1989

Intermediate-level Visual Processing And The Perception Of Inside/outside Spatial Relations

Richard David Wright

Recommended Citation

Wright, Richard David, "Intermediate-level Visual Processing And The Perception Of Inside/outside Spatial Relations" (1989). *Digitized Theses*. 1808.
https://ir.lib.uwo.ca/digitizedtheses/1808

National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service

Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

Intermediate-level Visual Processing and the Perception of Inside/Outside Spatial Relations

by
Richard David Wright

Department of Psychology

Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Graduate Studies
The University of Western Ontario
London, Ontario
March, 1989

ISBN 0-315-49336-4

ABSTRACT

The visual perception of inside/outside spatial relations often appears to be immediate and effortless. However, it is becoming increasingly apparent that the processes underlying this perception are numerous and complex. According to Ullman, these processes can be characterized as sequences of basic operations called visual routines. Each routine consists of different combinations of operations which allow the visual system to establish object shapes and to extract spatial relations between objects. The research described in this thesis involved manipulations of display parameters for tasks in which observers made judgements about whether a target was inside or outside of a bounding figure. It was found that variations of stimulus size, location, and relative temporal onsets of stimuli reliably and systematically affected inside/outside response times. Therefore, the nature of the basic operations involved in this perception can be investigated by examining the factors that affect it. In particular, the results suggest that operations like region colouring, contour tracing, and shifts of processing focus may have been invoked to determine the relation. This is consistent with Ullman's claim that the perception of spatial relations is mediated by a number of subprocesses operating together as a visual routine.

ACKNOWLEDGEMENTS

Many people have helped me in many different ways, large and small, throughout this project. I am grateful to each of you and would particularly like to thank: my supervisor, Zenon Pylyshyn, for his generous support and unerring advice, and for allowing me to work autonomously while always providing me with a clear sense of direction; Keith Humphrey, for his willingness to discuss any aspect of perception and cognition at virtually any time, for invaluable guidance and advice, and for providing the students at Western with an infectious supply of enthusiasm for vision research; Mike Dawson, for his encouragement and support throughout all stages of this thesis; Albert Katz, for sharing his keen sense of empirical design, and for the levity sometimes needed to keep a project like this in perspective; my brother, Ted Wright, for his helpful comments on an earlier draft of this thesis, for his advice on getting a postgraduate education productively while keeping a positive outlook, and for his support during some emotionally-trying times; Pierre Jolicoeur, Mel Goodale, Rick Gurnsey, and Ted Elcock for their suggestions about the final version of the manuscript; Jim Clark, Steve Lupker, and Regina Schuller for their help with many aspects of the data analyses; Jim Clark for help with programming the response key mechanisms; Inese Kramins and Aileen Orthner, two adminstrative officers who provided secretarial and administrative assistance in ways too numerous to mention; some special friends, Anne Krupka, Monica Mori, Chris Roney, and Jeff St. Pierre, who were always there when times were tough; and finally my parents, Bill and Claire Wright, for their unfailing support and love -- this thesis is dedicated to them.

TABLE OF CONTENTS

Page
CERTIFICATE OF EXAMINATION ii
ABSTRACT iii
ACKNOWLEDGEMENTS iv
TABLE OF CONTENTS v

CHAPTER 1: THEORETICAL PRELIMINARIES 1
  1.1 Overview of the Problem 1
  1.2 Low-level Vision 3
  1.3 Intermediate-level Vision 8
    1.31 Visual Routines 9
    1.32 Attentional Processes 17
  1.4 Developing Models of Intermediate-level Visual Processes 20
    1.41 Modeling the Perception of Inside/Outside Relations 21
  1.5 Outline of the Thesis 29

CHAPTER TWO: EFFECTS OF FIGURAL SIZE AND LOCATION ON THE PERCEPTION OF INSIDE/OUTSIDE RELATIONS 30
  2.1 Experiment 1 30
    2.11 Method 33
    2.12 Results 36
    2.13 Discussion 38
  2.2 Experiment 2 42
    2.21 Method 45
    2.22 Results 47
    2.23 Discussion 51
  2.3 Experiment 3 55
    2.31 Method 55
    2.32 Results 56
    2.33 Discussion 61
  2.4 General Discussion 63

