UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

THE NEURAL CORRELATES OF VISUAL SEARCH AND TARGET ACQUISITION

by

LINDA LUISE MEYER

DISSERTATION

submitted in fulfilment of the requirements for the degree of

MAGISTER SCIENTIA

In the Faculty of Health Sciences
Department of Physiology
University of Pretoria
Pretoria

Advisors: Dr. N. Claassen (Department of Physiology)
Prof. J. J. Hanekom (Department of Electrical, Electronic and Computer Engineering)

June 2004

ACKNOWLEDGEMENTS

I would like to sincerely thank the following people and institutions:

- My supervisors, Dr. N. Claassen and Prof. J. J. Hanekom, for their valuable input and guidance.
- Prof. P. Geerthsen and Mrs M. Smith for their help with statistical analyses.
- Mr J. H. Pretorius for software development.
- Mr J. Opperman for optometric evaluation of the participants.
- Mrs M. B. Bradley for language editing.
- University of Pretoria and the Council for Scientific and Industrial Research (CSIR) for financial support.
- My parents for their enthusiasm, support and giving me a sure foundation.
- Johan for his patience, love and support.

ABSTRACT

Visual target acquisition is performed during several daily tasks, often requiring time-dependent behavioural responses towards stimuli. Information processing during such tasks is subject to bottom-up as well top-down influences, which results in an integrated processing mechanism. It follows that if the underlying neural mechanisms can be elucidated, behaviour towards visual stimuli will be better understood, allowing for the development of visual environments that facilitates desired behavioural response. The current study aimed to develop a systems-level approach according to which the mechanisms that underlie visual target acquisition can be understood, by interpreting psychophysical data in terms of the structural and functional organization of the visual system. Empirical work entailed psychophysical experiments and elaborated on previous studies regarding conspicuity areas around and response time towards visual targets. The rationale was that these two measures can be used as an indication of the conspicuity of a target within a specific background, which in turn can be related to the nature of information processing during a target acquisition task. Results showed that a proportional relationship exists between the size of the conspicuity area and a target's perceived conspicuity, with the most conspicuous targets being associated with the largest conspicuity areas. Response time trends showed that target detectability at different positions within the conspicuity area is equal, but that detection performance at positions outside the conspicuity area is greatly influenced by the nature of the background surrounding the target. Interpretation of the results points to the importance of visual attention during target acquisition, which in turn is supported by the structural and functional organization of the visual system. Findings from the psychophysical study presented here, along with the proposed framework of information processing, emphasise that behavioural outcome during visual target acquisition cannot be explained without considering the structural and functional organization of the visual system.

Keywords: *visual search; target acquisition; visual cortex; conspicuity; visual attention; response time; psychophysics; visual pathways; bottom-up; top-down*

SAMEVATTING

Visuele teikenaansluiting vorm deel van verskeie daaglikse take en vereis dikwels tydsafhanklike reaksie ten opsigte van waargenome stimuli. Verwerking van inligting gedurende sulke take is ’n geïntegreerde proses wat deur opwaartse sowel as afwaartse faktore beïnvloed word. Dit volg dus dat indien die onderliggende neurale meganismes verklaar kan word, beter begrip aangaande gedrag ten opsigte van visuele stimuli tydens ’n teikenaansluitingtaak gevorm sal word. Sodoende kan visuele omgewings wat daarop gemik is om ’n spesifieke gedragsreaksie uit te lok, ontwerp word. Hierdie studie het gepoog om psigofisiese data aan die hand van die strukturele en funksionele organisasie van die visuele stelsel te verduidelik om sodoende ’n stelsel-vlak benadering daar te stel waarvolgens die onderliggende neurale meganismes van visuele teikenaansluiting verstaan kan word. Empiriese werk het psigofisiese studies behels en het op vorige werk aangaande die opsigtelikheidsarea rondom ‘n visuele teiken en die geassosieerde reaksietyd voortgebou. Die beginsel was dat ’n teiken se opsigtelikheid binne ’n gegewe agtergrond, waarvan hierdie twee parameters ’n aanduiding is, met die aard van inligtingprosessering in verband gebring kan word. Resultate het getoon dat daar ’n eweredigheid tussen die grootte van die opsigtelikheidsarea en die waargenome opsigtelikheid van die teiken bestaan, aangesien die mees opsigtelike teikens met die grootste opsigtelikheidsareas geassosieer is. Reaksietydtendense het getoon dat waarnemingsprestasie by verskeie posisies binne die opsigtelikheidsarea gelyk is, terwyl waarnemingsprestasie by posisies buite die area sterk deur die aard van die agtergrond beïnvloed word. Resultate dui op die belangrikheid van visuele aandagbepaling tydens visuele teikenaansluiting en word ondersteun deur die strukturele en funksionele organisasie van die visuele stelsel. Hierdie bevindings, tesame met die voorgestelde raamwerk van inligtingprosessering, beklemtoon dat gedrag tydens ’n visuele teikenaansluitingstaak nie sonder inagneming van die strukturele en funksionele organisasie van die visuele stelsel verklaar kan word nie.

Sleutelterme: *visual search; target acquisition; visual cortex; conspicuity; visual attention; response time; psychophysics; visual pathways; bottom-up; top-down*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF FREQUENTLY USED ABBREVIATIONS

AET: Attentional Engagement Theory

FEF: frontal eye fields

FIT: Feature Integration Theory

FST: fundus of the superior temporal area

GS: Guided Search

IOR: inhibition of return

IT: inferotemporal

LGN: lateral geniculate nucleus

LIP: lateral intraparietal

MST: medial superior temporal

MT: middle temporal

PFC: prefrontal cortex

PP: posterior parietal

RT: response time

SC: superior colliculus

SEF: supplementary eye fields

SOA: stimulus onset asynchrony

SNR: signal-to-noise ratio

VIP: ventral intraparietal

CHAPTER 1

INTRODUCTION

1.1 MOTIVATION FOR STUDY

Visual target acquisition is performed in several daily tasks. In some cases, such as finding one's car keys, successful completion is not of utmost importance and failing to detect the sought object rapidly, apart from causing frustration, will not have a severe impact on one's wellbeing. However, during tasks such as monitoring the instrument panel of an aeroplane or driving along a busy street, it is crucial to detect and react to visual targets timeously and appropriately.

Visual target acquisition can be defined as the process during which a predefined target needs to be found within a visual image, evaluated in terms of specific task requirements and then reacted to. The process thus includes searching for the target, deciding whether it meets requirements specified at the onset of the task and then initiating and executing a proper behavioural response based on that decision. The importance of proper visual target acquisition can be illustrated well against a defence background. Consider a situation where two parties are involved in an air-to-ground battle for possession of a specific area or object. The task is time-critical: both the defender and the attacker need to detect the target timeously, so that there will be sufficient time to prepare and elicit a response. The party that is able to detect its target first will be able to initiate a response earlier, thereby increasing its chances of acquiring the desired object or area. It follows that enhancing one's visual target acquisition ability, while simultaneously decreasing one's own detectability, will give an advantage over the enemy.

The situation sketched above indicates that if the detectability of a target is known, measures can be put in place to manipulate detectability, according to task requirements. Furthermore, if the detectability of the specified target can be determined, whether in its original state or after manipulation, estimates of the time necessary for detection can be made. These aspects can then be usefully applied during tactical planning.

A recent workshop on the effectiveness of detection avoidance techniques, as assessed by human observers, recommends that target detectability be reported in terms of its visual conspicuity (NATO, 2000). Visual target conspicuity refers to the discrepancy of a target within its background (Toet et al., 1998) and was operationally defined by Engel (1971) as "that combination of properties of a visible object in its background by which it attracts attention via the visual system, and is seen in consequence. "

An important indication of target conspicuity is the size of a peripheral area around the centre of fixation from which the target can be extracted during a single fixation and it is therefore useful to determine the size of the conspicuity area as indication of a target's detection probability (Toet et al., 1998; NATO, 2000). It has been shown that targets embedded in a complex background are associated with small conspicuity areas, while those that are located in a background from which they can easily be extracted, are associated with big conspicuity areas (Engel, 1971; Toet et al., 1998). Conspicuity area, also termed visual lobe or functional field of view, has been shown to be an elliptical to circular area around the target and its size is usually determined by finding the spatial border between detection and non-detection along meridians with the target as centre point (Engel, 1971; Bellamy & Courtney, 1981; Courtney & Chan, 1985b; Chan & Courtney, 1996; Toet et al., 1998).

Studies have furthermore shown that an inverse relationship exists between the size of the conspicuity area and search time - targets associated with large conspicuity areas are detected rapidly, while those associated with small conspicuity areas require long search times (Engel, 1971; Bellamy & Courtney, 1981; Kee et al., 1992; Toet et al., 1998, NATO, 2000). Toet and colleagues (Toet et al., 1998) subsequently showed that the mean search time for a specific target-background combination can be predicted once the target's relative conspicuity area has been determined. Combined results from conspicuity area determination and response time (RT) measurements can therefore be valuable for evaluation of a subject's search and target acquisition performance as well as design of visual environments where rapid detection of visual signals is important. Target conspicuity determinations as described here have also been shown to be easily performed in the laboratory, with results correlating well with similar field studies, while having the added advantage of being more time- and cost-effective to perform (Toet et al., 1998; NATO, 2000).

A large body of research, independent of any military applications, has focussed on investigating the principles of information processing governing visual target acquisition, in order to explain different behaviour trends observed for different target-background combinations. Information processing during visual target acquisition is regarded as proceeding according to a two-stage process, which consists of two hierarchically organized stages (Treisman & Gelade, 1980; Duncan & Humphreys, 1989; Theeuwes, 1993; Wolfe, 1994). Such an approach builds on Neisser's theory (1967) that distinguishes between an early pre-attentive mechanism and a later attention-dependent processing stage. It has been proposed that the pre-attentive stage corresponds to scanning of a visual image (i.e. the search-component of visual target acquisition), while the later attention-dependent stage involves decision-making and initiation and execution of a behavioural response (Wolfe, 2003). Although the two stages seem to be functionally distinct, a segmentation-and-selection step during which objects that may possibly correspond to the target description serves as the transition between the two stages (Wolfe, 2003). Since only limited processing capacity is available during the attentive stage, incoming information needs to be filtered, so that only relevant information is allowed to proceed from the first to the second stage of visual information processing.

From the preceding paragraphs it is clear that measures of conspicuity as described previously are behavioural outcomes of a two-stage perceptual process during which a visual image is searched for and relevant information extracted and passed forward for processing related to recognition, identification and evaluation of the selected target. Theories regarding the underlying mechanisms of information processing during visual target acquisition (e.g. Treisman & Gelade, 1980; Duncan & Humphreys, 1989; Wolfe, 1994;) suggest that the process is based on a highly structured functional organization. Psychophysics provides a valuable means to gain insight into the underlying structure of such a perceptual process non-invasively. Based on signal detection theory (Green & Swets, 1966), it allows an experimenter to determine the detection probability for a presented stimulus by designing an experiment, which clearly specifies response options to a given stimulus. Target detectability can subsequently be determined in order to determine how such stepwise manipulations affect the outcome of the perceptual process, which in turn can be used to infer the underlying mechanism of a perceptual process.

Response time studies further contribute to inferring the progression of information processing during a perceptual process. Since the functional organization seems to be highly structured, it is believed that manipulations of stimulus parameters will cause information processing to proceed along different routes, which in turn will be reflected in recorded response times (Luce, 1986). Combined response time results and detection probabilities are therefore valuable for developing models of information flow during a perceptual process (Smith, 1995).

The notion that information processing during visual target acquisition is highly structured is supported by neurophysiological and anatomical studies that show a highly structured neural organization to be involved during visual perception (Zeki, 1978; DeYoe & Van Essen, 1988; Hubel, 1988; Zeki & Shipp, 1988; Felleman & Van Essen, 1991; Bullier, 2001; Nobre et al., 2003). The visual system consists of approximately 30 cortical and subcortical areas that are hierarchically organized, forming a highly structured processing network (Felleman & Van Essen, 1991; Bullier, 2001). Within this network information pertaining to featural and spatial characteristics of the image is to a large extent segregated, resulting in two functional visual processing streams, namely the ventral "what" and dorsal "where" pathways (Goodale et al., 1994). At higher anatomical levels, however, information from the two pathways is recombined and subsequently projected to the same cortical areas. Anatomical organization supports the functional organization of the visual system, not only in that segregation of information is brought about by segregation of anatomical areas, but also in the hierarchical organization of the system, which results in more fundamental analyses (e.g. Zeki, 1978; Gallant et al., 1996; Pasupathy & Connor, 1999; Wilkinson et al., 2000; Kastner et al., 2000; Singh et al., 2000; Adams & Zeki, 2001; Pasupathy & Connor, 2002) being preformed earlier along the processing pathways, while analyses become more integrated and involved as information proceeds towards higher cortical areas (Miller et al., 1991; Chelazzi et al., 1993; Miller et al., 1993; Chelazzi, 1995; Desimone, 1996; Ungerleider et al., 1998; Schall, 2001).

From the brief background presented above it is evident that investigations into visual target acquisition can be approached from various perspectives, focussing either on its cognitive principles, the underlying neurophysiological mechanisms that govern the process or the applications of visual target acquisition. Thus far, findings from studies regarding these different aspects have been considered in isolation and it has not been established how such results can be linked in order to construct a comprehensive understanding of how visual target acquisition comes about. Since it is the behavioural outcomes of visual target acquisition that are used in industrial applications, it is important to understand how perceptual information is processed, which in turn should be understood in terms of the neurophysiology that underlies the perceptual process. It would furthermore be valuable if a comprehensive understanding of the process of visual target acquisition can be derived from behavioural outcomes, since psychophysical techniques, which are non-invasive, cost-effective and do not require specialized equipment, can then be used. The current study thus aimed to develop, based on the underlying structural and functional organization of the visual system, a systems-level approach of the different processing steps involved during visual target acquisition, according to which behavioural responses to a target acquisition task, as measured with psychophysical techniques, can be explained.

1.2 OBJECTIVES

In order to develop such a systems-level approach of visual target acquisition behaviour, the following objectives were identified:

(i) To identify the main anatomical regions involved in visual target acquisition.

(ii) To determine the functional importance of the identified regions with regard to visual target acquisition.

(iii) To investigate the extent to which specific manipulations of a visual image will influence target detectability as measured by means of determining the conspicuity area. Image manipulations involved manipulation of target and background parameters respectively, as well as combinations of altered targets and background configurations.

(iv) If image manipulations were to have any effect on target detectability, to determine how such manipulations will influence an observer's behavioural response during a visual target acquisition task, by means of measuring response time to the presented target.

(v) To determine under the same display conditions as earlier, how target position, relative to the conspicuity area, influences response time.

1.3 HYPOTHESES

The main hypothesis of the study can be formulated as follows:

Behavioural responses during a visual target acquisition task can be measured objectively with psychophysical techniques and interpreted according to the structural and functional organization of the visual system.

It is supported by the following subhypotheses:

(i) Manipulation of display conditions in which a visual target is presented, influences the associated conspicuity area.

(ii) Manipulation of display conditions will result in altered response time trends towards the target.

(iii) A relationship exists between response time and target position, relative to the conspicuity area.

(iv) Conspicuity area and response time trends can be applied successfully as psychophysical measures of target detectability.

(v) Neural correlates of psychophysical measures of target detectability can be identified.

1.4 METHODOLOGICAL APPROACH

The problem was approached from an empirical as well as theoretical viewpoint. The empirical component involved psychophysical measurement of behavioural responses during visual target acquisition tasks by means of conspicuity area determinations and response time measurements. The theoretical component entailed extensive literature study regarding the structural and functional organization of the visual system, as well as the principles of information processing during visual perception, leading to a conceptual model of the processing steps involved in visual target acquisition being developed. Findings from experiments were subsequently related to the model in order to explain the underlying mechanism of visual target acquisition behaviour.

Two types of experiments, namely conspicuity area determinations and response time studies, were performed with eight subjects under four different display conditions. Display conditions were designed to vary target detectability, by changing either target or background parameters or by using specific target-background combinations. All displays were generated by custom-designed software and presented on a computer screen without curvature.

During conspicuity area determinations subjects were instructed to fixate a fixation point and then indicate whether a predefined target was present or absent from a briefly presented display. Results were used to determine the area in which the target could be detected in a single fixation (the conspicuity area), which in turn served as indication of target detectability. The same display conditions as for the conspicuity area determinations were used during response time studies. Subjects were instructed to search for a predefined target in the display and to press a specific button as soon as the target was detected. Response time was recorded as elapsed time from onset of the display until activation of the button. Targets were presented at specific positions relative to the conspicuity area determined for the display condition and subject in question.

Prior to conducting these experiments, however, careful consideration was given to optimization of the proposed methodology. Pilot studies were conducted with preliminary display conditions to investigate the effect of different display manipulations on target detectability as well as to evaluate the usability of different techniques. Accuracy of response time measurement was also verified prior to conducting final experiments.

1.5 STRUCTURE OF THESIS

Since the study aimed ultimately to explain visual target acquisition behaviour according to a model that integrates behavioural outcomes with the structural and functional organization of the visual system, important concepts relating to these two components are introduced at an early stage. Structural and functional organization is dealt with in Chapter 2, discussing the initial segregation and later recombination of the two information processing streams within the visual system as well as the functional contribution of specific anatomical regions to visual target acquisition. A conceptual model, based on the hierarchical organization of the visual system, is introduced. In Chapter 3 it is shown that information processing pertaining to visual target acquisition is a hierarchically organized two-stage process that comprises an early pre-attentive stage followed by an attention-dependent stage. The process, together with important factors that influence behaviour during a visual target acquisition task, are discussed in terms of the proposed model introduced in Chapter 2. Chapter 4 explains the rationale for the methodological approach by providing the reader with some background on signal detection theory and psychophysics, conspicuity area determination and the relevance of using response time as an indication of the progression of information processing during a perceptual process. Optimization of the experimental design is also shown. Chapter 5 shows how the optimized experimental design was applied during the final empirical stage. Results from conspicuity area determinations show that the size of the conspicuity area depends on the level of target detectability, which in turn can be manipulated by changing display conditions. Response time results show a relationship between target detectability (as determined from conspicuity area sizes) and response time. It is further also shown that target position plays an important role in determining the difficulty level of a target acquisition task; target acquisition at positions outside the conspicuity area is highly susceptible to changes in target detectability, while at positions within the conspicuity area, changes to target detectability have little effect on response time. The results are explained in terms of the model introduced earlier. Chapter 6 provides a summary of the study and discusses its conclusions.

CHAPTER 2

STRUCTURAL AND FUNCTIONAL ORGANIZATION OF THE VISUAL SYSTEM

Chapter 2 serves to introduce the reader to the structural and functional organization of the visual system. After a brief history of the development of visual science, important stations along the visual processing pathway are discussed. Emphasis is on how structure and function of the individual areas contribute to processing of visual information along ventral and dorsal pathways, specialized for featural and spatial analyses respectively. It is followed by the introduction of a conceptual model of the progression of information pertaining to a visual target acquisition task, according to a systematic organization based on that of the visual system. Information progression from the retina to the early visual cortex, the bifurcation of the two processing streams and eventual recombination of their output at a high cortical level are shown. The functional relevance of each step is discussed.

2.1 A BRIEF HISTORY

Since the beginning of time, man has had to rely on sensory information to survive and through the ages, vision has evolved to become one of his most prominent senses. It is therefore not surprising that the human visual system has intrigued many a scientist, leading to remarkable progress regarding its structure and function.

Surprisingly, it was probably not the work of a neurobiologist that sparked interest in vision, but rather that of a physicist. Thomas Young made an almost accidental comment on the nature of colour analysis in the visual system during a lecture on the wave properties of light, which he delivered to the Royal Society in 1802. Young (in Zeki, 1993) proposed that so-called non-specific receptors, which were able to initiate a response to a general type of stimulus and then pass the information on to be analysed and processed elsewhere, had to be present on the retina. Von Helmholtz elaborated on Young's idea by proposing that perception is the result of "an act of judgment and not an act of sensation" (Von Helmholtz in Zeki, 1993). The importance of cortical involvement in visual perception was thus already suggested early in the development of visual science.

The suggestion of cortical involvement in visual perception was confirmed by clinical evidence from patients who suffered cortical lesions with resulting visual defects. The discovery of the primary cortical centre for vision can ultimately be accredited to Munk, who, in 1881, correctly described the area concerned with processing of visual information as being situated in the occipital lobe of the brain (Munk, in Glickstein, 1988). He was followed by Henschen, a Swedish neuropathologist, who refined Munk's work by showing that visual function was most prominent in the vicinity of the calcarine fissure located on the medial surface of the occipital lobe (Henschen, in Zeki, 1993). After almost a century, the existence of a cortical visual area was finally confirmed.

By the early 1950s, despite Talbot and Marshall mapping the early visual cortices in 1941 (Talbot and Marshall, 1941), little was known about the processing mechanism employed by the visual system. Kuffler's (1952) discovery of the activation patterns of retinal ganglion cells spurred exciting discoveries in the field. By shining discrete spots of light on a specific cell and recording its response with a microelectrode, he was able to describe the behaviour of retinal ganglion cells in terms of receptive field properties. Retinal ganglion cells exhibited a centre-surround organization and could be activated by presenting stimuli to either the centre or the surround of the receptive field. Those activated by stimuli presented in the centre of the receptive field were termed ON-centre, while those activated by stimuli appearing in the surround were termed OFF-centre. (Kuffler, 1952). Hubel and Wiesel took his findings one step further, investigating responses from cells in the striate cortex. Almost a decade later, they recorded their first successes when finding orientation-selective cells in the primary visual cortex of a cat (Hubel & Wiesel, 1959). Successive work with cats and monkeys indicated further specialization within the visual cortex – some cells responded to stimuli of specific orientation located at a certain position within their receptive fields, while others fired optimally when a moving stimulus of specific orientation crossed their receptive fields (Hubel, 1988).

The processing mechanisms employed beyond the striate cortex were, however, largely unknown. By the mid-1970s several areas in the extrastriate cortices of primates were thought to be involved with analysis and processing of distinct visual attributes (Zeki, 1978) on physiological, anatomical and cyto-architectural evidence, areas such as the V3-V3A complex, V4 and the middle temporal (MT) area were thought to be involved with processing of stationary form, colour and dynamic form respectively (Zeki, 1978). Given the specialization within the primary visual cortex, the occurrence of such functionally specialized areas beyond the primary visual cortex was not surprising.

As research continued, more cortical as well as subcortical areas, either directly or indirectly involved with visual processing, were identified. However, the question of interaction among these areas remained. Theories on potential mechanisms were controversial and through the years, thoughts on the nature of the processing mechanism employed by the visual system have shifted between hierarchical (e.g. Hubel & Wiesel, 1977) and parallel models (Bullier & Nowak, 1995). The hierarchical processing mechanism of the striate cortex, as proposed by Hubel and Wiesel (1977) led researchers to believe that a similar organization must hold for extrastriate areas, implying a hierarchical organization throughout the visual system. Discoveries of distinct types of retinal ganglion cells providing input to specific cortical areas, however, favoured a theory of parallel processing streams. Only a few years later, the pendulum swung back to the notion of serial processing mechanisms, owing to the identification of several anatomical connections between functionally distinct areas (Zeki & Shipp, 1988). Eventually a hybrid model was proposed, advocating that several parallel processing streams, with intrinsic hierarchical properties, operate concurrently (Zeki & Shipp, 1988; DeYoe & Van Essen, 1988).

David Hubel aptly summarized the ongoing search for explanations of visual processing in his Nobel Prize lecture in 1981: *"A few years ago, in a Ferrier lecture, Torsten (Wiesel) and I ended by saying that the striate cortex was probably, in broad outline, understood. This was done deliberately: one did not want the well (of knowledge) to dry up. When one wants rain, the best strategy is to leave raincoat and umbrella at home. So the best way to guarantee future employment was to declare the job finished"* (Hubel, 1982). In the following sections the objective is to deal comprehensively with current knowledge of the most important stations along the visual processing pathway. Topics such as the generation of visual information in the eye, its journey through the lateral geniculate nucleus (LGN) and its termination in the visual cortex will be discussed. In addition, analysis of specific attributes of a visual scene and the integration thereof will be discussed, followed by a discussion on the output pathways of processed visual information. It is hoped that this review will provide the reader with a sound understanding of the current state of affairs regarding the structural and functional organization of the human visual system.

2.2 STRUCTURAL AND FUNCTIONAL ORGANIZATION

2.2.1 The retina

The eye is often thought of as the human equivalent of a camera, since it enables the body to convert photons into meaningful images. Reflected light from objects in the environment first encounters the cornea, where most of the refraction takes place. As the refracted rays proceed, the lens finely adjusts the eye's refractive power by changing convexity, so that a clearly focused image can be projected onto the retina, the innermost layer of the eye. The retina can be thought of as an extension of the central nervous system, and considering its origin and specialized features, it is not surprising that the retina should be the point from where the visual journey departs.

Five major neuronal cell types are found in the retina (Figure 2.1), namely photoreceptors, horizontal cells, bipolar cells, amacrine cells and ganglion cells. The organization, together with functional specialization of these cells, allows for three retinal processing stages as discussed below. Visual information is introduced at one level, dispatched at the next and laterally distributed by the third.

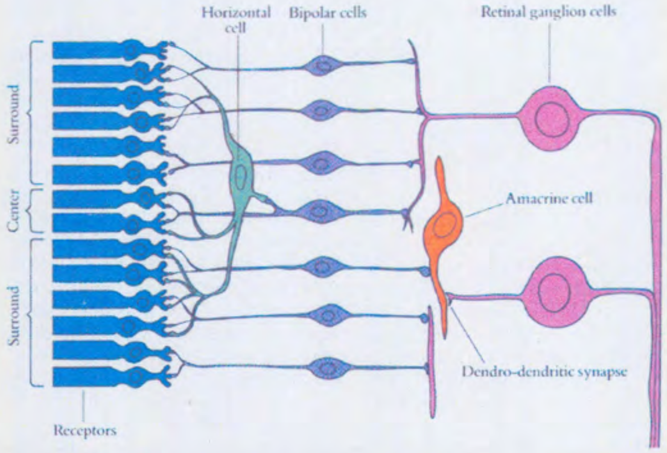


Figure 2.1: Schematic representation of retinal cell types (With permission from Hubel, 1988)

The first stage of retinal processing entails conversion of photic information to neural signals. Signals are then transferred to the next processing stage at the level of the bipolar cells, either via a direct connection onto the bipolar cells or via an indirect pathway, by first activating horizontal cells. Since horizontal cells form lateral connections between the synaptic bodies of receptors and dendrites of bipolar cells, they are able to activate or de-activate specific bipolar cells selectively, depending on the nature of their synapses. If an excitatory synapse should exist between a horizontal cell and a bipolar cell, signal transmission would continue, while an inhibitory synapse would result in the signal from a specific receptor being blocked. Horizontal cells thus provide a mechanism for lateral inhibition, thereby increasing the specificity with which visual signals are transmitted through the visual system.

The specificity with which visual signals are propagated within the retinal processing stage is further increased owing to the nature of the receptive field of the bipolar cell. Receptive fields of bipolar cells exhibit a centre-surround organization, with either the centre being excitatory and the surround being inhibitory, or the exactly opposite arrangement (Hubel, 1988). Such organization is similar to that of retinal ganglion cells (discussed below). Propagation of the signal along the retinal processing pathway thus depends on the part of the bipolar cell that is stimulated. This mechanism allows only signals originating in selected receptors to be passed on along the processing pathway, since axons of several receptors can converge onto a single bipolar cell. Thus, as with horizontal cells, bipolar cells contribute to specificity with which signals are transmitted during early stages of visual processing.

Upon leaving the bipolar cell, the visual signal proceeds to the final retinal processing stage, where ganglion cells generate action potentials, which are transmitted to cortical visual areas. Similar to the horizontal cells of the second processing stage, interneurons, namely amacrine cells, establish forward as well as lateral connections between bipolar and ganglion cells through dendritic presynaptic connections (Hubel, 1988)

From the foregoing discussion it is evident that modification and integration of visual signals occur at each successive retinal processing stage. At the level of the ganglion cells, further funnelling of information occurs as a result of the convergence of several bipolar cells onto a single ganglion cell (Sterling, 1999). When a ganglion cell is stimulated, either an ON- or an OFF-response can be elicited, which, as in the case of a bipolar cell, is a function of its centre-surround receptive field organization.

Since their discovery, functionally distinct classes of retinal ganglion cells have been observed in old world monkeys, with their respective roles in visual processing reflecting their different conduction velocities and receptive field sizes. Pβ- cells with small receptive fields and slow conduction velocities were found to be colour-selective and subsequently ascribed a role in pattern analysis (Zeki & Shipp, 1988). Pα-cells, however, were thought to be involved in fast transmission of information on changes in spatial location and movement of visual stimuli, owing to their large receptive fields, high conduction velocities and lack of colour selectivity (Zeki & Shipp, 1988). Pγ-ganglion cells are thought to contribute to analysis of directional movement of visual stimuli, because of their slow conduction speeds (Schiller, 1986; Burke et al., 1998). In view of the remarkable parallelism between the primate and human visual systems (Burke et al., 1998), these findings probably also apply to human retinal ganglion cells.

2.2.2 The lateral geniculate nucleus

Axons of the retinal ganglion cells are bundled to form the optic nerve, which proceeds to extraretinal stages of the visual system. At the optic chiasm, fibres from the two optic nerves partially cross over, so that each cerebral hemisphere will receive input from both the ipsi- and contralateral eye. However, before reaching the visual cortex, the majority of fibres from the optic nerve have to pass through a thalamic complex known as the LGN.

The primate LGN is a stratified structure, containing neatly organized populations of functionally different neurons. The two ventral layers are known as magnocellular (M) layers because of the prevalence of cells with large somata, while the four (or two, depending on the species) dorsal layers are referred to as parvocellular (P), since they contain cells with small nuclei. Intercalated between the M- and P-laminae, cells with nuclei resembling grains of sand, are the koniocellular (K) layers (Hendry & Reid, 2000).

The histological organization of the LGN is not accidental. It is believed that the layered structure of the LGN provides for visual information to be transferred along definite parallel pathways. Three parallel visual pathways, namely the P-, M- and K-pathways, are currently recognized (Cassagrande, 1994). These are thought to originate from the Pβ-, Pα- and Pγ-retinal ganglion cells respectively, owing to the striking similarities in response characteristics between the respective ganglion cells and neurons of the LGN. Across several primate species P-neurons of the LGN were found to have fairly small receptive fields and high response latencies as well as marked colour and contrast selectivity, while M-cells were found to have larger receptive fields, better temporal resolution and no colour selectivity (Zeki & Shipp, 1988; Merigan & Maunsell, 1993; Cassagrande, 1994; Usrey & Reis, 2000). It is therefore thought that the P- and M-pathways are equipped to handle information on chromatic and moving stimuli respectively.

It is thus evident that the different cellular populations of the LGN, as well as its organization, are functionally important during the processing of visual information, reinforcing the strong relationship between structure and function in the visual processing system.

2.2.3 The early visual cortex: Areas V1 and V2

The final destination of the visual information collected so far, as with all sensory information, is the cerebral cortex. Hermann Munk identified the occipital lobe of the cerebrum as the main cerebral visual area and dismissed suggestions, such as those made by British neurologist David Ferrier, that a visual area might be located in the parietal lobe (Glickstein, 1988). It emerged only later that both these observations were true, since multiple areas within the occipital, parietal and temporal lobes comply with criteria used to define visual cortical areas. These include (1) well-defined cytoarchitecture, (2) well-defined anatomical input, (3) retinotopic organization and (4) distinct functional properties (Zeki, 1978, Wandell, 1999). In fact, to date almost 30 visual areas have been identified in the primate cerebral cortex (Felleman & Van Essen, 1991).

Area V1

The primary visual cortex (V1), also called the striate cortex because of its striped appearance, is situated at the posterior end of the occipital lobe, most of it being buried in the calcarine sulcus on the medial surface (Hubel, 1988; Zeki, 1993; Wandell, 1999). Structurally and functionally different cell populations are neatly organized in six layers, with layer 6 situated closest to the white matter. Afferent fibres from the LGN thus enter V1 from the ventral layers, moving diagonally upwards before terminating in layer 4C, with fibres originating from the magno- and parvocellular layers of the LGN, terminating in layers 4Cα and 4Cβ respectively. The segregation between M- and P-pathway-derived information remains, seeing that signals from layer 4Cα are projected to layer 4B and those from layer 4Cβ terminate in layers 2 and 3. Since the different layers of V1 project to distinct cortical areas, layer 4C seems to act as a relay station, ensuring that incoming information proceeds to the correct destinations.

It seems that the farther one moves along the visual path, the more specialized cellular behaviour becomes. The well-organized cytoarchitecture of V1 is therefore not surprising. Similar to retinal ganglion and geniculate cells, stimuli must appear in the receptive field of a V1-cell before a response can be elicited, but with the distinction that V1 cells exhibit linear, rather than concentric, receptive fields (Hubel, 1988). This implies that only stimuli of the correct orientation will be able to evoke a response from such cortical cells. Further specialization, however, follows. Some cells, known as simple cells, need stimuli of precise orientation and position before responding, while others, known as complex cells, respond to appropriately oriented stimuli regardless of their position (Hubel, 1988). Apart from orientation selectivity, some complex cells also appear to have directional selectivity, since microelectrode studies showed that a pronounced response difference occurs when an optimally oriented stimulus is moved across a complex cell's receptive field in opposite directions (Hubel, 1988). Simple and complex cells are organized in vertical slabs with orientation shifts of 10 degrees every 50 μm. Such slab-like organization is related to the layered structure of V1, since vertical microelectrode penetrations from the upper layers downwards (except in layer 4) yield the same orientation preference in a specific 50 μm area. Patches of neurons, inserted in the orientation slabs at the level of layers 2 and 3 of V1, however, lack orientation selectivity, but exhibit pronounced wavelength selectivity (Hubel, 1988; Zeki, 1993). These cell clusters, as well as areas between them, are important during analyses of chromatic information and are commonly referred to as "blobs" and "interblobs" respectively (Zeki & Shipp, 1988; Zeki, 1993).

Functionally the situation can be summarized as follows: Information enters V1 from the LGN through layer 4C, with magnocellular fibres penetrating 4Cα and parvocellular axons reaching 4Cβ. Subsequently connections from layer 4Cα are sent to the simple and complex cells of layer 4B, while fibres from layer 4Cβ associate with the blobs and interblobs of layers 2 and 3. Cells of the striate cortex thus show selectivity towards information conveyed by both the M- and P-pathways.

Area V2

Cytochrome oxidase staining revealed that repeating patterns of dark- and light-staining stripes appear in V2, with dark stripes being either thin or thick. Subsequent analysis showed that thick dark stripes contain predominantly orientation-selective cells as well as orientation-and-direction selective cells, and that they receive major input from layer 4B. Thin dark stripes on the other hand, were shown to be wavelength-specific, and receive input from the blobs of V1. Similarly, light-staining areas of V2 (interstripes) receive projections from interblob regions of V1 (Zeki & Shipp, 1988).

Structurally as well as functionally, V2 thus appears very similar to V1, but it has the added functionality of cells being selective to stereoscopic cues (Hubel, 1988). This allows for contour completion of stimuli both within and outside the classic receptive field of the cell

(Barkin et al., 2000), effecting contour integration of localized events as well as creating integrated contour relationships for areas beyond their receptive fields (Bradley, 2001)

2.2.4 Beyond V2: Bifurcation between P- and M-pathway derived information

Beyond V2, the segregation between P- and M-pathway derived information becomes more pronounced, since information on different attributes of a visual scene continue along two distinct parallel routes. These differ functionally as well as anatomically - visual cortical areas in ventral positions are mainly concerned with analysis of features regarding object recognition and identification, while those in dorsal regions are mostly specialized for analysis concerning localization and spatial relationships.

Ventral processing pathway

The ventral pathway starts out from V1 and V2, passes through V4 and eventually terminates in the inferotemporal (IT) cortex. Neurons in these areas have been shown to be selective towards wavelength- and orientation-specific cues, as well as susceptible to modulatory effects of visual attention. Because of their neuronal specificity, these areas are thus ideally suited to contribute to analyses pertaining to object recognition and identification. The ventral processing route has consequently been termed the "what" pathway.

Situated anterior to V2 is area V4. One of the most outstanding features of neurons in V4 is their wavelength selectivity. Studies on primate V4 have shown a heavy concentration of wavelength-specific cells in V4, while such specificity was almost neglible in other extrastriate areas (Bartels & Zeki, 2000). Considering the connections between V2 and V4, wavelength-related information presumably reaches V4 either directly from the blobs of V1, or indirectly via the thin stripes of V2, which both contain wavelength-selective cells (Zeki & Shipp, 1988). These findings have led to the belief that V4 represents the main colour-processing module of the visual system. Recently a human homologue to this area has been identified and subsequently been named human V4 (Zeki et al., 1998).[1]

Apart from its role in colour perception, V4 has also been implicated in intermediate shape analysis (Pasupathy & Connor, 1999; Wilkinson et al., 2000) and texture segregation (Kastner et al., 2000), owing to its responsiveness to contour-related information. Recently, Pasupathy and Connor (1999) showed that neuronal responses in V4 are systematically tuned for contour features like curves and angles, which have subsequently been shown to be extracted at earlier stages (Hegde & Van Essen, 2000). This means a variety of shapes can be coded for by activating selective combinations of neurons tuned to different orientation-related features (Pasupathy & Connor, 1999). These can in turn be used during perception of complex objects such as faces in higher order areas, since a significant population of V4 cells respond strongly to polar stimuli such as concentric, radial and hyperbolic gratings (Wilkinson et al., 2000), which have earlier been shown to be important components of various complex objects of the type analysed in the IT cortex (Kobatake & Tanaka, 1994; Wilkinson et al., 2000). Together, these findings have been taken as evidence that V4 represents an intermediate shape analysis stage during object recognition. V4's role in texture segregation can possibly be attributed to its involvement in extracting complex contour information that defines texture boundaries, which can then be projected back to earlier processing stages where it can be used to modulate responses to these attributes (Kastner et al., 2000).

