a parsing technique could still be used to examine whether perceptual re-organisation occurred during the course of learning a discrimination task, by requiring observers to parse the stimuli before and after learning. Any changes in element grouping following learning, as revealed through observers' parsing, might indicate the operation of perceptual re-organisation.

6.4 Low-level learning

If the features required to make a discrimination cannot be detected at all, then low-level learning will be required before a successful discrimination can be accomplished. Thus, this type of learning may also underlie improvements in performance on some discrimination tasks. Recent research has indicated that our sensitivity to basic visual features can be improved through experience (see section 1.5.5). This type of learning is perhaps the best example of true perceptual learning, as defined in this thesis. The experiments reported in chapter three further examined low-level learning for stereoaucity and contrast sensitivity. In general, these experiments illustrated the use of transfer of learning paradigms that could indicate when learning was low-level in nature, and that further, could provide some indication of the likely neural locus of learning. The results of one study indicated that improvements in stereoaucity were not localised at a low-level, and it was suggested instead, that the specificity effects observed in previous studies on improvements in stereoaucity, may have resulted from the operation of selective spatial attention mechanisms. However, the results of studies
on contrast sensitivity indicated bi-directional plasticity for contrast sensitivity localised at an early stage of visual processing. This finding is of the utmost importance for many basic visual discrimination tasks, since the detection of contrast boundaries represents possibly the most fundamental aspect of visual analysis. Overall, these findings provided further support for the accumulating evidence that adult vision is considerably more plastic than had previously been supposed.

Following the findings on low-level learning, two studies, reported in chapter four, were conducted to investigate the types of learning which occurred in the naturalistic task of detecting features in X-rays. It might have been expected that, given the plasticity for contrast sensitivity observed in chapter three, some of the improvement on this contrast sensitivity dependent task would be low-level in nature. However, the results of these studies indicated that although observers improved their detection performance this was not contingent upon low-level learning (the implications of this finding are discussed in section 6.9). Instead, some type of search learning seemed to have occurred, and this contributed to the design of the final experiment which examined learning in a basic visual search task.

6.5 Search learning

Learning in visual search tasks has been investigated in a variety of different contexts ranging from investigations of whether a serial search process can become parallel with practice (e.g. Treisman, Vieira and Hayes, 1992; Sireteanu and Rettenbach, 1994) through to investigation of the specificity of learning for texture segregation tasks (e.g. Karni and Sagi, 1991, 1993; Ahissar and Hochstein, 1993). As such, this taxonomic category may represent something of a mixed bag of processes, some of which are localised at a low-level of visual processing (e.g. Karni and Sagi, 1991, 1993) and some of which are localised at higher levels of processing (e.g. Sireteanu and
All of these processes have in common the fact that they indicate improvements in search speed on tasks where the required detection can already be made (i.e. it is above the appropriate threshold). Chapter five of this thesis made a preliminary investigation of learning in visual search tasks. The results indicated that an improvement in search speeds occurred, but that unlike Sireteanu and Rettenbach (1994) and in agreement with Treisman, Vieira and Hayes (1992) search never became truly parallel. It was proposed that the general improvement in search speeds resulted from a general learning mechanism such as optimisation of search rules. In addition to improvements in search speed, an improvement in search accuracy was also observed. This result supported the proposition that changes in search process could also underlie improvements in discrimination performance, which had been attributed to perceptual learning.

6.6 Implications for the taxonomy of studies of perceptual learning

Although the empirical work of this thesis has supported the operational distinctions made, it is not argued that the taxonomy proposed is necessarily the only classificatory scheme which could be adopted. For instance, it could be argued that the category which contains low-level learning studies should be sub-divided into studies of fast learning and studies of slow learning. When considering further sub-division, future work needs to establish whether it really is important to separate out the additional processes. At the extreme, one could continue to form new taxonomic categories until there was one for each learning phenomenon studied. However, this would considerably complicate matters for no real gain in terms of the theoretical framework. The important point is to form categories at a level where the phenomena placed into them are induced by equivalent circumstances, have a similar time course, durability, and likely locus. At this level, useful distinctions can be made, and implications for
training people to perform different types of discrimination task are the most readily apparent.

It could also be argued that some studies should be allocated to a different taxonomic category altogether. An example of this possibility concerns the studies considered to show search learning. It could be that the same detectors which show improvements in sensitivity, as revealed in low-level learning studies, are also those which are used in visual search. For example, Fahle (1994) indicated both that improvements in vernier discrimination could occur as the result of low-level learning, and that the detection of vernier breaks could be accomplished by pre-attentive processes. The use of common detectors for these different types of task could account for the similar findings regarding specificity of learning which have been observed: for instance, the similar types of specificity observed in search learning studies (e.g. Karni and Sagi, 1991, 1993) and low-level learning studies (e.g. Fahle and Edelman, 1993). If it was the case that these different types of learning occurred in the same populations of neural detectors, then it might be argued that these studies should be considered to show learning which in a fundamental sense is contingent on the same process, even though the nature of the learning in search learning tasks is qualitatively different to that observed in low-level learning studies. It would be of interest to examine this possibility in future research perhaps by examining transfer of learning from threshold tasks to supra-threshold tasks performed on the same visual dimension. If the same populations of detectors and in some sense a common learning process were involved in both types of task, then we might expect that improvements in sensitivity at threshold could translate into improvements in detection speed for supra-threshold stimuli.
6.7 Implications for a general theory of perceptual learning