CHAPTER THREE: PERCEIVING INSIDE/OUTSIDE RELATIONS BETWEEN ASYNCHRONOUSLY PRESENTED STIMULI 66
  3.1 Experiment 4 68
    3.11 Method 68
    3.12 Results 69
    3.13 Discussion 73
  3.2 Experiment 5 79
    3.21 Method 81
    3.22 Results 81
    3.23 Discussion 86
  3.3 Experiment 6 86

    - 3.31 Method 86
    - 3.32 Results 90
    - 3.33 Discussion 93
  - 3.4 Experiment 7 93
    - 3.41 Method 93
    - 3.42 Results 94
    - 3.43 Discussion 96
  - 3.5 General Discussion 102

CHAPTER FOUR: CONCLUDING REMARKS 104

REFERENCES 108

CURRICULUM VITAE 115

CHAPTER ONE: THEORETICAL PRELIMINARIES

1.1 OVERVIEW OF THE PROBLEM

The determination of spatial relations among objects and features appears to be an important part of object perception. In particular, some spatial relations may play a role in the creation of visual descriptions of objects. For example, an object with parts representing eyes, a nose, and a mouth may be clearly perceived as a schematic picture of a face. Moreover, relational predicates such as "inside of" may play a role in creating visual descriptions of this object (e.g., the nose is "inside of" the contour representing the head). However, if the spatial arrangement of these parts is changed, it will probably be much more difficult to perceive the object as a face. Hence, the determination of spatial relations among facial parts plays an important role in the perception of this object. Also, the determination of spatial relations appears to be essential for visually-guided object manipulation. That is, in order to manipulate an object under visual guidance, relational properties such as "beside", "on top of", and "inside of" must be determined. Thus, knowing how spatial relations are determined appears to be an important part of understanding the processes involved in object perception and manipulation.

It is becoming increasingly apparent that the perception of spatial relations involves processes that are numerous and complex. Ullman (1984) has attempted to describe these processes by proposing the notion of *visual routines*. These routines are comprised of visual operations that are part of what could be considered a basic "instruction set." Different combinations of these basic operations are said to be assembled in order to establish the shapes of objects and their spatial relations. Arguments supporting Ullman's proposal will be discussed in more detail later in this chapter. On the basis of these arguments, it appears that in order to understand how spatial relations are determined, it is necessary to discover what types of operations are involved and how they are combined.

The purpose of this thesis was to investigate the nature of the operations involved in the perception of inside/outside spatial relations. One reason that inside/outside may play an important role in the creation of object descriptions is that it remains the same over many different viewpoints. Throughout the thesis, "inside of" is meant in the sense of a topological relational property of a two-dimensional scene in which an object is completely surrounded by the contour of another object. There are also other senses of insideness. For example, an object could be *inside* another object from a two-dimensional perspective but *above* from a three-dimensional perspective. However, this raises questions of a semantic nature about what it is that constitutes the property of insideness. These questions were not addressed because the primary concern was with the operations involved in determining insideness in a two-dimensional scene.

Very little empirical evidence is available concerning the nature of the basic operations underlying visual routines. However, researchers (Jolicoeur, 1988; Jolicoeur, Ullman, & MacKay, 1986; Pringle & Egeth, 1988) have had some success in this direction by manipulating the display parameters of spatial relations tasks and examining the resulting response time differences. A similar procedure was used in the experiments in this thesis. In each study, observers were asked to judge as quickly and as accurately as possible whether a target was inside or outside a bounding figure. In addition, display parameters such as figural size and target location were manipulated. It was expected that the nature of the perceptual operations involved would be evident if the manipulations of the parameters led to systematic response time differences.