---

[1] Much controversy about such a colour area's exact location and name exists (Zeki et al., 1998; Hadjikani et al., 1998), but it will be referred to here as V4.

V4's importance in the ventral object recognition pathway is further demonstrated by its susceptibility to attentional effects. Results from primate studies indicate that when attended and unattended stimuli appear simultaneously in the same receptive field, the response to the unattended stimulus is significantly reduced (Moran & Desimone, 1985). Furthermore, cellular activity preceding the onset of the behavioural response seems to be almost completely governed by the attended stimulus (Chelazzi et al., 2001). Such effects in V4 were observed both when the task demanded spatial attention (Moran & Desimone, 1985) and feature attention (Motter, 1994; McAdams & Maunsell, 2000; Chelazzi et al., 2001). On account of its involvement in analyses of several aspects pertaining to feature processing, V4 seems to represent an integral component of the ventral processing pathway.

The IT cortex is generally regarded as the last exclusively visual area of the ventral pathway and serves as an important link between information coming from lower cortical areas (bottom-up) and that related to higher cognitive functions projecting downwards from higher cortical areas (top-down). In terms of bottom-up analysis, neurons of the IT cortex show considerable specialization and use information from earlier analysis stages along the ventral pathway to perform object identification. Studies have shown that complex stimuli can be systematically broken down in order to determine the feature or combination of features which caused maximal activation of IT neurons (Kobatake & Tanaka, 1994; Tanaka, 1996). Although such critical features are more complex than simple textures, orientation, size or colour, they are not complicated enough to enable a single cell to code for a complete object. Responses from multiple neurons thus need to be pooled in order to create a representation of the complete object (Tanaka, 1996). This is probably facilitated by cellular organization, since IT reportedly consists of columnar modules where cells of similar but not identical selectivity overlap (Tanaka, 1996). It has also been shown that IT is the major area for face recognition, since separate populations of cells have been found to respond selectively to faces, with some neurons responding to parts of faces or specific spatial arrangements of facial components (Desimone et al., 1984).

Because of the progressive increase in receptive field sizes in successive stages of the visual processing pathway, massive information convergence occurs at the level of the IT cortex. Efficient bottom-up analysis is aided by IT neurons' involvement in visual attention (Chelazzi et al., 1993; Chelazzi, 1995) and visual memory (Miller et al., 1991; Miller et al., 1993; Desimone, 1996). Such functionality facilitates selection of behaviourally important stimuli and thus ensures optimal allocation of processing capacity. However, these two processes seem so closely linked that it is difficult to decide whether selection depends on attended stimuli being held in memory, or remembered stimuli being attended to (Chelazzi, 1995; Desimone, 1996). Nevertheless, the IT cortex seems to be suitably equipped for both processes. In the case of visual attention, studies have shown that, as in V4, neuronal responses to ignored stimuli are suppressed, so that the activity in these areas reflects only properties of attended stimuli (Chelazzi et al., 1993; Chelazzi, 1995;). This modulatory effect can be guided either by spatial attention or by attention to object features alone.

When selection of behaviourally relevant stimuli is memory-driven, the current stimulus is evaluated against a sample of the relevant stimulus retained in memory, in order to decide on the correct response (Desimone, 1996). Studies where monkeys performed a delayed match-to-sample task suggested that two parallel mechanisms mediate memory-driven selection in the IT cortex, namely a suppressive mechanism, as well as an enhancement mechanism (Miller et al., 1991; Miller & Desimone, 1994;). In the suppressive mode, responses to stimuli similar to the sample seem to be attenuated in a subpopulation of IT cells, which then probably serves as a signal for selection when the stimulus reappears. Furthermore, this suppression is proportional to the similarity between sample and current stimulus, so that the stronger the suppression, the more confident the selection. Such response suppression was also observed when a non-matching stimulus appeared repeatedly, which may be interpreted as a memory mechanism for unfamiliar objects (Desimone, 1996). In contrast to the suppressive mechanism, responses from another population of IT cells were enhanced by a matching stimulus, but not by repetition of non-matching stimuli, thereby creating a unique selection signal for the relevant stimulus. These two mechanisms thus co-exist and preference for either possibly depends on specific task demands, training history or other individual variables (Desimone, 1996). However, maintaining the memory of the sample does not seem to result within the IT cortex itself (Miller et al., 1993), but appears to be rather mediated by reciprocal connections between IT and the prefrontal cortex (PFC) (Miller & Desimone, 1994; Desimone, 1996).

The dorsal processing stream

Similar to the ventral pathway, the dorsal processing stream arises in V1 and V2, and then diverts to area V3 and the MT area to terminate eventually in the posterior parietal (PP) cortex. Its main output relates to spatial analyses, motion perception, visuomotor behaviour and direction of visual attention. As in the case of the ventral pathway, these functions are supported by neuronal specialization.

Functional organization of the two streams appears to be very similar. Processing of spatial features occurs at intermediate levels of the pathway, while top-down modulation of spatial perception and activity related to visuomotor output are functions of higher areas in the pathway, possibly facilitated by their connections with the specialized areas of the PFC and premotor cortex.

An important intermediate processing station along this pathway is area MT. On account of the prevalence of direction-sensitive cells, as well as the connections between area MT and layer 4B of V1 and the thick stripes of V2, area MT has been identified as the main motion-sensitive region of the visual processing pathway (Zeki, 1978; Zeki & Shipp, 1988). Subsequent studies have shown that motion perception, especially speed sensitivity, in MT also depends on other stimulus parameters such as contrast and colour (Op de Beeck et al., 2001). Its importance in motion processing is further supported by extensive connections to other areas, such as the medial superior temporal area (MST) and the fundus of the superior temporal area (FST).

Area V3 also appears to be an important contributor to spatial analysis, especially in view of the abundance of cells selective for orientation, direction and stereoscopic depth (Zeki, 1978; Singh et al., 2000; Adams & Zeki, 2001). These properties are thought to be significant for processing of three-dimensional form later on in the motion pathway (Adams & Zeki, 2001), as well as dynamic form analysis effected within V3 (Gegenfurtner & Hawken, 1996). However, observed specificity towards chromatic and contrast properties has indicated that V3 might also act as an important integrator between stimulus attributes destined respectively for the ventral “object” and dorsal “spatial” pathways (Gegenfurtner & Hawken, 1996).

The areas discussed above ultimately project to the posterior parietal cortex, which can be functionally subdivided into four regions, each specialized for higher order analysis of spatial information (Rolls & Deco, 2002). These include the lateral intraparietal area (LIP), ventral intraparietal area (VIP), MST and area 7a.

Neurons of the two intraparietal regions appear to be involved in transforming spatial coordinates into head-centred spatial representations. LIP neurons have been shown to respond to specific combinations of visual eccentricity on the retina and eye position, resulting in remapping of retinotopic spatial information into head-centred coordinates (Zipser & Andersen, 1988; Andersen et al., 1990; Rolls & Deco, 2002). This in turn affects guiding of eye movements, facilitated by connections with subcortical and prefrontal areas involved in saccade generation (Rolls & Deco, 2002). LIP neurons have also been implicated in creating a spatial memory trace for absent objects, which, together with the ability to create an anticipatory response just before a saccade is made to the specific receptive field in question, contributes to generating a head-centred internal representation of the visual environment (Duhamel et al., 1992; Rolls & Deco, 2002). MST seems to contribute to processing of global pattern motion owing to its motion and directional selectivity (Rolls & Deco, 2002). The PP cortex is also involved in visual attention, effected mostly by the neuronal activity in area 7a. These neurons have been shown to be sensitive to stimuli that appear at one location, while attention is directed elsewhere (Steinmetz & Constantinidis, 1995). This suggests that area 7a may be involved with redirecting attention, probably by providing information on the spatial locations of novel stimuli (Steinmetz & Constantinidis, 1995). Similar to the IT cortex in the ventral visual pathway, the PP cortex thus seems to fulfil an integratory function.

2.3 A CONCEPTUAL MODEL OF INFORMATION PROCESSING DURING A VISUAL TARGET ACQUISITION TASK

Based on the structural and functional organization of the visual system as discussed in the foregoing section, a conceptual model of the processing of information during a visual target acquisition task was developed (Figure 2.2). It shows that beyond the early visual cortex information is processed along two parallel streams up to the level of the PP and IT cortices respectively, after which they recombine. Within each stream feedforward connections between lower- and higher-order cortical areas facilitate a hierarchical structure. Such organization not only allows for forward-processing, but owing to extensive feedback projections between higher and lower cortical areas, also facilitates modulation of responses in earlier areas. The organization is related to visual target acquisition by showing how each component's function and connectivity contribute to information processing during a visual target acquisition task (also see Figures 3.1 and 3.2).

2.3.1 Discussion

The retina represents the first stage of information processing. Ultimately its acts as a **receiver** and remarkable selectivity towards different kinds of visual information, facilitated by the specificity of the two main types of retinal ganglion cells, namely the Pβ- and Pα-cells, is already evident. Thus, a definite distinction between colour-related and motion-related information occurs and is amplified in later processing stages where the division between object identification features and spatial characteristics becomes more pronounced.

From the retina, information is sent forward to the LGN, which represents a transforming relay stage. The LGN is a stratified structure, with ventral layers populated by parvocellular neurons, while the dorsal layers contain magnocellular neurons, innervated by Pβ- and Pα-retinal ganglions cells respectively. Apart from retinal input, the LGN also receives input from brainstem structures such as the superior colliculus (SC) and parabigeminal nucleus (PN), which contribute to eliciting eye movements (Hubel, 1982; Cassagrande, 1994)

Output from the LGN proceeds to the early visual cortex (V1 and V2) as well as higher-order areas within the dorsal processing pathway, where it *inter alia* influences deployment of visual attention. On account of neuronal characteristics, V1 and V2 act as local analysers (Bullier, 2001) and thus function as the main **information segregators** of the visual system.

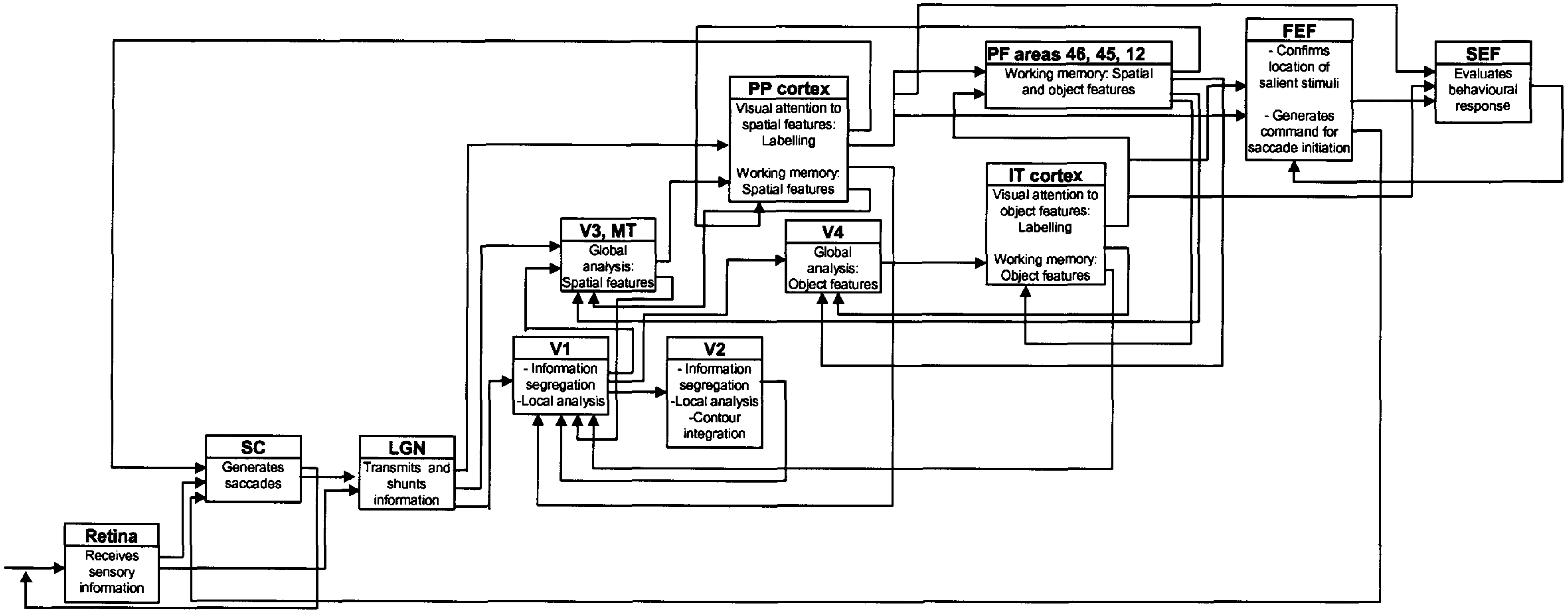


Figure 2.2: A conceptual model of information processing during a visual target acquisition task.

V1 receives converging input from both the magno- and parvocellular LGN pathways, but is neatly organized to handle the different kinds of information appropriately, as shown by several histological studies (Hubel, 1982; Livingstone & Hubel, 1984; Schiller, 1996). Via feedforward connections, information is passed on to V2 with similar organization (Hubel, 1982; Zeki & Shipp, 1988) and shunted to the correct higher-order areas for more complex analysis from there on.

Apart from its role as **segregator**, V2 is also involved with **contour integration** (Barkin et al., 2000). Resulting information can, however, be retroinjected to V1 via feedback connections. This probably serves to improve selectivity of V1 neurons (Payne et al., 1996). Local analysis done in the early visual cortex is, however, modulated by output from higher areas of the dorsal pathway, especially area MT. It has been shown that neural activity driven by magnocellular neurons of the LGN reaches V1 approximately 20 ms earlier than signals conveyed by the parvocellular pathway (Bullier, 2001). This leads to a rapid activation of the dorsal processing pathway and area MT subsequently seems to be activated almost simultaneously with V2. Because of the speed with which input reaches area MT as well as its rapid feedback, area MT is able to create a rapid **global analysis** of the visual scene (Bullier, 2001). It has furthermore been shown that without V1 input, MT neurons cannot drive V2 activity. This indicates that the feedback connections between MT and early cortical visual areas serve to enhance already present activity, rather than activating silent neurons (Hupe et al., 1998). Taken together, these findings suggest that computational output from higher-order dorsal areas is used to adjust computations in V1 and V2 dynamically.

Area V4 seems to be the functional homologue of area MT in the ventral object recognition pathway, since it integrates information on elemental features sent on from V1 and V2. Apart from its important role in colour perception, V4 has also been implicated to be actively involved in extracting form associated with colour (Zeki, 1978), intermediate shape analysis (Gallant et al., 1996; Pasupathy & Connor, 1999; Wilkinson et al., 2000) and texture segregation (Kastner et al., 2000). This can possibly be attributed to the connections between V4 and areas within both V1 and V2 that contain wavelength-specific as well as orientation-selective cells (Zeki & Shipp, 1988).

Because of the increasing convergence of information that occurs during the integration stages, a mechanism whereby irrelevant information can be ignored, and relevant information can be selected, should exist. Such **selective filtering** seems to occur at the next hierarchical level and has subsequently been attributed to visual attention. By implication visual attention, whether towards object or spatial features, thus serves to label relevant stimuli according to either identity or location, for later processing (Lamme & Roelfsema, 2000). Selective visual attention is currently believed to act according to a biased competition model, whereby stimuli within a visual scene compete for limited available processing capacity (Desimone & Duncan, 1995). At neural level, this process seems to result from the suppression of some neural responses, allowing other neurons to be primed for properties of relevant stimuli. This leads to enhanced responses being elicited when a target stimulus is detected, while activity from neurons responding to irrelevant stimuli is inhibited (Moran & Desimone, 1985; Chelazzi et al., 1993; Desimone & Duncan, 1995).

Area V4 and the IT cortex have been assigned important roles in selective visual attention towards object features, while the PP cortex is actively involved with deploying visual attention towards spatial aspects of an image (Constantinidis & Steinmetz, 2001). Activity related to keeping properties of selected stimuli available has also been observed in the abovementioned regions. Resulting computational output from these areas in turn serves as modulatory input to direct previous areas of both ventral and dorsal pathways, as well as to

V1 and V2. This contributes significantly to dynamic processing in earlier areas (Bullier, 2001).

Processed signals from the IT and PP cortices, however, not only serve to fine-tune feature analysis in lower-order cortical areas, but also contribute to the process concerned with behavioural response towards the visual stimulus. Once the current visual signals have perceptually been processed, it needs to be decided whether the information is adequate and appropriate for use in the task at hand, or whether additional data should be acquired. These decision processes and the initiation of subsequent action reside in the PFC. Dorsolateral area 46, anteroventral areas 12 and 45, the frontal eye field (FEF) and the supplementary eye field (SEF) all receive feedforward connections from IT and PP cortices (Lamme & Roelfsema, 2000). Activity in areas 46, 12 and 45 relates to working memory for spatial and object features respectively (Ungerleider et al., 1998). Location and object characteristics of the attended stimulus are thus kept available for evaluation purposes (Schall, 2001), and output is sent forward to the FEF, where it is combined with input from the IT and PP cortices. Here the combined information is used to discriminate between appropriate and inappropriate information for the task at hand and also to confirm the location of the salient stimulus, possibly facilitated by input from area 7a of the PP cortex (Constantinidis & Steinmetz, 2001). Depending on the outcome of this discrimination process, a command to elicit an eye movement is generated (Schall, 2001). The SEF, however, contributes to evaluation of the behavioural decision and can, via feedback projections, modulate activity of the FEF.

Output from the FEF is destined for SC via feedback projections (Schall, 2001) where, combined with signals from the PP cortex, it drives the generation of a saccade. This leads to new sensory information being projected onto the retina. The loop is thus closed and the entire processing mechanism is activated anew.

2.4 SUMMARY

Chapter 2 served as an introduction to the functional and structural organization of the visual system. It provided a background against which a model regarding information processing during a visual target acquisition task could be developed. In the following section findings from perceptual and computational studies will be used to discuss visual target acquisition from a theoretical perspective, showing how the perceptual process is believed to progress. It will relate particular steps during the progression of the perceptual process to the proposed model.

CHAPTER 3

VISUAL TARGET ACQUISITION

The relevance of visual target acquisition for several industrial applications has already been explained. It was followed by a discussion of the structural and functional organization of the visual system, yet one seldom stops to think how it relates to the ability to perform a visual target acquisition task so seemingly effortlessy. The objective of this chapter is therefore to discuss information processing during a visual target acquisition task from a theoretical perspective. Theories that regard visual target acquisition as a two-stage process, consisting of an early parallel processing phase, followed by a later serial-like stage, are considered in terms of findings from perceptual and computational studies. The organization of the perceptual process as presented here is followed by a discussion on the contribution from aspects related to stimulus characteristics (bottom-up influences) as well as those related to user-driven control (top-down influences). The step-wise progression of information processing will be related to the structural and functional organization of the visual system as described in the model proposed in Chapter 2.

3.1 BACKGROUND

The process during which relevant information from a visual scene is extracted and processed can be described as consisting of two functionally independent stages which are hierarchically organized (Theeuwes, 1993). Accordingly, output from the first stage serves as input to the second. This approach derives from Neisser's theory (1967) that distinguishes between an early pre-attentive mechanism and a later attention-dependent processing stage. Although the terms pre-attentive and attentive have become controversial over the years, several influential models of visual information processing assume a two-stage mechanism (Hoffman, 1978; Treisman & Gelade, 1980; Duncan & Humphreys, 1989; Wolfe, 1994), describing the first stage as a parallel process that operates across the entire visual scene, while the second phase is believed to act in a rather serial manner at subsections of the display.

In terms of visual target acquisition, both stages contribute to the efficiency and accuracy with which the relevance of an object from the display can be evaluated in terms of the current task requirements. According to the two-stage theory, the visual scene is scanned in order to locate positions that contain possible target items, followed by further analysis to determine whether an item at the indicated location agrees with target specifications. This active process during which information is gathered and processed for visual target acquisition thus constitutes *visual search*, which can be defined as the *process whereby spatial uncertainty of a scene is reduced, in order to facilitate efficient allocation of capacity-limited resources necessary to distinguish a target from non-target stimuli.*

Scanning of the visual scene precedes selection of locations eligible for further processing. Such scanning possibly represents the pre-attentive phase of visual search, while evaluation of selected information constitutes the attentive stage of visual information processing. Since this latter stage depends on the allocation of constrained processing resources, efficient visual search requires possible target locations to be identified with high certainty, so that any unnecessary resource expenditure can be prevented.

Selection serves as a filtering mechanism, which allows appropriate information to be transferred from the first to the second stage of information processing, while irrelevant information is discarded. However, selection results in a bottleneck between the two stages - items can be selected for further processing every 25-50 ms (Wolfe, 1998), but considerably more time is required to process incoming information up to the point of identification (Nakayama & Silverman, 1986; Thorpe et al., 1996; Müller-Plath & Pollmann, 2003). This delay can possibly be attributed to the fact that in order to decide on an item's status (i.e. whether it is a target or a distractor), either of two decision thresholds, representing target or distractor status respectively, should be reached. The attention-dependent recognition/identification stage is believed to progress according to a gradual information accumulation process (Wolfe et al., 2002).

3.1.1 Theories of visual information processing during visual search

During the past 25 years, much research has been conducted in an attempt to unravel the questions regarding the mechanism of visual search (e.g. Treisman & Gelade, 1980; Duncan & Humphreys, 1989; Wolfe, 1994; Wolfe, 2003). Subsequently, several theories have been proposed. Discussed below are three that have prevailed, despite several revisions, providing the most influential models according to which current thinking about visual search has been shaped.

Feature Integration Theory

One of the earliest formally described theories of visual information processing, namely the Feature Integration Theory (FIT) of Treisman and Gelade (1980), was based on the behavioural differences observed during visual perception of objects consisting of single features as opposed to objects consisting of conjunctions of features. In brief, it proposes that the nature of the mechanism responsible for perception of single features differs from the one employed during perception of conjunctions of features. According to this theory, single features, defined as particular values of a perceptual dimension (e.g. red is a feature of the dimension colour), are extracted in parallel across the visual scene during the early stage of processing, while combinations of these features require attention to be correctly perceived. This notion thus supports Neisser's (1967) theory of pre-attentive and attentive stages during visual information processing.

To evaluate their theory, Treisman and Gelade tested it according to, amongst others, a visual search paradigm, since it allows targets to be defined either by only a single feature or by a specific combination of such attributes. It was reasoned that if single features are indeed perceived in parallel, without the need to deploy focal attention, search patterns should be independent of the number of distractors in the display. On the other hand, if objects defined by combinations of single features require attention to be located, search patterns should reflect serial scanning that depends on the number of distractor items in the display.

Experiments compared search for feature and conjunction targets by relating RT to display size (set size). Subjects were asked to search for a pre-defined target amongst a set of distractors, of which the number varied during different trials. Observers were asked to respond as quickly as possible upon detecting the target. In both feature and conjunction searches, distractors were brown T's and green X's. In the feature condition, the target was either a blue letter or an S-shaped figure and thus differed from distractors only according to a single attribute. In the conjunction searches, however, the target was a green T – thus a combination of two distinctive features of the distractors. Results showed that, for feature searches, RT was independent of set size, producing functions with linear, near-zero slopes. During conjunction searches, however, RT increased linearly with increasing set size, resulting in functions with linear, positive slopes. These results were interpreted as evidence that qualitatively different mechanisms mediate feature and conjunction searches. It was suggested that when a target was defined by a single distinct feature, the entire visual scene could be searched in parallel, without necessitating visual attention, while targets defined by conjunction of features required attention to be deployed serially to several locations within the scene to be located successfully.

Further elaboration of their work (Treisman, 1985; Treisman & Gormican, 1988) resulted in the two stages of information processing to be described according to an abstract framework that assigns each feature of a perceptual dimension (as defined above) to a specific location in a perceptual space. These spaces were termed "activation maps". It was argued that during the first stage of visual processing, regions conforming to specific features within these spaces are activated in parallel across the visual field, signalling the presence of certain features. However, the relationship between separate activated regions could only be determined when attention was focussed on a specific area. Attention is therefore needed to bind separate features together accurately to form complete objects. Such binding was suggested to occur serially during the second stage of visual processing.

Guided Search Model

The Guided Search (GS) model (Cave & Wolfe, 1990; Wolfe, 1994;) originated as a corollary to Treisman and Gelade's theory. After its introduction, data inconsistent with the search patterns suggested by FIT were presented. Of special interest were cases where conjunction searches produced flat RT vs. set size functions (Nakayama & Silverman, 1986) as well as cases during which search for a specific conjunction target elicited performance patterns resembling neither strict serial nor parallel search (Cave & Wolfe, 1990).

Many of the properties of FIT are preserved in the GS model. Similar to FIT, it also involves an early parallel stage during which features are extracted along each dimension across the entire visual field, followed by a serial stage that allows intricate analysis to be performed across limited parts of the field successively. The difference, however, is that processing completed during the parallel stage assists serial search, by directing processing resources to locations with high probability of containing the target. Information extracted during the parallel stage therefore improves efficiency of the serial stage by allowing limited-capacity resources to be deployed in a structured manner, instead of the fairly random deployment suggested by FIT.

The GS model proposes that the parallel stage identifies elements in each feature dimension that resemble target specifications for that dimension most closely and that differ sufficiently from other elements in the display. Based on the featural difference between the candidate object and other elements in the display, as well as the level of similarity between the candidate object and pre-defined target specifications, each identified candidate is assigned a value that represents its target probability. Target probability is thus determined from stimulus-driven (bottom-up) as well as user-driven (top-down) input, referring to featural differences between display elements and the candidate, and featural differences between the candidate object and the specified target, respectively.

Output from the parallel stage contributes to efficient serial processing by guiding the deployment of limited-capacity processing resources to eligible locations. When the serial stage commences, it starts processing candidate locations in descending order of probability. If an element matches target specifications, search is terminated; otherwise the element is discarded and processing resources are deployed at the location with the next highest probability of containing the target. Processing efficiency is further enhanced by the fact that search need not proceed systematically through every location – only locations with associated probabilities exceeding a certain threshold are included in the list of potential target locations. Even with imperfect guidance, serial processing thus proceeds more efficiently than proposed by FIT.

Shortcomings of the original version of GS were addressed in a later revision (Wolfe, 1994). These related on the one hand to the amount in which different locations contributed to bottom-up input, and on the other hand, to the interaction between bottom-up and top-down inputs during calculation of a location's target probability. The original GS suggested that each location contributed equally to calculation of the bottom-up component, without using any information regarding spatial relationships between different locations. In the revised model, however, the strength with which a certain location influenced target probability of another location was also considered in terms of its distance from that location, so that nearby locations influenced probability values more than more distant locations. Combination of bottom-up and top-down components was also revised – in the new version the activation value was computed as a weighted average, with relative weight based on task demands, instead of equal contribution from bottom-up and top-down components as proposed in the original version.

Attentional Engagement Theory

The Attentional Engagement Theory (AET) of Duncan and Humphreys (1989) has a similar approach to visual information processing as the models discussed earlier, and in fact shares many of the qualities of the GS model (Cave & Wolfe, 1990). It also proposes an early parallel component that operates across the entire visual field, followed by a capacity-limited mechanism that only acts in designated parts of the visual field. As with the other models, early processing generates a structured representation of the entire visual field, from which areas worthy of further analysis can then be selected. If stimuli within these areas meet task requirements, they are used to drive behavioural responses.

Duncan and Humphreys (1989) proposed that during the first stage of processing, the entire visual field is segmented into structural units. Such segmentation involves perceptual grouping of items, which in turn is influenced by the amount of similarity as well as the spatial separation between items. Each structural unit is thus associated with a set of descriptive properties, which serves as input to subsequent processing stages.

Once a description of the visual scene has been generated, a stage of comprehensive analysis, during which input is evaluated in terms of task demands, ensues. By implication, input should be matched to predefined specifications. This stage is governed by processing resources of limited capacity, which can only be allocated to units that exhibit high target probability. Each unit thus has to compete to be selected for further processing, with only those contending strongest being considered. According to AET, each segment is therefore associated with some selection "weight" which is proportional to the degree of match between segment properties and those of the target.

Duncan and Humphreys further proposed the concept of weight linkage, which refers to the association of selection weights of units in the same perceptual group. It was argued that perceptual grouping results in selection weights being relative entities, so that any change in selection weight for a specific unit will alter selection weights of all linked units proportional to the strength of the association between them. It follows that weight linkage contributes to search efficiency, especially through efficient rejection of inappropriate items. Since linked units' weights will be reduced accordingly if one specific segment's weight is reduced, the possibility of any of these being selected for further processing is diminished. This allows processing resources to be allocated economically to possible target locations.

Such an approach strongly suggests that the degree of similarity between target specifications and those of non-target elements (distractors), as well as the level of similarity between various distractors themselves, substantially influence search efficiency. Moreover, the interaction between target-distractor and distractor-distractor similarities determines the nature of the search pattern. They therefore proposed that search efficiency should be modelled along a continuum, with no strict division between parallel and serial search behaviour.

Current views

Current theories agree that visual information processing functionally fits a two-stage model with serial as well as parallel components (Theeuwes, 1993; Thorpe Davis et al., 2003; Wolfe, 2003; Müller-Plath & Pollmann, 2003). The first stage is considered as preattentive and seems to proceed without processing constraints across the entire visual field. It generates a crude representation of the entire visual scene and identifies locations with high target probability, which in turn serve as input to the second processing stage (Wolfe, 2003). The second processing stage, however, involves sequential deployment of limited-capacity processing resources. Selection facilitates effective allocation of these resources by providing the second processing stage with a manageable subset of information extracted during early processing. Within these subsets, processing resources can be deployed in parallel, so that a number of items can be evaluated simultaneously during a single allocation of resources (Müller-Plath & Pollmann, 2003). However, such information subsets are provided sequentially and therefore processing resources can only be deployed sequentially. This constitutes the serial component of visual information processing.

The current view of visual information processing thus encompasses ideas from all three theories discussed earlier and presents a general mechanism according to which visual information processing proceeds. As shown recently, serial search and parallel search are not qualitatively different (Müller-Plath & Pollmann, 2003). Visual search should thus be considered as proceeding along a continuum, with search efficiencies that resemble strictly serial or parallel mechanisms representing the extremes of such a continuum.

3.2 VISUAL SEARCH MECHANISMS

3.2.1 The serial/parallel dichotomy

The nature of the search mechanism employed during a visual search task is often investigated by manipulating display properties in a manner that allows temporal parameters to be used to express search performance. During RT studies, the slope of the function relating RT to the number of items in the search display (set size) is used to determine the nature of the search strategy evoked by the search display. Apart from RT studies, accuracy methods are also used to infer search strategies (Wolfe, 1998). Typically, the search display is presented only briefly and then replaced by a mask, which serves to terminate the search. The time between the onset of the stimulus and that of the mask is varied and accuracy is consequently expressed as a function of stimulus onset asynchrony (SOA).

Early studies on visual search performance regarded functions with near-zero slopes as indications of parallel search, while those with steep, positive slopes were taken as support for serial capacity-limited search (Neisser, 1967; Treisman & Gelade, 1980). Moreover, parallel search was believed to proceed independently of attention, while serial search was thought to rely strongly on the deployment of visual attention. This led to the inception of the serial/parallel dichotomy, which proposed that the nature of the search display determines the employed search mechanism. Furthermore, only one of either mechanism was believed to be employed throughout an entire search task. Support for such a serial/parallel dichotomy was provided by findings that feature searches (where the target differs from distractor objects only according to a single attribute) usually produced RT vs. set size slopes indicative of parallel search, leaving the observer with a feeling that the target had "popped out" from the background. Searches where the target was a combination of distractor attributes (e.g. finding a red vertical object among green vertical objects and red horizontal items), on the other hand, were usually perceived as "difficult" by observers and produced RT vs. set size slopes indicative of serial processing. Feature searches were therefore often labelled as parallel, while a conjunction search was seen as a serial process.

Function slopes should, however, be interpreted with caution, and preferably not in isolation (Cheal & Lyon, 1992). Studies conducted during the past 15 years have proposed that search occurs along a continuum, rather than according to a strict division between serial and parallel search (Duncan & Humphreys, 1989; Cheal & Lyon, 1992; Nothdurft, 1992; Wolfe, 1998; Müller-Plath & Pollmann, 2003). Further evidence against a strict distinction is provided by findings that serial search can become parallel and vice versa, depending on the search conditions and task demands (Nothdurft, 1993a; Nothdurft, 1993b). Pure pop-out and difficult search behaviours thus appear to represent the extremes along such a continuum.

These findings imply that some kind of interaction or overlap should exist between the different search modes. Evidence in support of such a hypothesis was recently provided when

Olds and colleagues (Olds et al., 2000; Olds & Punambolam, 2002) conducted studies in which they interrupted initial pop-out search tasks by changing the search display shortly after onset from a pop-out display to a difficult display. Tasks could thus only be completed by mechanisms mediating difficult search. Results showed that partially completed pop-out search assists difficult search, possibly by transferring spatial information. It was found that initial pop-out mechanisms extracted sufficient information to either steer subsequent processing *towards* locations with a high possibility of containing the target, or to prevent locations *unlikely* to contain a target, or where it has already been determined not to be, from receiving further processing. These findings are in line with theories (discussed earlier) which propose that early parallel search guides allocation of processing resources during serial search.

Support for possible interaction between the mechanisms mediating pop-out and difficult search stems from studies that found attention to be an important contributor to search performance in both search modes. Nothdurft (1999) found that both search strategies involve dynamic attentional control and that different performance trends observed for pop-out and difficult search possibly result from the extent to which the search conditions allow this attentional control to exert its influence. Salient targets immediately attract attention (Wolfe, 1994; Nothdurft, 1999) and thus, if there is only a single salient object in the display (i.e. pop-out conditions), attention will be allowed to exert maximum processing control at this position. However, if there are various items of similar saliency, several positions will be equally eligible for attentional control, resulting in the performance trends characteristic of difficult search. A study aimed at determining the subprocesses of visual feature search (Müller-Plath & Pollmann, 2003) showed that items from a visual image are scanned in groups, of which the size is determined by the strength of attentional control. Within these groups, all items are processed simultaneously, so that attention is immediately directed to the target if it is located within that group. If none of the items of that group coarsely fit target requirements, attention is moved to the next group of items, resulting in a serial search.

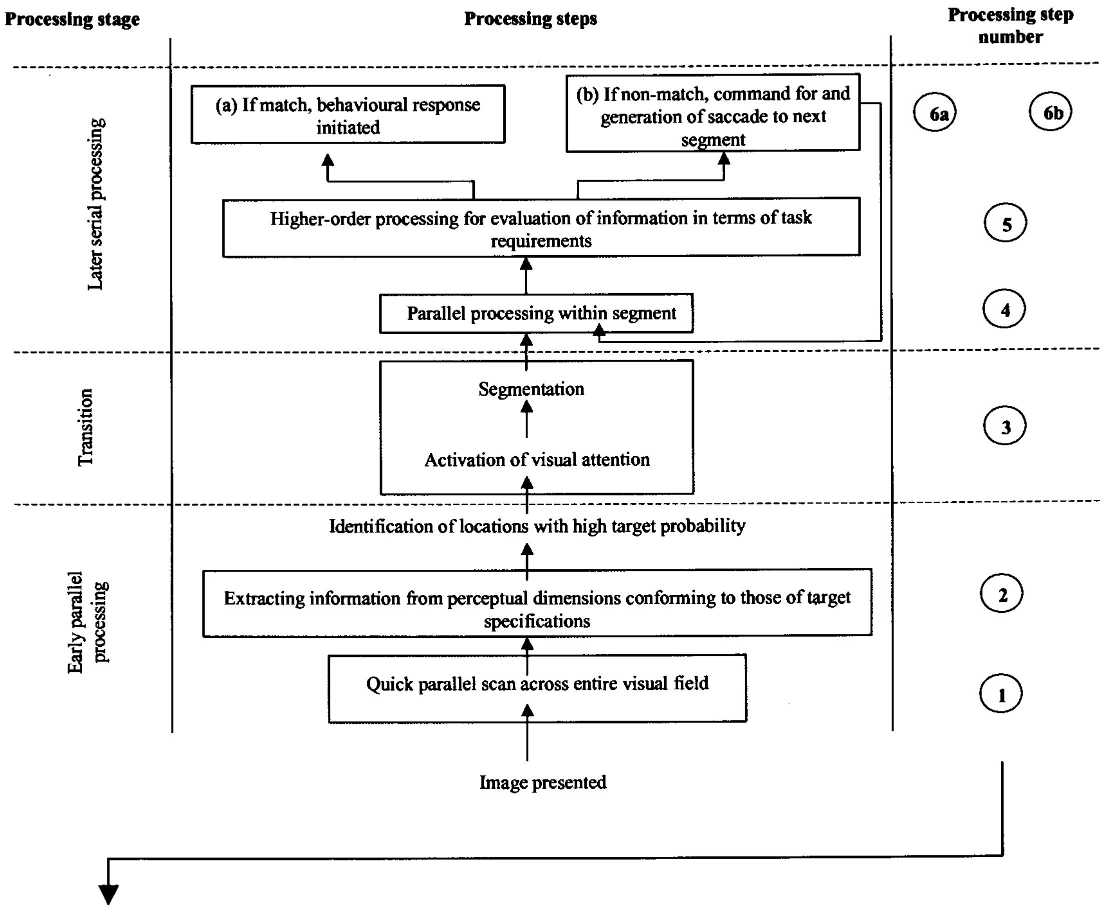
Taken together, these findings argue against a search task being governed by strictly serial or parallel mechanisms, but rather suggest that both operate at different stages during a search task. The extent to which they interact, however, is determined by display characteristics, with the predominant mechanism being reflected in behavioural observations.