As has been discussed, it had often been assumed that improvements in discrimination necessarily indicated perceptual learning. When considering improvements in performance at this surface level, it might have seemed that a unitary learning process was in operation, and thus, that a general theory of perceptual learning, such as that attempted by Gibson (1967), was possible. However, under the definition of perceptual learning adopted in this thesis, only some studies can still be considered to indicate perceptual learning, and even these can be subdivided as illustrating various distinct processes (e.g. perceptual re-organisation vs. low-level learning). Hence, in the light of the theoretical framework established in this thesis, a general theory of perceptual learning would now seem to be an unrealistic goal. Instead, it is suggested that a more realistic and useful aim is the development of a more precise understanding of how each of the different learning processes identified occurs. In some instances this understanding may have to be developed at the level of the individual phenomena (for instance, improvements in stereoacuity may rest upon a unique learning process) and indeed some researchers have been working towards this goal (e.g. the work on vernier hyperacuity; Poggio, Edelman and Fahle, 1992; Poggio, Fahle and Edelman, 1992; Fahle, Edelman and Poggio, 1995). By working at this more precise level, a richer understanding of the complexity of human perception and perceptual learning can be achieved. However, although it is no longer possible to derive a general theory of perceptual learning, the processes which have been distinguished can be integrated into a model of discrimination learning. The nature of such a model has already been strongly implicated in the previous review of the empirical work conducted in this thesis. However, it is of use to make explicit the way in which the different processes may each underlie improvements in discrimination performance. An understanding of the many sources from which improvements in discrimination can arise will have
important implications for training individuals to perform complex discrimination tasks as discussed in sections 6.9 and 6.10.

6.8 Modelling discrimination learning

In the work of this thesis it has been argued that the use of the single term 'perceptual learning' to refer to all instances of improvement in discrimination has been misleading on two counts. First, rather than one process, a number of learning processes have underlain the improvements in discrimination. Second, not all of these processes indicate the observer learning to perceive anything; i.e. they are not perceptual learning. However, the distinguishing of these processes is not intended to suggest that they should be viewed in isolation. Although separate processes can be identified, they can all underlie improvements in discrimination, and they are likely to interact when an observer is required to perform a discrimination/detection task (discrimination necessarily involves detection of relevant features). It is important to understand how this interaction might occur since, by modelling discrimination learning, we can begin to distinguish between the possible sources of improvement in a discrimination, and from an applied point of view, better train an individual to make that discrimination. Figure 6.1 presents an outline of the way in which the learning processes may interact when an observer has to conduct a discrimination task.

For an observer to complete successfully a visual discrimination task, there are a number of critical hurdles which they must pass. In order to do so, various forms of learning may be required (sites of the different types of learning are indicated by the numbered boxes in Figure 6.1). First, the observer must develop a successful discrimination rule through a process of hypotheses generation and rule testing as indicated in the studies reported in chapter two.
Figure 6.1 Diagram to show relationship between processes involved in conducting a visual discrimination task
Hypotheses could be derived either externally from the visual stimulus via a preliminary search (involving a first glimpse and possibly an attentive search (see Gale, 1992, p. 120)) or internally from prior knowledge. The latter could have been acquired from instruction/training or from previous experience with similar tasks (for instance, the observer may hold relevant visual schemata). The process of feature selection, hypothesis generation and rule testing is likely to feed into the refinement of existing schemata or lead to the generation of new ones (concept formation). Thus, the first type of learning which may be involved in a discrimination task is discovery of a successful discrimination rule, and concept formation (1). Second, the feature/s on which a successful discrimination can be based must be perceptible (i.e. above a theoretical detection threshold). If the features are not perceptible then a discrimination will not be possible unless low-level perceptual learning occurs (2). The operation of low-level learning processes was indicated in the studies reported in chapter three. Third, if the discriminative feature/s are perceptible then they must also form a good group (i.e. they must constitute a distinct figure). If they do not then some form of learning in the shape of perceptual re-organisation (3) will be required before a discrimination can successfully be accomplished. The elaboration of the experiment conducted by Biederman and Shiffrar (1987), reported in chapter two, indicated some scope for this type of learning in the task of chick-sexing and illustrated a method for identifying when it is required. Fourth, the discriminative feature must be successfully detected by visual search processes. If the feature does not pop-out, then it must be detected by an attentive search, and on some occasions, the feature will be missed (external factors such as background noise and internal factors such as the observer's expectations and motivation will have an impact on the probability of a miss). Repeated searching for specific features may lead to search learning (4) such that the features increase in distinctiveness, and thus, are less likely to be missed; evidence for this possibility was reported in chapter five. Features may even come to be detected
automatically and to pop-out. Finally, a decision making process must be engaged in which can result in a number of outcomes including a decision to generate additional hypotheses and select further features for examination, a decision that a non-target has been mistakenly identified as a target, a decision that a discriminative feature is present, or a decision that a discriminative feature is not present. Detection of a discriminative feature and the consequent decision of difference can have two possible outcomes, either a correct detection (true positive) or a false positive. Similarly, non-detection of a discriminative feature and a consequent decision that two stimuli are the same, can have two possible outcomes, either a correct rejection or a miss of a target.