It was necessary to take an *exploratory* approach to this research because little is known about the nature of the visual operations involved in performing these tasks. Therefore, the manner in which the experiments were conducted and reported deviated somewhat from that usually associated with a traditional "hypothesis testing" approach. For example, some of the experimental design decisions were made without the benefit of information provided by previous research of this type. In addition, decisions about which parameters to manipulate also were made in the spirit of exploratory research. It was expected that the results of these studies would form a database of information about factors affecting the perception of inside/outside relations. On the basis of this information, it also was expected that some inferences could be made about the operations involved in this perception.

The remainder of this chapter is a discussion of the processes thought to be involved in low- and intermediate-level vision. The discussion of the latter includes a detailed description of Ullman's (1984) visual routines hypothesis. This is followed by a brief discussion of the role that attentional processes may play in intermediate-level vision, and an outline of models of the perception of inside/outside relations.

1.2 LOW-LEVEL VISION

A number of researchers (e.g., Marr, 1982; Pylyshyn, 1988; Ullman, 1984) have proposed that visual perception is carried out in a series of stages. These stages are characterized to some degree by how early they are thought to occur along the visual pathway from the retinae to the visual cortex, and to what extent high-level, cognitive factors such as beliefs, expectations, and specific world knowledge affect the processes involved. The initial stage of processing, *low-level vision*, is thought to be driven by information in the visual scene, and to be carried out independently of cognitive factors. Marr (1982) has referred to this as pure, autonomous perception. In contrast, *high-level vision* is thought to be knowledge-dependent. According to Pylyshyn (1984), processing at this stage can be influenced by an observer's beliefs and expectations about what the proximal stimulus is. It has been proposed (Pylyshyn, 1988; Ullman, 1984) that there is an intermediate level of processing between low- and high-level vision. In some cases, operations at this level are primarily data-driven, whereas in others, they are influenced by top-down, cognitive processes. Thus, low-level vision can be thought of as the "hard-wired" operations that constitute the functional architecture of the visual system. Intermediate-level operations serve as the interface between this functional architecture and the higher level, cognitively penetrable processes that constitute the knowledge-dependent part of visual processing (see Pylyshyn, 1984 for a more detailed discussion of this view of mental architecture). In addition, intermediate-level operations are thought to be involved in the determination of object shapes and their spatial relations. These operations appear to be extremely rapid like those of low-level vision, but are sometimes sequential and knowledge-dependent like those of high-level vision.

The function of low-level vision appears to be the transformation of physical intensity differences in the visual field produced by distal stimuli into descriptions of visible surfaces and their orientations. One approach to low-level vision proposed by Marr (1982) holds that the physical stimulation of the retinae is immediately converted into a symbolic representation of the visual field called the *raw primal sketch*. This is a data structure containing information about primitive visual elements such as edge segments, blobs, and line terminators. Moreover, Marr has proposed that a set of independent, modular processes operate on the primal sketch in order to compute "structure from X" where X is some property of the representation. For example, Ullman (1979) developed a modular algorithm that produces a description of image structure on the basis of the motion of primal sketch elements. Marr and Poggio (1976, 1979) proposed two stereopsis algorithms that take disparities of primal sketch elements from pairs of retinal images and fuse these elements to yield a description of surface depth and orientation. According to Marr, these modular, low-level processes produce outputs that are combined to form a relative depth map of visible surfaces called a *2.5-D sketch*. Marr has claimed that this is the output representation of low-level vision.

There is some disagreement about the nature of the processes and representations involved in low-level vision. Marr (1982) has argued that one of the initial steps carried out by the visual system is the conversion of differences in physical intensity signals across the retinae into a *symbolic* representation of the visual field. In contrast to this proposal, others (e.g., Hinton & Anderson, 1981; McClelland, Rumelhart, & the PDP

Research Group, 1986) have argued that symbolic labeling of features is not a property of low-level vision. Instead they claim that visual computations with unlabeled features are more similar to the computations performed by neurons at the early stages of visual processing, and are less complex than tasks involving symbolic labeling. Zucker (1987, p. 79) has also stated that many low-level visual processes and parameters are continuously distributed as opposed to discretely symbolic. However, there is a consensus about some aspects of this stage of processing. Most agree that operations are predominantly local and parallel, and that they are data-driven. While some researchers claim that a depth map is unnecessary for many visual processes (Lowe, 1985; Witkin & Tenenbaum, 1983), many others claim that the output representation of low-level vision is a surface orientation relative depth map (e.g., a representation like the 2.5-D sketch -- see Brady, 1981, p. 7).