In terms of the model presented in Chapter 2, it is possible that neural activity at the level of area MT is responsible for the initial scanning of the entire visual field, mediated by feedforward connections from the LGN. Simultaneous to MT activation, some featural analysis already occurs at the level of areas V1 and V2, although processing efficiency is enhanced by input from MT facilitated by feedback connections. Locations with high target probability can consequently be identified, possibly facilitated by feedback received from prefrontal, posterior parietal and inferotemporal cortices, which are involved in working memory. These steps possibly represent the early parallel processing stage of visual search.

It is followed by activation of visual attention, which results in a section of the image within the identified region being segmented. These steps appear to represent the platform for interaction between the mechanisms involved in early parallel processing and later serial processing. It is possibly facilitated by neural activity at the level of V4 and PP and IT cortices, owing to their involvement in visual attention. The search trends observed for a given visual search task will thus reflect the extent to which neural activity from these structures contributed to completion of the task.

Once segmented from the image, parallel processing ensues. As during the early parallel processing phase, a global impression of that specific segment is created at the level of MT. The information can then be fed back to early cortical areas which have already started extracting elementary information. Analysed information is gradually fed forward to higher-order areas such as the PFC, FEF and SEF, where it is evaluated in terms of the target description specified at the onset of the task. If the acquired information matches the target description, a behavioural response can be initiated by activation of cortical areas responsible for planning and execution of motor actions. If, however, the acquired information does not correspond to the target description, a command for and generation of a saccade to the next segment commences, facilitated by neural activity at the level of the PFC and SC respectively. The eye thus moves onto the location of next highest target probability, as determined during the early processing stage, where processing of the new segment starts from anew. This stage represents the serial stage of visual search, since segments from the visual scene are processed sequentially until information corresponding to the target description is found. It should be noted that although this processing mechanism appears very systematic, processing in fact occurs in a massively parallel manner, as is evident from the numerous feedforward and feedback connections between the different cortical and subcortical areas (Figure 2.2).

The mechanism can be summarized as in Figure 3.1. The top section illustrates the progression of information processing steps followed by an explanation in terms of the conceptual model proposed earlier and repeated here as Figure 3.2.



Processing step	Description in terms of conceptual model neural organization
1	Quick parallel scanning of entire visual field possibly results from neural activity at the level of area MT. It is mediated by feedforward connections from the LGN.
2	Simultaneous to activation of area MT, activity in V1 and V2 facilitates featural analysis. Processing efficiency is enhanced by feedback from area MT, ensuring dynamic adjustment of featural analysis. Neural activity possibly results in extraction of information from perceptual dimensions conforming to those of target specifications. Feedback from PFC, PP and IT cortices allows for locations of high target probability to be indentified, owing to their involvement in working memory. Steps 1 and 2 therefore possibly represent the early processing stage during a visual target acquisition task.
3	Segmentation possibly results from activation of visual attention, owing to neural activity in V4, PP and IT cortices. These areas not only receive input from both lower and higher-order cortical areas, but also send output to lower and higher-order areas. Activation of visual attention therefore appears to represent the transition between the two processing stages of a visual target acquisition task.
4	Parallel processing occurs within each image section segmented earlier. As with early parallel processing, it possibly results from neural activity in area MT., which sends output to early cortical areas, which are already performing featural analysis.
5	Analysed information is gradually fed forward to prefrontal area where it is evaluated in terms of the target description specified at onset of the task.
6a	If the acquired information corresponds to the target decription, output is sent forward to cortical areas responsible for planning and execution of motor actions.
6b	If the acquired information does not match target specifications, a command for and generation of a saccade to the segment with next highest target probability commences. This is facilitated by activity in the PFC and its output to the SC. This represents the serial component of processing during target acquisition, as the loop will continue to execute until an object matching target requirements is found.

Figure 3.1: Diagrammatic representation of the mechanism of information processing during a visual target acquisition task. Numbered circles in the top section correspond to specific processing steps, which are explained in the lower section, in terms of the structural and functional organization of the visual system proposed earlier.

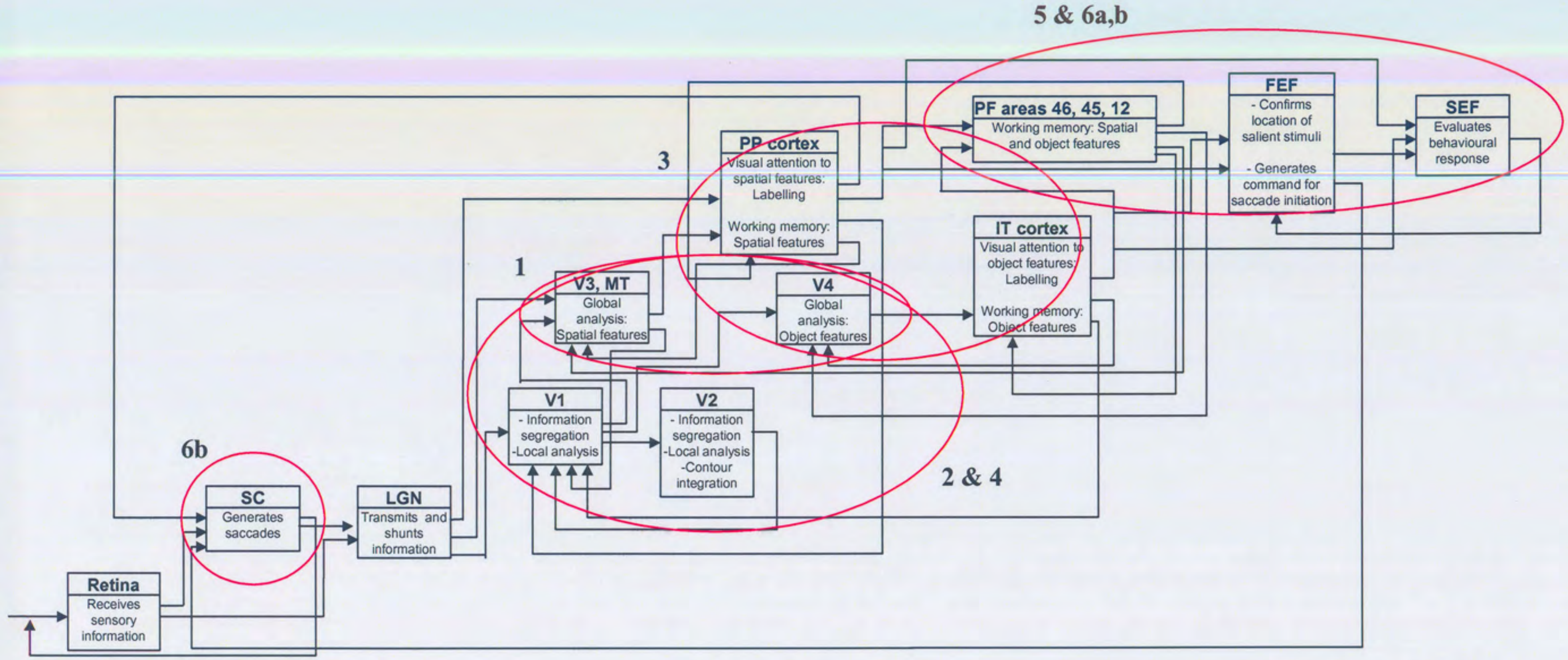


Figure 3.2: A conceptual model of information processing during a visual target acquisition task, showing the functional relevance of specific anatomical regions. Encircled areas correspond to processing steps explained in the lower section of Figure 3.1.

3.3 FACTORS INFLUENCING SEARCH BEHAVIOUR

Recent investigations into their respective contributions showed that bottom-up and top-down processes interact in order to produce optimal visual search performance (Wolfe et al, 2003). The bottom-up component of a feature search task was isolated by randomly varying either target feature or target dimension from trial to trial. Observers were subsequently instructed to search for the odd item within the display. Random change of target characteristics as well as the fact that target characteristics could become distractor characteristics on subsequent trials introduced substantial uncertainty about target identity and ensured that the search for specific items, which would skew results, could be eliminated. Results showed that bottom-up guidance was adequate to complete the tasks in the absence of top-down information, but that search efficiency was impaired. Further evidence for the interplay between the two types of information derives from observations where decreasing featural contrast between target and distractors reduced the contribution of bottom-up information. Results from one set of experiments where red targets of varying orientation and size were placed among heterogeneous green distractors were compared to those where the same targets were placed among heterogeneous orange distractors. It was found that the less distinct the difference between targets and distractors, the harder the search task became, possibly because of the increased difficulty of deciding on the identity of the display items. Even though sufficient top-down information was available, it was of little use since the bottom-up contribution was too weak (Wolfe et al., 2003).

3.3.1 Bottom-up information during visual search

A number of factors combine to produce the overall bottom-up contribution from a certain display. The following section will discuss the influence of stimulus similarity, eccentricity effects and spatial interactions on search behaviour.

Stimulus similarity

A vertical line segment among horizontal line segments of the same length is easily detected (Figure 3.3a). Searching for the same vertical line segment among line segments of which some are orientated horizontally, some tilted 45 degrees to the left and some tilted 45 degrees to the right, however, becomes more difficult (Figure 3.3b). Similarly, searching for a vertical line among line segments, which are all just slightly tilted, is effortful (Figure 3.3c) This effect can be attributed to stimulus similarity. It is evident from the example that stimulus similarity refers not only to target-distractor similarity, but also to similarity between distractors.

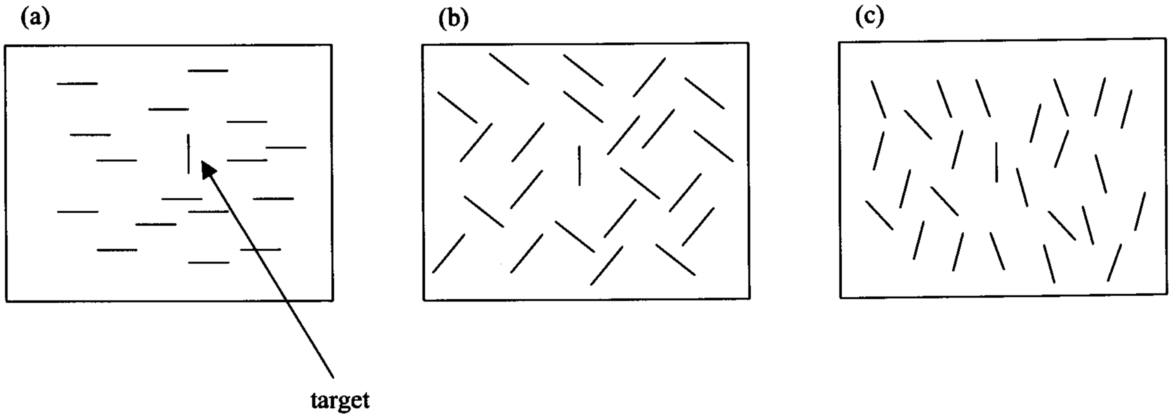


Figure 3.3: Examples of search displays of varying difficulty. The target is always the vertical line at the centre of the display.

Duncan and Humphreys (1989) investigated the effect of stimulus similarity on search behaviour and found a relationship between search efficiency and target-distractor similarity as well as distractor-distractor similarity. They subsequently summarized the relationship as a three-dimensional "search surface", shown below in Figure 3.4.

Distractor-distractor similarity and target-distractor similarity are plotted on the z- and x-axes respectively. Search efficiency, as indicated by the slope of the RT vs. set size function, is represented on the y-axis. According to this proposal, search will always be efficient when the target and distractors are highly dissimilar, regardless of distractor-distractor similarity. Search will also be efficient when distractors are homogeneous, with target-distractor similarity having only a marginal impact on search behaviour. However, when distractor-distractor similarity decreases simultaneously to increasing target-distractor similarity, search efficiency is affected adversely.

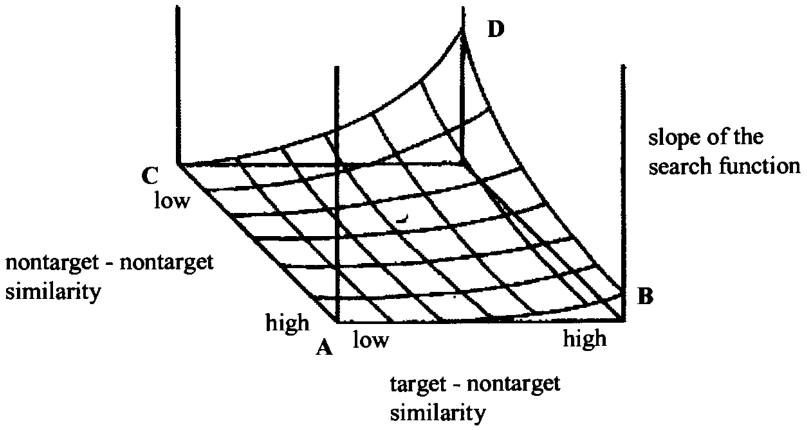


Figure 3.4: Duncan and Humphreys' "Search surface", describing search efficiency in terms of target-nontarget similarity and nontarget-nontarget similarity (With permission from Duncan & Humphreys, 1989).

This relationship has been attributed to feature contrast (Nothdurft, 1993a; Nothdurft, 1993b). Feature contrast advocates that salience is not derived from a specific feature per se, but rather from the context of that feature within the display. This might explain why feature searches are generally more efficient than conjunction searches – during a feature search, the relative difference between target and distractor features is large, whereas during conjunction search the relative difference is diminished owing to greater heterogeneity of background items. Furthermore, it has recently been shown that salience from feature contrast adds across dimensions, but that the resulting salience increase is not linear (Nothdurft, 2000). These findings suggest that separate yet partially overlapping mechanisms underlie saliency perception. This may also be responsible for search asymmetry effects.

Eccentricity effects

Much of the initial work on visual search behaviour was done without taking the effect of target location into account. The usual design of search experiments involves targets appearing at random positions across the display. During analysis, RT from all these positions is averaged and expressed as a function of set size. However, various studies have shown that it is more difficult to accurately detect and localize targets that appear in the periphery of the visual field, opposed to when they appear at or near the fixation point (Carrasco & Chang, 1995; Carrasco et al., 1995; Wolfe, 1998; Scialfa & Joffe, 1998). This has been termed the eccentricity effect and has been documented for feature as well as conjunction searches by Carassco and colleagues (Carrasco & Chang, 1995; Carrasco et al., 1995).

Early investigations into eccentricity effects showed that set size alone cannot explain RT results as determined during search studies; set size effects should also be considered in terms of target eccentricity. Moreover, eccentricity effects seem to become more pronounced as set size increases. It was therefore argued that typical set size effects observed during RT studies are in fact due to an interaction between target eccentricity and set size (Carrasco & Chang, 1995; Carrasco et al., 1995).

A systematic exploration of eccentricity effects pointed to visual factors such as spatial resolution (Carrasco et al., 1995; Carrasco et al., 1998) and lateral interactions (Carrasco & Chang, 1995; Carrasco et al., 1995) as the main determinants of eccentricity effects. Spatial resolution decreases from the fovea towards the periphery of the retina, possibly owing to a decreasing number of receptor cells and increasing receptive field sizes towards the periphery. Therefore, when a target appears among distractors, it is subject to more lateral interference at large eccentricities than near the fixation point. Support for the visual account of eccentricity effects was also gained from a study during which a cortical magnification factor was employed to equate spatial resolution. According to this factor, stimulus dimensions can be scaled so that the amount of cortical activation evoked by a stimulus is equated, irrespective of retinal eccentricity (Carrasco & Frieder, 1997). It was found that when stimuli were scaled according to this factor, eccentricity effects were effectively neutralized (Carrasco & Frieder, 1997).

Wolfe and colleagues proposed an alternative to a purely visual account of eccentricity effects (Wolfe et al., 1998). Results from their studies showed that when display conditions such as set size, the level of interference from distractors and organization of display characters are equal, a significant eccentricity effect is observed even when no distractors flank the target. This indicates that with all other display conditions being equal, items are processed in order of increasing eccentricity. Furthermore, scaling stimuli according to the cortical magnification factor produced variable results depending on whether eye movements were allowed or not, indicating that cortical magnification does not reliably explain eccentricity effects. Further experiments showed that when target and distractors appeared at equal eccentricities, search performances were similar at various eccentricities, irrespective of whether the stimuli appeared close to or far from the fixation point. On the other hand, search performance was worse when the target was farther from the fixation point than distractors, than when the target was closer to the fixation point than distractors. This indicates that eccentricity exerts its influence relative to the fixation point, rather than according to absolute target position (Wolfe et al., 1998).

In light of these findings, eccentricity effects can rather be explained in terms of an attentional account, which proposes that eccentricity of stimuli modulates allocation of attention, so that central stimuli will be processed earlier than more peripherally located items (Wolfe et al., 1998). Such an approach does not negate the contribution of visual factors proposed previously (Carrasco & Chang, 1995; Carrasco et al., 1995; Carrasco & Frieder, 1997), but suggests that they should be considered as contributing to guiding attention for efficient allocation of processing resources.

Spatial interaction

Target salience should not be considered only in terms of feature properties, but also in terms of spatial organization of the display (Casco & Campana, 1999; Olds et al., 1999). One of the most important determinants is inter-element distance – i.e. the distance between target and neighbouring distractors, as well as the distance between adjacent distractor items (Kooi et al., 1994). It has been shown that the ability to identify a target accurately deteriorates in a cluttered environment, especially when the separation between flanking objects is small (Bachall & Kowler, 1999). This effect is referred to as "crowding", "spatial interaction" or "contour interaction" (Tripathy & Cavanagh, 2002).

The degree of spatial interaction is in turn influenced by display properties. Recent studies found that crowding impairs search performance to a greater extent at large eccentricities than at positions close to fixation. Target size also affects the strength of spatial interaction– large targets can be accurately identified at small separations, as opposed to smaller targets that require larger separations to achieve the same level of performance, although this relationship is not proportional (Tripathy & Cavanagh, 2002). Other factors that affect the extent to which spatial interaction influences search performance are feature similarity between target and distractors (Kooi et al., 1994), stimulus duration (Kooi et al., 1994; Tripathy & Cavanagh, 2002) and background organization (Olds et al., 1999; Casco & Campana, 1999). It has been found that with increasing similarity between target and nearby distractor items, the crowding effect becomes more pronounced, whereas increased stimulus duration weakens its impact. With regard to the influence of background configuration on the degree of spatial interaction, it has been shown that regularity of the background often increases target saliency, possibly by facilitating effortless target segmentation (Casco & Campana, 1999).

Crowding can best be explained in terms of an attentional framework. According to such a proposal, crowding exerts its effect at the level of target selection, with attentional resolution that is much coarser than visual acuity. This is in line with findings that attentional selection creates an inhibitory surround around the selected item (Caputo & Guerra, 1998), which in effect serves as the minimum area available for selection. This region has a specific constant size at different eccentricities (Tripathy & Cavanagh, 2002), so that it becomes increasingly difficult to attend to two items as the spatial separation between them decreases, even though they are still within the limits of visual acuity. Preferential processing of the selected item is thus ensured by simultaneously enhancing processing at the chosen location and suppressing processing of nearby distractors (Bachall & Kowler, 1999).

3.3.2 Top-down contributions to visual search

Apart from bottom-up information derived from the image itself, successful visual search also requires top-down guidance. Top-down information allows user-driven control during the search task and can therefore contribute to faster and more efficient search, by modulating processing of bottom-up information. Two of the most pronounced top-down influences during visual search are selective attention and memory. Since selective attention and visual search are so intertwined, the role of visual attention during the search process has already been discussed extensively. The following section will therefore only discuss memory as a top-down influence during visual target acquisition.

Memory

The role of memory during successful visual search has been controversial. Some researchers claim that visual search proceeds independent of memory-driven guidance (Horowitz & Wolfe, 1998), while others suggest that memory contributes to efficient search by preventing re-allocation of attention to previously examined (and rejected) locations (Klein, 2000; Kristjansson, 2000; Paul & Tipper, 2003). Current evidence suggests, however, that mnemonic strategies do contribute to visual search and that their benefit can be explained according to a model somewhere between a perfect-memory model and a no-memory model (Gilchrist & Harvey, 2000; Wolfe, 2003).

Evidence for memory-related influences on visual search are often inferred from studies examining inhibition of return (IOR) or eye-movement patterns. IOR is a mechanism that prevents re-examination of a location already attended to up to 300 ms earlier (Klein, 2000) by inhibiting a return of attention to that location. It functionally contributes to efficient visual search by enabling the visual system to keep track of locations already examined and found to be task-irrelevant. Such tracking ability seems to be limited to four or five previously attended loci (Snyder & Kingstone, 2000). A recent study by Paul and Tipper (2003), however, showed that IOR not only contributes to spatial memory, but also facilitates feature memory. Studies regarding eye movement patterns further support the notion of some degree of memory assistance during visual search – it was shown that refixations at previously examined locations were fewer than predicted by a memoryless model (Gilchrist & Harvey, 2000; Peterson et al., 2001).

Although many unresolved issues remain, it seems that memory definitely does have a role during successful visual search. Given that limited processing capacity is available for an immense amount of bottom-up information, it appears that any top-down influence that prevents unnecessary and inefficient allocation of processing resources would be beneficial to the search process.

3.4 SUMMARY

Chapter 3 dealt with the perceptual organization of information processing during a visual target acquisition task. It was discussed in terms of a two-stage model, where the first stage operates in a parallel manner across the entire visual scene followed by a later, serial processing stage during which segments from the scene are processed sequentially in order of their probability of containing the target. Segmentation and selection involve activation of visual attention and constitute a transition between the two stages. The relationship between the perceptual organization of the process and the organization of the visual system was illustrated according to the model proposed in Chapter 2, showing that behavioural outcome of a visual target acquisition task derives from an organized perceptual system, which is the result of structural and functional organization at neural level. Contributions from bottom-up and top-down information were explained according to this framework.

Chapters 2 and 3 provided the reader with background according to which the relationship between behavioural outcome of a visual target acquisition task and the visual system's structural and functional organization can be understood. Coming chapters will deal with experimental application of these principles. Chapter 4 will deal with the approach to the study. It will introduce tools used during the empirical application as well as showing how experimental design and methodology were optimized prior to the final experiments.

CHAPTER 4

APPROACH

The previous two chapters presented a background against which the relationship between the structural and functional organization of the visual system and information processing during a visual target acquisition task can be understood. This provided the framework for presentation of the empirical stage of the study (Chapters 4 and 5). Chapter 4 deals with the approach according to which the empirical applications were designed and developed. It introduces the reader to the principles of signal detection theory as applied in psychophysics, in order to explain the rationale behind using measures of search performance such as conspicuity area and RT during investigation of visual target acquisition behaviour. It is followed by a discussion of the optimization of experimental design and measurement techniques pertaining to the study.

4.1 INTRODUCTION

Much of the research in cognitive neuroscience aims to describe progression of a perceptual process, from reception of the signal down to when a behavioural response is made. In order to create such models, scientists often work "from the outside in" - collecting behavioural data which can then be used to infer the underlying mechanism of a certain perceptual process. Psychophysical techniques have proved effective in this regard, since they allow a researcher to monitor behavioural outcome of a perceptual process and how it is affected by changing stimuli or task requirements non-invasively. It therefore provides an "observation window" into the mechanism underlying a perceptual process, which in turn is a function of neural organization.

Psychophysics is based on signal detection theory, which combines principles from mathematical statistics used in probability theory and signal detection as applied in communication systems, to determine the probability of an observer deciding on either of two response options (i.e. "signal detected" or "signal not detected") presented by a perceptual task. The focus is therefore on investigating the reason as to *why* that specific choice was made for the specific presentation, rather than determining *how* the observer perceived the presented stimulus. It thus provides an objective manner in which the effect of changing stimulus characteristics on stimulus detectability can be determined.

Given that stimulus detectability can be determined objectively, it also needs to be expressed in a way that allows easy reporting and comparison of results. In visual psychophysics conspicuity area determination is a valuable tool to this end, since the conspicuity area represents the limit of peripheral target detection for a given target-background combination.[2] In terms of signal detection theory, its border can be seen as the division between the two decision options and conspicuity area can therefore be used to translate abstract detection probability into an objective and more comprehensible measure of behavioural outcome during a perceptual task, allowing stimulus detectability to be easily reported and compared across different stimulus conditions.

Similar to conspicuity area determinations, application of the principles of signal detection theory also renders RT measurements a valuable and objective performance marker of behavioural outcome of a perceptual process. As already shown in Chapters 2 and 3, structural and functional organization of the visual system is closely related to perceptual organization of a visual target acquisition task. It follows that any task manipulation, whether

---

[2] Conspicuity area is discussed comprehensively in a following section.

it relates to bottom-up or top-down information, that will affect information progression along the processing route will be reflected in recorded RT (Luce, 1986). Although RT measurements only allow for gross organization of the process to be inferred, the results become valuable when interpreted together with findings from other techniques, such as conspicuity area determinations, especially for developing models of the information progression underlying a perceptual process (Smith, 1995).

4.2 PSYCHOPHYSICS AND SIGNAL DETECTION THEORY

According to signal detection theory, a sensory event arises either from a signal presented within background noise or from background noise alone (Swets, 1964; Gelfand, 1990). Each hypothesis is associated with a specific probability of being true and, depending on the amount of bottom-up and top-down information available, the observer has to set a criterion according to which he will decide to either accept or reject one of the hypotheses. The observer may respond that a signal was present or that it was absent, which can be either a correct or incorrect answer for the specific event. Therefore, a sensory event will have any of four possible outcomes, each associated with a specific conditional probability: hit, miss, correct rejection or false alarm. Graphically, the situation can be portrayed as probability density functions (Figure 4.1), where the abscissa represents the energy contained in the stimulus, while the ordinate indicates the probability associated with the sensory event arising from either of the two possibilities mentioned above.

It follows that the greater the energy difference between the two possibilities, the lower the sensitivity required to distinguish between the two situations. The level of sensitivity required,

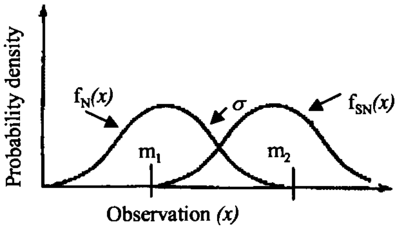


Figure 4.1: Example of probability density functions for events arising from noise alone ($f_N(x)$) and signal presented within noise ($f_{SN}(x)$). The means of the functions when the signal is absent (m_1) and when the signal is present (m_2) are shown. The standard deviation is represented by σ.

expressed as the parameter d', is represented by the separation between the two functions and can be calculated by dividing the difference between the mean energy of the two distributions by the standard deviation (Swets, 1964; Gelfand, 1990):

$$d' = \frac{m_1 - m_2}{\sigma} \quad , \qquad (1)$$

with m = mean energy of distribution and σ = standard deviation of mean.

As can be seen from Figure 4.1, the more overlap between the two functions, the smaller d', resulting in a greater likelihood of an erroneous response.

The most common paradigms during psychophysical studies include those using yes/no methods, two- (or multiple-) interval forced choice methods (2IFC or nIFC) and confidence rating methods (Green & Swets, 1966). All three approaches employ the basic principles of signal detection theory, but differ slightly with regard to stimulus presentation and/or response options. During a yes/no task, a single observation interval is presented, which may or may not contain the signal. This is followed by a response interval during which the observer has to indicate whether or not the signal was present. With the 2IFC procedure, two observation intervals, with one always containing the signal, are presented. During the response interval the observer has to indicate during which observation interval he observed the signal. In both cases, it is the experimenter's choice whether or not to provide feedback after a response was made. Rating methods employ the same sequence of events as the yes/no task, but differ in the nature of the observer's response. Instead of a binary response, a number of response options are available. Each response option is associated with a certainty level, with absolute certainty that the signal was present at one extreme and absolute certainty that it was absent at the other extreme (Green & Swets, 1966). Of course every procedure has its own advantages and disadvantages and procedure choice is usually determined by time constraints, required accuracy level and ultimate aim of the study. This issue is dealt with in more detail elsewhere (Green & Swets, 1966), but falls beyond the scope of the current discussion.

4.3 VISUAL CONSPICUITY

Engel (1971) operationally defined visual conspicuity as "that combination of properties of a visible object in its background by which it attracts attention via the visual system, and is seen in consequence." This agrees with subsequent studies that show that target-distractor as well as distractor-distractor similarity influences search efficiency (Brown & Monk, 1975; Monk & Brown, 1975; Jenkins & Cole, 1982; Duncan & Humphreys, 1989). Therefore, the more conspicuous a target, the easier it will "pop out" from its background and the less effort will be required to detect it. It thus follows that search efficiency can be used to determine the visual conspicuity of a certain target-distractor combination, which in turn can be used as a measure according to which a visual image can be classified in terms of detectability.

4.3.1 The relevance of the conspicuity area and the relationship to signal-to-noise ratio

The detectability of a visual target depends on its conspicuity. Conspicuity, however, is an abstract concept and needs to be expressed in a way that is easy to interpret and comparable across several target-background combinations. Determining the conspicuity area, also termed visual lobe size or functional field of view, answers to these requirements (Engel, 1971; Bellamy & Courtney, 1981; Courtney & Chan, 1985b; Chan & Courtney, 1996; Toet et al., 1998), since it represents the limit of peripheral sensitivity for a specific target-background combination. The rationale is that the size of the conspicuity area expresses the probability of detecting the target during a single glance when a central point is fixated (Engel, 1971). The smaller the area around the fixation point, the less conspicuous a target is in that specific background and the more effortful search will become. Depending on the goal of the search task, either background or target properties can then be manipulated to alter the conspicuity of the target.

In terms of signal detection theory, determination of the conspicuity area requires that the detection threshold for a target within a specific target-background combination be found. As explained previously, signal detection theory advocates that a sensory event arises either from a signal embedded in noise, or noise alone. Both components add to the sensory impact of the event and the ratio (signal-to-noise ratio) according to which they contribute determines the relative detectability of the signal. It follows that for a sensory event with a high signal-to-noise ratio (SNR), the signal will be more detectable than for an event with a low SNR.[3]. During a visual search and target acquisition task, the target serves as signal, while the

---

[3] For a detailed discussion of SNR, refer to Green & Swets, 1966.

distractors constitute noise. By determining the conspicuity area, the spatial extent to which the target is detectable in a particular background is determined. That is, the conspicuity area serves as a visual representation of the detection threshold for that particular target-background combination. The conspicuity area can therefore serve as a measure, akin to SNR, according to which the relative sensory contributions from target and distractors of a given image and how they influence target detectability can be reported and compared.

4.4 RESPONSE TIME DETERMINATION DURING STUDIES OF A PERCEPTUAL PROCESS

The fundamental purpose of RT tasks during perceptual studies is to determine how fast a subject can initiate and execute a simple, predetermined response to a simple triggering signal (Luce, 1986). At the core is the assumption that observed RT is the sum of two different times, namely decision time and residual time (Luce, 1986). Decision time refers to time required for evaluating the internal representation of the input signal in order to decide whether the signal was presented. Residual time represents all the time that elapses after the decision on presence of the signal was made until the response has been executed, and is believed to contribute little to the overall variability in simple RT tasks (Smith, 1995). For visual perception such serial decomposition is plausible, considering the general hierarchical character of the structural and functional organization of the visual system and visual target acquisition process, as discussed in Chapters 2 and 3. However, despite the essential hierarchical organization of a visual target acquisition task, RT results should be interpreted and applied with great care, since some temporal overlap, possibly facilitated by feedback connections, does exist between specific functions within the system. This is shown in the model and discussions presented earlier.

4.4.1 Relationship between RT and visual conspicuity

Applications of visual search and target acquisition tasks, such as industrial inspection, monitoring of control panels and military target acquisition, often require tasks to be completed in a predetermined time. Successful completion of such tasks depends to a large extent on the conspicuity of the target in its background, which, as already discussed, can be expressed in terms of the target's conspicuity area. When results from conspicuity area studies were related to RT measurements, a strong inverse relationship was observed between mean search time (expressed as RT) and size of the conspicuity area (Engel, 1977; Bellamy & Courtney, 1981; Kee et al., 1992). Toet and colleagues (Toet et al., 1998) subsequently showed that mean search time for a specific target-background combination can be predicted once the target's relative conspicuity area has been determined. Combined results from conspicuity area determination and RT measurements can therefore be valuable for evaluation of a subject's search and target acquisition performance, design of visual environments where rapid detection of visual signals is important, as well as making a contribution to the generation of models of the underlying processing strategy.

In order to evaluate the validity of the proposed model regarding structural and functional organization of the visual system and its relevance in terms of perceptual organization during visual target acquisition, results from RT measurements and conspicuity area determinations were combined in the current study. It elaborates on earlier work in that it applies principles of signal detection theory to explore the relationship between RT and target detection probability at specific display positions, relative to the conspicuity area, for different target-background combinations, in order to infer the mechanism underlying visual target acquisition (Chapter 5). Prior to commencement, however, several methodological issues had to be dealt with, as discussed in the following section.

4.5 OPTIMIZATION OF EXPERIMENTAL DESIGN

Accurate measurement of response variables is important during psychophysical studies. In the course of the current study much time and effort were spent on optimising the design and execution of the experiments. The following section will show results from preliminary studies regarding parameters that contribute to visual search and target acquisition behaviour, methods for determining the conspicuity area and RT measurements.

4.5.1 Preliminary study regarding parameters contributing to visual search and target acquisition behaviour

Background

The study reported on here was based largely on the premise that conspicuity area and RT measurements can be used to express a target's visual conspicuity, which in turn will serve as indication of expected search performance (Engel, 1971; Engel, 1977; Bellamy & Courtney, 1981; Toet et al., 1998). It has previously been shown that target as well as background properties, and especially the contextual relationship between them, influence search performance (Duncan & Humphreys, 1989; Nothdurft, 1992). This allows for numerous relevant target-background combinations to be investigated and can easily lead to approaching the problem in a disorganised manner. It was therefore worthwhile to conduct various pilot studies in order to determine which aspects of target-background composition warrant further exploration.

Display conditions were subsequently changed to manipulate SNR by changing either target or distractor properties. It was argued that if such changes had a successful impact on SNR, it will be reflected in conspicuity area data, given that the conspicuity area is believed to serve as indication of a target's visual conspicuity. However, prior to commencing with any investigations into which target or distractor manipulations would effectively alter SNR, any display-related factors that could possibly bias search performance had to be investigated. It was therefore important to establish the extent to which (i) the nature of the background organization as well as (ii) presentation time would influence search performance.

Early studies have shown the shape of the conspicuity area to be approximately ovaloid, with the longer axis horizontal (Chaikin et al., 1962; Engel, 1971; Engel, 1977), but later investigations found the shape to be much more irregular (Courtney & Chan, 1985a; Courtney & Chan, 1985b). The pilot studies reported here thus also aimed to provide insight into the general shape of the conspicuity area and whether aspects that influence its size also have an impact on its shape. This would justify using methods that estimate only the border of the conspicuity area, since full field mapping is extremely time-consuming because several observations are needed at each location of the visual field (Bellamy & Courtney, 1981).

Method

All experiments were conducted in a dimly lit room with minimal noise disturbance. Subjects were seated 44 cm away from a computer screen (without curvature), which subtended 36.5° horizontally and 29° vertically. Viewing position was maintained by means of a height-adjustable chin rest, so that gaze was directed at the centre of the screen when looking straight ahead. An externally situated computer, with custom-designed software, was used to generate displays and control progression of the experiments. The experimenter could view progression of the experiment through a small window.

The standard display was designed according to similar stimuli dimensions, nature and density as used by Engel (1971). It consisted of 747 white lines (distractors) of different orientations against a black background and positioned in a random manner, among which a single target (a white line drawing of a square) was hidden. The target could appear at any of 748 positions and although distractors were allowed to cross, they never crossed the target. Target and distractors subtended a visual angle of 1.2° and 0.6° respectively. Figure 4.2 shows a part of the standard display. Four stimulus parameters that could possibly influence SNR were investigated by varying only a single aspect of the standard display and then comparing results to those from the standard display (also see Table 4.1).

At the start of an experiment subjects fixated on a white fixation dot at the centre of the screen. After 1 000 ms, the fixation screen was replaced by the test screen for 75 ms, followed again by the black fixation screen for 1 000 ms. The fixation dot was present for the duration of the test screen. Subjects were asked to maintain fixation and to activate the left mouse button if they detected the target. A response could be registered for up to 500 ms after the test screen disappeared, but if no response was registered within 575 ms from the onset of the test screen, it was recorded that the target had not been detected. Target coordinates were always recorded in the result output, regardless of whether the target was detected or not. The importance of rapid, yet accurate responses was emphasised. All subjects used their dominant hand for mouse button activation. No feedback was provided, in order to ensure that subjects always responded objectively.

A complete testing session for a single condition consisted of 450 trials, divided into three identical experiments of 150 trials each, with 5 minutes rest after every 150 trials. Depending on the subject's time schedule, either two or three conditions were tested per day, but total testing time never exceeded 1 hour per subject per day. For the very first set of experiments, practice trials were performed, but since the tasks were so similar for following experiments, it was argued that previous sessions would be adequate to eliminate familiarization effects. Subjects were informed of the nature of the display prior to commencement of the experiment.