The three learning processes of low-level learning, perceptual re-organisation and search learning may lead to further rule testing and concept formation (1) by feeding into the hypothesis generation process (the observer will be able to generate new hypotheses around the newly discovered features) and also into the knowledge store (the newly identified features may lead to the refinement of existing schemata).

Thus, in summary, conducting a discrimination task can be seen as an iterative process which may involve one or all of the different types of learning process distinguished in the taxonomy of learning and investigated in the work of this thesis. In some cases it may be that only identification of a suitable rule or an improvement in decision making processes is required for a discrimination to be successfully completed; in these cases perceptual learning is not required for a discrimination to be learned. In other cases, one or all of low-level learning, perceptual re-organisation, and search learning may be required before a suitable rule can even be formulated, and the former two are considered in the present thesis to be indicative of perceptual learning. Given the model of learning a complex discrimination task developed here, it is now of interest to
examine the implications that this view has for the learning of real world discrimination tasks, in which different processes can be involved in improvements in discrimination.

6.9 Implications for examining the learning of ‘real world’ discriminations

It might be noted that, at an operational level, the paradigms used to examine the learning processes distinguished in this thesis are generally fundamentally different, and that it is thus quite a simple matter to distinguish between the different types of learning. Indeed it could be further argued that the different types of learning observed are almost certainly a function of the type of learning which can be induced by the particular task with which an observer is presented. This raises the important question of to what extent is it likely that in a real world discrimination task, which is not so tightly constrained, could these different types of learning still be expected to occur? The findings of the two X-ray experiments, in which, despite prior expectations (e.g. see Davies, Sowden, Hammond and Ansell, 1994), the pattern of results indicated that the learning observed did not occur at a low-level of visual processing, highlight this issue. In other words, although low-level learning can be induced and demonstrated to exist in a precisely controlled laboratory experiment to what extent can it be induced by the types of circumstances that occur in everyday adult life? This question is in many senses an empirical one to be answered by examining ‘real-world’ discrimination tasks. In such tasks the types of learning occurring are not of course readily apparent from the experimental paradigm. Thus, it is in the examination of such tasks that the techniques demonstrated in this thesis are of the utmost importance. If the learning of a discrimination task relies upon particular types of learning, then by identifying the types of learning required, individuals can be better trained to perform that task; components of training programmes can then be designed that are specifically targeted at inducing the different types of learning. Training programmes designed in this manner could have two main advantages. First, they could lead to improvements in
discrimination performance over and above that achieved as a result of traditional training approaches. This could occur because by having components of a training programme targeted at specific learning processes they are refined to a higher degree, or even because additional learning process are stimulated that would not have been induced by traditional training approaches, and that would not have occurred during the course of everyday experience. Second, training programmes designed to target specific learning processes are likely to facilitate the more rapid learning of a task. The above points will perhaps be best illustrated through a relatively brief examination of a complex real-world discrimination, that performed by the expert radiologist in mammographic screening programmes.

6.10 Implications for the learning of complex real-world discrimination tasks: the importance of task analysis in the design of more effective training programmes for mammographic film reading.

Mammographic film reading is a complex, real world, visual discrimination task the learning of which has often been regarded as, in part, an example of perceptual learning. The expert film reader is required to examine a pair of greyscale images (approximately 400 shades of grey, Pauli (1993), pp. 42), one for each breast, for the presence of certain types of feature which can be indicative of malignant disease. At a conceptual level these features are often considered to fall into a number of broad categories. These categories include micro-calcifications, which appear as white irregular flecks in the mammogram, masses, which appear as discrete areas of density greater than the background breast tissue pattern and which can be either regular or irregular in outline, distortions, which can appear as a disturbance in the typical structure of the breast, and asymmetries, which may be visible as a proliferation of white fluff in one breast as compared with the other. As can be appreciated from the model developed in section 6.8, a number of visual processes are likely to be involved in the detection of abnormal features in a mammogram. In the first instance, a film
reader obviously requires the conceptual knowledge about the types of features he/she is looking for. Second, the features must be of 'good form' in order for a film reader to detect them. Third, the features the film reader seeks to detect must also be above various detection thresholds for that reader. For instance, the contrast between a mass and the background must be above a film-reader's contrast threshold in order for the mass to be detected. Fourth, a film reader must also use visual search processes in order to detect features. Any failure of these search mechanisms will result in features indicative of malignant disease being missed. Finally, in addition to the operation of visual processes, once an abnormal feature has been detected the film reader is required to decide whether that feature requires further investigation (further discussion of decision making processes is beyond the scope of the present thesis).