The main problem to be solved by low-level vision is the underdetermination of the distal stimulus by the proximal stimulus. In particular, there is an infinite number of three-dimensional stimuli of different configurations that can give rise to the same two-dimensional pattern of retinal stimulation (the proximal stimulus). To select the correct spatial configuration of the distal stimulus the visual system must take additional information into account. "New Look" theorists (e.g., Bruner, 1957; Gregory, 1970) have dealt with this problem by assuming that perceptual processes are the same as cognitive processes. That is, these processes have access to both contextual information and higher-level, general knowledge of the world and can use these types of information to select the correct configuration of the distal stimulus. For example, a brief glimpse of a black object may be all that is required to determine that it is a telephone if the object is resting on a telephone table. Hence, contextual information facilitates perceptual processing. This approach has been challenged, however, on several grounds (e.g., Fodor, 1975; Fodor & Pylyshyn, 1980; Pylyshyn, 1984). One criticism is that early perceptual processes differ from cognitive processes because the former are belief-independent and therefore not cognitively penetrable. For example, information about an imminent threat must always be delivered by the perceptual system despite beliefs that the occurrence of such a threat is extremely unlikely or impossible. Hence, the "New Look" claim that higher-level processes are involved in constraining the problem of underdetermined retinal images is questionable.

The *natural computation* approach to low-level vision (Marr, 1982) is based on the assumption that problems of visual underdetermination can be constrained by general properties of the physical world that almost always hold true in *our* environment. Many researchers (e.g., Dawson & Pylyshyn, 1986, 1988; Hildreth, 1983; Marr & Poggio, 1976, 1979; Ullman, 1979) have developed algorithms based on these *natural constraints* that generate the surface structures of distal stimuli without recourse to higher-level influences. For example, the Marr and Poggio (1976, 1979) stereo algorithms are only constrained by the assumptions that the physical world consists primarily of smooth surfaces as opposed to edges (the continuity constraint), and that a feature can only exist at one location in space at one time (the uniqueness constraint). Ullman's (1979) structure-from-motion algorithm is constrained only by the assumption that many objects in our world are rigid. Natural constraints also provide an account of how unfamiliar objects that are *not* part of an observer's general knowledge of the world can still be perceived. Hence, this approach posits low-level visual processes that operate independently of higher-level influences and yet can solve the problem of underdetermined retinal images.

Marr (1982) has claimed that in order to understand fully low-level vision and visual processing in general, one must be prepared to contemplate explanations at three different levels of description that are linked, at least in principle, into a cohesive whole. The first level of description is the *computational theory* of visual processing. It accounts for what is to be computed and why. The others are the *algorithmic level* which is concerned with how computations are carried out, and the *level of physical instantiation* (the neural implementation). As an example, Marr has described the information processing carried out at these three levels by a supermarket cash register. The computational theory of the cash register is the theory of addition (e.g., principles of associativity and commutativeness). This level is unrelated to the algorithm used to implement it (e.g., base 10 arithmetic, binary arithmetic, etc.). The computational and algorithmic levels are also unrelated to the level of physical instantiation of the addition (e.g., silicon chips, mechanical gears, etc.). Thus, in order to understand fully how the cash register works, it must be understood at each of the three levels of explanation. By analogy, the same holds true for the visual system. Moreover, Marr has claimed that we are more likely to understand how the visual system works by first understanding the nature of the problem being solved than by understanding the hardware of this system. If this is the case, then the primary goal of low-level vision research should be to isolate constraints that are powerful enough to allow a process to be defined.