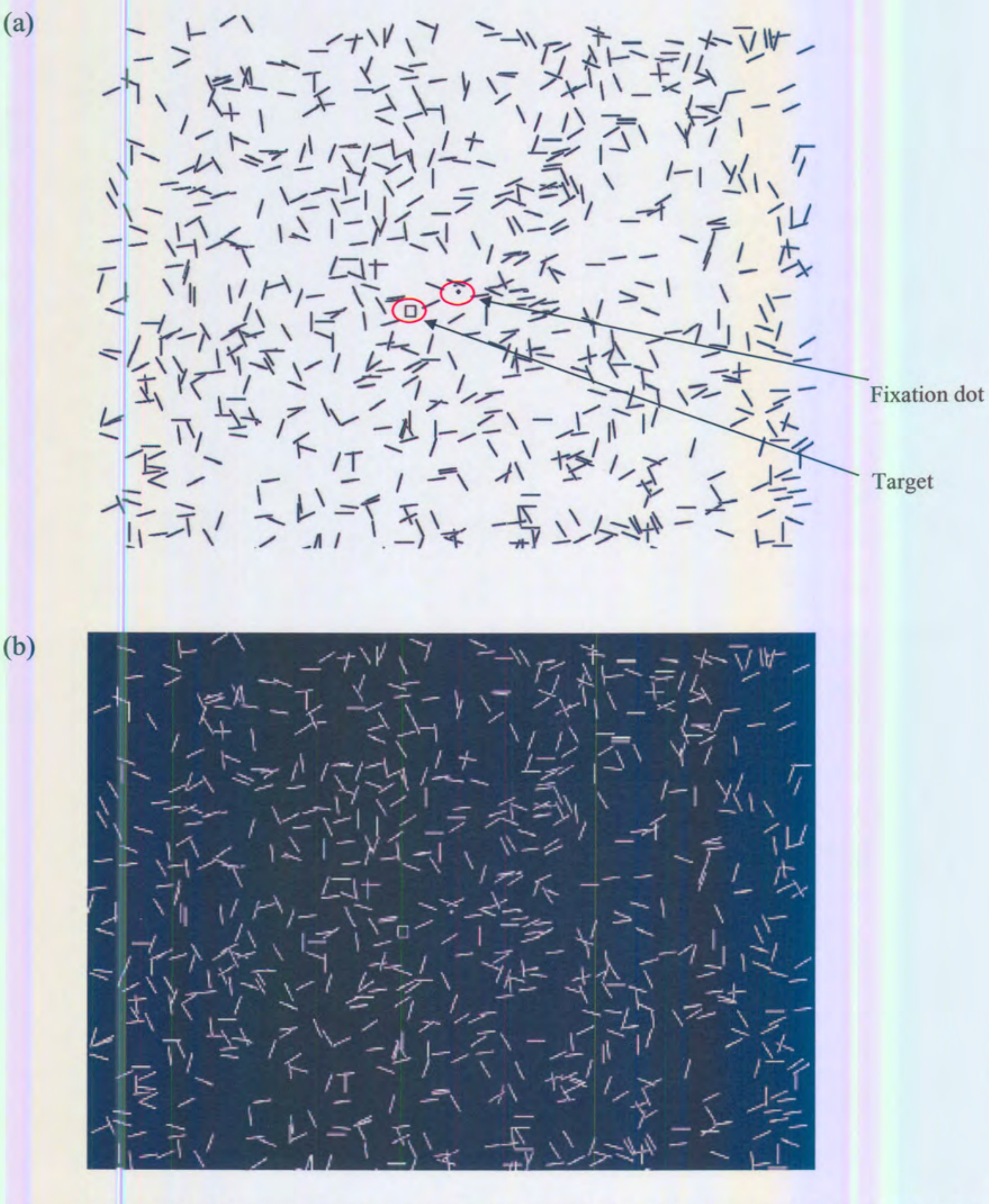


Figure 4.2: (a) Example of the standard display design. The target is positioned just left of the fixation dot, both circled for easy reader recognition. All further examples of displays will be portrayed as in Figure 4.2a, although the colour scheme of actual displays was always as in Figure 4.2b.

Two male university students, aged 25 and 28 respectively, participated in the pilot studies. Both had normal visual acuity (1.0 arcmin^{-1} or better) and normal contrast sensitivity (Vision Contrast Test System, Vistech Consultants Inc.) as determined by a qualified optometrist prior to commencement of the study. Neither had previous experience in tasks of the type reported here. Written informed consent was given prior to commencement, in accordance to ethical approval granted by the Student Research Ethical Commission, Faculaty of Health Sciences, University of Pretoria.

Results and discussion

Results from pilot studies investigating the size and shape of the conspicuity area, as well as the extent to which it is affected by different parameters, are shown in Figure 4.4 and Table 4.1. Results are always compared to those obtained with the standard display. Coordinates of detected targets were used to generate scatterplots, which served as initial estimates of shape and size of the conspicuity areas. Areas were sufficiently regular to justify use of only the border coordinates for further analyses and coordinates were subsequently estimated by visual inspection of all the obtained data. The shapes resembled ellipses and non-linear least squares regression was performed to provide initial estimates of ellipse parameters. Curve fitting was performed in Matlab[4], using code adapted from N. Baroni and S. Grobler, which also contains code developed by K. Pankratov (Grobler, 2000). Since only the general shape and size of the area were of interest at this stage, parameter estimates and a visual representation of the function that best described the data were adequate.

---

[4] Matlab is a software package intended for mathematical programming. Also see www.mathworks.com.

Results from the standard display show that the shape of the conspicuity areas is described well by the function of an ellipse, with the horizontal axis being longest and a large part of the area being situated in the upper hemifield:

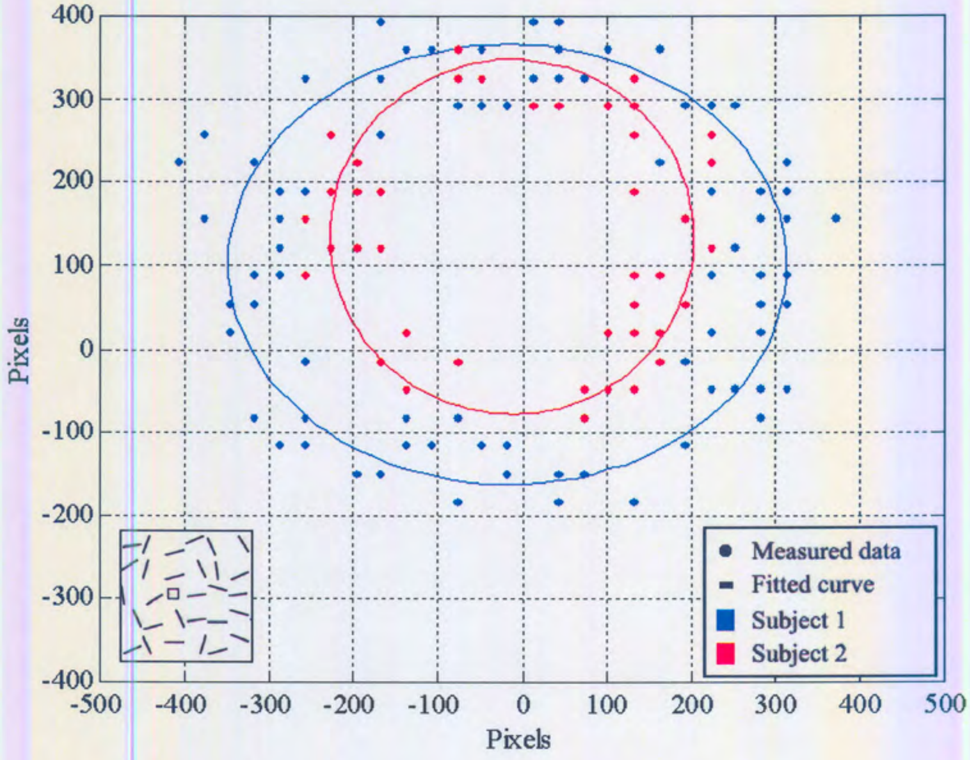
$$\frac{(x-h)^2}{a^2}+\frac{(y-k)^2}{b^2}=1 \qquad \ldots\ldots\ldots\ldots\ldots 1$$

The distances from the centre *(h,k)* to the boundary on the major and minor axes are represented by a and b respectively, with $a>b$, $h \approx 0$ and $k \approx 100$.

Table 4.1: Conspicuity area results for different display conditions.

Influencing factor	**Display condittion**	**Ellipse parameter**	**Subject 1**	**Subject 2**
Not applicable	• Target: Square • Background: Distractors positioned randomly, allowed to cross one another • Target-distractor intersection: No • Presentation time: 75 ms *(Also see Figures 4.3b – 4.6b)*	h (pixels)	1	-24
		k (pixels)	95	126
		a (pixels)	253	195
		b (pixels)	207	91
		a:b	1.22	2.15
		area (deg^2)	284.31	97.02
Background regularity	• Target: Square • Background: Distractors positioned in regular grid, not allowed to cross one another. • Target-distractor intersection: No • Presentation time: 75 ms *(Also see Figure 4.3a)*	h (pixels)	-18	-13
		k (pixels)	102	135
		a (pixels)	331	216
		b (pixels)	265	213
		a:b	1.25	1.01
		area (deg^2)	470.34	250.53
Target-distractor intersection	• Target: Square • Background: Distractors positioned randomly, allowed to cross one another • Target-distractor intersection: Yes • Presentation time: 75 ms *(Also see Figure 4.4a)*	h (pixels)	-22	13
		k (pixels)	100	126
		a (pixels)	285	104
		b (pixels)	183	103
		a:b	1.56	1.00
		area (deg^2)	283.26	59.22
Target contour closure	• Target: Four lines, perpendicular to each other • Background: Distractors positioned randomly, allowed to cross one another • Target-distractor intersection: No • Presentation time: 75 ms *(Also see Figure 4.5a)*	h (pixels)	-7	32
		k (pixels)	129	164
		a (pixels)	266	90
		b (pixels)	187	109
		a:b	1.43	0.82
		area (deg^2)	270.97	53.99
Presentation time	• Target: Square • Background: Distractors positioned randomly, allowed to cross one another • Target-distractor intersection: No • Presentation time: 200 ms *(Also see Figure 4.6a)*	h (pixels)	16	26
		k (pixels)	111	128
		a (pixels)	268	170
		b (pixels)	214	151
		a:b	1.25	1.13
		area (deg^2)	310.83	140.81

(a) Distractors arranged according to regular grid



(b) Distractors arranged irregularly across display

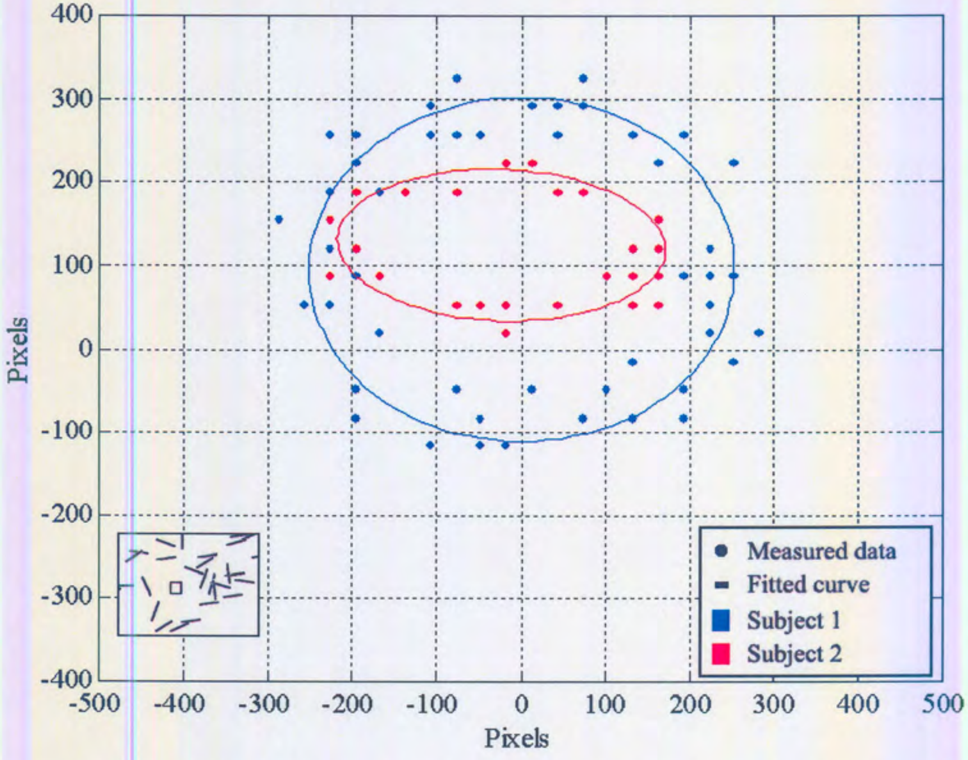
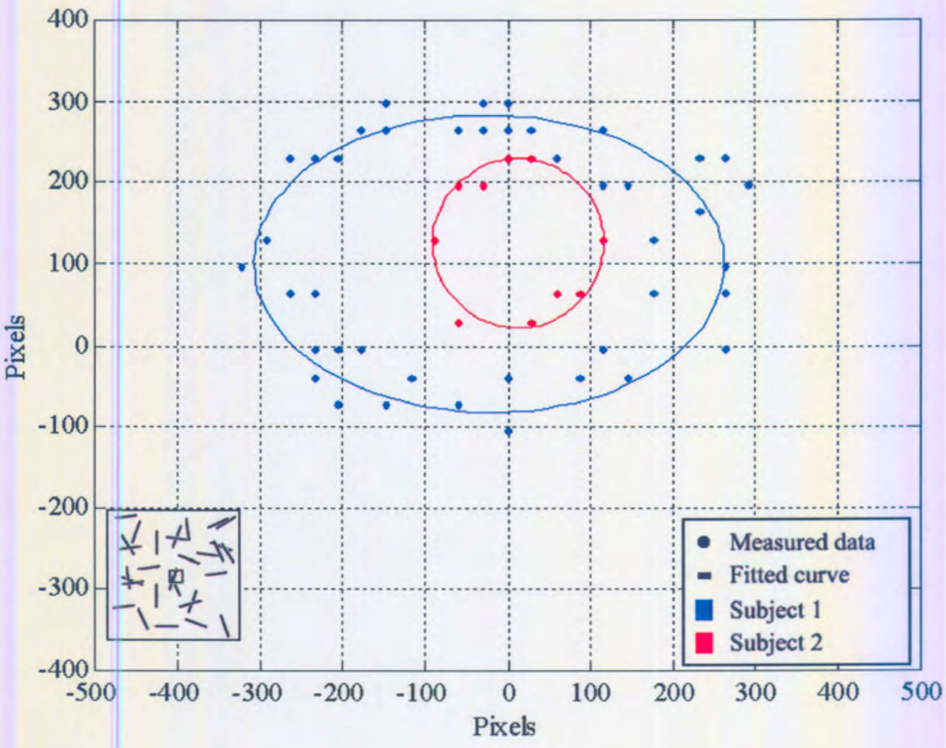


Figure 4.3: Conspicuity area as determined when (a) distractors were arranged according to a regular grid, compared to (b) the standard display, when distractors were placed irregularly across the display. Examples of the displays are provided in the insets.

(a) Target-distractor intersection



(b) No target-distractor intersection

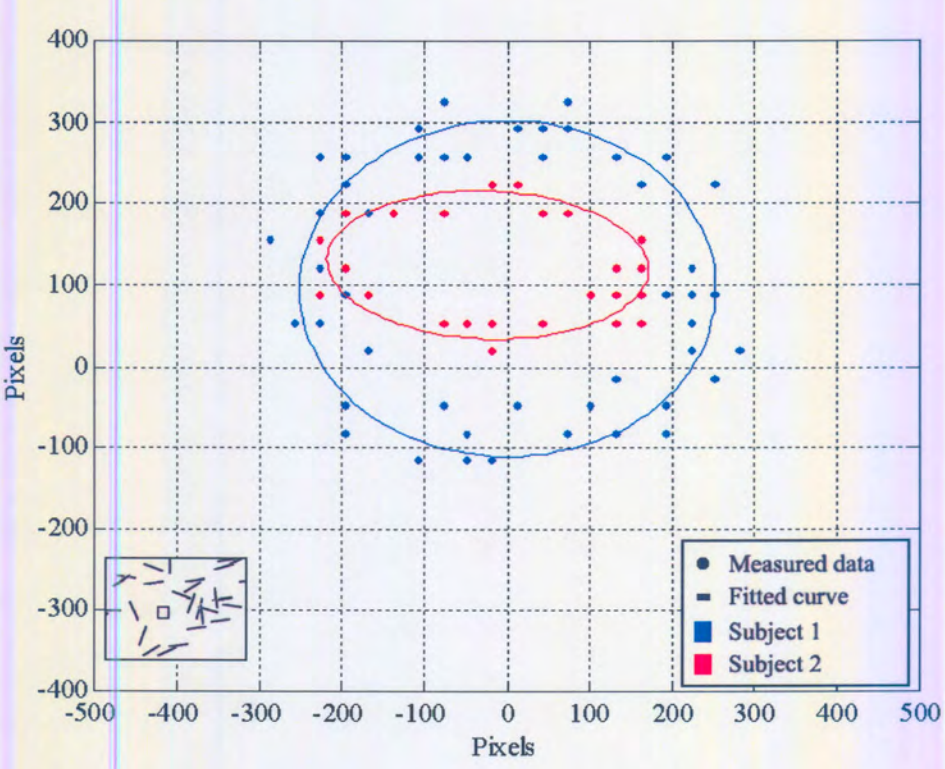
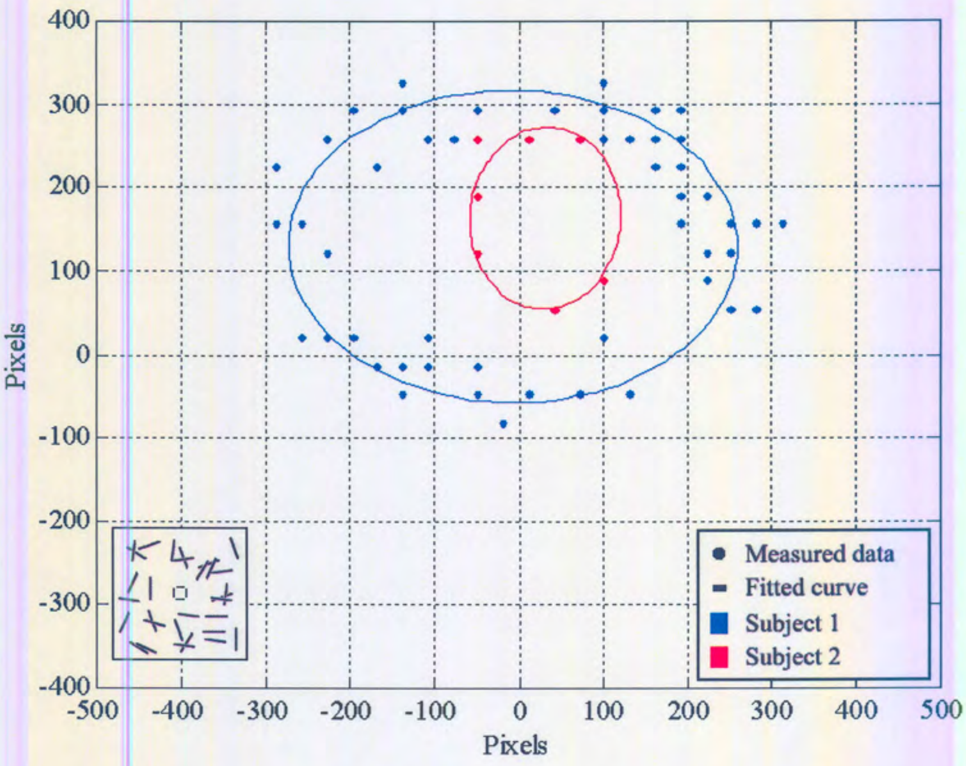


Figure 4.4: Conspicuity area as determined when (a) target was intersected by distractors compared to (b) the standard display, when no distractors crossed the target. Examples of the displays are provided in the insets.

(a) Incomplete square used as target



(b) Complete square used as target

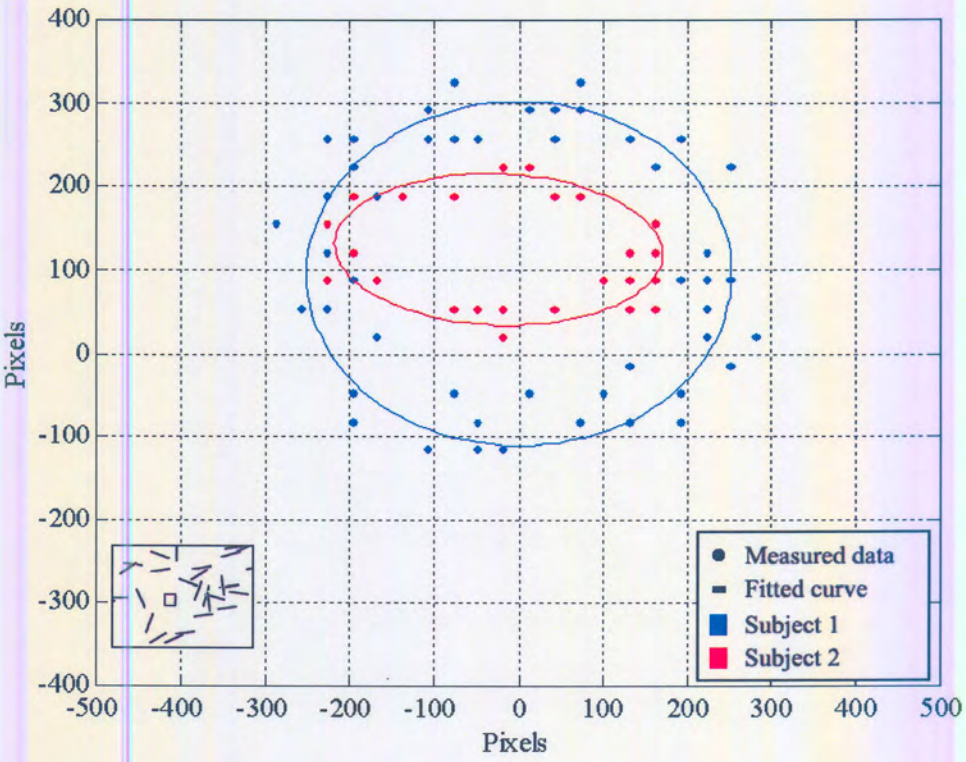
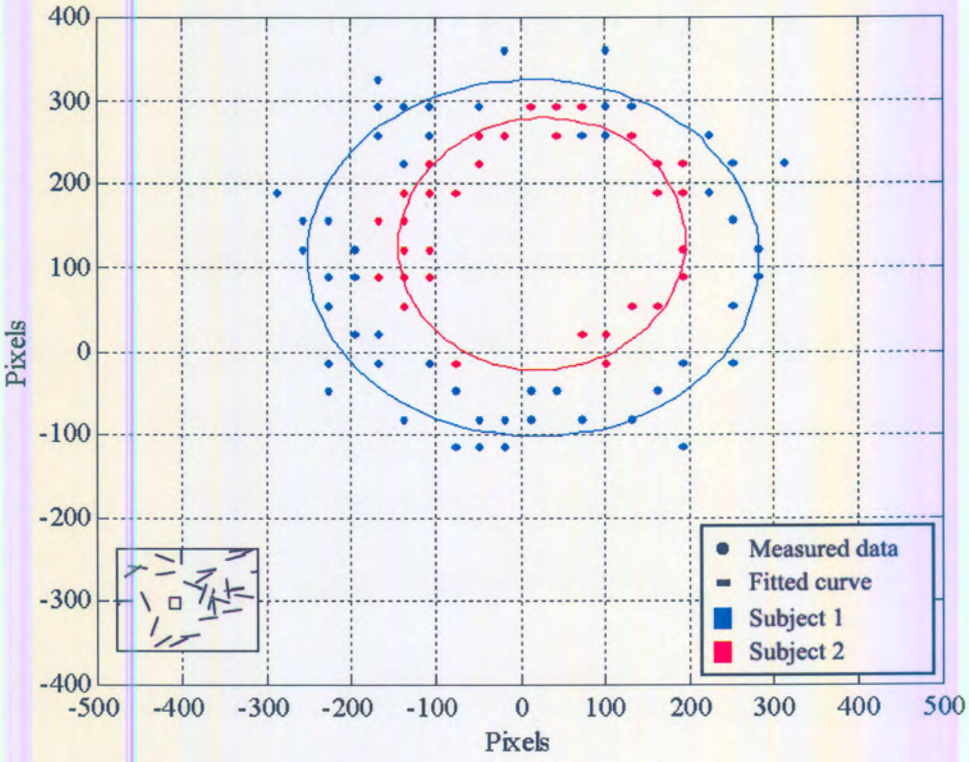


Figure 4.5: Conspicuity area as determined when (a) target was an incomplete square, compared to (b) the standard display, when a completed square was used as target. Examples of the displays are provided in the insets.

(a) 200 ms display duration



(b) 75 ms display duration

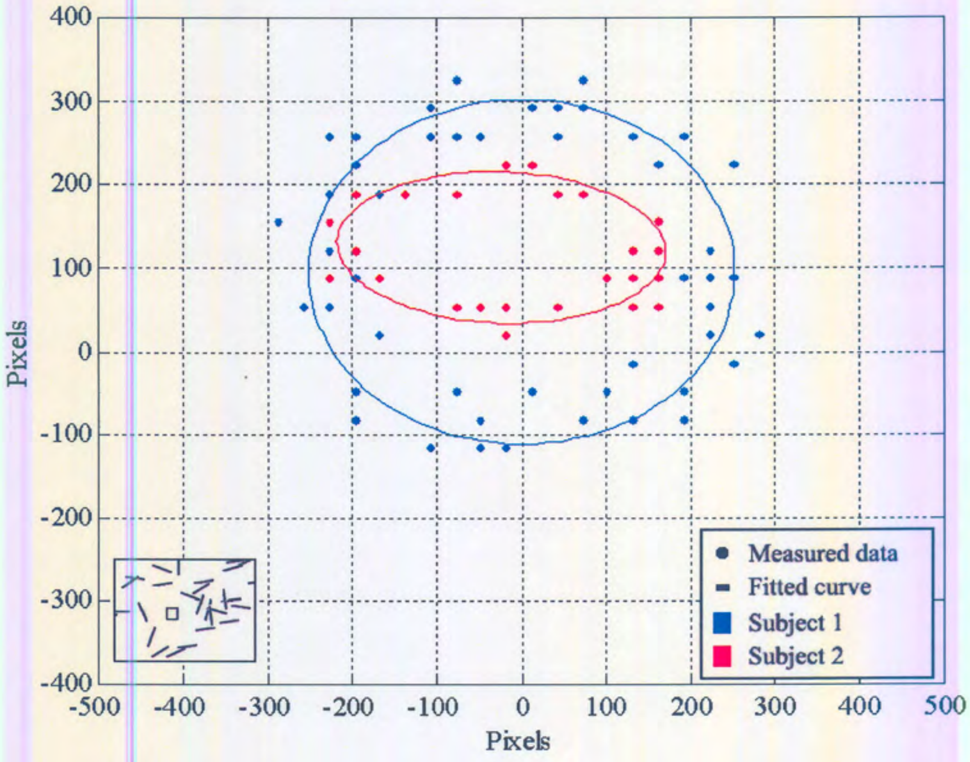


Figure 4.6: Conspicuity area as determined when (a) display duration was 200 ms compared to (b) the standard display, with display duration at 75 ms. Examples of the displays are provided in the insets.

The size of the conspicuity areas, reflected in the area and expressed in terms of square degrees, of Subject 1 (blue curves) was notably larger than that of Subject 2, possibly indicating that the size of the conspicuity area is subject to individual variability across subjects. Should the conspicuity area be used as measure of target detectability, it would thus be important to determine to which extent such variability would affect interpretation of results.

For conditions where display properties were altered, results from Subject 1 show that conspicuity areas always resemble ellipses, similar in shape to the conspicuity area obtained with the standard display. In cases where conspicuity areas showed little variation in terms of size (i.e. size changes of less that 10% compared to that of the standard display), shape was affected by the parameter changes, as reflected in the a:b ratio. In one case where substantial size increase (approximately 60% compared to the standard display) was observed, shape was unchanged.

Conspicuity areas obtained from Subject 2, however, showed more variation when compared to that of the standard display. The condition where target-distractor intersection was allowed, as well as the condition where distractors were arranged according to a regular grid across the display, produced conspicuity areas tending to be more circular than elliptical, with the a:b ratio approximating 1. The condition where target contour closure was not allowed produced an elliptical conspicuity area, but with an a:b ratio < 1. Size also varied considerably more than for Subject 1 - changes of 38%, -40%, -45% and 69% relative to the standard display were observed, compared to changes of 60%, 0%, -5% and 10% for the same conditions observed in results of Subject 1.

Parameter changes applied in this study were chosen on the premise that they would manipulate SNR. As shown in earlier studies (Duncan & Humphreys, 1989; Nothdurft, 1992) increasing the level of similarity between target and distractors in an already heterogeneous background will result in impaired search efficiency. Increasing the level of distractor heterogeneity to compile the background would also hamper search performance (Duncan & Humphreys, 1989). Display conditions where an "open" square (i.e. no contour closure) was used, as well as conditions where target-distractor intersection was allowed, aimed to increase target-distractor similarity by manipulating target and distractor properties respectively. The display condition where distractors were not arranged according to a regular grid across the screen, but allowed to intersect one another, aimed to increase distractor heterogeneity in the background. Conspicuity area results indicate that these manipulations successfully affected SNR, as reflected in changes in conspicuity area sizes and shapes. Results therefore confirm that conspicuity area determination can be valuable for evaluating the extent to which different target-background combinations affect target conspicuity, as proposed previously (Toet et al., 1998). Results, however, also indicated that a single property of the conspicuity area (e.g. size) cannot be used in isolation as a measure of target detectability – in cases where parameter manipulations resulted in shape changes, size was left relatively unaffected and vice versa. This points to the importance of considering the conspicuity area as a whole when using it as measure of target detectability.

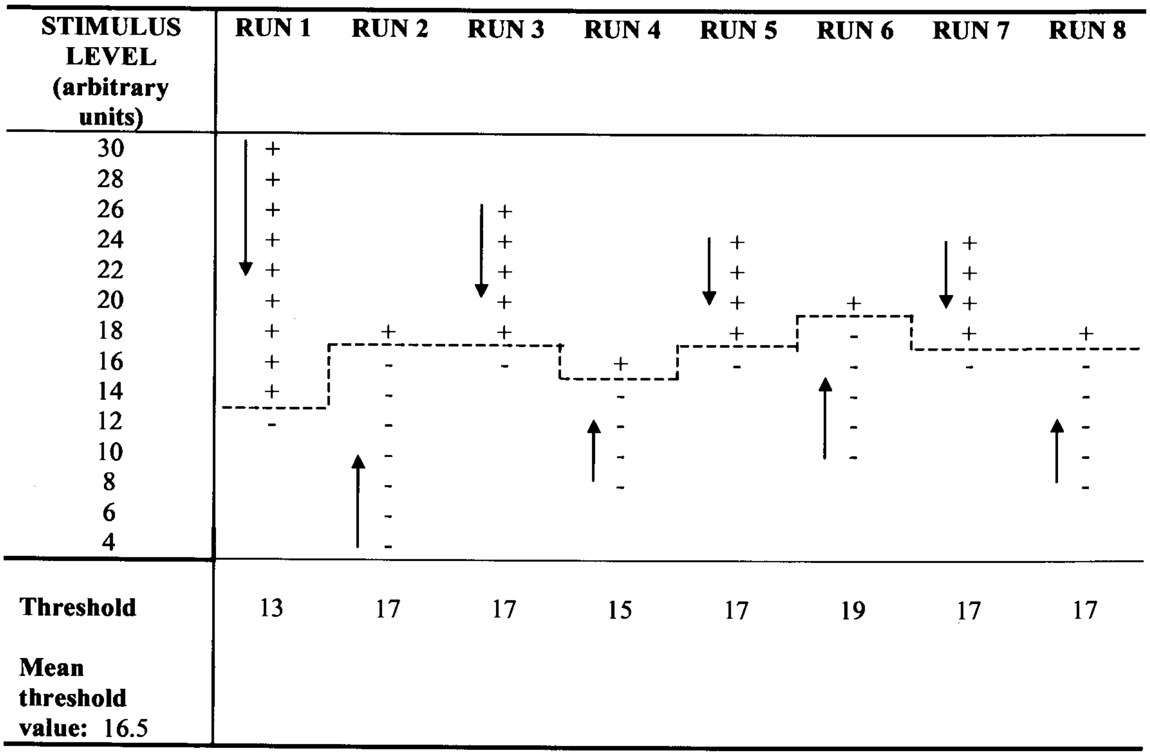
Conspicuity areas obtained with full field mapping proved sufficiently regular to justify the use of border coordinates to evaluate the shape and size of conspicuity areas. This refutes earlier findings that suggest that regular conspicuity area shapes are the result of measuring boundary coordinates only along selected meridians, instead of fully mapping the field of view (Bellamy & Courtney, 1981; Courtney, 1989). The results furthermore confirm that the shape of a target's conspicuity area is adequately described by an ellipse or circle, as proposed by Engel (1971; 1977).

4.5.2 Different methods of conspicuity area determination

Background

Signal detectability is usually described in terms of the threshold value that separates detection from non-detection. The same principles apply when determining the conspicuity area of a target, since the boundary of the conspicuity area represents the eccentricity threshold at which a target can be detected in a single fixation. It follows that if the conspicuity area is to serve as a valid indication of target detectability, its boundary should be determined reliably and accurately.

In order to describe performance on psychophysical tasks in terms of a particular stimulus characteristic, it is useful to construct a psychometric function, with stimulus level on the abscissa and performance (usually reported as percentage correct responses) on the ordinate (Green & Swets, 1966; Levit, 1971). Two parameters of the function are of interest: (1) the location of the function along the stimulus axis, which indicates the threshold value and (2) the slope of the curve, indicating the rate at which performance changes relative to changes in stimulus levels (Levitt, 1971; Leek, 2001). To this end, several methods can be used. The classical Method of Constants is well suited to characterize the psychometric function fully, since several observations are placed, with equal probability, across the complete range of stimulus levels (Levitt, 1971; Leek, 2001). If, however, only a particular value along the psychometric function is of interest, adaptive procedures prove to be more efficient (Levitt, 1971; Leek, 2001; Marvit et al., 2003), since each successive stimulus presentation is governed by the subject's response at the preceding presentation (Figure 4.7). Stimuli are presented stepwise, with the direction of the sequence reversing as soon as the subject's response reverses. This allows the threshold value to be approached rapidly and then converge accurately on the exact value, especially if numerous runs are performed. The Method of Limits[5] and the Transformed Up-Down method are examples of adaptive techniques.



STIMULUS LEVEL (arbitrary units)	RUN 1	RUN 2	RUN 3	RUN 4	RUN 5	RUN 6	RUN 7	RUN 8
30	+							
28	+							
26	+		+					
24	+		+		+		+	
22	+		+		+		+	
20	+		+		+	+	+	
18	+	+	+		+	-	+	+
16	+	-	-	+	-	-	-	-
14	+	-		-		-		-
12	-	-		-		-		-
10		-		-		-		-
8		-		-				-
6		-						
4		-						
Threshold	13	17	17	15	17	19	17	17
Mean threshold value: 16.5								

Figure 4.7: Schematic representation of the design of an adaptive method used during psychophysical studies. Signal detection is indicated by +, while failure to detect the signal is indicated by -. Dashed lines represent the threshold for the specific run.

The current study aimed to compare the three abovementioned methods in terms of accuracy, reliability and speed with which the boundary of the conspicuity area (i.e. eccentricity threshold) could be determined, in order to find the most time-effective, yet accurate method for conspicuity area determination.

---

[5] The Method of Limits is conventionally classified as a classical psychophysical technique, but is essentially adaptive (Levitt, 1971).

Method

Subjects and experimental conditions were the same as for the previous pilot study. The standard display (Figure 4.2) was used for all the experiments of the current study. Three methods with which to determine the conspicuity area were evaluated in terms of accuracy, reliability and speed of convergence to threshold. Subjects were informed of the nature of the display prior to commencement of the experiment. Since the same subjects participated in the previous study, no training was provided.

Testing sessions employing the Method of Constants proceeded as for standard display conditions, described in Paragraph 4.5.1. A schematic representation is provided in Figure 4.8a.

The Method of Limits employed the same display, but differed from the Method of Constants in terms of stimulus presentation (Figure 4.8b). At the start of the testing session subjects fixated a white fixation dot at the centre of the screen for 1 000 ms, after which it was replaced by the test screen for 75 ms, followed again by the fixation screen for 1 000ms. The target moved across the screen in a traversing manner, approaching the fixation dot from the far periphery and then moving outwards again. Subjects were required to activate the left mouse button once they detected the target, and then once more when they could not detect the target any longer, maintaining fixation throughout. The target moved along four directions: left to right, right to left, top to bottom and bottom to top. Starting positions and directions were determined randomly by the software which controlled experiment progression. Target coordinates, as well as the direction in which the target moved, were recorded each time the mouse button was activated. A complete testing session lasted approximately 1.5 hours and consisted of three identical experiments of 300 trials each. Short rest periods were allowed after each set of 300 trials.