Given the variety of processes identified as potentially involved in mammographic film reading there are obviously a number of potential sources for error, and a number of distinct processes which training needs to target in order to eliminate this error. The desirable goal of training is an effective, cost-efficient screening programme. In other words, one which maximises the number of detections whilst minimising the number of false positives. In order to achieve this goal it would seem logical to design training programmes which systematically target each of the different potential learning processes involved in film reading. At present, training programmes used in the U.K. do not follow such a systematic approach. As such it may be that certain types of learning which would be of benefit to the film reading process (e.g. low-level learning) are not even induced, or are induced to a less than optimum level.

Currently, typical training programmes involve the trainee spending two weeks at one of the four U.K. training centres. During this time trainees attend teaching sessions on specific features which indicate malignant disease and take part in all aspects of the
daily screening routine. The trainee first observes an expert and then themselves reads ‘rollers’ (a conveyor belt display device) of screening mammograms, discusses cases with the resident radiologists, and attends assessment clinics and pathology meetings. Trainees also spend a considerable amount of time engaged in home study examining teaching material such as Tabar’s Mammography Teaching Atlas (Tabar and Dean, 1983). Thus, training adopts what is essentially a tutorial approach, and the assumption is made that this small amount of training combined with further support during initial screening practise will lead to optimal performance. This assumption is questionable on many grounds. What follows is a description of one way in which a more systematic training programme might be designed, based on what is known about the different types of learning process identified in this thesis and which might be involved in the task of mammographic film reading.

The first question which arises concerns the broad conceptual categories for abnormal features that trainees are taught about (this corresponds to the development of prior knowledge and subsequent learning at stage 1 in Figure 6.1). For each mammogram a radiologist examines, he or she fills in a form indicating whether there is an abnormality present and if so what type of abnormality it is (the categories used on the forms vary across screening centres). Thus, from the outset, film-readers are required to consider abnormalities in terms of conceptual categories. Whilst it is not unlikely that these categories are derived from the natural structure present in mammograms and their features, this has not been systematically established and, further, perceived structure is likely to vary across individuals (see Garner, 1974). If the structure of the images, as perceived by novice film readers, does not map onto that used by radiologists in the categorising of abnormal features, then it may be that training programmes need to target specifically the establishment of these alternative perceptual groups and consequent conceptual categories (i.e. induce learning at stage 3 in Figure 217.
6.1). An examination of the way in which novice film-readers parse mammograms, using a method such as that outlined in chapter two of this thesis, would provide some information regarding their perceptual segmentation of the images, and thus, whether training should target perceptual grouping processes. In all probability it is this aspect of film reading that is best targeted by current approaches to training, which operate in a way not dissimilar to the approach employed by Biederman and Shiffrar (1987).

A second area which could be targeted in the design of training programmes, and which is almost certainly not adequately addressed by current approaches, is improving film readers detection of features through the inducement of low-level learning (i.e. learning at stage 2 in Figure 6.1). Most of the evidence on the occurrence of low-level learning suggests a process which occurs as a result of mass stimulus exposure (e.g. the experiments reported in chapter three); studies have typically revealed a slow but steady decrease in thresholds with repeated exposure (e.g. Fahle and Edelman, 1993). In routine mammography screening the film reader receives a very small amount of exposure to features indicative of malignant disease since the incidence of breast cancer in the screening population is only between 0.4% and 0.7% (Pauli, 1993, pp. 46). Since a large part of training in the U.K. is 'on the job' this means that a film reader typically gets little of the extensive and concentrated exposure which laboratory experiments suggest is necessary to induce low-level learning. Further, in the U.K., a relatively large number of cancers which are not detected during routine screening are detected during screening intervals (the 'interval' cancers) via self examination etc. It is possible that with inducement of low-level learning the incidence of interval cancers could be reduced through the film reader developing the ability to detect previously undetectable features. Consequently, it is suggested that training programmes might usefully recreate the circumstances conducive to low-level learning, namely mass exposure to features indicative of abnormality.
An additional issue, related to detection accuracy, concerns the provision of feedback on screening performance. The results of the X-ray experiments conducted in this thesis and the work of Roling (1995) suggests that, without the provision of suitable feedback, improvements in detection accuracy are unlikely to occur (it is unlikely that motivational incentives are lacking in the film reading context given the gravity of detection failures). In routine training and screening practise there is little or, in the latter case, no provision for regular assessment and individual feedback about screening performance. Thus, it would seem unlikely that training programmes and routine screening procedure provide adequate conditions for improvements in detection accuracy to occur. Whilst findings regarding the necessity of feedback in laboratory studies of low-level learning have been somewhat equivocal, feedback would seem to be necessary in the much noisier and more complex task of film reading. In addition to incorporating feedback into training programmes, it might be important to incorporate regular monitoring and feedback into routine screening practise.