According to Ullman (1984), the perception of spatial relations among objects should also be studied at different levels of explanation. More specifically, at the level of computational theory, the goal should be to determine which spatial relations are most useful for the creation of visual descriptions of objects. At the algorithmic level, the primary concern should be with the nature of the basic operations involved, and with the nature of the routines (algorithms) that are assembled from them. The latter was the approach I have adopted for studying the perception of inside/outside relations. As mentioned previously, there is some reason to believe that the relational predicate "inside of" plays an important role in object perception. Thus, the eventual goal of this type of research at the computational level is to determine the importance of perceiving "inside of" to object perception in general. The focus of the current research was at the algorithmic level -- i.e., to determine the nature of the operations involved in this perception.

1.3 INTERMEDIATE-LEVEL VISION

Intermediate-level vision was described previously as a stage of processing at which spatial relations are determined among objects. This section is a description of the operations thought by some researchers (e.g., Pylyshyn, 1988; Ullman, 1984) to be carried out at this stage. In addition, the nature of object representations involved and shifts of attentional focus at this stage will be discussed.

Perhaps more than any other aspect of low-level vision, information about the nature of its output is of critical importance to those carrying out research on vision at the intermediate-level. This output representation can be thought of as a "starting point" for intermediate-level processing. Many researchers agree that this representation is a description of an image's visible surfaces and the orientations of these surfaces in terms of a viewer-centred coordinate frame. However, it appears that information about spatial relations among objects is not yet explicitly encoded during low-level processing. In particular, Marr (1982, p. 350) has stated that the angle between two lines is not explicit in the primal sketch, and the angle between two surfaces is not explicit in the 2.5-D sketch. This suggests that in order to infer an object's nonvisible surfaces and perceive them in terms of an object-centred coordinate frame, higher-level, general knowledge of the world must be brought into play. Moreover, there is empirical evidence that information about spatial relations between objects is not yet explicitly encoded during preattentive processing. That is, Treisman and Gormican (1988, p. 36) have found that perceiving the spatial relation "inside of" is not a preattentive process like those of low-level vision (see also Treisman, 1986, p. 119). Thus, the determination of spatial relations does not appear to be initiated until the intermediate level of visual processing.

There are also other reasons for proposing that spatial analysis is carried out, at least in part, at the intermediate level. The speed at which simple relations such as "longer than" are perceived is often extremely fast (e.g., 400 ms), and therefore is like the speed of local, parallel processing of low-level vision. However, some spatial analysis can begin only with the availability of higher level information about which relation is to be determined. This information about the computational goal is top-down, as opposed to the information driving low-level vision which is thought to be bottom-up. Moreover, it appears that in some cases determining spatial relations involves operations that are essentially sequential (see Ullman, 1984, p. 121) and therefore beyond the capacity of the local, parallel operations of low-level vision. Hence, the determination of spatial relations shares properties of both low- and high-level vision and therefore appears to be initiated at an intermediate stage between these two levels.

1.31 Visual Routines

Ullman (1984) has attempted to describe the organization of operations that may be carried out at the intermediate level by proposing a set of processes called *visual routines*. These routines are composed of certain combinations of basic operations that allow the system to establish object shapes, and to extract spatial relations among objects. Basic operations are selected as components of visual routines in accordance with the visual task to be carried out. Jolicoeur (1988) has compared basic operations to computer machine language operations, and visual routines to short, compiled programs comprised of these operations. A hypothetical example of a simple visual routine is the following procedure for determining whether a figure is open or closed: The routine begins with the marking of some point on the figure's contour with a place token. The contour is then traced starting at this location and a decision about whether the figure is closed or not is contingent upon the contour being traced back to the marked point. Hence, it is a visual routine composed of two basic operations -- contour marking and tracing.