The third method under investigation was an adaptive method that employed Levitt's Transformed Up-Down technique (Levitt, 1971). Display properties were the same as during the two other method evaluations, but progression of the experiment was adapted (Figure 4.8c). For evaluation of the current method, the target followed for 75 ms after the fixation screen was displayed for 1 000 ms. Once the target disappeared, it was replaced again by the fixation screen for 1 000 ms, but the fixation dot changed from white to red. This indicated that the subject's response was awaited. During these experiments, catch trials, i.e. trials during which no target was present, were also included. Subjects were instructed to activate the left mouse button if they detected the target and the right mouse button if they thought it was absent. Fixation had to be maintained throughout. Levitt's Transformed Up-Down technique states that the stimulus level (in this case eccentricity) should be increased after two successive positive responses, while it should be decreased after a single negative response or a positive response followed by a negative one. Since the previous pilot studies confirmed that the conspicuity area's shape corresponds to either a circle or an ellipse, targets were presented radially along eight meridians, spanning 360°. The first target along each meridian was always presented close to the fixation dot, located at the centre of the screen. Targets were placed more peripherally according to fixed step sizes during subsequent trials. In order to avoid anticipation of the target location, meridians were visited in random manner. Software was, however, designed to keep track of the previous position on each meridian, as well as the subject's response for the preceding presentation on that specific meridian. The test terminated after 15 response reversals were recorded in both directions of each meridian. Recorded results included target coordinates, meridian moved along, step size and correctness of response.

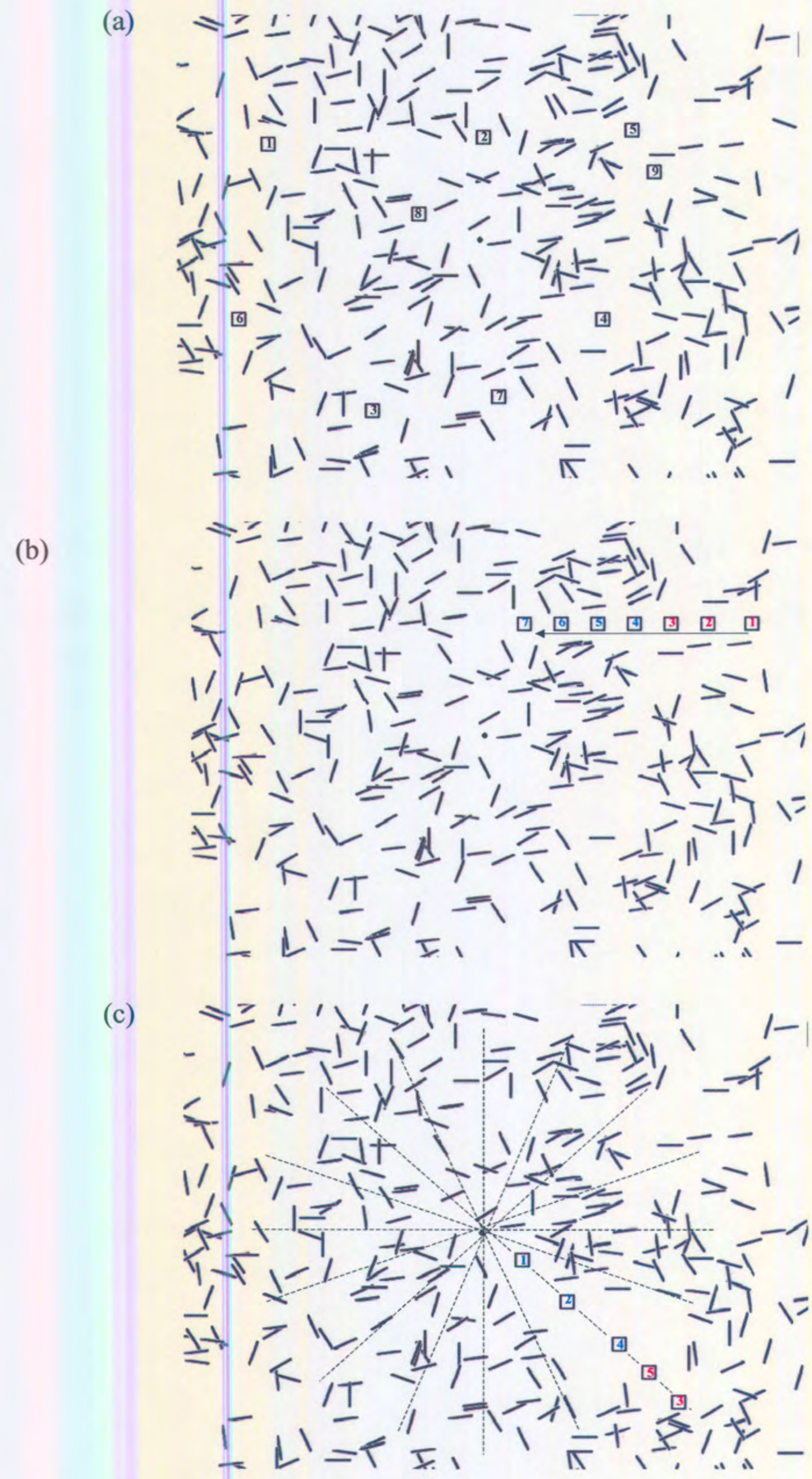


Figure 4.8: Schematic representations of the methods evaluated for conspicuity area determination, with (a) the Method of Constants, (b) the Method of Limits and (c) the Transformed Up-Down method. For Figure 4.8a, the numbered squares represent locations of successive target appearances. In Figures 4.8b-c, blue and red digits represent positive and negative responses respectively.

Results and discussion

Results from conspicuity area determinations using the three methods are shown in Figure 4.9 and Table 4.2. Nonlinear least squares regression was performed to provide initial estimates of ellipse parameters. Curve fitting was performed in Matlab, according to code adapted from that developed by N. Baroni and S. Grobler and also K. Pankratov (Grobler, 2000). Since only general properties of the conspicuity area were of interest at this stage, parameter estimates and a visual representation of the function that best described the data, were adequate.

Table 4.2: Comparison of different methods for conspicuity area determination.

Method	Presentation mode and task requirements	Ellipse Parameter	Subject 1	Subject 2
Method of Constants	Target presented randomly at any of 748 positions. Subject to press left mouse button upon detection. *Also see Figures 4.8a and 4.9a.*	h	1	-24
		k	95	126
		a	253	195
		b	207	91
		a:b	1.22	2.15
		area (degree^2)	284.31	97.02
Method of Limits	Target traversing across screen, along any of four directions. Subject to press mouse button upon first and last detection. *Also see Figures 4.8b and 4.9b.*	h	-33	12
		k	136	141
		a	418	286
		b	299	212
		a:b	1.40	1.35
		area (degree^2)	660.91	327.71
Transformed Up-Down Method	Target moving from centre point outwards towards periphery, along any of 16 meridians. Direction of movement reverses when response reverses. Subject required to press left mouse button upon target detection or right mouse button in case of no target detection. *Also see Figures 4.8c and 4.9c.*	h	-22	-4
		k	-9	-3
		a	254	179
		b	209	142
		a:b	1.21	1.26
		area (degree^2)	289.33	140.10

The results comment on several aspects regarding conspicuity area determination. Conspicuity areas are elliptical, with the longer axis horizontal, irrespective of whether border contours are determined from full-field mapping or from only measuring along selected meridians. It is interesting to note that results from the Transformed Up-Down method show the conspicuity areas to be symmetrical around the centre point of the screen. This is unlike conspicuity areas determined by methods that do not force the target through the centre point

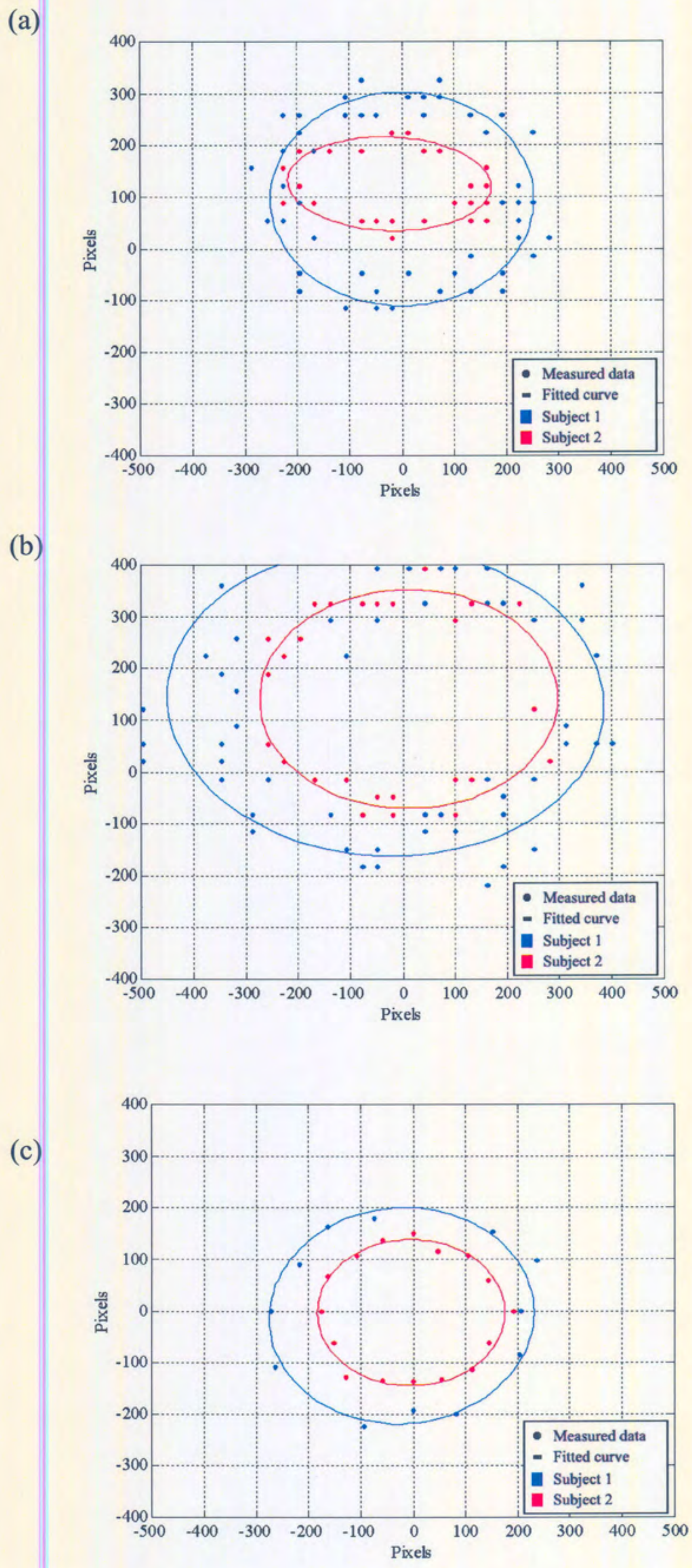


Figure 4.9: Conspicuity areas as determined with (a) Method of Constants, (b) Method of Limits and (c) Transformed Up-Down method.

of the screen, where a significantly larger portion of the areas is situated in the upper hemifield. Courtney and Chan (1985b) attributed such areas of insensitivity to more pronounced background interference occurring in those areas. It is, however, a doubtful explanation in this case, since the same display conditions applied during all three method evaluations.

The Method of Limits produced areas approximately twice to three times larger than obtained with either the Method of Constants or the Transformed Up-Down method. This may be attributed to the target traversing across the screen, which could facilitate unconscious tracking of the target, despite the importance of maintaining fixation being stressed. Conspicuity areas of similar size were obtained with the Method of Constants and the Transformed Up-Down method. This may be attributed to the fact that subjects could not anticipate target location for successive trials (Levitt, 1971), since targets were either located randomly at any of 748 possible positions (Method of Constants) or located along any of the specified meridians, considering the previous target position and response along that specific meridian (Transformed Up-Down method).

Despite obtaining usable results from all these methods, the Transformed Up-Down method seems to be most suitable for conspicuity area determination. Since the boundary of the conspicuity area is homed in on by terminating testing along a particular direction only after several reversals, reliability of the procedure increases, leading to results that are appreciably more accurate than for the other methods. This is in line with a recent study that found the Transformed Up-Down method to be superior during a comparison of psychophysical procedures for determining detection thresholds (Marvit et al., 2003). Ease of execution also contributes to the attraction of the method - both subjects commented that they preferred this method to the two others, as it was less cumbersome and time-consuming and demanded more attentive viewing, causing less fatigue.

4.5.3 Development of an accurate and reliable RT measurement technique

Accurate and reliable timing was important for experiments reported on in this thesis, since both display presentation and RT measurement relied on execution of time-dependent procedures. It was therefore necessary to determine the accuracy with which these functions could be performed as well as the margin of error associated with the measurements.

An application program for creating test displays as well as measuring RT to an external event was developed under Microsoft Windows 2000 (Meyer et al., in preparation; also see Appendix A). Time measurements were verified by means of an independent external timer device. Results showed that when the application's timer controlled presentation time, the display was consistently terminated 8 ms earlier than it was supposed to. By using different timing routines, the resolution of the application's timer could be improved, so that display presentation time, when controlled by the application's timer, was only 1 ms shorter than it was supposed to be. Results from the independent timer device, however, showed that displays were actually active for approximately 26 ms longer than measured by the application's timer. The difference between measurements from the application's timer and the independent timer was consistent. These results were discussed in terms of operating system architecture. It was concluded that time-dependent applications can be implemented successfully in a general-purpose operating system, provided that application design is structured correctly.

4.6 SUMMARY

Chapter 4 served as introduction to the empirical stage of the study. It explained the rationale behind the approach to the study, showing the relevance and value of methodological tools used during completion of the experiments. Pilot studies provided insight into the validity and efficiency of using the conspicuity area as measure of target detectability, showing that manipulations to SNR are readily reflected in the altered nature of the area. It was also shown that the observed ovaloid to circular shape of the conspicuity area is not an artefact of the method used for its determination and that it can therefore reliably be determined by using methods that estimate border coordinates along only selected meridians, instead of performing full-field mapping. The Transformed Up-Down method proved to be most effective. Lastly it was shown that with correctly structured application design, time-dependent experiments of the kind conducted in this study, can successfully be performed in a Windows 2000 environment. Findings from the pilot studies are applied in the final empirical stage, which is presented in Chapter 5.

CHAPTER 5

CONSPICUITY AREA AND RESPONSE TIME AS MEASURES OF VISUAL TARGET ACQUISITION BEHAVIOUR

5.1 INTRODUCTION

In the foregoing chapters the emphasis has been on introducing important principles pertaining to the study of visual target acquisition behaviour, as well as describing the course of method development. It is thus appropriate now to attend to experiments during which optimized psychophysical techniques and experimental conditions, as described in Chapter 4, were used to investigate how visual conspicuity, expressed in terms of the conspicuity area, and associated response time, can be related to visual target acquisition behaviour. Results are interpreted in terms of the framework developed in previous chapters, in order to explore possible mechanisms underlying visual target acquisition behaviour.

The experiments described here were aimed at investigating the relationship between conspicuity area and response time, for application as psychophysical measures of target detectability and associated visual target acquisition behaviour. Conspicuity areas for four different target-background combinations, found to influence target detectabilty to different extents in pilot studies, were determined according to an optimized method described earlier. Response times were subsequently measured at predefined positions relative to the conspicuity area for each display, in order to relate the observed behaviour to the framework proposed for information processing during a visual target acquisition task.

5.2 METHODS

5.2.1 Conspicuity area determination

For conspicuity area experiments four target-background combinations, using parameters found to influence target detectability successfully during pilot studies, were investigated. The target could either be a complete square or one consisting of four separate lines perpendicular to one another, while the background could consist either of distractors intersecting one another but not the target, or distractors intersecting the target as well as each other. In order to prevent open spaces around the target unduly directing attention towards the target location, between three and five distractors were always in close proximity to the target. Examples of the four displays are shown in Figure 5.1. The nature and dimensions of the stimuli, as well as all other experimental conditions (i.e. testing facilities, viewing distance etc.) were as described earlier (Paragraph 4.5.1).

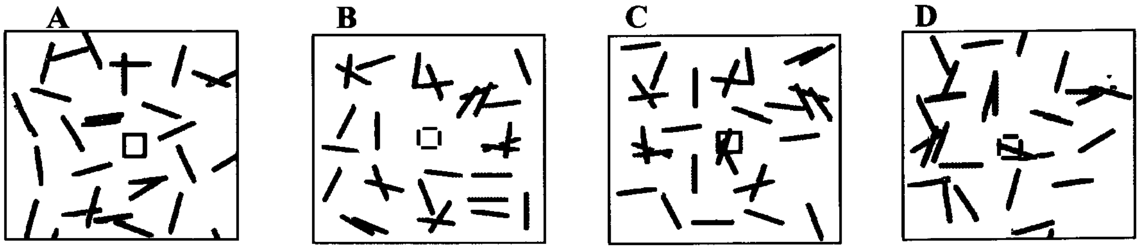


Figure 5.1: Examples of display conditions A-D. Targets could either be a complete square or four lines of the same lengths, orientated perpendicular to one another (as in A and B respectively). The background could consist either of distractors intersecting one another but not the target or distractors intersecting one another as well as the target (as in A and C respectively).

Experiments proceeded according to the principles of Levitt's (1971) Transformed Up-Down method (Paragrpah 4.5.2), found to be the most effective of the three methods evaluated during pilot studies. Each subject had to complete four sessions, dedicated to each of the four display conditions, over the course of several days; however, a subject was only scheduled for a single session per day. A testing session consisted of approximately 1 600 trials and lasted approximately 1.5 hours. Short rest periods were allowed after every 400 trials. No feedback or training was provided.

Eight male subjects (mean age = 24.4 years) participated in the study. All had normal (or corrected to normal) visual acuity (1.0 arcmin^{-1} or better) and normal contrast sensitivity (Vision Contrast Test System, Vistech Consultants Inc.) as determined by a qualified optometrist prior to commencement of the study. Two of the subjects participated in the pilot studies. Written informed consent was given prior to commencement of the study.

RT measurements

The second set of experiments aimed to determine the relationship between RT and target location, relative to the conspicuity area. Experimental conditions, target-background combinations, display properties and participants were the same as for conspicuity area determinations.

Each trial consisted of a fixation screen and a test screen. The fixation screen was displayed for 1 000ms and was followed by the test screen, which the subject had to search for the presence of a predefined target. The test screen was terminated only once either of two mouse buttons was activated; activation of the left mouse button indicated target detection, while activation of the right button signalled that the target could not be detected. Fixation was required only during display of the fixation screen in order to ensure that search during display of the test screen always commenced from the same location. RT was recorded as elapsed time from the onset of the test screen until user input was received. Approximately half of all trials were catch trials, but were assigned randomly in order to avoid anticipation from subjects. Subjects were informed about the nature of the task, as well as the possibility of catch trials (but not their proportion), prior to commencement of the experiment. The importance of fast yet accurate responses was stressed, but no feedback was provided. Four blocked test sessions, each lasting approximately 1 hour, had to be completed over the course of several days. The sequence of experiments was the same as for the conspicuity area determinations.

During RT measurements targets were presented at predetermined positions across the screen, but subjects perceived them as appearing at random positions. Locations were determined with regard to conspicuity area results. Targets were presented at five locations along each of 16 meridians, originating from the centre of the conspicuity area and together spanning 360°. Each meridian was divided into five sections of different eccentricity, resulting in targets being placed at two positions inside the conspicuity area, on the border of the conspicuity area and at two peripheral positions along each meridian (Figure 5.2). It should be stressed that these positions were determined according to each subject's individual conspicuity area for a given display condition. Targets were presented five times at each position, resulting in 400 target presentations. With catch trials included, a single session consisted of approximately 800 trials.

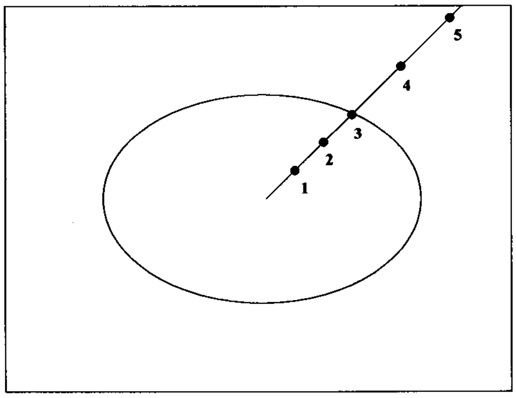


Figure 5.2: Schematic representation of the five eccentricity positions, relative to the conspicuity area, represented by the ellipse. The first and second eccentricity positions are within the conspicuity area, while the third is always on the border of the conspicuity area. The fourth and fifth eccentricity positions are outside the conspicuity area.

5.3 RESULTS

5.3.1 Conspicuity area determination

In order to determine the nature of the conspicuity area for a given target-background combination, the spatial detection boundary was determined. Border coordinates along each of the 16 meridians were calculated from the mean radius around which the last eight reversals concentrated. Coordinates of which the standard deviation exceeded 25% of the calculated value were excluded from the final dataset, since the nature of the conspicuity area could still be estimated adequately from as few as ten data points. Non-linear least squares curve fitting was performed in Matlab, according to code adapted from that developed by N. Baroni and S. Grobler and also K. Pankratov (Grobler, 2000). Since pilot studies showed that conspicuity areas tend to be elliptical, data were fitted according to a function that describes the nature of an ellipse with a generic origin and inclination with regard to the coordinate system of the input data:

$$f = \frac{[(x-h)\cos\phi - (y-k)\sin\phi]^2}{a^2} + \frac{[(x-h)\sin\phi - (y-k)\cos\phi]^2}{b^2} - 1 \quad , \qquad (2)$$

with *(x, y)* representing input data, *(h,k)* centre coordinates, *a* and *b* radii of the major and minor axes respectively and ϕ the inclination of the ellipse.

Since results from curve fitting showed $\phi \approx 0$ (see later), the radius of each sector of the estimated ellipse could be determined by expressing the ellipse equation in polar coordinates:

$$r^2 = \frac{a^2 b^2}{a^2 \sin^2 \vartheta + b^2 \sin^2 \vartheta} \quad . \qquad (3)$$

Goodness of fit is reported as the square root of the mean squared error (RMSE) associated with the measured radius:

$$RMSE = \sqrt{\frac{\Sigma(r - r')^2}{n - m}} \,, \qquad (4)$$

with r representing the measured radius, r′ the estimated radius, n the number of data points and m the number of fitted coefficients.

Results are shown in Figures 5.3a – h and summarized in Table 5.1. Graphs, grouped according to each subject, depict conspicuity areas for each display condition.

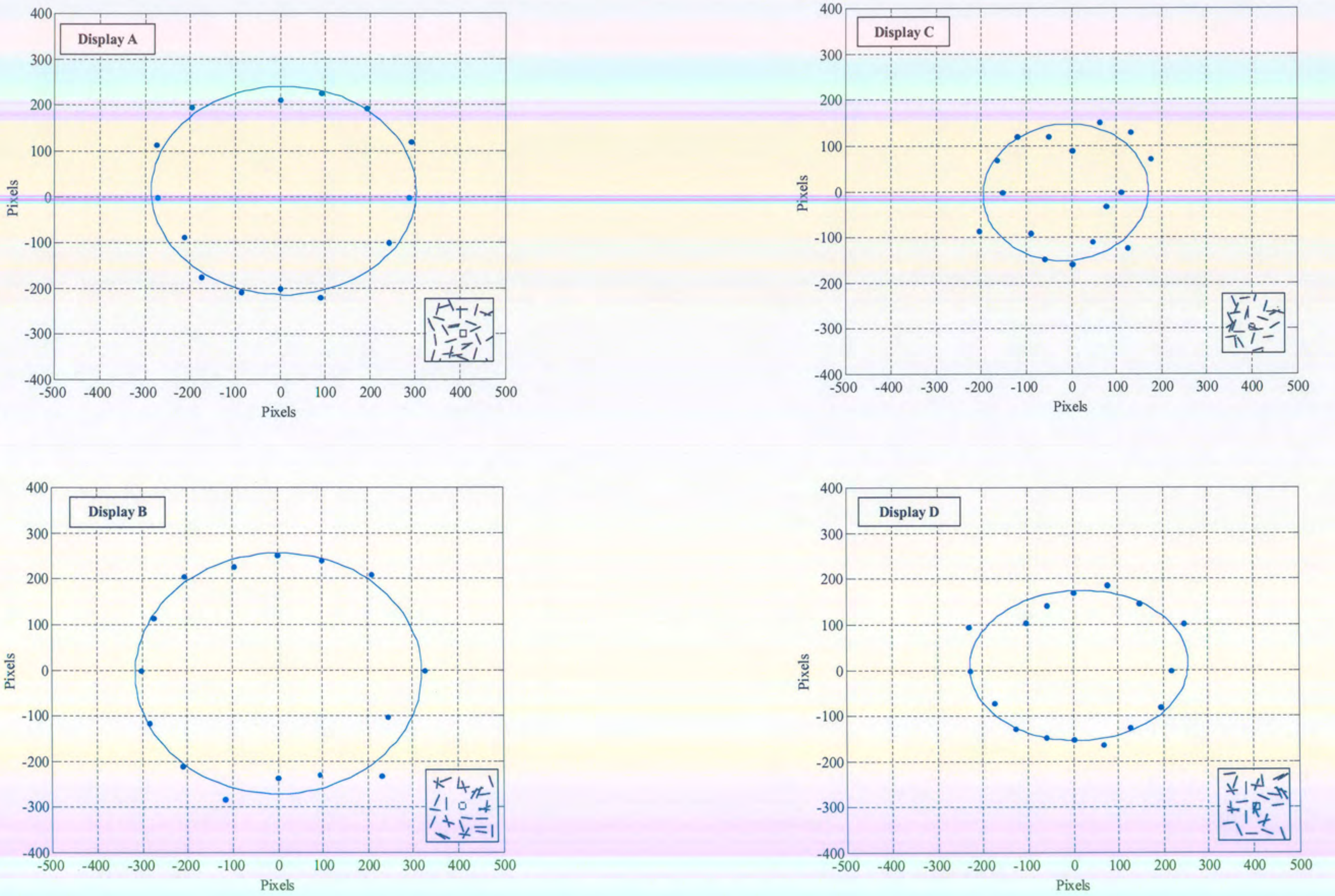


Figure 5.3a Conspicuity areas of Subject 1 for display conditions A–D. Dots represent measured data, while the solid lines represent the fitted curves. Examples of the display conditions are given in the insets.

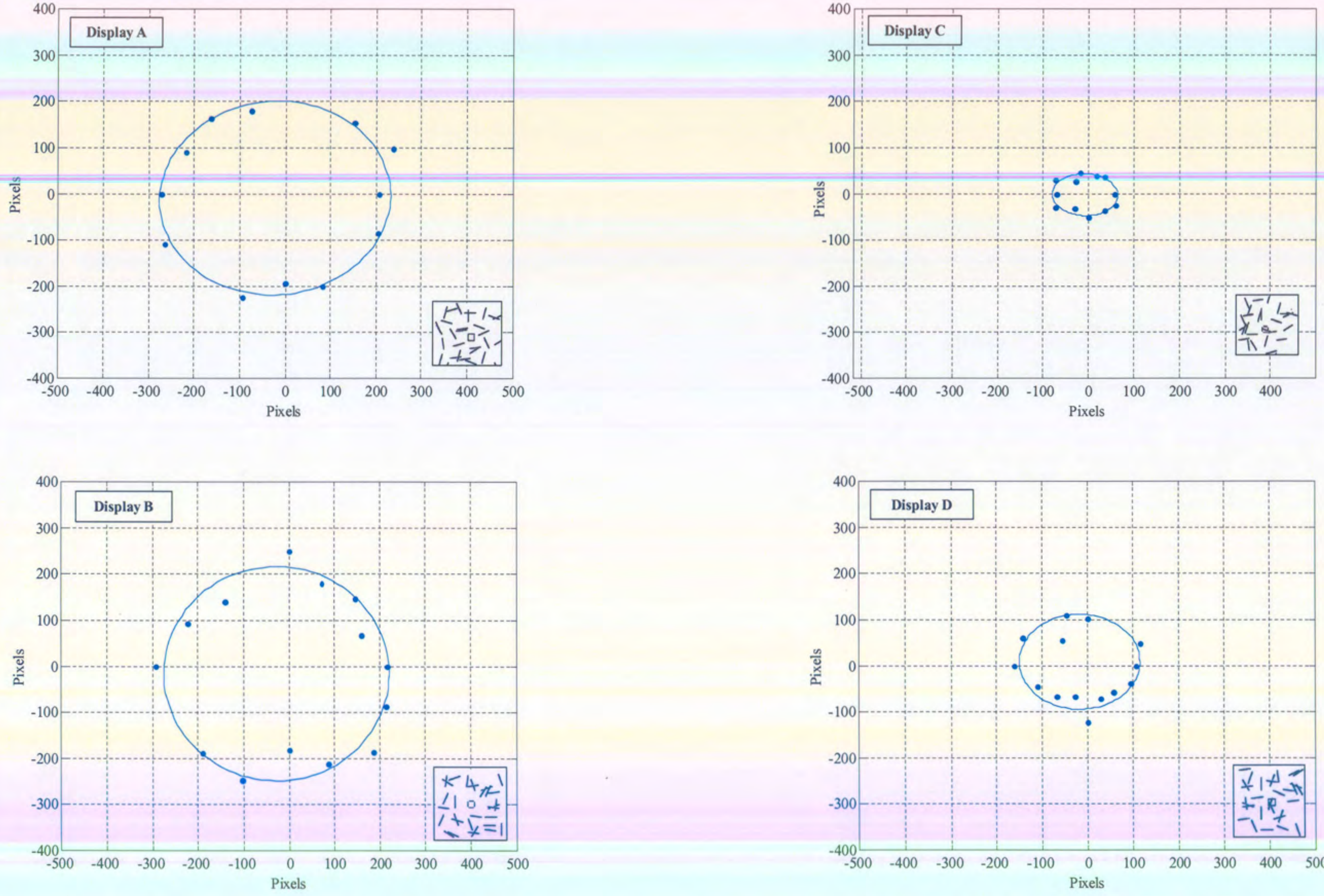


Figure 5.3b: Conspicuity areas of Subject 2 for display conditions A –D. Dots represent measured data, while the solid lines represent the fitted curves. Examples of the display conditions are given in the insets.

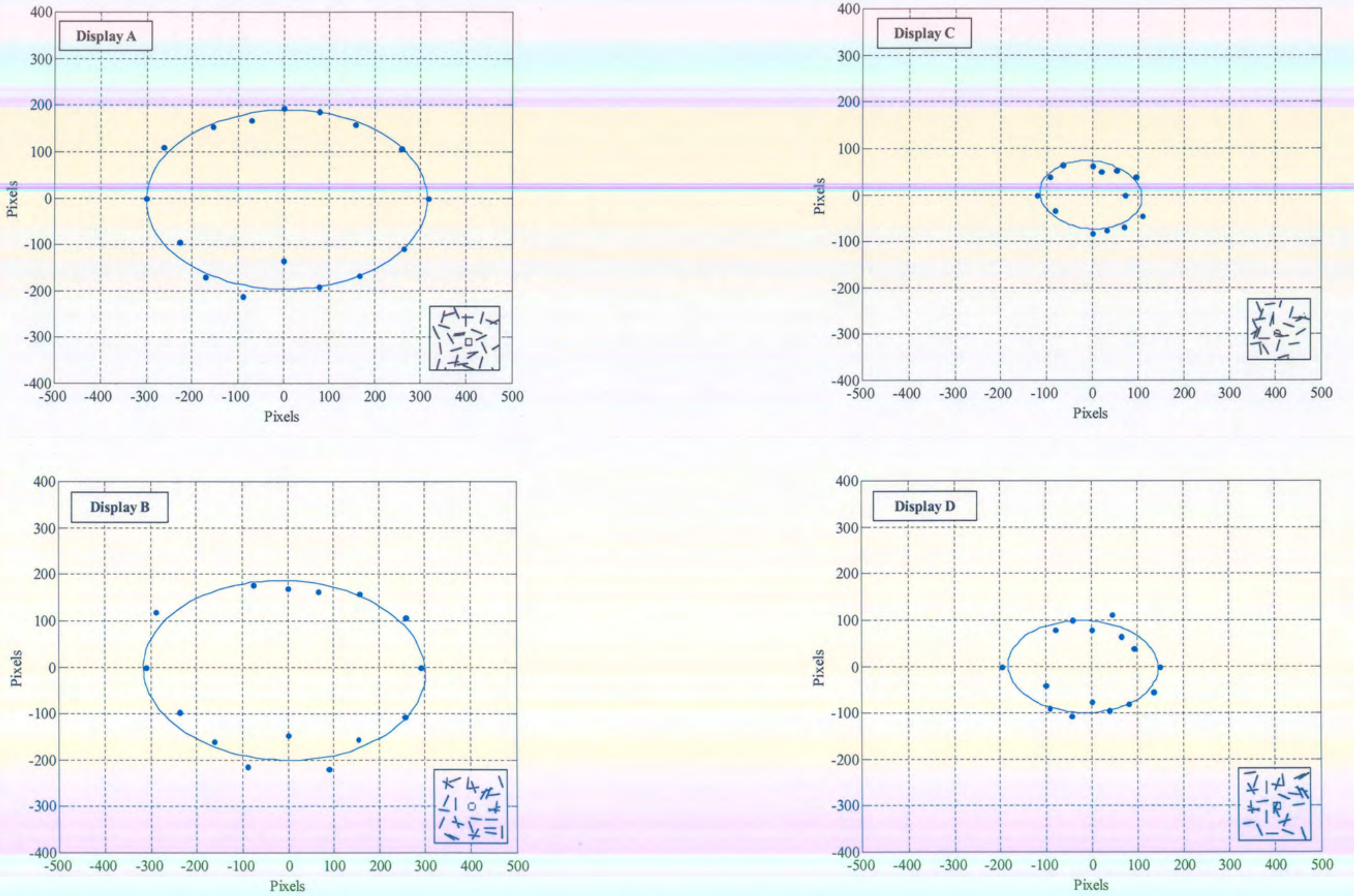


Figure 5.3c: Conspicuity areas of Subject 3 for display conditions A –D. Dots represent measured data, while the solid lines represent the fitted curves. Examples of the display conditions are given in the insets.

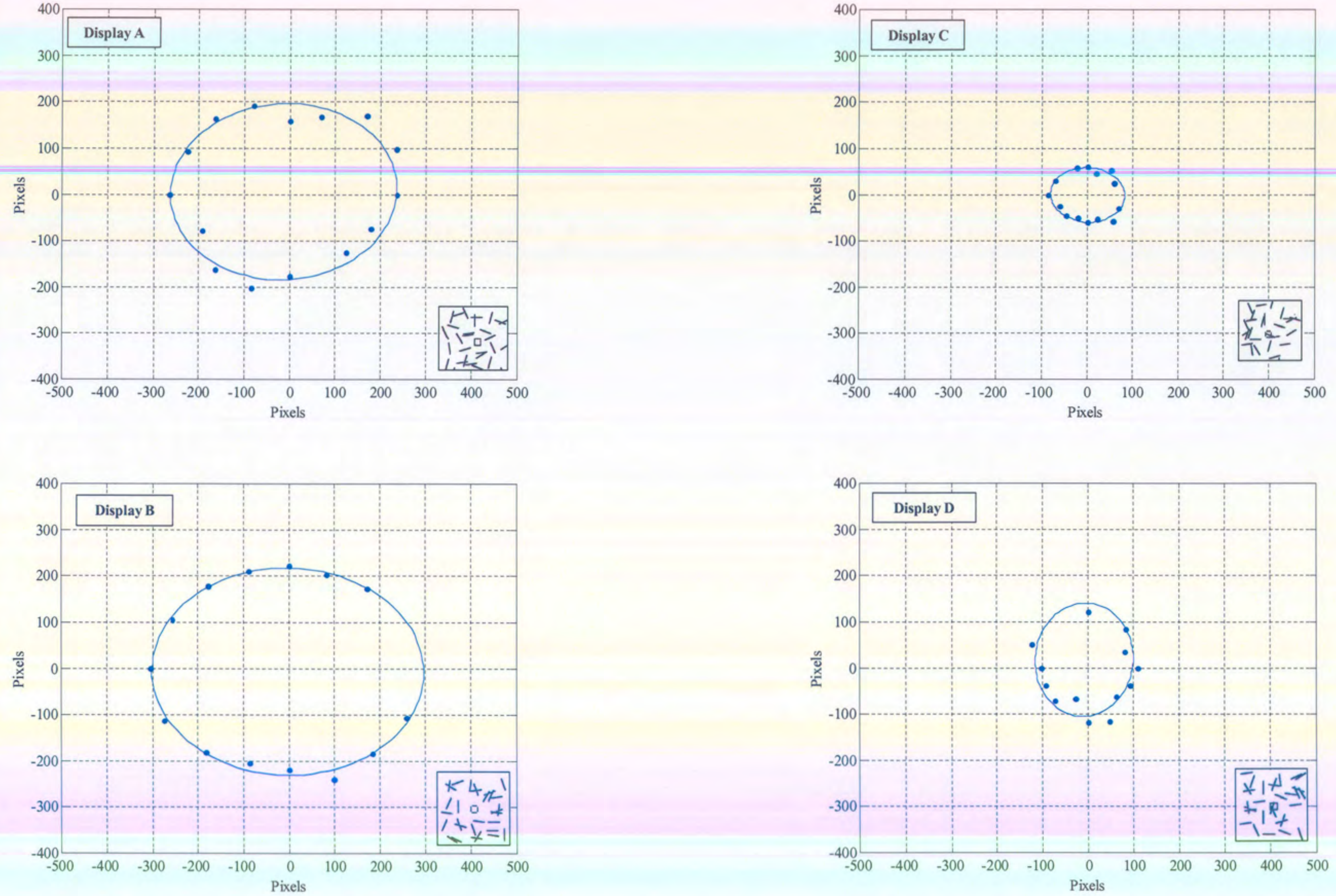


Figure 5.4d: Conspicuity areas of Subject 4 for display conditions A –D. Dots represent measured data, while the solid lines represent the fitted curves. Examples of the display conditions are given in the insets.