A final learning process which is potentially important for film reading is search learning (i.e. learning at stage 4 in Figure 6.1). As has been discussed, improvements in visual search may lead to a reduction in the number of missed targets. Previous research has suggested that the expert film reader may detect some abnormalities via pre-attentive detection processes (e.g. Hendee, 1993, p. 140). However, it has not been established to what extent this reflects learning on the part of the expert and to what extent it reflects the operation of detection mechanisms for already existing perceptual primitives. Recent work by Mugglestone, Gale, Cowley and Wilson (1995) has indicated that detection performance with briefly presented mammograms is worse than when unlimited viewing is allowed, suggesting that, as might be expected, not all abnormalities are detected by pre-attentive processes. Chapter five of this thesis revealed that, as a result of practice, observers' detection accuracy for a basic visual
search task improved in addition to the expected improvements in search speed (cf. Treisman, Vieira and Hayes, 1992; Sireteanu and Rettenbach, 1994). Further, the amount of improvement observed was contingent on the amount of practice an observer had. This finding underlines the potential of introducing mass practice at detecting abnormalities into training programmes for mammographic screening. Further, as outlined in section 6.6, if at a fundamental level some types of search learning share a common process with low-level learning, then it might be expected that massed practice on detection of threshold level abnormalities would have additional benefits on the detection rate for supra-threshold abnormalities.

Thus, to summarise, in the design and implementation of U.K. training programmes, mammographic film reading has in many ways been treated as a unitary, global process for which a uni-dimensional approach to training could be adopted. However, the brief analysis conducted in the present section, set within the theoretical framework and model of discrimination learning developed in this thesis, and based upon previous research on mammographic film reading, would suggest that a multifaceted approach to training which systematically targeted each of the potential learning processes could better lead to optimal film reading performance.

In examining the specific task of mammographic film reading, the intention has been to illustrate the importance of separating out the different learning processes which could underlie improvements in discrimination performance at a practical level, as well as at the theoretical level, with which the majority of this thesis has been concerned. Not only is it important to realise, at a theoretical level, that many different processes have previously been considered simply to be examples of a single process which was referred to as perceptual learning, but it is important to distinguish these processes in...
real world tasks if performance on them is to be optimised through appropriate training.

6.11 Limitations and suggestions for future work

Specific limitations of the experiments conducted in this thesis and recommendations for related future work have been discussed in the experiments' respective chapters. However, there are two general points which can be made regarding the work of this thesis. First, as was discussed in section 6.6, the way in which studies have been classified together in this thesis has been based upon distinctions made at an operational level. It is likely that future work will suggest alternative classifications of individual studies, the formation of whole new categories based on the discovery of new learning processes and changes in the way in which some processes are conceptualised. On this latter point, as was suggested in section 6.6, one potentially fruitful route for investigation would be examination of whether there is transfer of learning from threshold to supra-threshold tasks and vice versa. Second, although this thesis has shown that perceptual learning as it had previously been considered represented a variety of different learning processes, this thesis has not investigated in any detail the precise mechanisms by which each occurs (i.e. the neural processes underlying low-level learning). Understanding these mechanisms is an important goal for future work and one which is currently being addressed in the work of others.
References


Canfield, T.H. (1941) Sex determination of day-old chicks II. Type variations. *Poultry Science*, 20, 327-328


Geiger, G. and Helmle, H. (1994) Forms are detected visually long before they are
recognised in random dot stereograms. Investigative Ophthalmology and Visual
Science, 35, 2113.

Perception, 11, 655-661.

Spatial Vision, 1, 103-112.


Gibson, E.J. (1953) Improvement in perceptual judgements as a function of
controlled practice or training. Psychological Bulletin, 50, 401-431.

Gibson, E.J. (1967) Principles of Perceptual Learning and Development. New
York: Appleton-Century-Crofts.

Houghton Mifflin.

presented patterns on learning to discriminate them. Journal of Comparative and
Physiological Psychology, 49, 239-242.

Gibson, J.J. and Gibson, E.J. (1955) Perceptual learning: differentiation or


Appendix A

Random-dot anaglyph, stereo stimulus

A circle should be seen floating in front of the background
Appendix B

Instructions for stereo experiment

In this experiment you will be required to make judgements about the relative depth of two squares. On the first and last days of the experiment you will complete 120 trials per day. On the intervening three days you will complete 200 trials a day.

To start the experiment press the spacebar. The screen will clear and a dot will be displayed in the centre of the screen. Above and below this dot there will be a green line. Make sure that when you stare at the dot the two lines appear to be in direct alignment with it and with each other. This indicates that you are properly fixated.

Once you are properly fixated on the dot press the spacebar again to start the first trial. The stimuli will flash up immediately. On each trial two squares will be displayed very rapidly. Both of the squares will be either above the fixation dot or below the fixation dot. The squares will always appear at a different depth to the fixation spot. It may take a lot of practice for you to see this. One of the squares will be closer to you than the other. Your task is to press a mouse button to indicate which square appears to be closest. Press the left mouse button if the left-hand square appears closest and the right mouse button if the right-hand square appears closest.

Make sure that you answer as quickly and as accurately as possible. Both your reaction time and whether you were correct or not are being recorded.

After completing the trial make sure you are properly fixated again. When you are sure that you are fixating correctly press the spacebar to begin the next trial.