Ullman (1984) has proposed a number of reasons why spatial analysis appears to be mediated by different combinations of a small set of basic operations. First, hierarchical feature detectors have been applied with some success to object perception but do not seem suitable for determining spatial relations. In particular, when a relation such as "inside of" is determined, it is often the case that the entire bounding contour must be analyzed, and in a sequential rather than a parallel manner. Also, Minsky and Papert (1969) demonstrated that simple operations such as determining the connectedness of a contour requires a serial rather than a parallel processing algorithm. Hence, a series of operations may be involved rather than a one-step operation carried out by something like an "inside of" feature detector. In addition, the number of different shapes that share the property of *closure* is unbounded. Therefore, the number of different shapes that can be involved in perceiving "inside of" is unbounded. Moreover, the size of the support set required for determining a single spatial relation such as "inside of" is also large. In conjunction with the fact that the number of *different* spatial relations that can be determined is prohibitively large and perhaps unbounded, this effectively rules out a template matching scheme. Rather than positing the storage in memory of a separate mechanism for performing each type of analysis, Ullman has argued that applying visual routines assembled from a small set of basic operations is a more parsimonious alternative. In this way, only a small number of operations are required and, by altering how they are combined, they can be "tailored" to carry out the visual task at hand.

Basic Operations. Few claims have been made about the nature of basic operations. In the previous example of a visual routine for determining whether a figure is open or closed, contour marking was carried out by allocating a place token to a particular location within the visual representation. This type of operation is necessary for keeping track of locations that have already been activated by operations such as contour tracing. Pylyshyn (1988) has proposed that there is a set of resource-limited place tokens called FINSTs (for fingers of instantiation) that serve as markers and indexes for visual routines. According to Pylyshyn, four or five features within a representation can be marked preattentively and in parallel by FINST tokens. Moreover, he has maintained that this marking or indexing of feature clusters (or places) in the visual field must occur before spatial analysis of an image can begin. In other words, before it can be determined that "this" is inside of "that", there must be a way to, in effect, "point to" the two features to which the inside relation will apply. In the case of inside/outside problems, the internal/external point and the bounding contour would be indexed. FINSTs do not themselves make an encoding of the location of features available to higher level processes such as visual routines that compute spatial relations. They simply serve as markers and index tokens. By doing so, they make it possible to obtain information about a feature at a particular location by referring to the FINST at that location. Pylyshyn has suggested that one can characterize the idea of FINSTs in a concrete way by imagining that each of one's fingers is placed on a different object (or feature token) in a scene. Hence, one may not know what each finger is touching or its exact location, but finger contact gives one a way to, in effect, directly *refer* to the objects so that some further processing of them can be undertaken.

According to Pylyshyn (1988), a direct, mechanical indexing procedure analogous to finger contact makes it possible for a FINST to keep pointing to a feature cluster as it moves across the retina. This claim has been empirically supported by Pylyshyn and Storm (1986). Subjects were shown a field of identical, randomly arranged stimuli and were required to keep track of a subset of them as they moved independently and in random directions. After a period of time, one of them changed shape. Subjects were asked to state whether or not that stimulus belonged to the subset of stimuli being visually tracked. It was found that subjects could perform this task when the size of the subset was as large as four and sometimes five stimuli. Moreover, the velocity of each stimulus was fast enough (and the proximity of the non-targets small enough) to rule out a sequential location encoding procedure in which attention moved to each of four tracked stimuli in turn to update their location codes. Instead, it appeared that subjects were able to simultaneously keep track of four or five distinct, moving features without encoding their location relative to a global frame of reference. This supports the claim that there is a primitive referencing mechanism for pointing to certain kinds of features, thereby maintaining their distinctive identity without either recognizing them (in the sense of categorizing them), or explicitly encoding their features.