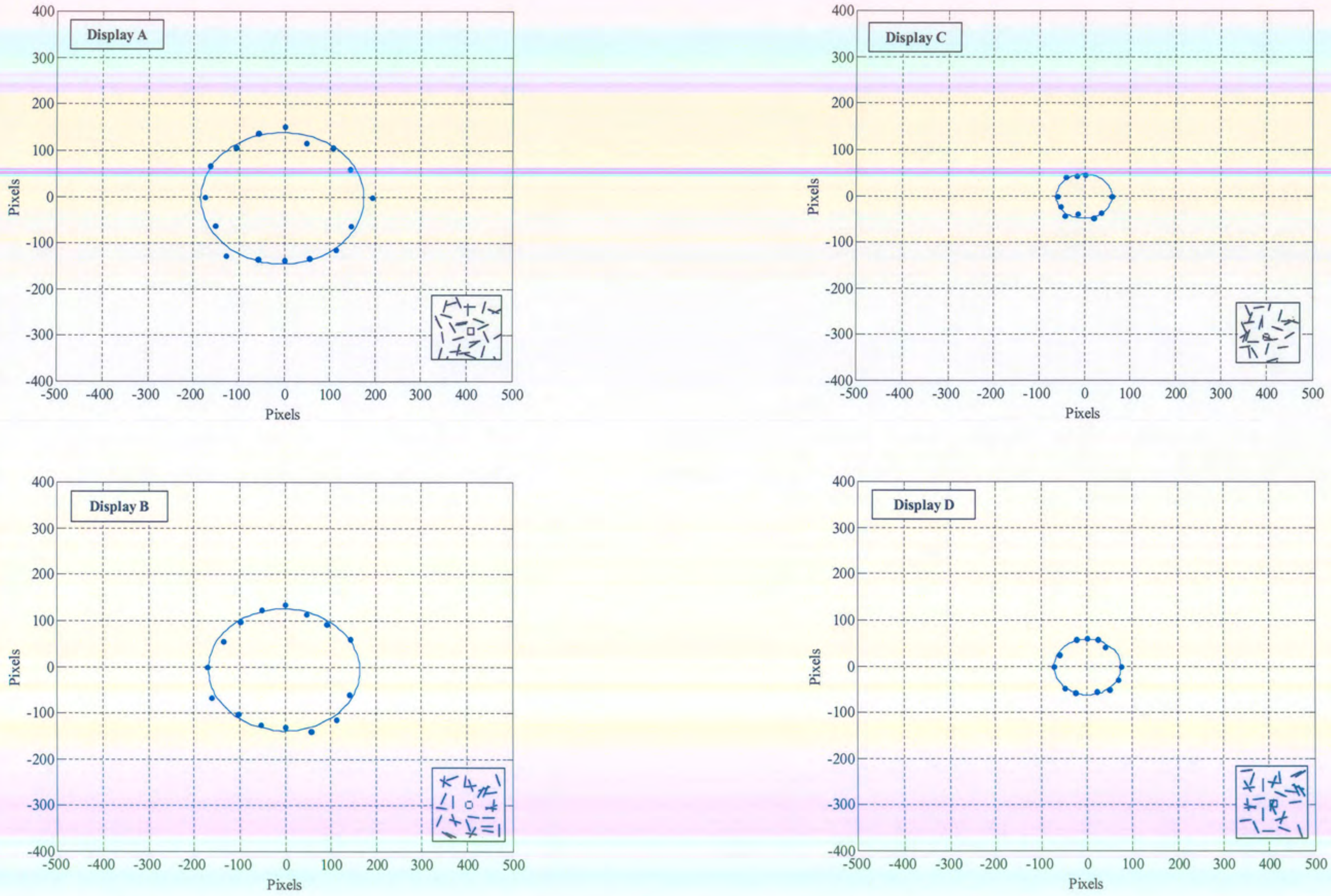


Figure 5.3e: Conspicuity areas of Subject 5 for display conditions A –D. Dots represent measured data, while the solid lines represent the fitted curves. Examples of the display conditions are given in the insets.

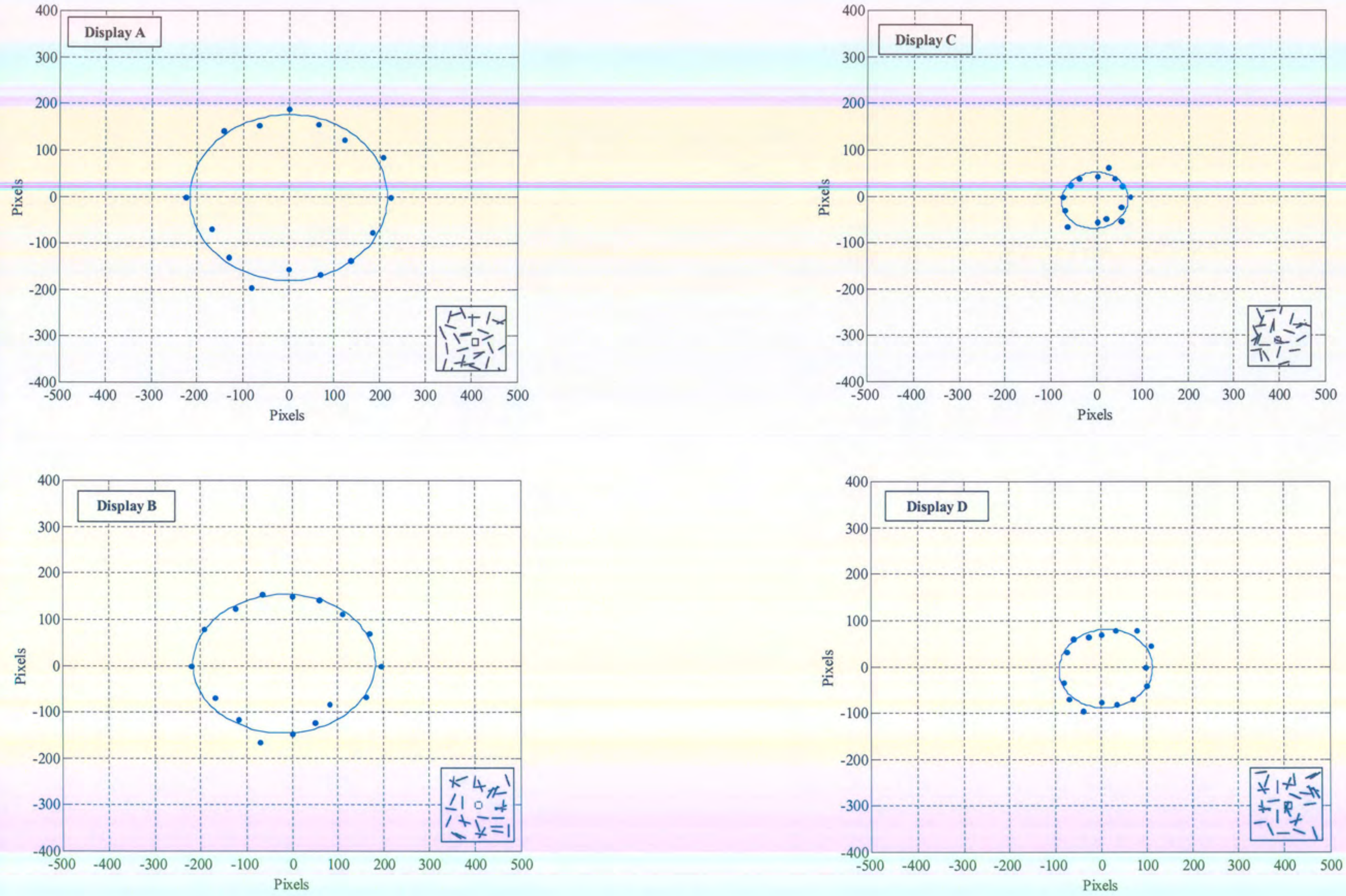


Figure 5.3f: Conspicuity areas of Subject 6 for display conditions A –D. Dots represent measured data, while the solid lines represent the fitted curves. Examples of the display conditions are given in the insets.

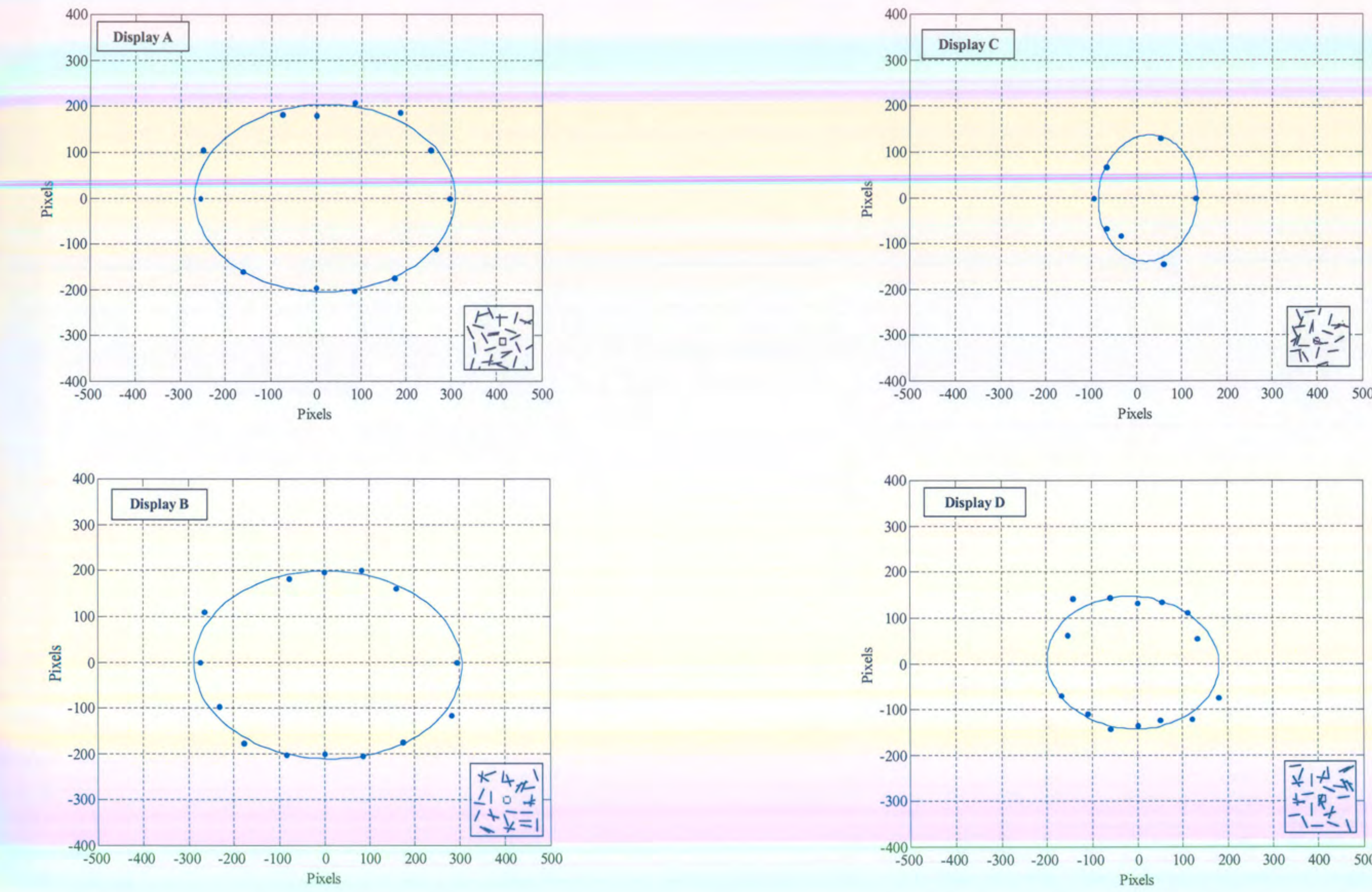


Figure 5.3g: Conspicuity areas of Subject 7 for display conditions A –D. Dots represent measured data, while the solid lines represent the fitted curves. Examples of the display conditions are given in the insets.

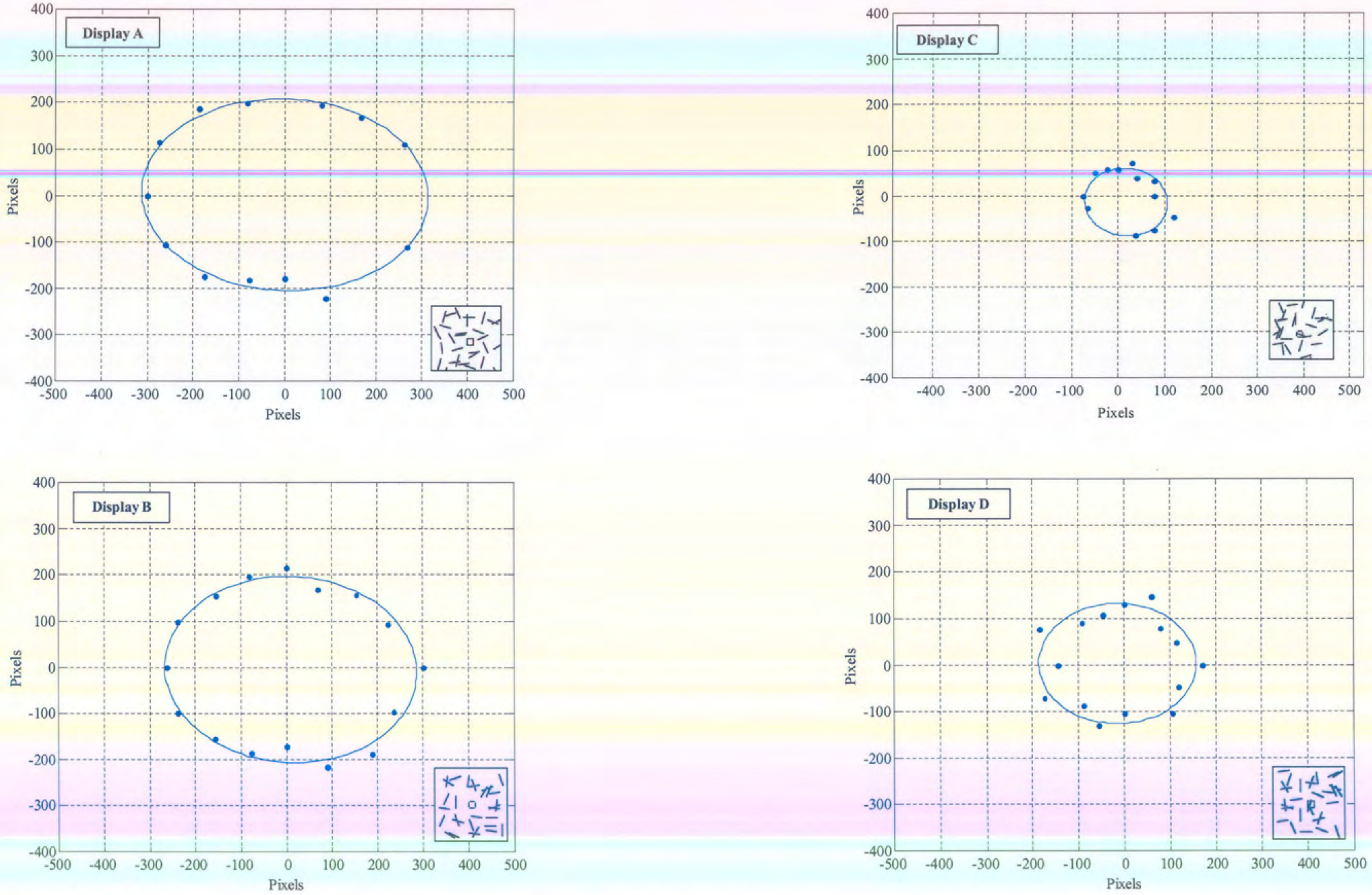


Figure 5.3h: Conspicuity areas of Subject 8 for display conditions A –D. Dots represent measured data, while the solid lines represent the fitted curves. Examples of the display conditions are given in the insets.

Table 5.1: Summary of conspicuity area determination results. Display conditions are as depicted in Figure 5.1.

Display condition	Ellipse Parameter	Subject							
		1	2	3	4	5	8	7	8
A	h (pixels)	9	-22	8	-13	-4	1	21	2
	k (pixels)	13	-9	-3	8	-3	-1	-1	1
	a (pixels)	294	254	307	249	179	218	288	313
	b (pixels)	227	209	193	190	142	178	204	208
	a:b	1.29	1.21	1.59	1.31	1.26	1.21	1.41	1.52
	Area (degree^2)	361.68	289.33	321.22	258.15	140.10	211.15	319.29	349.26
	Detection probability	0.34	0.27	0.30	0.24	0.13	0.20	0.30	0.33
	RMSE (pixels)	28.07	30.19	23.66	28.58	13.07	19.45	21.81	17.57
B	h (pixels)	3	-27	-7	-2	-1	-16	9	11
	k (pixels)	-7	-16	-7	-7	-7	4	-6	-5
	a (pixels)	318	247	308	299	167	200	296	277
	b (pixels)	264	233	194	224	132	149	205	202
	a:b	1.20	1.06	1.59	1.33	1.26	1.34	1.44	1.37
	Area (degree^2)	450.56	313.15	323.20	362.61	120.71	163.62	329.82	303.22
	Detection probability	0.42	0.30	0.30	0.34	0.11	0.15	0.31	0.29
	RMSE (pixels)	27.17	44.59	27.42	12.86	11.46	21.07	18.64	22.14
C	h (pixels)	-13	-8	-3	2	-1	-4	25	17
	k (pixels)	-2	-2	0	0	-1	-8	-1	-14
	a (pixels)	184	71	111	82	61	72	109	89
	b (pixels)	149	46	73	59	48	60	138	75
	a:b	1.23	1.57	1.53	1.38	1.27	1.19	0.79	1.20
	Area (degree^2)	150.21	18.01	44.79	26.66	16.34	24.09	82.85	36.93
	Detection probability	0.14	0.02	0.04	0.03	0.02	0.02	0.08	0.03
	RMSE (pixels)	47.45	10.90	19.65	9.13	5.91	13.58	47.17	9.02
D	h (pixels)	13	-17	-19	-8	2	11	-10	-14
	k (pixels)	11	9	0	18	-2	-4	2	3
	a (pixels)	242	132	164	107	73	102	190	171
	b (pixels)	164	103	100	123	60	85	144	129
	a:b	1.47	1.28	1.65	0.87	1.22	1.20	1.32	1.32
	Area (degree^2)	217.16	75.40	89.80	72.96	24.48	47.83	150.19	121.86
	Detection probability	0.20	0.07	0.08	0.07	0.02	0.05	0.14	0.11
	RMSE (pixels)	28.49	30.52	26.94	23.84	4.90	14.98	21.01	29.94

Conspicuity areas generally seem to be elliptical with negligible inclination, centred approximately on the fixation point and the major axis being horizontal, as shown by (h,k) values close to (0,0) and a:b values generally greater than 1. The observations are supported by small errors associated with the measured data. These findings agree with earlier reports in terms of the general shape of the conspicuity area (Chaikin et al., 1962; Engel, 1971), but differ with regard to its location, since Chaikin reported a significantly larger part of the ellipse to be located in the upper hemifield. The results also show the border of the conspicuity area to be regular, supported by small errors, reflected in RMSE values, associated with measured data. This finding contradicts irregular boundaries reported previously (Courtney & Chan, 1985; Courtney, 1989). However, the results reported by

Courtney and Chan could have been influenced by their technique used to estimate target location, since targets were located at discreet positions using a fairly large step size and location then had to be indicated on a similar display showing all possible target locations only after the display disappeared. Subjective error could therefore also contribute to the observed boundary irregularities. Results from the current study, on the other hand, are more objective, since boundary coordinates are accurately homed in on by using an adaptive psychophysical method.

Size of the conspicuity area is reported as the area in square degrees visual angle. Despite inter-subject variation, general trends in terms of size of the conspicuity area (Figure 5.4) show conspicuity areas for display condition B to be the biggest, while display condition C produced the smallest conspicuity areas (paired t-test, $p<0.0001$). Significant differences were also observed when the size of conspicuity areas for display condition A was compared to those of display conditions C and D (paired t-tests, $p<0.0001$). It is interesting to note that conspicuity areas produced by display conditions A and B are of similar sizes, as determined by a paired t-test, despite the anticipated change to SNR by changing the target from a complete square to one consisting of four separate lines, orientated perpendicular to one another. These observations are supported by the calculated detection probability (Table 5.1), expressed as the ratio between area of the conspicuity area and that of the entire display (Engel, 1977). The results therefore show that allowing distractors to intersect the target, significantly reduces target detectability, probably by increasing target-distractor similarity. However, when distractors were not allowed to intersect the target, only separating the four constituting lines of the square, while preserving their relative orientations, target-distractor similarity was probably not increased sufficiently to result in altered target detectability. This is in line with Duncan and Humphreys' (1989) "search surface", which suggests that search performance during a visual target acquisition task will be altered sufficiently to be reflected in behavioural outcomes when (i) either target-distractor similarity or distractor heterogeneity is high or when (ii) target-distractor similarity nd distractor heterogeneity are increased simultaneously.

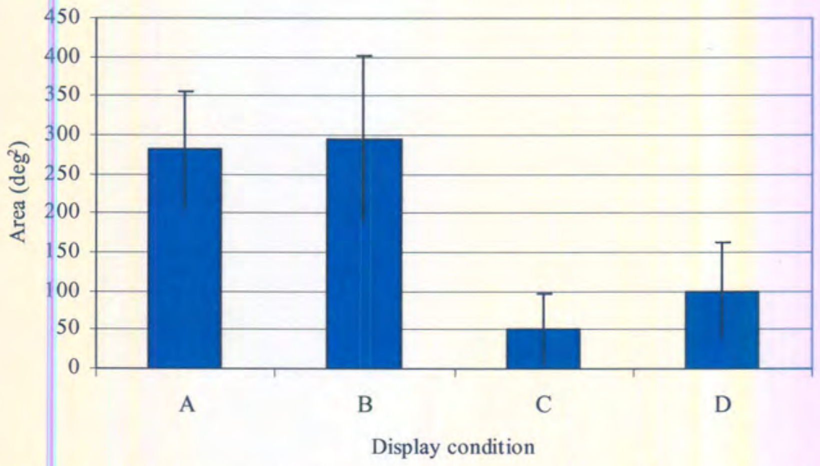


Figure 5.4: General trend of conspicuity area sizes as determined from pooled data for four different display conditions (A- D).

5.3.2 RT measurements

In order to investigate the relationship between target position and a target's visual conspicuity, RT were measured at specific eccentricity positions, relative to the conspicuity area. Each meridian was divided in to five sectors, resulting in targets being located at equal distances well within the conspicuity area, at two equidistant peripheral positions as well as on the boundary of the conspicuity area. It should be noted that these locations are determined relative to each subject's resulting conspicuity area for a given display condition, as shown in Figure 5.2. General RT trends for the four display conditions, depicting mean RT across five eccentricity positions, are shown in Figure 5.5 and summarized in Table 5.2.

Table 5.2: Mean RT, measured in milliseconds, at each of the five eccentricity positions. Slope (m) between positions 1 and 3, as well as between positions 3 and 5 is also shown.

Display condition	Eccentricity position (EP)					$m_{(1-3)}$ (ms/EP)	$m_{(3-5)}$ (ms/EP)	$m_{(3-5)} / m_{(1-3)}$ (ms/EP)
	1	2	3	4	5			
	RT (ms)	RT (ms)	RT (ms)	RT (ms)	RT (ms)			
A	444	472	571	902	1164	64	297	4.66
B	437	453	519	808	1099	41	290	7.08
C	478	506	550	876	1588	36	519	14.55
D	434	482	546	890	1461	56	457	8.16

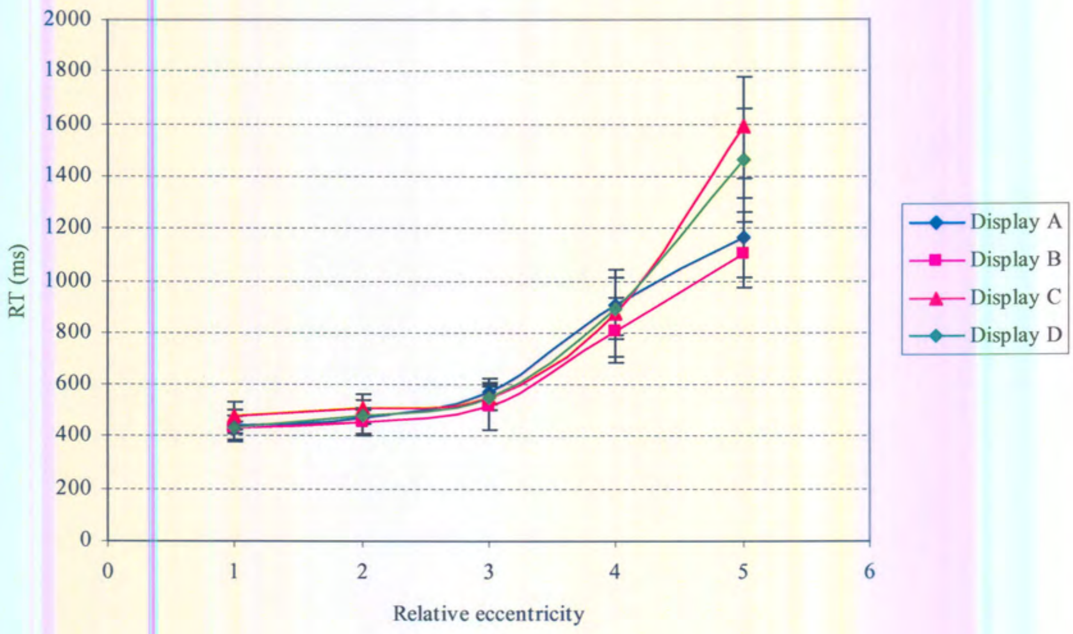


Figure 5.5: General RT trends observed for the four different display conditions.

Graphic representations of results of each subject's mean RT at specific eccentricity positions are shown in Figures 5.6a-h and summarized in Table 5.3. Errors associated with the measurements are reported as standard error of the mean:

$$MSE = \sqrt{\frac{s}{n}} \quad , \qquad (5)$$

where s representes standard deviation from the mean and n the number of observations. Results of variance analysis, performed with the statistical software package GenStat (2000), are shown in Table 5.4a-b.

Table 5.3: RT results obtained at different eccentricity positions for display conditions A – D.

Display condition	Eccentricity	Subject 1		Subject 2		Subject 3		Subject 4		Subject 5		Subject 6		Subject 7		Subject 8	
		RT (ms)	MSE	RT (ms)	MSE	RT (ms)	MSE	RT (ms)	MSE	RT (ms)	MSE	RT (ms)	MSE	RT (ms)	MSE	RT (ms)	MSE
A	1	432	11.26	418	7.05	416	9.23	550	7.63	369	4.66	499	9.38	391	9.52	475	15.00
	2	477	19.55	449	9.71	433	8.61	581	11.70	395	8.58	503	10.00	407	10.39	531	23.85
	3	578	23.31	545	27.66	589	28.59	639	16.99	521	21.10	579	13.42	494	30.53	623	43.83
	4	868	36.99	838	52.06	945	43.33	1139	66.32	828	33.94	959	58.80	804	41.78	837	44.35
	5	1004	55.36	1315	90.08	1229	82.84	1415	50.67	998	47.72	1223	72.13	1092	41.65	1037	45.31
B	1	395	10.62	399	8.02	439	10.95	534	9.95	410	8.64	452	8.92	454	12.22	411	11.01
	2	424	16.53	424	11.90	451	10.39	567	13.22	405	7.89	453	8.82	470	10.41	433	13.54
	3	572	33.70	432	12.90	648	32.83	627	22.87	395	8.11	470	9.26	516	13.88	492	14.74
	4	762	22.95	674	45.46	1057	67.28	909	44.42	696	51.25	771	47.88	823	47.73	774	38.91
	5	936	34.45	996	46.94	1344	60.09	1118	37.24	1088	74.49	1181	59.77	1086	34.43	1045	54.15
C	1	495	18.98	413	10.07	472	17.61	515	9.54	445	16.28	429	9.40	574	36.36	483	14.99
	2	543	20.80	427	11.93	505	15.99	523	10.80	459	11.60	448	9.33	592	22.88	550	40.49
	3	557	48.68	539	75.15	582	18.62	547	8.79	463	13.97	509	18.80	634	30.06	565	26.97
	4	903	59.96	695	39.59	1192	69.78	831	63.14	758	43.86	930	84.71	1000	53.13	701	29.15
	5	1541	81.75	1592	103.21	1685	80.89	1758	93.98	1276	60.89	1692	135.50	1822	71.46	1336	65.46
D	1	437	11.87	451	18.82	424	7.04	450	8.61	394	10.16	422	4.61	476	7.07	420	5.87
	2	494	26.29	503	50.17	482	13.58	489	8.58	424	11.18	448	9.76	519	20.84	494	16.70
	3	551	15.99	559	38.41	568	29.77	516	15.86	467	17.46	527	31.60	635	25.22	547	18.19
	4	860	34.99	763	37.84	1067	60.78	852	45.79	763	52.45	905	51.50	1063	66.68	851	51.92
	5	1288	50.02	1443	94.07	1489	112.08	1558	65.64	1332	40.83	1800	117.99	1610	68.22	1169	53.09

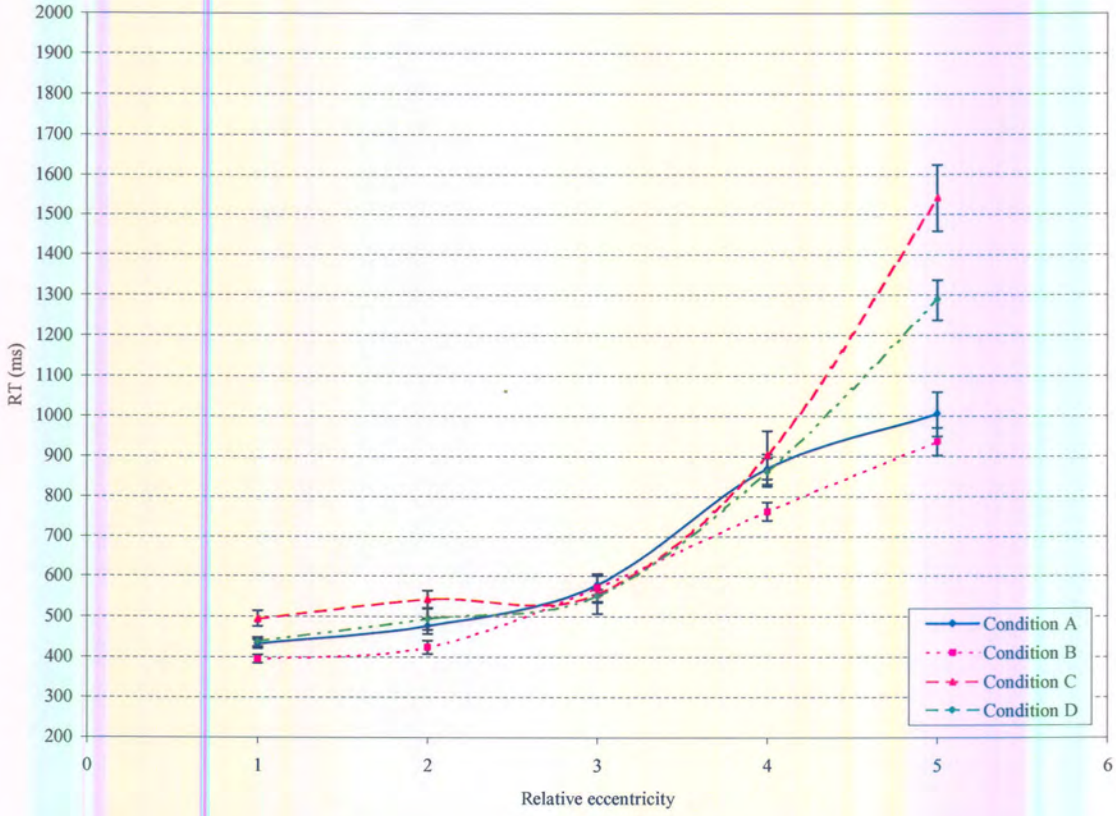


Figure 5.6(a): RT vs. relative eccentricity for Subject 1. Eccentricity position 3 represents the border of the conspicuity area as depicted in Figure 5.2.

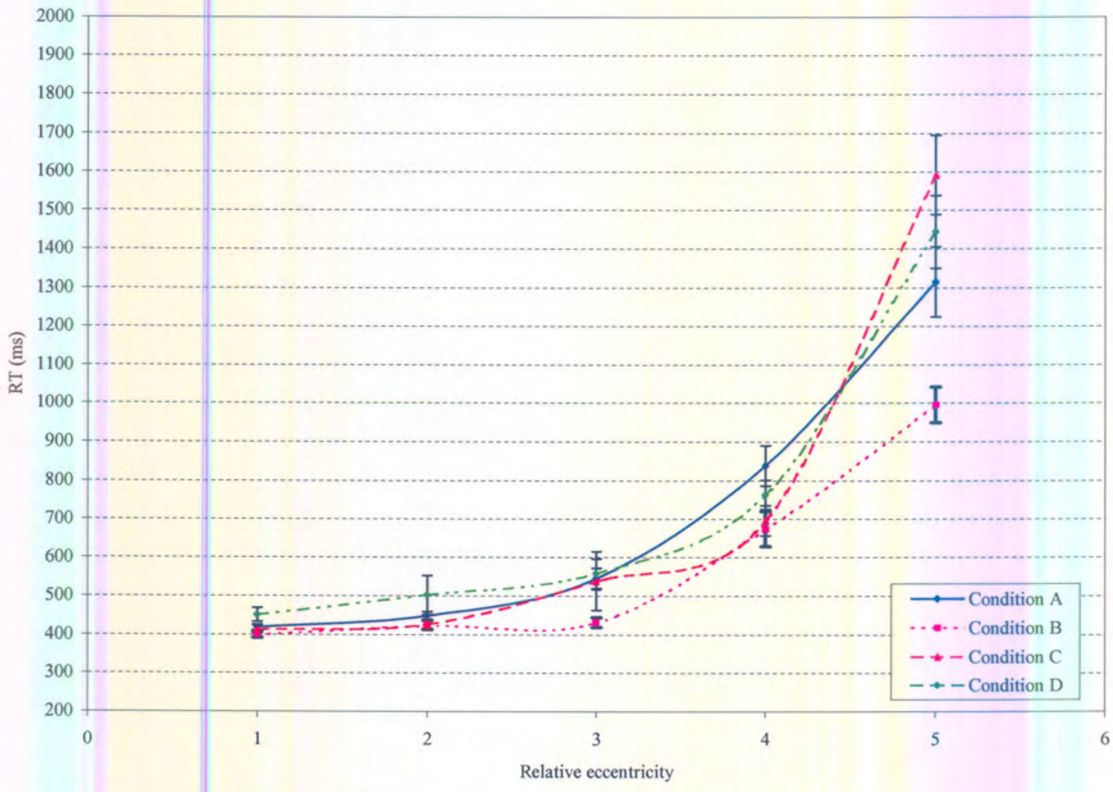


Figure 5.6(b): RT vs. relative eccentricity for Subject 2. Eccentricity position 3 represents the border of the conspicuity area as depicted in Figure 5.2.

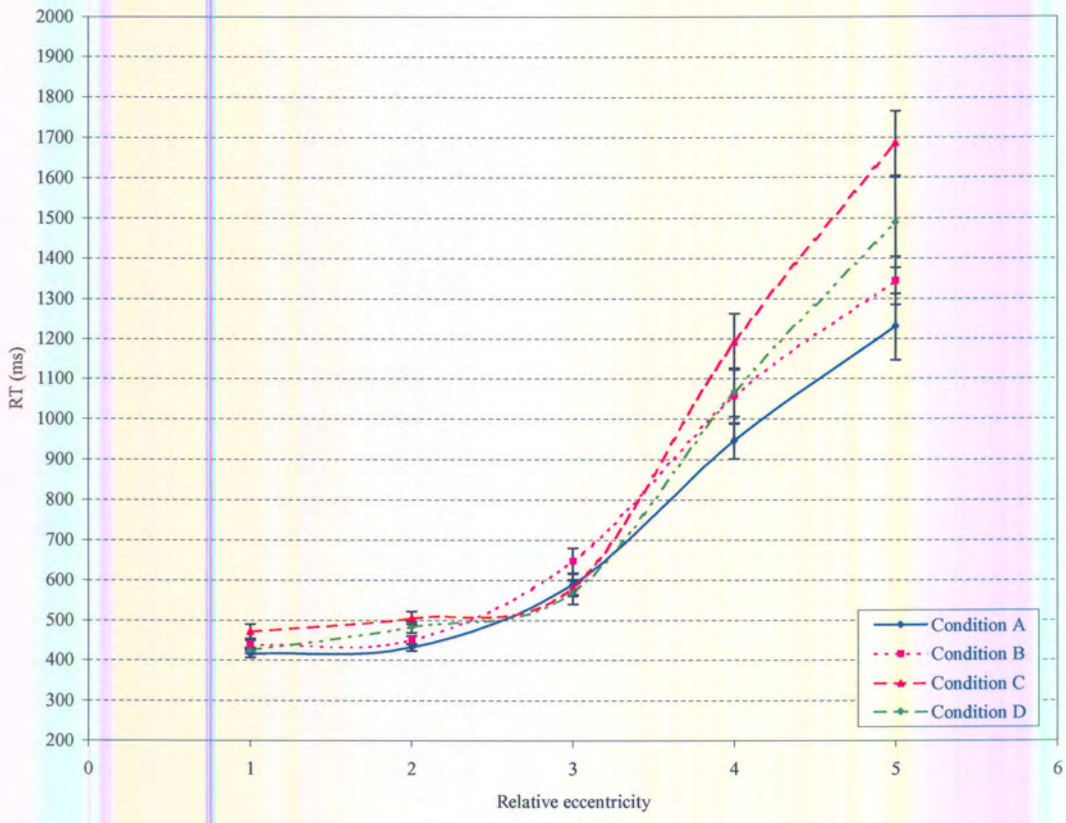


Figure 5.6(c): RT vs. relative eccentricity for Subject 3. Eccentricity position 3 represents the border of the conspicuity area as depicted in Figure 5.2.