The experiment will stop automatically when you have completed the correct number of trials.

It is very important that you are properly fixated before you push the spacebar to start a trial.

Please remember that both accuracy and speed are important. Obviously if both the squares appear to be at the same depth then you should guess. You will probably find that your ability to see the squares in depth improves over the course of the experiment.
## Appendix C

### Raw scores for observers in stereo study

Raw scores (% correct) across days, for observers in stereo study.

<table>
<thead>
<tr>
<th>Observer</th>
<th>Pre</th>
<th>Training 1</th>
<th>Training 2</th>
<th>Training 3</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>46.30</td>
<td>53.16</td>
<td>59.31</td>
<td>51.29</td>
<td>58.26</td>
</tr>
<tr>
<td>2</td>
<td>71.92</td>
<td>78.14</td>
<td>72.19</td>
<td>73.38</td>
<td>60.33</td>
</tr>
<tr>
<td>3</td>
<td>59.84</td>
<td>76.82</td>
<td>76.65</td>
<td>83.17</td>
<td>89.91</td>
</tr>
<tr>
<td>4</td>
<td>46.66</td>
<td>53.72</td>
<td>53.69</td>
<td>48.67</td>
<td>49.69</td>
</tr>
<tr>
<td>5</td>
<td>60.89</td>
<td>55.76</td>
<td>58.19</td>
<td>68.68</td>
<td>58.81</td>
</tr>
<tr>
<td>6</td>
<td>61.00</td>
<td>52.93</td>
<td>48.82</td>
<td>59.96</td>
<td>49.25</td>
</tr>
<tr>
<td>7</td>
<td>58.84</td>
<td>67.96</td>
<td>67.04</td>
<td>64.21</td>
<td>60.14</td>
</tr>
<tr>
<td>8</td>
<td>55.17</td>
<td>58.58</td>
<td>64.43</td>
<td>70.52</td>
<td>75.40</td>
</tr>
<tr>
<td>9</td>
<td>49.75</td>
<td>57.30</td>
<td>50.63</td>
<td>49.20</td>
<td>51.06</td>
</tr>
<tr>
<td>10</td>
<td>51.30</td>
<td>46.29</td>
<td>48.20</td>
<td>51.20</td>
<td>54.60</td>
</tr>
<tr>
<td>11</td>
<td>46.82</td>
<td>75.79</td>
<td>80.84</td>
<td>80.04</td>
<td>88.87</td>
</tr>
<tr>
<td>12</td>
<td>57.26</td>
<td>68.06</td>
<td>61.82</td>
<td>58.35</td>
<td>63.49</td>
</tr>
<tr>
<td>13</td>
<td>47.27</td>
<td>53.28</td>
<td>66.30</td>
<td>57.28</td>
<td>67.29</td>
</tr>
<tr>
<td>14</td>
<td>65.15</td>
<td>71.24</td>
<td>67.74</td>
<td>65.05</td>
<td>49.26</td>
</tr>
<tr>
<td>15</td>
<td>50.46</td>
<td>50.94</td>
<td>51.79</td>
<td>60.50</td>
<td>47.60</td>
</tr>
</tbody>
</table>

Raw scores (reaction times, msec) across days, for observers in stereo study.

<table>
<thead>
<tr>
<th>Observer</th>
<th>Pre</th>
<th>Training 1</th>
<th>Training 2</th>
<th>Training 3</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1007</td>
<td>1086</td>
<td>1121</td>
<td>782</td>
<td>787</td>
</tr>
<tr>
<td>2</td>
<td>561</td>
<td>535</td>
<td>429</td>
<td>466</td>
<td>410</td>
</tr>
<tr>
<td>3</td>
<td>1053</td>
<td>739</td>
<td>723</td>
<td>741</td>
<td>615</td>
</tr>
<tr>
<td>4</td>
<td>616</td>
<td>641</td>
<td>775</td>
<td>724</td>
<td>533</td>
</tr>
<tr>
<td>5</td>
<td>616</td>
<td>683</td>
<td>588</td>
<td>509</td>
<td>471</td>
</tr>
<tr>
<td>6</td>
<td>817</td>
<td>572</td>
<td>540</td>
<td>499</td>
<td>508</td>
</tr>
<tr>
<td>7</td>
<td>894</td>
<td>761</td>
<td>659</td>
<td>734</td>
<td>684</td>
</tr>
<tr>
<td>8</td>
<td>691</td>
<td>690</td>
<td>691</td>
<td>708</td>
<td>631</td>
</tr>
<tr>
<td>9</td>
<td>594</td>
<td>565</td>
<td>527</td>
<td>555</td>
<td>479</td>
</tr>
<tr>
<td>10</td>
<td>1056</td>
<td>822</td>
<td>890</td>
<td>991</td>
<td>847</td>
</tr>
<tr>
<td>11</td>
<td>1418</td>
<td>1194</td>
<td>1034</td>
<td>979</td>
<td>941</td>
</tr>
<tr>
<td>12</td>
<td>688</td>
<td>800</td>
<td>757</td>
<td>718</td>
<td>622</td>
</tr>
<tr>
<td>13</td>
<td>750</td>
<td>1549</td>
<td>1049</td>
<td>948</td>
<td>883</td>
</tr>
<tr>
<td>14</td>
<td>1202</td>
<td>1027</td>
<td>825</td>
<td>816</td>
<td>689</td>
</tr>
<tr>
<td>15</td>
<td>1105</td>
<td>965</td>
<td>783</td>
<td>749</td>
<td>738</td>
</tr>
</tbody>
</table>
Appendix D