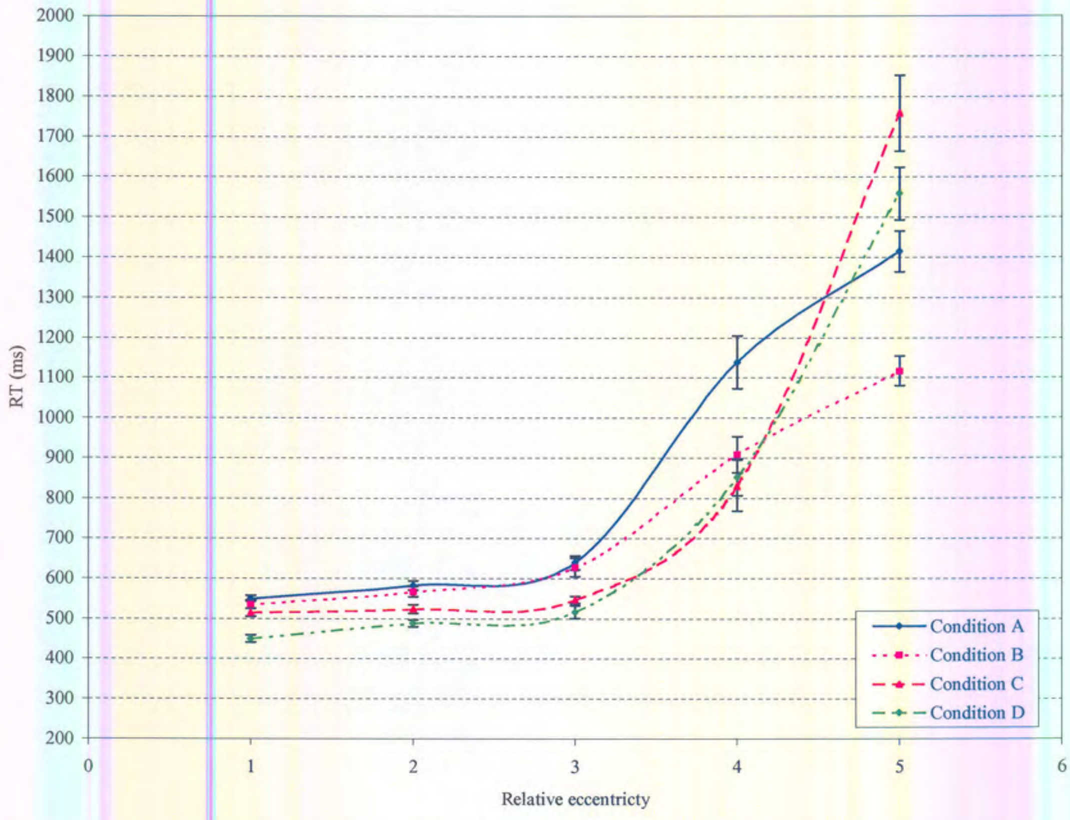


Figure 5.6(d): RT vs. relative eccentricity for Subject 4. Eccentricity position 3 represents the border of the conspicuity area as depicted in Figure 5.2.

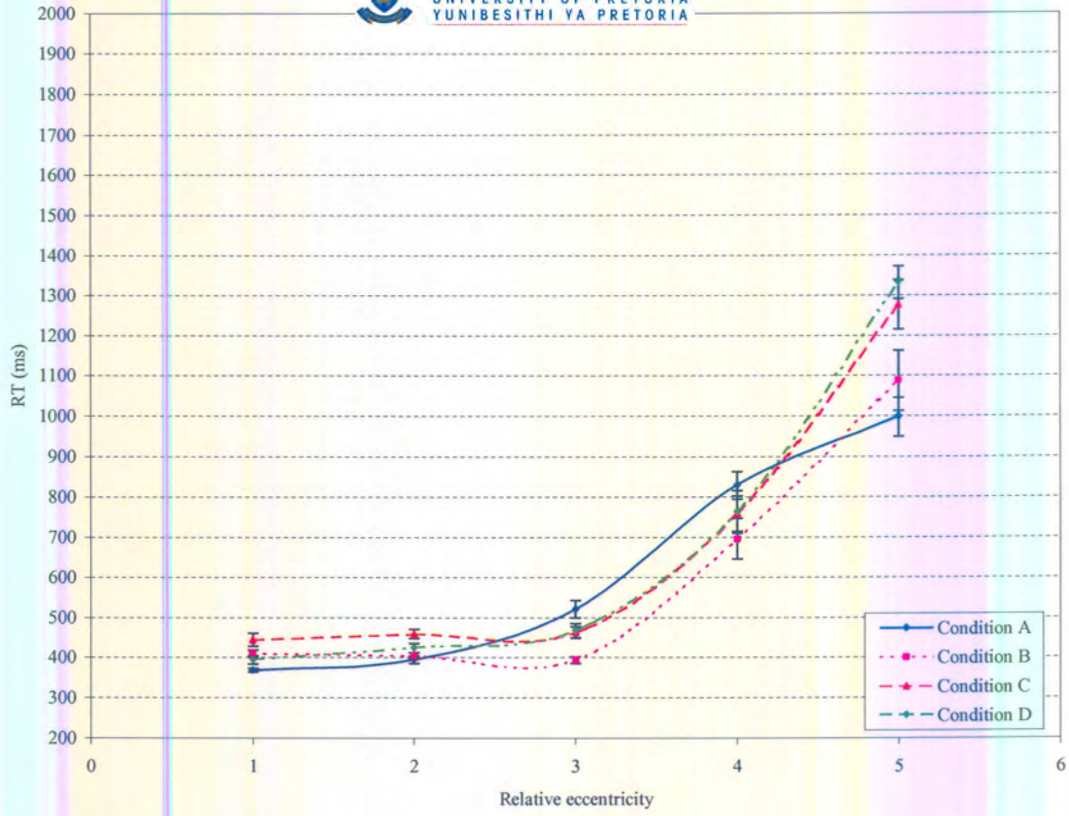


Figure 5.6(e): RT vs. relative eccentricity for Subject 5. Eccentricity position 3 represents the border of the conspicuity area as depicted in Figure 5.2

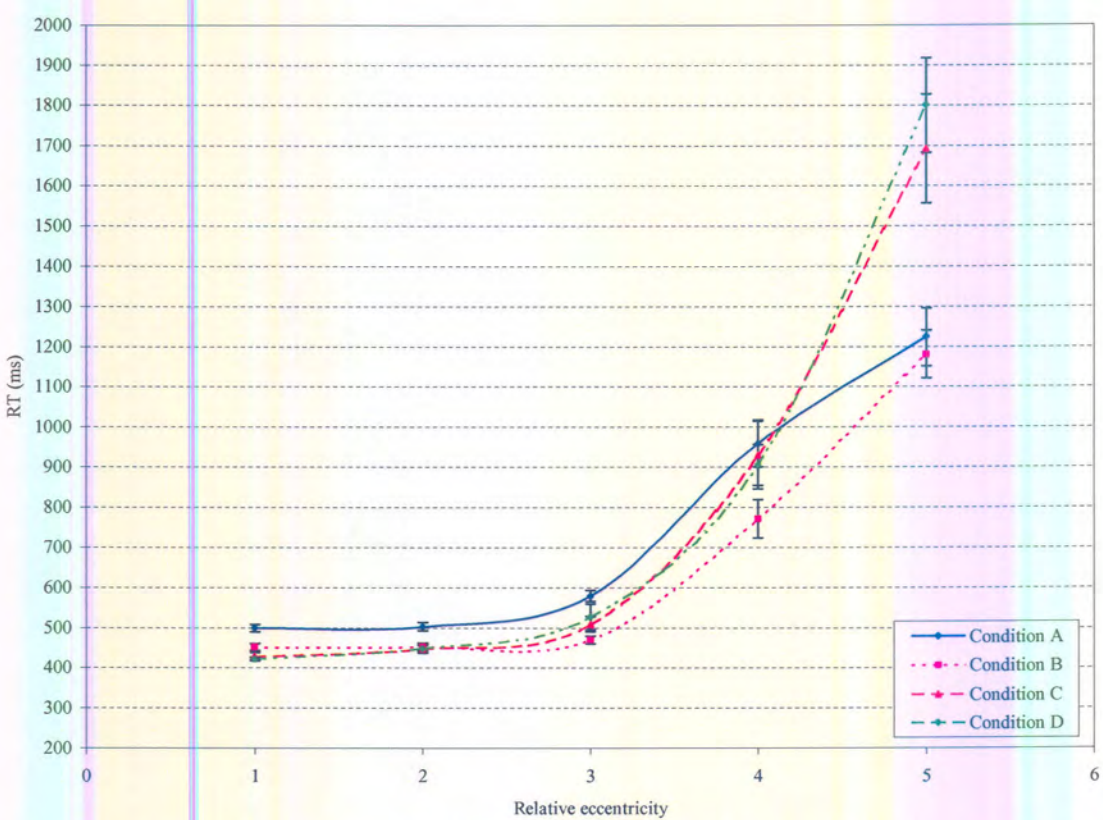


Figure 5.6(f): RT vs. relative eccentricity for Subject 6. Eccentricity position 3 represents the border of the conspicuity area as depicted in Figure 5.2

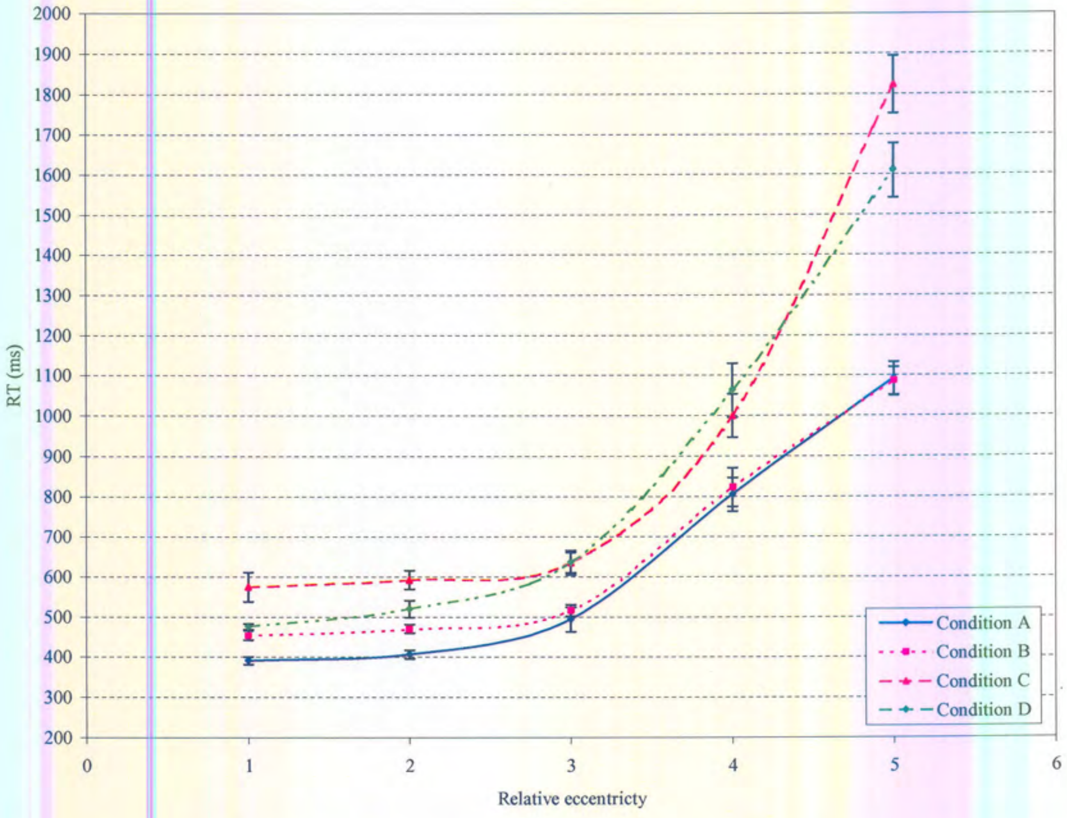


Figure 5.6(g): RT vs. relative eccentricity for Subject 7. Eccentricity position 3 represents the border of the conspicuity area as depicted in Figure 5.2

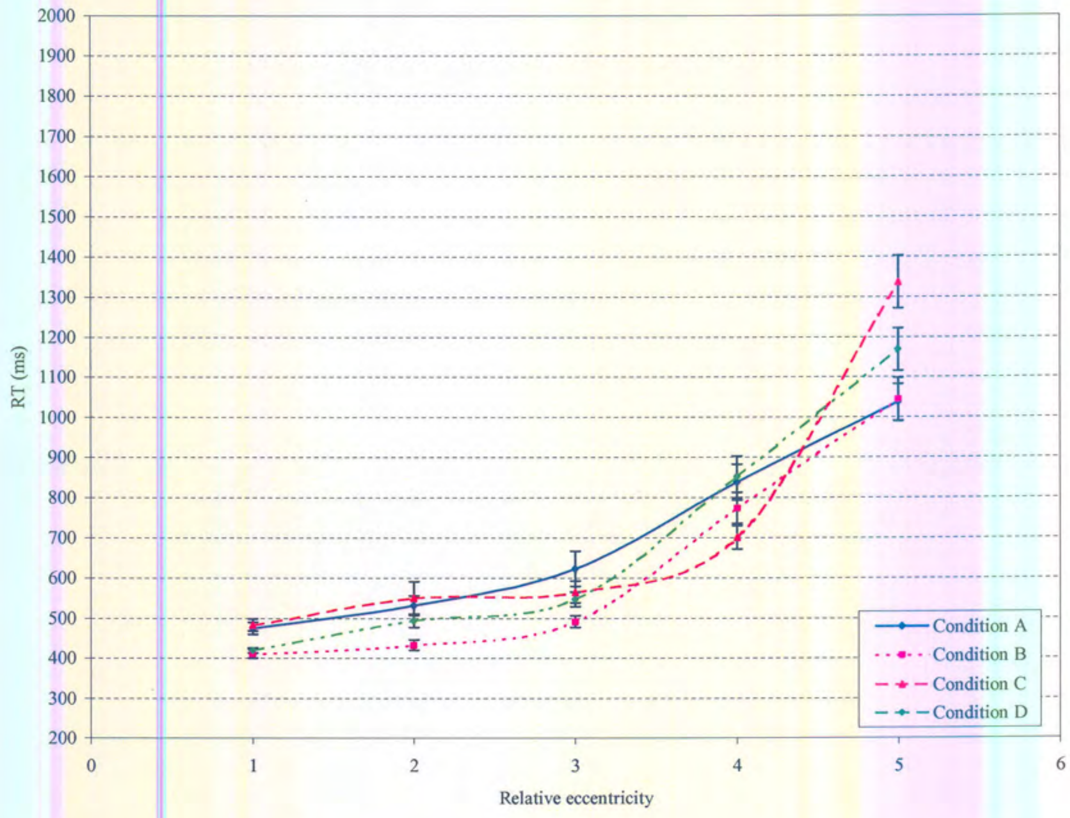


Figure 5.6(h): RT vs. relative eccentricity for Subject 8. Eccentricity position 3 represents the border of the conspicuity area as depicted in Figure 5.2

Table 5.4 (a): Results of variance analysis performed on RT results. Interaction between effects is represented by *x*.

Source of variation	*F*-value	*p*-value	Contribution to variance (%)
Subject	51.07	<0.001	1.9
Display condition	104.86	<0.001	1.7
Eccentricity position	3238.76	<0.001	70.6
Angle	10.14	<0.001	0.8
Subject x display condition	9.91	<0.001	1.1
Subject x eccentricity position	12.72	<0.001	1.9
Subject x angle	2.37	<0.001	1.4
Display condition x eccentricity position	59.17	<0.001	3.9
Display condition x angle	1.07	0.356	0.3
Eccentricity position x angle	6.09	<0.001	2.0
Subject x display condition x eccentricity position	2.67	<0.001	1.2
Subject x display condition x angle	1.05	<0.001	1.8
Subject x eccentricity position x angle	1.45	<0.001	3.3
Display condition x eccentrcity position x angle	1.20	<0.001	1.1

Table 5.4 (b): Results of variance analysis of RT results, showing effects' contribution to variance when effect of eccentricity position is not considered. Interaction between effects is represented by *x*.

Source of variation	*F*-value	*p*-value	Contribution to variance (%)
Subject	51.07	<0.001	6.6
Display condition	104.86	<0.001	5.9
Angle	10.14	<0.001	2.8
Subject x display condition	9.91	<0.001	3.8
Subject x eccentricity position	12.72	<0.001	6.6
Subject x angle	2.37	<0.001	4.8
Display condition x eccentricity position	59.17	<0.001	13.4
Display condition x angle	1.07	0.356	1.0
Eccentricity position x angle	6.09	<0.001	6.9
Subject x display condition x eccentricity position	2.67	<0.001	4.1
Subject x display condition x angle	1.05	<0.001	6.2
Subject x eccentricity position x angle	1.45	<0.001	11.4
Display condition x eccentrcity position x angle	1.20	<0.001	3.8

Figures 5.5 and 5.6a-h show that the relationship between RT and eccentricity position resemble exponential functions. Flat slopes are observed up to eccentricity position 3 for all conditions, with steep increases in RT thereafter. Across all conditions RT at eccentricity position 1 is within the same range as RT at eccentricity position 3. Results from one-way variance analysis (ANOVA) further showed that no significant differences exist between slopes between positions 1 and 3 for any of the display conditions. Considering that eccentricity position 3 represents the border of the conspicuity area, these results indicate that

response to targets presented within the conspicuity area is similar across display conditions of varying target detectability.

The slope of the RT vs. set size function is traditionally seen as an indication of search efficiency (e.g. Treisman & Gelade, 1980; Wolfe, 1998), with flat slopes representing easy, pop-out search, while steep slopes are indicative of effortful search. The steep increase in RT between position 3 and 5 compared to that between positions 1 and 3, observed for all display conditions, indicates that search becomes effortful when targets are located outside the conspicuity area. Variance analysis (Table 5.4) also indicated that eccentricity position is a highly significant ($p<0.001$) contributor to observed variance of the RT data, thereby emphasising the strong relationship between RT and eccentricity position (Carrasco et al., 1995; Scialfa & Joffe, 1998).

The exponential character of the RT vs. eccentricity position curve further indicates that already difficult target acquisition tasks, as indicated by conspicuity area results, are substantially more difficult at more peripheral target locations. Results from Table 5.2 and subsequent paired t-tests showed that search performance (indicated by the slope between positions 3 and 5) was significantly lower ($p<0.001$) for display condition C than for display condition B, as well as for display condition D compared to display condition A. No statistical differences were observed when search performance of display condition B was compared to that of display condition A. The results are in line with findings from conspicuity area determinations, which showed that target detectability is reduced when distractors are allowed to intersect the target, but not when the target's four constituting lines are separated while preserving their original orientations. The results confirm the inverse relationship between conspicuity area size and search performance (Engel, 1977; Bellamy & Courtney, 1981; Kee et al., 1992), *provided that difficult search is induced*, which as shown here, can be achieved by positioning targets outside of the conspicuity area.

The RT results obtained at peripheral target locations can be explained in terms of the effective set size. Recall that the eccentricity positions were determined relative to the conspicuity area. For conditions associated with small conspicuity areas, and thus low target detectability, the effective peripheral set size would be larger than for display conditions associated with large conspicuity areas. This would result in more possible target locations, which in turn would make search more effortful and require more time to complete. It thus appears that increasing target eccentricity, while simultaneously decreasing target detectability places strain on the information processing mechanism. The results are supported by significant interaction between display condition and eccentricity position ($p<0.001$), as shown by variance analyses (Table 5.4). These findings are in line with suggestions that search efficiency should not be inferrred from RT slopes without taking target eccentricity into account (Carrasco et al, 1995; Wolfe et al., 1998).

Variance analysis also showed that target acquisition behaviour is significantly different ($p <0.001$) for the different subjects, as already suggested by the different conspicuity area sizes observed for a given target acquisition task. The findings are supported by the significant interaction ($p < 0.001$) between subject and eccentricity position. However, the interaction between subject and display condition is almost negligible, indicating that for a given display condition, subjects' behaviour trends were very similar. This can also be seen from the similar trends of the RT vs. eccentricity position functions observed for all subjects across the four conditions. It should be noted, however, that these results indicate that the current group of participants was homogeneous. Further investigation will thus be necessary to determine if these trends are valid for a more heterogeneous group of participants, especially if a general evaluation test of target acquisition behaviour is to be developed.

5.4 DISCUSSION

5.4.1 Why are pronounced difficulty levels for different display conditions observed?

Results showed that difficulty associated with target acquisition in the four display conditions was noticeably different, especially for conditions where distractors intersected the target as opposed to non-intersecting conditions. These findings can be explained according to saliency effects, as reflected in detection probabilities associated with the different conditions.

Salience[6], or conspicuity, of an object has previously been shown to be a function of the feature contrast between display elements, rather than being determined solely by the presence of a specific feature with specified associated salience (Nothdurft, 2000). This means that target salience is a relative rather than absolute concept, based on the relations between display elements, which can in principle be specified by any stimulus attribute (Duncan & Humphreys, 1989). Local as well as global feature contrasts contribute to overall salience of a target, since the similarity between target and distractors as well as the similarity between distractors themselves is determined during visual search (Duncan & Humphreys, 1989; Olds et al., 1999; Nothdurft, 2000). This relationship has been summarized by Duncan and Humphrey's "search surface" (Duncan & Humphreys, 1989), which proposes that target salience is optimal for conditions of low target-distractor similarity together with high distractor-distractor similarity. Local feature contrast probably also explains why isolated

---

[6] The terms "salience" and "conspicuity" both refer to target detectability, but the researcher prefers to use the latter when referring to target detectability within a complete display, as opposed to "salience" which is used to refer to a target's ability to attract attention.

targets are detected more easily than those closely surrounded by distractors, as shown previously by Brown and Monk (1975). It has further also been found that feature contrasts add across dimensions (also see Chapter 3). However, such additivity is non-linear, indicating that the mechanisms responsible for determining feature contrasts of the various dimensions present in the display are not completely independent (Nothdurft, 2000). Resulting salience of a target is therefore the weighted sum of differences between features along several dimensions.

The above principles indicate that target salience differed for the various display conditions. By allowing distractors to intersect the target, local feature contrast between target and distractors decreased since target isolation was abolished. When component lines of the target square were disconnected while simultaneously allowing distractors to intersect, salience decreased even further. This is probably due to decreasing local feature contrast, as well as decreasing global feature contrast. The same principle applies to non-intersecting background conditions: disconnecting component lines of the target square increases the amount of isolation around the target, thereby increasing local feature contrast.

The question as to why saliency differences result in different target acquisition probabilities can be explained according to an attentional account. According to the two-stage process of target acquisition, explained in a previous chapter, target candidates are selected during a pre-attentive processing stage, followed by an attention-dependent processing step during which each selected candidate's status is evaluated. Selection occurs by segmenting possible targets from the background. It follows that the more complex the background, or the less the feature contrast between target and distractors, the more error-prone this segmentation step becomes, since attention needs to be directed at several objects of similar salience (Nothdurft, 2000; Wolfe et al., 2002). The possibility of allowing distractors mistaken for the target to progress to the attention-dependent evaluation stage thus increases. As soon as an item is segmented from the background, information on its target status starts to accumulate, with a decision being made once either the target or distractor threshold is reached (Wolfe et al., 2002). Thus, the more items presented for evaluation owing to ineffective segmentation, the slower the rate of information accumulation, resulting in more cumbersome search.

Taken together, these arguments explain why target acquisition in display condition C proved to be most difficult, while it was easiest in display condition B.

5.4.2 Why is a conspicuity area observed?

The conspicuity area represents the area around a central fixation point in which a target can be detected during a single fixation (Engel, 1971). Results of the present study have shown that the size of the conspicuity area decreases along with decreasing target salience. As explained in the foregoing section, target salience reflects the target's ability to attract visual attention. The conspicuity area can therefore aptly be defined as the normal range of focal attention elicited by a given target during visual target acquisition (Motter, 1998; Steinman et al., 1995). It follows thus that smaller conspicuity areas result from less salient targets (Engel, 1971; Motter, 1998), which are only able to attract visual attention to a limited extent.

At processing level, the relevance of such a zone of focal attention has been shown to be that processing of stimuli within this area is enhanced, while processing of peripheral stimuli is simultaneously inhibited (Steinman et al., 1995; Bachall & Kowler, 1999; Cutzu & Tsotsos, 2003). As reflected in the results from the conspicuity area determination experiments, the zone of focal attention is circular or elliptical (Pan & Eriksen, 1993) and, as shown by Cutzu and Tsotsos (2003), associated with an annular inhibitory surround with finite spatial extent.

In terms of the conspicuity area results, it can thus be argued that when target salience decreases, the region of enhanced processing, surrounded by inhibition of distractors, is reduced. Similarly, for display conditions which produced large conspicuity areas due to high target salience, a larger area benefits from enhanced processing coupled to inhibition in the surrounding region. An attentional account of the conspicuity area, as proposed here, thus also helps to explain the observed difficulty levels associated with the different display conditions.

5.4.3 Why is such a significant eccentricity effect observed?

During investigations into visual target acquisition behaviour, with RT trends as indication of search efficiency, results are usually averaged across all display conditions. It is thus assumed that for a given display condition, target acquisition is equally efficient across the entire display. Carrasco and colleagues (Carrasco et al., 1995), however, reported reduced accuracy and increasing RT as targets were located further from the fixation point. The term "eccentricity effect" was introduced and subsequent studies reported similar results (Wolfe et al., 1998; Scialfa & Joffe, 1998). The existence of a conspicuity area around the fixation point further strengthens the argument against equal target acquisition performance across the entire display and the current study therefore specifically considered the influence of target eccentricity on target acquisition behaviour.

It was initially argued that eccentricity effects arise owing to spatial resolution constraints and lateral inhibition during early processing stages (Carrasco et al., 1995). However, Wolfe and colleagues (Wolfe et al., 1998) later proposed that eccentricity effects should be explained according to an attentional account, seeing that when two targets of equal salience are presented simultaneously, the one closer to the fixation point is attended to, and subsequently processed, first. Support for an attentional account of eccentricity effects also comes from eye movement data. A strong coupling between visual attention mechanisms and saccade programming has been reported (Kowler et al., 1995; Deubel & Schneider, 1996), and findings of an increasing number of fixations together with increasing target eccentricity, underline attentional involvement in eccentricity effects (Scialfa & Joffe, 1998).

Results from the current study are in line with an attentional account of eccentricity effects. The most pronounced eccentricity effects occur at eccentricities beyond the border of the conspicuity area. Given that the conspicuity area represents a zone of focal attention, with finite boundaries and surrounded by an inhibitory region, it follows that targets within the conspicuity area will be processed before targets located at more peripheral locations. This is supported by findings that attention is concentrated at or near the point of fixation at the beginning of a fixation, disengaging if the target was not detected (Scialfa & Joffe, 1998). This strengthens the proposal for an attention-dependent information accumulation phase just after targets were segmented from the background (Wolfe et al., 2002).

It was previously found that eccentricity effects covary with task difficulty (Wolfe et al., 1998). The marked interaction between target eccentricity and task difficulty observed during the current study supports these findings and suggests that target eccentricity becomes more critical as target salience decreases. In terms of the attentional basis of the conspicuity area, it can be argued that focal attention has a narrower distribution in conditions of low target salience and that several attentional shifts will be necessary for detection of more peripherally located targets, as indeed shown by the fact that the number of fixations increases along with increasing target eccentricity (Scialfa & Joffe, 1998). However, the increase in eye movements for peripheral targets is less than would have been expected if each item was saccaded to individually, supporting the notion of a zone of focal attention that shifts across the display.

5.4.4 What do the results mean in terms of processing of visual information?

The mechanisms underlying visual target acquisition behaviour are often inferred from RT trends observed when set size or target-distractor discriminability is varied. Functions not displaying a pronounced gradient are considered to indicate parallel processing, while steep increasing functions are taken as evidence of serial processing. During the current study, however, it was aimed to investigate these mechanisms by relating RT to target eccentricity, relative to the nature and size of observed conspicuity areas. In terms of the arguments presented earlier regarding attentional involvement in the conspicuity area as well as eccentricity effects, it is believed that slopes of RT vs. eccentricity functions can also be used to infer the nature of the process underlying target acquisition behaviour.

Results from the present study (Figures 5.5 and 5.6a-h) show that for all display conditions, RT vs. eccentricity functions for targets located within the conspicuity area have near-zero slopes, while for targets located outside the border of the conspicuity area, steep positive slopes are observed. Target locations within the conspicuity area thus all seem to be processed simultaneously, while peripheral locations need to be processed serially. The fact that the curves' inflection points occur at eccentricity position 3, indicates that the border of the conspicuity area represents the transition between parallel and serial components. These trends are in line with models that propose visual search and target acquisition to consist of a fast, parallel component, followed by a slower, serial processing step (Theeuwes, 1993; Wolfe, 1994; Müller-Plath & Pollmann, 2003). Furthermore, intersubject slope variances emphasise that search and visual target acquisition occur along a continuum (Duncan & Humphreys, 1989; Müller-Plath & Pollmann, 2003), rather than search for a given target requiring either strictly parallel or strictly serial processing, as proposed in early reports on visual target acquisition behaviour (Treisman & Gelade, 1980).

The serial nature of processing of location outside the conspicuity area indicates that the zone of focal attention shifts if the target is not found to be located within that area, consistent with the suggestion that attention is not moved to individual items during the serial component of visual target acquisition (Scialfa & Joffe, 1998), but rather to groups of items encompassed by a range of attention (Müller-Plath & Pollmann, 2003). It follows from previous discussions that the smaller this zone, owing to reduced salience, the more attentional shifts will be necessary, resulting in longer search times and more fixations. Previous studies have shown that when that target is not located within the zone of focal attention, the parallel processing stage transmits valuable information on possible target locations to the serial stage (Olds et al., 2000; Olds & Punambolam, 2002), thereby increasing search efficiency. These results thus underline the central role of attentional guidance during visual target acquisition.

5.4.5 What do the results mean in terms of neural organization of the visual system?

Results from the current empirical application indicate that visual attention fulfils a central role during successful visual target acquisition. In terms of the proposed model, it is not surprising; cortical areas which are most concerned with attentional deployment occur at a processing level between early and late processing stages. Furthermore, they not only receive input from lower- as well as higher-order processing stages, but also send output to these areas. The functional importance of such a fronto-parietal network has already been proposed (Nobre et al., 2003). As shown by the tight link between mechanisms responsible for attentional shifts and oculomotor responses, output to higher-order areas directly influences computations pertaining to behavioural responses towards visual information (Bullier et al., 1996; Corbetta, 1998; Mitchell & Zipser, 2003). Output to lower-order areas, on the other hand, contributes to dynamic fine-tuning of computations during early processing stages.

Such an interpretation is in line with a recent study regarding information flow during visual processing by Hayakawa and colleagues (Hayakawa et al., 2003). They found that after activation in the early visual cortex, activity can be recorded in the parietal and temporal cortices, followed by activation of early cortical areas once more. Early activation is believed to contribute to quick feature analysis, while later activation constitutes target segmentation from the background. These information flow patterns, which are possibly facilitated by feedforward as well as feedback connections between cortical areas involved in early and intermediate processing stages, are in line with proposals that attentional guidance is necessary for successful visual target acquisition (Wolfe, 1994; Wolfe et al., 2002).

The importance of visual attention and cortical areas responsible for its deployment, during visual target acquisition, is further illustrated by findings that cortical activation patterns are similar for both efficient and inefficient search (Hayakawa et al., 2003). This provides neurophysiological evidence for proposals that the two search types are not qualitatively different (Müller-Plath & Pollmann, 2003), but rather represent the extremes along a continuum (Duncan & Humphreys, 1989; Müller-Plath & Pollmann, 2003), brought about by the same processing mechanism.

5.5 SUMMARY

Chapter 5 investigated the use of psychophysical measures such as conspicuity area and response time to establish the relationship between a target's conspicuity and resulting target acquisition behaviour. It was found that the size of the conspicuity area is proportional to the SNR, which in turn determines target detectability. Results from RT experiments confirmed that the conspicuity area constitutes the area within which a target can be detected with a single fixation, with associated efficient search behaviour. Search for targets located outside the conspicuity area, however, is effortful and even more so for conditions of low target detectability. The importance of taking eccentricity position into account when interpreting RT results obtained during visual target acquisition studies was confirmed.

Results were discussed in terms of the framework of visual information processing and associated neural organization developed earlier. It was shown that processing during a visual target acquisition task involves early parallel processing, followed by later serial processing. Visual attention was shown to be an important contributor to the progression of information processing, a finding which was supported by the neural organization of the visual system. It is therefore concluded that psychophysical measures, such as conspicuity area and response time, can be used as valid indicators of visual target acquisition behaviour.

CHAPTER 6

CONCLUSION

Visual target acquisiton depends on bottom-up as well as top-down influences and should thus be studied according to an integrated approach. Comprehensive neurophysiological knowledge regarding the functional organization of the visual systems exists. Similarly, extensive research has been conducted into the nature of the mechanisms underlying visual target acquisition behaviour. However, knowledge from the two disciplines has not been integrated to propose a systems-level explanation of visual target acquisition behaviour. The current study therefore aimed to show how behavioural data can be explained according to the structural and functional organization of the visual system.

6.1 DISCUSSION OF HYPOTHESES

The main hypothesis of the study stated that behavioural response during a visual target acquisition task can be objectively measured with psychophysical techniques and interpreted according to the structural and functional organization of the visual system. The rationale was that behavioural outcomes during a perceptual task are a function of the underlying processing mechanism, which in turn is a function of the underlying neural organization of the perceptual system. It follows that in order to derive a systems-level explanation of visual target acquisition behaviour it is not only necessary to have insight into neural and perceptual organization and associated behavioural outcome, but also to be able to relate these three components to one other.

As a first step the structural and functional organization of the visual system was investigated (Chapter 2). It was shown that the visual system consists of, apart from the eyes, several cortical and subcortical structures, each specialized to fulfil an important function during visual perception. Early functional differentiation results in the bulk of the information being processed along two functionally different processing streams, specialized to analyse spatial and featural information respectively. A conceptual model of the structural and functional organization of the visual system was subsequently compiled, showing that despite the structures within these streams being essentially hierarchically organized, they do not function independently – feedforward and feedback connections ensure integration of processing activity, resulting in meaningful behavioural response during each visual target acquisition task.

Chapter 3 discussed information processing during a visual target acquisition task from a theoretical perspective. It was concluded that visual target acquisition proceeds according to a two-stage process consisting of an early parallel processing stage, followed by a later serial-like stage, unlike early views that regarded the process as being of either parallel or serial nature, depending on stimulus characteristics. The influences of bottom-up as well as top-down factors on information processing were discussed and it was shown that both contribute to visual target acquisition behaviour, with the strength of the respective contributions depending on task demands. The step-wise progression of information processing was related to the structural and functional organization of the visual system as proposed in Chapter 2. Taken together, Chapters 2 and 3 provided a framework in terms of which visual target acquisition behaviour as measured in later empirical applications could be explained.

During empirical applications psychophysical techniques were optimized to determine conspicuity areas and response time towards visual targets presented in different display conditions. Conspicuity areas were shown to be elliptical to circular, with manipulations of SNR being reflected in the size of the area. Weaker SNR subsequently resulted in smaller conspicuity areas, whereas stronger SNR caused an increase in conspicuity area size. Such manipulation resulted in similar changes to RT. In order to determine how target detectability as determined by conspciuity area determination relates to RT, RT was measured at specific eccentricity positions relative to the conspicuity area. It was found that RT is not affected much by eccentricity when targets are located within the conspicuity area. However, beyond the border of the conspicuity area, target eccentricity greatly influences RT. It is also at these peripheral positions that the influence of SNR manipulations are most pronounced. It can therefore be concluded that conspicuity area and RT can be applied successfully as psychophyscial measures of target detectability. Results were subsequently interpreted in terms of the framework proposed in earlier chapters, which showed that objective measurement of behavioural responses can be used to infer the underlying processing mechanism during a visual target acquisiton task.

6.2 CONTRIBUTIONS AND IMPLICATIONS

The main contribution of the study has been to show that psychophysical measures of target detectability can be used to explain visual target acquisition behaviour according to a systems-level approach. Results were in line with current views that regard the mechanism underlying visual target acquisition as always consisting of a fast parallel processing stage, followed by a later, serial processing stage. Progression of information processing corresponds to the structural and functional organization of the visual system. It is supported by the finding that visual attention, which is important for the transition between the two processing stages, is the function of specialized cortical areas, situated at an intermediate anatomical level, where they are not only subject to input from lower-order areas as well as feedback from higher order areas, but also send output forward to higher-order areas and back to lower-order areas.

Other contributions relate to elaboration of previous work regarding the measurement of the conspicuity area, the relationship between conspicuity area and RT and the use of psychophysical measures to infer the underlying processing mechanism during visual target acquisition.

(i) It was shown that the shape of the conspicuity area corresponds to an ellipse or circle. Earlier reports of irregular conspicuity area are thought to be the result of presenting stimuli at random positions across the display. By using an adaptive psychophysical method, the border of the conspicuity area could be determined accurately and it was shown that it is sufficiently regular to justify presentation of stimuli only on selected meridians, rather than performing full-field mapping. In terms of industrial application, the finding implies that conducting conspicuity area experiments can be an easy and relatively fast method for determining target detectabilty by human observers.

(ii) The inverse relationship between size of the conspicuity area and RT was confirmed, but it was shown that the relationship is most pronounced when targets are located outside of the conspicuity area. In terms of industrial application the finding indicates that for situations where fast detection of targets should be prevented, visual environments should be designed such as to minimize the size of the conspicuity area, while simultaneously positioning targets peripherally with regard to the conspicuity area.

(iii) It was shown that RT data can be used to infer the underlying mechanism of information processing duirng a visual target acquistion task, provided that the effect of target position is taken into account. When interpreted in relation to conspicuity area data, the division between the parallel and serial components of visual target acquisition becomes evident. The finding can be useful for the design of automated target detection systems.

6.3 FUTURE WORK

The study provided a systems-level approach according to which the underlying mechanism of visual target acquisition behaviour can be explained. At processing level, the conspicuity area has been argued to represent the normal range of focal attention elicited by a target and visual attention has subsequently been shown to fulfil an important role during the progression of visual information processing. In light of the close relationship between visual attention and eye movements at both a functional and structural level (Scialfa & Joffe, 1998; Gitelman et al., 2002), eye movement measurements will be valuable for validating and expanding the proposed model.

It can further be recommended that for purposes of industrial application, studies regarding the influence of bimodal stimulus presentation on visual target acquistion behaviour will be beneficial. To this end the proposed model will provide a baseline from which to work, while findings from such studies will help to improve and expand the model.

A further recommendation from the current study relates to the use of imaging techniques such as functional magnetic resonance imaging. Although several studies have recently started to use imaging techniques during studies of visual target acqusition behaviour (e.g.

Singh et al., 2000; Nobre et al., 2003; Pollman et al., 2003), they focussed either on only specific aspects or only certain anatomical regions involved in the processing mechanism. An approach as the one proposed in this study will allow for results from imaging studies to be interpreted according to a systems-level approach.

REFERENCE LIST

Adams, D. H. & Zeki, S. 2001. Functional organization of macaque V3 for stereoscopic depth. *Journal of Neurophysiology*. 86: 2195-2203.

Andersen, R. A., Bracewell, R. M., Barash, S., Gnadt, J. W., Fogassi, L. 1990. Eye position effects on visual, memory, and saccade-related activity in areas lip and 7a of a macaque. *Journal of Neuroscience*. 10: 1176-1196.

Bachall, D. O. & Kowler, E. 1999. Attentional interference at small spatial separations. *Vision Research*. 39: 71-86.

Barkin, J., Nakayama, K. & Gilbert, C. 2000. Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. *Journal of Neuroscience*. 20: 8188-8198.

Bartels, A. & Zeki, S. 2000. The architecture of the colour centre in the human visual brain: new results and a review. *European Journal of Neuroscience*. 12: 172-193.

Bellamy, L. J. & Courtney, A. J. 1981. Development of a search task for the measurement of peripheral visual acuity. *Ergonomics*. 24: 497-509.

Bradley, D. 2001. Early visual cortex: Smarter than you think. *Current Biology*. 11: R95-R98.

Brown, B. & Monk, T. H. 1975. The effect of local target surround and whole background constraint on visual search times. *Human Factors*. 17: 81-88.

Bullier, J. 2001. Integrated model of visual processing. *Brain Research Review*. 36: 96-107.

Bullier, J. & Nowak, L. 1995. Parallel versus serial processing: New vistas on the distributed organization of the visual system. *Current Opinion in Neurobiology*. 5: 497-503.

Bullier, J., Schall, J. D. & Morel, A. 1996. Functional streams in occipito-frontal connections in the monkey. *Behavioural Brain Research*. 76: 89-97.

Burke, W., Dreher, B. & Wang 1998. Selective block of conduction in Y optic nerve fibres: Significance for the concept of parallel processing. *European Journal of Neuroscience*. 10: 8-19.

Caputo, G. & Guerra, S. 1998. Attentional selection by distractor suppression. *Vision Research.* 38: 669-689.

Carrasco, M. & Chang, I. 1995. The interaction of objective and subjective organizations in a localization search task. *Perception & Psychophysics.* 57: 1134-1150.

Carrasco, M. & Frieder, K. S. 1997. Cortical magnification neutralizes the eccentricity effect in visual search. *Vision Research.* 37: 63-82.

Carrasco, M., Evert, D. L., Chang, I. & Katz, S. M. 1995. The eccentricity effect: Target eccentricity affects performance on conjunction searches. *Perception & Psychophysics.* 57: 1241-1261.

Carrasco, M., McLean, T. L., Katz, S. M. & Frieder, K. S. 1998. Feature asymmetries in visual search: Effects of display duration, target eccentricity, orientation and spatial frequency. *Vision Research.* 38: 347-374.

Casco, C. & Campana, G. 1999. Spatial interactions in simple and combined-feature visual search. *Spatial Vision.* 12: 467-483.

Cassagrande, V. 1994. A third visual pathway to primate area V1. *Trends in Neurosciences.* 17: 305-310.

Cave, K. R. & Wolfe, J. M. 1990. Modeling the role of parallel processing in visual search. *Cognitive Psychology.* 22: 225-271.

Chaikin, J. D., Corbin, H. H. & Volkmann, J. 1962. Mapping a field of short term visual search. *Science.* 138: 1327-1328.

Chan, A. H. S. & Courtney, A. J. 1996. Foveal acuity, peripheral acuity and search performance: A review. *International Journal of Industrial Ergonomics.* 18: 113-119.

Cheal, M. L. & Lyon, D. R. 1992. Attention in visual search: multiple search classes. *Perception & Psychophysics.* 52: 113-138.

Chelazzi, L. 1995. Neural mechanisms for stimulus selection in cortical areas of the macaque subserving object vision. *Behavioural Brain Research.* 71: 125-134.

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

Chelazzi, L., Miller, E. K., Duncan, J. & Desimone, R. 2001. Responses of neurons in macaque area V4 during memory-guided visual search. *Cerebral Cortex.* 11: 761-772.

Constantinidis, C. & Steinmetz, M. A. 2001. Neuronal responses in area 7a to multiple-stimulus displays: I. neurons encode the location of the salient stimulus. *Cerebral Cortex.* 11: 581-591.

Corbetta, M. 1998. Frontoparietal cortical networks for directing attention and the eye to visual locations: Identical, independent, or overlapping neural systems. *Proceedings of the National Academy of Science of the United States of America.* 95: 831-838.

Courtney, A. J. & Chan, H. S. 1985a. Eight-axis visual-field mapping for targets embedded in a homogenous competing background. *Perceptual and Motor Skills.* 60: 655-663.

Courtney, A. J. & Chan, H. S. 1985b. Mapping the binocular visual field for a target embedded in a regular background. *Perceptual and Motor Skills.* 61: 1067-1073.

Courtney, A. J. 1989. Continuation of visual-field mapping for a target embedded in a regular background. *Perceptual and Motor Skills.* 69: 843-849.

Cutzu, F. & Tsotsos, J. K. 2003. The selective tuning model of attention: Psychophysical evidence for a suppressive annulus around an attended item. *Vision Research.* 43: 205-219.

Davis, E.T., Shikano, T., Peterson, S. A. & Michel, R. K. 2003. **See** Thorpe Davis, E., Shikano, T., Peterson, S. A. & Michel, R. K. 2003.

Desimone, R. & Duncan, J. 1995. Neural mechanisms of selective visual attention. *Annual Review of Neuroscience.* 18: 193-222.

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

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

Deubel, H. & Schneider, W. X. 1996. Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*. 36: 1827-1837.

DeYoe, E. & Van Essen, D. 1988. Concurrent processing streams in monkey visual cortex. *Trends in Neurosciences*. 11: 219-226.

Duhamel, J. R., Colby, C. L. & Goldberg, M. E. 1992. The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*. 255: 90-92.

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

Duncan, J., Ward, R. & Shapiro, K. 1994. Direct measurement of attention dwell time in human vision. *Nature*. 369: 313-314.

Engel, F. L. 1971. Visual conspicuity, directed attention and retinal locus. *Vision Research*. 11: 563-576.

Engel, F. L. 1977. Visual conspicuity, visual search and fixation tendencies of the eye. *Vision Research*. 17: 95-108.

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

Gallant, J. L., Connor, C. E., Rakshit, S., Lewis, J. W. & Van Essen, D. C. 1996. Neural responses to polar, hyberbolic, and Cartesian gratings in area V4 of the macaque monkey. *Journal of Neurophysiology*. 76: 2718-2739.

Gegenfurtner, K. R. & Hawken, M. J. 1996. Interaction of motion and color in the visual pathways. *Trends in Neurosciences*. 19: 394-401.

GenStat for Windows. 2000. Release 4.2. Fifth Edition. Oxford: VSN International.

Gilchrist, I. D. & Harvey, M. 2000. Refixation frequency and memory mechanisms in visual search. *Current Biology*. 10: 1209-1212.

Gitelman, D. R., Parrish, T. B., Friston, K. J. & Mesulam, M.-M. 2002. Functional anatomy of visual search: regional segregations within the frontal eye fields and effective connectivity of the superior colliculus. *NeuroImage*. 15: 970-982.

Glickstein, M. 1988. The discovery of the visual cortex. *Scientific American*. 259: 84-91.

Goodale, M. A; Meenan, J. P., Bülthoff, H. H., Nicolle, D. A., Murphy, K. J. & Racicot, C. I. 1994. Separate neural pathways for the visual analysis of object shape in perception and prehension. *Current Biology*. 4: 604-610.

Green, D. & Swets, J. 1966. *Signal detection theory and psychophysics*. Huntington, New York: Robert E. Krieger Publishing Company.

Grobler, S. 2000. *Re: circle/eliptical curve fit?* Google Groups: comp.soft-sys.matlab. Available from: http://groups.google.com/groups [Accessed Feb. 2004]

Hadjikani, N., Liu, A. K., Dale, A. M., Cavanagh, P. & Tootell, R. B. 1998. Retinotopy and color sensitivity in human visual cortical area V8. *Nature Neuroscience*. 1: 235-241.

Hayakawa, T., Miyauchi, S., Fujimaki, N., Kato, M. & Yagi, A. 2003. Information flow related to visual search assessed using magnetoencephalography. *Cognitive Brain Research*. 15: 285-295.

Hegde, J. & Van Essen, D. C. 2000. Selectivity for complex shapes in primate visual area V2. *Journal of Neuroscience*. 20: RC61

Hendry, S. & Reid, R. 2000. The koniocellular pathway in primate vision. *Annual Review of Neuroscience*. 23: 127-153.

Hoffman, J. E. 1978. Search through a sequentially presented visual display. *Perception & Psychophysics*. 23: 1-11.

Horowitz, T. S. & Wolfe, J. M. 1998. Visual search has no memory. *Nature*. 394: 575-577.

Hubel, D. H. & Wiesel, T. N. 1959. Receptive fields of single neurons in the cat's striate cortex. *Journal of Physiology*. 148: 574-591.

Hubel, D. H. & Wiesel, T. N. 1977. Functional architecture of macaque monkey visual cortex (Ferrier Lecture). *Proceedings of the Royal Society of London B*. 198: 1-59.

Hubel, D. 1982. Exploration of the primary visual cortex, 1955-78. *Nature*. 299: 515-524.

Hubel, D. 1988. *Eye, brain, and vision*. New York: Scientific American Library.

Hupe, J., James, A., Payne, B., Lomber, S., Girard, P. & Bullier, J. 1998. Cortical Feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature*. 394: 784-787.

Jenkins, S. E. & Cole, B. L. 1982. The effect of density of background elements on the conspicuity of objects. *Vision Research*. 24: 1241-1252.

Jones, M.B. & Regehr, J. The problems you're having may not be the problems you think you're having: Results from a latency study for Windows NT. Available from: http://www.cs.utah.edu/~regehr/papers/hotos7/hotos7.pdf [Accessed Feb. 2004].

Johnson, A. & Macauley, M. 2001. High-precision timing within Microsoft Windows: threads, scheduling and system interrupts. *Microprocessors and Microsystems*. 25: 297-307.

Johnson, A., Horsemann, B., Macauley, M. & Barnes, W. 2002. PC-based visual stimuli for behavioural and electrophysiological studies of optic flow field detection. *Journal of Neuroscience Methods*. 114: 51-61.

Kastner, S., De Weerd, P. & Ungerleider, L. G. 2000. Texture segregation in the human visual cortex: a functional MRI study. *Journal of Neurophysiology*. 83: 2453-2457.

Kee, D., Jung, E. S. & Chung, M. K. 1992. Isoresponse time regions for the evaluation of visual search performance in ergonomic interface models. *Ergonomics*. 35: 243-252.

Klein, R. M. 2000. Inhibition of return. *Trends in Cognitive Sciences*. 4: 138-147.

Kobatake, E. & Tanaka, K. 1994. Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology*. 71: 856-867.

Kooi, F. L., Toet, A., Tripathy, S. P. & Levi, D. M. 1994. The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision.* 8: 255-279.

Kowler, E., Anderson, E., Dosher, B. & Blaser, E. 1995. The role of attention in the programming of saccades. *Vision Research.* 35: 1897-1916.

Kristjansson, A. 2000. In search of remembrance: Evidence for memory in visual search. *Psychological Science.* 11: 328-332.

Kuffler, S.W. 1952. Neurons in the retina: Organization, inhibition and excitatory problems. *Cold Spring Harbour Symposia on Quantitative Biology* 17: 281-292.

Lamme, V. & Roelfsema, P. 2000. The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences.* 23: 571-579.

Leek, M. 2001. Adaptive procedures in psychophysical research. *Perception & Psychophysics.* 63: 1279-1292.

Levitt, H. 1971. Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America.* 49: 467-477.

Livingstone, M. & Hubel, D. 1984. Anatomy and physiology of a color system in the primate visual cortex. *Journal of Neuroscience.* 4: 309-356.

Luce, R. 1986. *Response times: Their role in inferring elementary mental organization.* New York: Oxford University Press.

Marvit, P., Florentine, M. & Buus, S. 2003. A comparison of psychophysical procedures for level-discrimination thresholds. *Journal of the Acoustical Society of America.* 113: 3348-3361.

McAdams, C. J. & Maunsell, J. H. 2000. Attention to both space and feature modulates neuronal responses in macaque area V4. *Journal of Neurophysiology.* 83: 1751-1755.

Meyer, L. L., Hanekom, J. J. & Pretorius, J. H. Developing a time sensitive application in Windows 2000. *(In preparation).*

Merigan, W. & Maunsell, J. 1993. How parallel are the primate visual pathways? *Annual Review of Neuroscience*. 16: 369-402.

Microsoft Corporation. Resource Kits Windows 2000. Available from: http://www.microsoft.com/windows2000/techinfo/reskit/ [Accessed Feb 2002]

Miller, E. K. & Desimone, R. 1994. Parallel neuronal mechanisms for short-term memory. *Science*. 263: 520-522.

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

Miller, E. K., Lin, L. & Desimone, R. 1993. Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *Journal of Neuroscience*. 13: 1460-1478.

Mitchell, J. F. & Zipser, D. 2003. Sequential memory-guided saccades and target selection: a neural model of the frontal eye fields. *Vision Research*. 43: 2669-2695.

Monk, T. H. & Brown, B. 1975. The effect of target surround density on visual search performance. *Human Factors*. 17: 356-360.

Moran, J. & Desimone, R. 1985. Selective attention gates visual processing in the extrastriate cortex. *Science*. 229: 782-784.

Motter, B. C. 1994. Neural correlates of feature selective memory and pop-out in extrastriate area V4. *Journal of Neuroscience*. 14: 2190-2199.

Motter, B. C. 1998. The zone of focal attention during active visual search. *Vision Research*. 38: 1007-1022.

Müller, H. J., Heller, D. & Ziegler, J. 1995. Visual search for singleton feature targets within and across feature dimensions. *Perception & Psychophysics*. 57: 1-17.

Müller-Plath, G. & Pollmann, S. 2003. Determining subprocesses of visual feature search with reaction time models. *Psychological Research*. 67: 80-105.

Nakayama, K. & Silverman, G. H. 1986. Serial and parallel processing of visual feature conjunctions. *Nature*. 320: 264-265.

NATO. 2000.Workshop of the RTO Systems Concepts and Integration (SCI) Panel. Organized by NATO Research and Technology Agency, The Netherlands, 21-23 June 1999. Available from http://www.rta.nato.int/Rdp.asp?RDP=RTO-MP-045 (Executive summary).[Accessed May 2004]

Neisser, U. 1967. *Cognitive Psychology.* New York: Appleton-Century-Crofts.

Nobre, A., Coull, J. T., Walsh, V. & Frith, C. D. 2003. Brain activations during visual search: Contributions of search efficiency versus feature binding. *NeuroImage.* 18: 91-103.

Nothdurft, H. C. 1992. Feature analysis and the role of similarity in pre-attentive vision. *Perception & Psychophysics*. 52: 355-375.

Nothdurft, H. C. 1993a. Saliency effects across dimensions in visual search. *Vision Research.* 33: 839-844.

Nothdurft, H. C. 1993b. The role of features in preattentive vision: Comparison of orientation, motion and color cues. *Vision Research.* 33: 1937-1958.

Nothdurft, H. C. 1999. Focal attention in visual search. *Vision Research.* 39: 2310

Nothdurft, H. C. 2000. Salience from feature contrast: additivity across dimensions. *Vision Research.* 40: 1183-1201.

Olds, E. & Punambolam, R. J. 2002. The decay and interruption of interactions between search mechanisms. *Vision Research.* 42: 747-760.

Olds, E., Cowan, W. & Jolicoeur, P. 1999. Spatial organization of distractors in visual search. *Canadian Journal of Experimental Psychology.* 53: 150-159.

Olds, E., Cowan, W. & Jolicoeur, P. 2000. The time-course of pop-out search. *Vision Research.* 40: 891-912.

Op de Beeck, H., Wagemans, J. & Vogels, R. 2001. Can neuroimaging really tell us what the human brain is doing? The relevance of indirect measures of population activity. *Acta Psychologica.* 107: 323-351.

Pan, K. & Eriksen, C. W. 1993. Attentional distribution in the visual field during same-different judgments as assessed by response competition. *Perception & Psychophysics*. 53: 134-144.

Pasupathy, A. & Connor, C. E. 1999. Responses to contour features in macaque area V4. *Journal of Neurophysiology*. 82: 2490-2502.

Pasupathy, A. & Connor, C. E. 2002. Population coding of shape in area V4. *Nature Neuroscience*. 5: 1332-1338.

Paul, M. A. & Tipper, S. P. 2003. Object-based representations facilitate memory for inhibitory processes. *Experimental Brain Research*. 148: 283-289.

Payne, B. R., Lomber, S. G., Villa, A, E. & Bullier, J. 1996. Reversible deactivation of cerebral network components. *Trends in Neurosciences*. 12: 535-542.

Peterson, M. S., Kramer, A. F., Wang, R. F., Irwin, D. & McCarley, J. S. 2001. Visual search has memory. *Psychological Science*. 12: 287-292.

Pollman, S., Weidner, R., Humphreys, G. W., Olivers, C. N., Muller, K. Lohmann, G. Wiggins, C. J. & Watson, D. G. 2003. Separating distractor rejection and target detection in posterior parietal cortex – an event-related fMRI study of visual marking. *NeuroImage*. 18: 310- 323.

Rolls, E. & Deco, G. 2002. *Computational neuroscience of vision*. Oxford: Oxford University Press.

Schall, J. 2001. Neural basis of deciding, choosing and acting. *Nature Reviews Neuroscience*. 2: 33-42.

Schiller, P. 1986. The central visual system. *Vision Research*. 26: 1351-1386.

Schiller, P. 1996. On the specificity of neurons and visual areas. *Behavioural Brain Research*. 76: 21-35.

Scialfa, C. T. & Joffe, K. M. 1998. Response times and eye movements in feature and conjunction search as a function of target eccentricity. *Perception & Psychophysics.* 60: 1067-1082.

Singh, K. D., Smith, A. T. & Greenlee, M. 2000. Spatiotemporal frequency and direction sensitivities of human visual areas measured using fMRI. *NeuroImage.* 12: 550-564.

Smith, P. L. 1995. Psychophysically principled models of visual simple reaction time. *Psychological Review.* 102: 567-593.

Snyder, J. J. & Kingstone, A. 2000. Inhibition of return and visual search: How many separate loci are inhibited? *Perception & Psychophysics.* 62: 452-458.

Steinman, B. A., Steinmam, S. B. & Lehmkuhle, S. 1995. Visual attention mechanisms show a center-surround organization. *Vision Research.* 35: 1859-1869.

Steinmetz, M. A. & Constantinidis, C. 1995. Neurophysiological evidence for a role of posterior parietal cortex in redirecting visual attention. *Cerebral Cortex.* 5: 448-456.

Sterling, P. 1999. Deciphering the retina's wiring diagram. *Nature Neuroscience.* 2: 851-853.

Swets, J. A. 1964. *Signal detection and recognition by human observers.* New York: John Wiley & Sons.

Talbot, S. A. & Marshall, W. H. 1941. Physiological studies on neural mechanisms of visual localization and discrimination. *American Journal of Ophthalmology.* 24: 1255-1264.

Tanaka, K. 1996. Representation of visual features of objects in the inferotemporal cortex. *Neural Networks.* 9: 1459-1475.

Theeuwes, J. 1993. Visual selective attention: A theoretical analysis. *Acta Psychologica.* 83: 93-154.

Thorpe, S., Fize, D. & Marlot, C. 1996. Speed of processing in the human visual system. *Nature.* 381: 520-552.

Thorpe Davis, E., Shikano, T., Peterson, S. A. & Michel, R. K. 2003. Divided attention and visual search for simple versus complex features. *Vision Research.* 43: 2213-2232.

Timmerman, M. & Van Beneden, B. 2000. Executive summary of the evaluation report on Windows NT 4.0 Workstation. *Dedicated Systems Magazine. Special Issue:*

Timmerman, M. 1997. Windows as a real-time OS? *Real-Time Magazine.* 97: 6-13.

Toet, A., Kooi, F. L., Bijl, P. & Valeton, J. M. 1998. Visual conspicuity determined human target acquisition performance. *Optical Engineering*. 37: 1969-1975.

Treisman, A. 1985. Preattentive processing in vision. *Computer Vision, Graphics, and Image Processing*. 31: 156-177.

Treisman, A. & Gelade, G. 1980. A feature-integration theory of attention. *Cognitive Psychology*. 12: 97-136.

Treisman, A. & Gormican, S. 1988. Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*. 95: 15-48.

Treisman, A. & Souther, J. 1985. Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*. 114: 285-310.

Tripathy, S. P. & Cavanagh, P. 2002. The extent of crowding in peripheral vision does not scale with target size. *Vision Research*. 42: 2357-2369.

Tweed, D. 1997. Three-dimensional model of the human eye-head saccadic system. *Journal of Neurophysiology*. 77: 654-666.

Ungerleider, L., Courtney, S. & Haxby, J. V. 1998. A neural system for human visual working memory. *Proceedings of the National Academy of Science of the United States of America*. 95: 883-890.

Usrey, M. & Reis, R. 2000. Visual physiology of the lateral geniculate nucleus in two species of New World monkey: *Siamiri scireus* and *Aotus trivirgatis*. *Journal of Physiology*. 523: 755-769.

Wandell, B. 1999. Computational neuroimaging of the human visual cortex. *Annual Review of Neuroscience*. 22: 145-173.

Wilkinson, F., James, T. W., Wilson, H. R., Gati, J. S., Menon, R. S. & Goodale, M. A. 2000. An fMRI study of selective activation of human extrastriate form vision areas by radial and concentric gratings. *Current Biology.* 10: 1455-1458.

Wolfe, J. M. 1994. Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review.* 1: 202-238.

Wolfe, J. M. 1998. Visual Search. In Pashler, H. (Eds), *Attention.* (pp. 13-56). London: University College London Press.

Wolfe, J. M. 2003. Moving towards solutions to some enduring controversies in visual search. *Trends in Cognitive Sciences.* 17: 70-76.

Wolfe, J. M., O'Neill, P. & Bennett, S. C. 1998. Why are there eccentricity effects in visual search? Visual and attentional hypotheses. *Perception & Psychophysics.* 60: 140-156.

Wolfe, J. M., Oliva, A., Horowitz, T. S., Butcher, S. J. & Bompas, A. 2002. Segmentation of objects from backgrounds in visual search tasks. *Vision Research.* 42: 2985-3004.

Wolfe, J. M, Butcher, S. J., Lee, C. & Hyle, M. 2003. Changing your mind: On the contributions of top-down and bottom-up guidance in visual search for feature singletons. *Journal of Experimental Psychology: Human Perception and Performance.* 29: 483-502.

Zeki, S. 1978. Functional specialization in the visual cortex of the rhesus monkey. *Nature.* 274: 423-428.

Zeki, S. 1993. *A vision of the brain.* Chicago: Blackwell Scientific Publications.

Zeki, S. & Shipp, S. 1988. The functional logic of cortical connections. *Nature.* 335: 311-317.

Zeki, S., McKeefry, D. J., Bartels, A. & Frackowiak, R. S. 1998. Has a new color area been discovered? *Nature Neuroscience.* 1: 335-336

Zipser, D. & Andersen, R. 1988. A backprojection programmed network that stimulates response properties of a subset of posterior parietal neurons. *Nature.* 331: 679-684.

APPENDIX A

ACCURATE AND RELIABLE TIMING IN WINDOWS 2000

Timing in the multi-tasking Windows 2000 environment

Despite multi-tasking operating systems not originally being designed with low-latency response to events and predictable time-based scheduling in mind (Johnson & Macauley, 2001), they have become increasingly convenient to use for time-dependent studies relating to aspects of visual perception (Tweed, 1997; Johnson et al., 2002). Similarly, the application program for the experiments described here also runs under the Microsoft Windows operating system and since the execution of the code is time-dependent, careful consideration had to be given to achieve accurate and reliable timing.

A real-time operating system (RTOS) can be defined as one that responds in a timely predictable manner to unpredictable external input (Timmerman, 1997). In short, this means that (i) there is no tolerance for failing to elicit a response to an event within a set time limit and (ii) that simultaneous processing should be possible without compromising time-restricted action to events (Timmerman, 1997). With these considerations in mind, Windows 2000 cannot be considered a true RTOS, but rather a general purpose operating system with the ability to service events relatively fast and within fairly consistent time constraints (Microsoft Corporation, 2000).

When attempting to develop an application that is able to meet predefined deadlines within the Windows environment, it is important to keep in mind the way the operating system handles processes that run simultaneously. Each process consists of at least one thread, which is executed by the processor. Associated with threads are execution priorities. These can be divided into three classes, ranging from highest to lowest execution importance. According to the priority associated with a particular thread, the operating system's scheduler determines when it should be executed, as well as how much processing time can be allocated to it. This is a weighted scheduling system, with high priority threads being executed earlier and allowed more processor time than those of lower priority. However, processing of messages derived from external events (interrupts) is not controlled by the scheduler and thus usually has higher priority than threads. This means that thread execution can be suspended in favour of an interrupt, although the system attempts to minimize their impact on thread processing by shunting them to Deferred Procedure Call routines (Microsoft Corporation, 2000; Johnson & Macauley, 2001).

Experimental procedure and application design

In order to evaluate the accuracy and reliability of Microsoft Windows 2000 to perform time-dependent tasks, experiments were run under Windows 2000, Service Pack 2, on a PC using a Celeron 1000 MHz processor with 256 MB RAM. The PC was also fitted with a SiS 630/730 graphics card and the monitor (LG Flatron 775FT) was set at a resolution of 1024 x 768 pixels at 75 Hz. An application program, to create the test display as well as measure RT to an external event, was developed in Borland's Delphi 6. In brief, the test program was required to generate visual output, which would terminate either owing to a mouse event, or in the absence of such an event, after a predefined amount of time had elapsed. Elapsed time is then read by a high-resolution counter, as explained below. A single thread that waits for a command to end the generation of the display, controls the execution of visual output instructions. In the absence of a mouse event this command is generated by expiry of the predefined amount of time controlled by a timer that runs on a separate thread. In the event of mouse input terminating visual output, the high resolution counter value is read by calling the QueryPerformanceCounter() routine. The difference between this value and the value taken at the start of visual output generation depicts the elapsed time from onset to termination of the visual display.

In order to verify time measurements, an independent external timer device was used. Once the image is displayed, a photodiode is activated, while termination of visual output causes deactivation of the photosensor. The sensor's amplified signal is fed to a 4 MHz crystal clocked microcontroller with 1 μs instruction cycles, after which two timers on the microprocessor can measure the amount of time a light source is at a specific level and establish a reference light source. Results are written to the PC via an EIA232 port after measurements have been made. Figure A1 shows a schematic representation of the method.

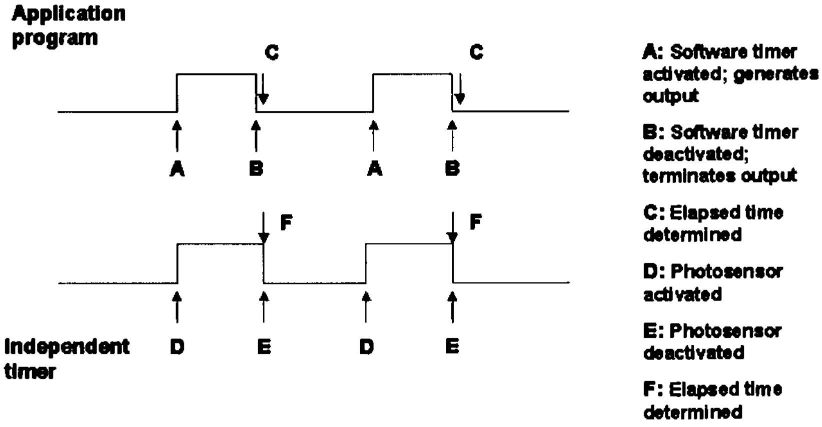


Figure A1: Schematic representation of experimental procedure during evaluation of timing accuracy and reliability.

Results and discussion

A comparison between the time measurements made by the test program and the independent timer device is shown in Figure A2. These results depict the average time (over approximately 120 display cycles) after which visual output was terminated in the absence of mouse input. Screen display time was supposed to be 75, 150, 250, 500, 750 and 1 000 ms respectively. As determined by the high-resolution counter, the threaded timer persistently signalled visual output to terminate 8 ms earlier than supposed to. When termination of visual output was expected after 75 ms, a deviation of only 3 ms was recorded.

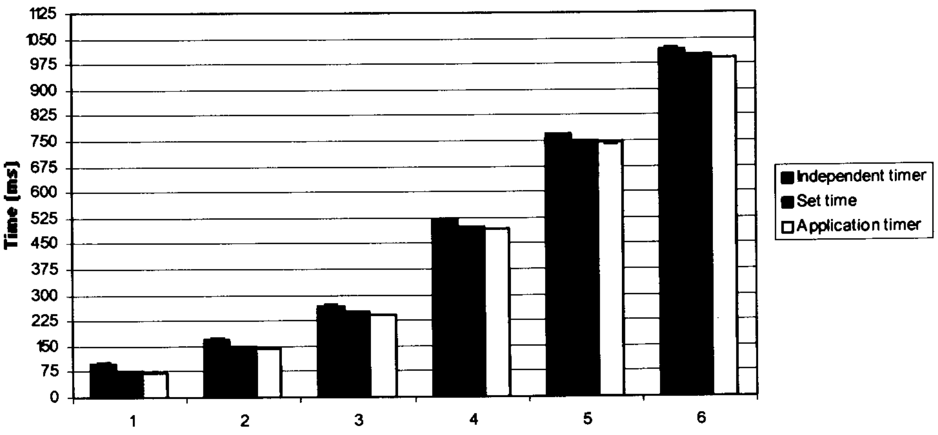


Figure A2: Comparison between screen display time as measured by an independent timer device and application program's timer.

These results should, however, be interpreted with the resolution of timers operated under Windows 2000 in mind. Under normal circumstances, the threaded timer's resolution is reported to be approximately 10 ms in Windows 2000, which means that it is only able to handle time in approximately 10 ms intervals. In order to determine the boundaries of these intervals and how in-between values are rounded, the QueryPerformanceCounter() API call was used. It revealed that these intervals do not proceed according to factor-wise steps, but rather seem to start two time units into the increment. For example, instead of increments being at 250, 260 or 270 ms, they were at 252, 262 and 272 ms. It was furthermore found that the division for rounding is not half-way, but that the first 8 ms into the increment is rounded to the lower boundary, while measurements exceeding this value are rounded tot the upper limit.

The fact that the timer consistently came 8 ms short of the goal, caused visual output always to be terminated 8 ms earlier than it was supposed to. Thus, because of the way the increments are positioned and the manner in which rounding is handled, the values were always rounded to the same limit. This argument also holds when time for ending the display is set at 75 ms and the deviation is only 3 ms below the expected time, since it is within 8 ms of the lower boundary.

As shown in Figure A3 the resolution of the test program's timer can be significantly improved by using the TimeBeginPeriod() and TimeEndPeriod() routines. If the TimeBeginPeriod(1) and TimeEndPeriod(1) routines are called prior to the start and at the end of application execution respectively, it was found that the timer's resolution improves to approximately 1 ms, with a standard deviation never exceeding 0.24 ms. These results are in line with those from previous studies (Jones & Regehr, 1999). As a result, the timer terminated visual output within 1 ms of the intended time. It is furthermore shown that the independent timer measured screen display time to be 25.90 ± 1.34 ms and 25.86 ± 1.21 ms longer under normal and improved resolution circumstances, respectively. This indicates that the threaded timer truly supports a 1 ms resolution, and that, although timers operated under Windows 2000 are said to have a resolution of only 10 ms, it is possible to adjust their resolution to be more accurate by designing the application differently.

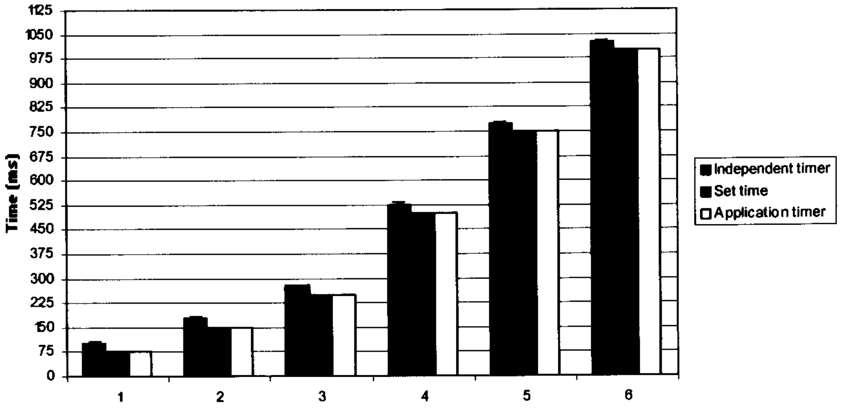


Figure A3: Results of using the application program's timer with improved resolution to time screen display.

One should wonder, though, what the cause of the difference between independently measured and application-dependently measured screen display time is. Since the independent timer employs a photosensor for activation, it is subject to the rate at which the light source (i.e. the screen) is deactivated. The monitor used during experiments operated at a refresh rate of 75 Hz. This means that 13 ms was required to switch the screen from bright to dark, and that the independent timer would thus report screen display time to be up to 13 ms longer.

However, even with the time needed to repaint the screen accounted for, a delay of about 13 ms is left. In order to see whether this delay can be attributed to code execution, the QueryPerformanceCounter() API call was again used. This revealed that the code executes within a few microseconds, and that it therefore cannot be the cause of the delay. It is assumed that operating system and hardware-associated factors could be the cause of the delay. The consistency of the delay, however, indicates that the program performs adequately when referenced with a hard real time process such as the independent timer.

When a mouse event signals termination of the visual output the difference in measurements made by the independent timer and the high performance counter is 27.16 ± 4.70 ms. This variation is very close to that encountered when screen output was controlled by the threaded application timer. This shows that the application is reliable in handling time-dependent tasks and that it is possible to measure RT to an external event, such as mouse input, accurately.

Conclusion

At the outset of designing an application that should be able to perform tasks within certain time constraints, it is important to define the accuracy with which these deadlines should be met. If meeting deadlines is crucial, Windows cannot be used to operate such applications. Windows was in fact never designed to function as an RTOS and despite good overall architecture will never be able to do so (Timmerman, 1997; Timmerman & Van Beneden, 2000). Despite these shortcomings, the easy-to-use API calls that allow the developer freedom in writing code that meets user-defined requirements, make the Windows operating system very appealing for time-sensitive studies in the biomedical sciences.

This study has shown that a time-dependent application can be operated under a general purpose operating system such as Windows. Timers that run with accurate time intervals can successfully be created in Windows 2000 and their resolution can be improved to 1 ms when code is structured correctly. It has thus been demonstrated that visual output can be generated for a predefined time. Furthermore, by carefully considering how the operating system assigns scheduling priority and deals with interrupts, it is possible that response to an interrupt, such as a mouse event, can be accurately timed. On the whole, Windows offers a convenient platform to create applications for research studies of this nature.