Raw scores (contrast threshold %) for observers in contrast sensitivity experiment one

<table>
<thead>
<tr>
<th>Observer MS</th>
<th>Baseline</th>
<th>Transfer</th>
<th>19 days</th>
<th>50 days</th>
<th>80 days</th>
<th>195 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trained</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 cpd</td>
<td>2.17</td>
<td>1.60</td>
<td>1.55</td>
<td>1.26</td>
<td>1.35</td>
<td>1.25</td>
</tr>
<tr>
<td>untrained</td>
<td>1.88</td>
<td>1.70</td>
<td>-</td>
<td>-</td>
<td>1.80</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Observer PS</th>
<th>Baseline</th>
<th>Transfer</th>
<th>19 days</th>
<th>50 days</th>
<th>80 days</th>
<th>195 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trained</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 cpd</td>
<td>1.94</td>
<td>1.57</td>
<td>1.43</td>
<td>1.28</td>
<td>1.31</td>
<td>-</td>
</tr>
<tr>
<td>untrained</td>
<td>1.86</td>
<td>1.67</td>
<td>-</td>
<td>-</td>
<td>1.89</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Observer ID</th>
<th>Baseline</th>
<th>Transfer</th>
<th>19 days</th>
<th>50 days</th>
<th>80 days</th>
<th>195 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trained</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 cpd</td>
<td>5.49</td>
<td>6.12</td>
<td>5.01</td>
<td>4.06</td>
<td>4.33</td>
<td>-</td>
</tr>
<tr>
<td>untrained</td>
<td>3.59</td>
<td>4.23</td>
<td>-</td>
<td>-</td>
<td>4.19</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Observer GH</th>
<th>Baseline</th>
<th>Transfer</th>
<th>19 days</th>
<th>50 days</th>
<th>80 days</th>
<th>195 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trained</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 cpd</td>
<td>2.12</td>
<td>2.53</td>
<td>1.65</td>
<td>1.24</td>
<td>1.70</td>
<td>-</td>
</tr>
<tr>
<td>untrained</td>
<td>1.91</td>
<td>1.76</td>
<td>-</td>
<td>-</td>
<td>1.76</td>
<td>-</td>
</tr>
</tbody>
</table>
Appendix E

Raw scores (contrast threshold %) for observers in contrast sensitivity experiment three

### Observer JH

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Transfer</th>
<th>19 Day</th>
<th>50 Day</th>
<th>130 Days</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dominant eye</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>1.302</td>
<td>1.160</td>
<td>1.411</td>
<td>1.108</td>
<td>0.976</td>
</tr>
<tr>
<td>Transfer</td>
<td>1.095</td>
<td>0.880</td>
<td>1.460</td>
<td>1.347</td>
<td>1.287</td>
</tr>
<tr>
<td><strong>90°</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>0.660</td>
<td>1.121</td>
<td>1.034</td>
<td>1.363</td>
<td>0.890</td>
</tr>
<tr>
<td>Transfer</td>
<td>1.201</td>
<td>1.173</td>
<td>1.547</td>
<td>1.970</td>
<td>1.529</td>
</tr>
<tr>
<td><strong>135°</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>0.890</td>
<td>0.976</td>
<td>1.147</td>
<td>2.452</td>
<td>1.147</td>
</tr>
<tr>
<td>Transfer</td>
<td>1.529</td>
<td>1.776</td>
<td>1.494</td>
<td>1.411</td>
<td>1.287</td>
</tr>
<tr>
<td><strong>3 cpd</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>3.086</td>
<td>2.538</td>
<td>2.452</td>
<td>2.815</td>
<td>3.232</td>
</tr>
<tr>
<td>Transfer</td>
<td>1.860</td>
<td>2.751</td>
<td>1.477</td>
<td>2.111</td>
<td>2.016</td>
</tr>
<tr>
<td><strong>8 cpd</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>2.751</td>
<td>5.616</td>
<td>4.618</td>
<td>3.841</td>
<td>3.384</td>
</tr>
</tbody>
</table>

### Observer PTS

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Transfer</th>
<th>19 Day</th>
<th>50 Day</th>
<th>130 Days</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Right Eye</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>2.947</td>
<td>2.751</td>
<td>2.016</td>
<td>2.087</td>
<td>1.993</td>
</tr>
<tr>
<td>Transfer</td>
<td>1.317</td>
<td>1.860</td>
<td>1.583</td>
<td>1.134</td>
<td>1.070</td>
</tr>
<tr>
<td><strong>90°</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>1.512</td>
<td>1.547</td>
<td>1.411</td>
<td>1.797</td>
<td>1.565</td>
</tr>
<tr>
<td>Transfer</td>
<td>1.317</td>
<td>1.925</td>
<td>2.314</td>
<td>1.529</td>
<td>1.547</td>
</tr>
<tr>
<td><strong>135°</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>0.821</td>
<td>0.999</td>
<td>1.034</td>
<td>1.058</td>
<td>1.201</td>
</tr>
<tr>
<td>Transfer</td>
<td>1.583</td>
<td>1.427</td>
<td>1.797</td>
<td>1.160</td>
<td>1.173</td>
</tr>
<tr>
<td><strong>3 cpd</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>2.538</td>
<td>1.494</td>
<td>1.970</td>
<td>1.583</td>
<td>2.210</td>
</tr>
<tr>
<td>Transfer</td>
<td>2.288</td>
<td>2.719</td>
<td>2.880</td>
<td>2.982</td>
<td>3.158</td>
</tr>
<tr>
<td><strong>8 cpd</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>4.217</td>
<td>5.122</td>
<td>3.754</td>
<td>6.302</td>
<td>4.360</td>
</tr>
</tbody>
</table>

### Observer IR

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Transfer</th>
<th>19 Day</th>
<th>50 Day</th>
<th>130 Days</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Right Eye</strong></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>1.121</td>
<td>*</td>
<td>1.121</td>
<td>1.121</td>
<td>0.880</td>
</tr>
<tr>
<td>Transfer</td>
<td>1.095</td>
<td>*</td>
<td>1.046</td>
<td>1.148</td>
<td>0.784</td>
</tr>
<tr>
<td><strong>90°</strong></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>1.034</td>
<td>*</td>
<td>1.347</td>
<td>0.954</td>
<td>1.347</td>
</tr>
<tr>
<td>Transfer</td>
<td>1.070</td>
<td>*</td>
<td>0.987</td>
<td>1.215</td>
<td>1.494</td>
</tr>
<tr>
<td><strong>135°</strong></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>1.022</td>
<td>*</td>
<td>1.160</td>
<td>1.215</td>
<td>1.229</td>
</tr>
<tr>
<td>Transfer</td>
<td>1.657</td>
<td>*</td>
<td>2.210</td>
<td>1.860</td>
<td>2.210</td>
</tr>
<tr>
<td><strong>3 cpd</strong></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>3.086</td>
<td>*</td>
<td>2.982</td>
<td>3.423</td>
<td>3.797</td>
</tr>
<tr>
<td>Transfer</td>
<td>1.257</td>
<td>*</td>
<td>1.215</td>
<td>1.427</td>
<td>1.696</td>
</tr>
<tr>
<td><strong>8 cpd</strong></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>1.229</td>
<td>*</td>
<td>1.925</td>
<td>1.736</td>
<td>2.063</td>
</tr>
</tbody>
</table>

* Data omitted because of high voltage during measurement session

248
Observer JR

<table>
<thead>
<tr>
<th>Right Eye</th>
<th>Baseline</th>
<th>Transfer</th>
<th>19 Day</th>
<th>50 Day</th>
<th>130 Days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trained</td>
<td>2.396</td>
<td>1.696</td>
<td>3.463</td>
<td>3.976</td>
<td>4.726</td>
</tr>
<tr>
<td>90°</td>
<td>1.993</td>
<td>2.627</td>
<td>4.260</td>
<td>2.538</td>
<td>6.158</td>
</tr>
<tr>
<td>135°</td>
<td>2.751</td>
<td>2.947</td>
<td>9.873</td>
<td>17.157</td>
<td>16.198</td>
</tr>
<tr>
<td>3 cpd</td>
<td>1.583</td>
<td>1.121</td>
<td>3.711</td>
<td>3.423</td>
<td>4.022</td>
</tr>
<tr>
<td>2 cpd</td>
<td>2.016</td>
<td>1.547</td>
<td>3.384</td>
<td>2.480</td>
<td>2.914</td>
</tr>
<tr>
<td>1 cpd</td>
<td>2.452</td>
<td>1.696</td>
<td>1.817</td>
<td>2.063</td>
<td>2.314</td>
</tr>
<tr>
<td>6 cpd</td>
<td>4.022</td>
<td>3.232</td>
<td>5.488</td>
<td>11.734</td>
<td>7.404</td>
</tr>
<tr>
<td>8 cpd</td>
<td>6.675</td>
<td>7.753</td>
<td>9.873</td>
<td>12.007</td>
<td>14.603</td>
</tr>
</tbody>
</table>

*Data omitted because of damage sustained to right eye following accident
Appendix F

Photograph of an X-ray used in the dot detection tasks
Appendix G

Examples of stimuli used in the visual search tasks

- Familiarisation set
- Novel dimensions set TA
- Dark circles and light ellipses TA
- Dark ellipses TP
- Dark circles and dark ellipses TP
- Dark circles TA
- Dark circles TP
- Dark circles and light ellipses TP

TP = target present
TA = target absent