Ullman (1984) has claimed that one step in the process of indexing a location within a visual representation is the shift of processing focus to that location by data-driven processes. An example of indexing is the "pop out" effect that occurs when observers visually search an array of objects for a particular target object (see e.g., Neisser, 1963; Treisman & Gelade, 1980; Treisman & Schmidt, 1982). If one of the features of the target object differs from the features of the rest of the objects in the array (e.g., the target is red and the distractor objects are green), the time required to find the target will be very rapid and will not be affected greatly by increases in the number of distractors in the array. In other words, the target pops out from the distractors as if the array is searched in parallel at all locations, and attention is immediately drawn to the location of the object with the unique feature. The location of the pop-out object is assumed to be indexed (Pylyshyn, 1988). However, if the target does *not* possess a unique feature (e.g., the target is a red vertical line in a field of green vertical lines and red horizontal lines), the time required to find it is considerably slower and increases when there are more distractors. In this case it appears as if the array is searched serially until a target is found, presumably because no indexing of the location of a unique feature takes place. Koch and Ullman (1984) have attempted to model the indexing of a pop-out location with a "winner-take-all" algorithm. According to Ullman, indexed locations serve as anchor points for the initiation of visual routines.

Pylyshyn (1988) has claimed that FINSTs can be assigned in parallel to a small number of points and, as a result of this indexing, the cardinality of the points can be determined without visually visiting them. Only the number of active FINSTs needs to be counted. Thus, indexing may underlie the phenomenon of rapid counting referred to as subitizing. Trick and Pylyshyn (1988) have demonstrated empirically that subitizing does not occur if the targets to be counted are embedded in a field of distractors that these targets do not pop out of. However, subitizing does occur if the targets to be counted do pop out of the distractor field. These results indicate that the location of the pop-out objects are indexed, and that these index tokens are limited in number (the subitizing range is usually from one to four objects).

Previous research (Ingleton, 1987; Jolicoeur, 1988; Jolicoeur, Ullman, & MacKay, 1986; Pringle & Egeth, 1988) has also provided empirical evidence that contour tracing is a basic operation invoked by visual routines. Jolicoeur et al. presented subjects with displays consisting of two nonintersecting curves and two Xs lying on either the same curve or on different curves. The direct retinal distance between the Xs was always the same (1.8°) but the distance along the curve between them varied. It was found that the time required to determine the property "two Xs on the *same* curve" was a positive monotonic function of the distance along the curve between them. Because the exposure duration of each display was short enough to preclude eye movements (250 ms), it appears that a rapid contour tracing operation was carried out independently of eye movements.

Other possible basic operations of visual routines suggested by Ullman (1984) are *shifting the location of processing focus*, and *bounded activation* of a region. For most spatial relations tasks, processing focus must be shifted from one location to another within the representation being operated on. For example, in a task such as "is the dot above or below the horizontal line?", presumably either the dot or the line will be focused on first and then the focus of processing will be shifted to the other stimulus. When tasks such as subitizing or counting features are carried out, presumably the focus of processing is shifted to each of the indexed locations of the objects in question as they are counted. Pylyshyn (1988) also claims that the location of processing focus can be shifted by following an index assigned to a feature cluster moving across the retina. Hence, shifting the location of processing focus appears to be of critical importance to visual routines.

Bounded activation, like contour tracing, is an activation operation in which successive locations within a visual representation are "activated" and remain in this state as this operation is carried out. When this occurs, each of the immediately neighbouring locations of a particular seed point are activated. Then, each of the neighbouring locations of these points are activated and so on. This spread of activation is analogous to paint flowing outwards in all directions from the seed point. Activation is bounded in the sense that it will stop spreading when a bounding contour is reached. This operation has been described in the computer vision literature as a useful way of determining the internal regions of closed figures and a related process has been used to compute figure-ground relationships (Kienker, Sejnowski, Hinton, & Schumacher, 1986; Sejnowski & Hinton, 1987). Ullman (1984) has also suggested that it may be one of the operations involved in the determination of inside/outside spatial relations. This will be discussed in more detail in the next section.

Universal Routines and Incremental Representations. One of the basic assumptions of the visual routines hypothesis is that routines are initially applied to the output representation of low-level vision (the base representation). The selection of routines is knowledge-dependent because the nature of the visual routine invoked depends in part on the computational goal of the perceiver. For example, if, in one case, the goal was to determine whether two Xs were the same size and in another case it was to determine whether these Xs were inside the same circle, two different routines would be invoked. However, other spatial relations appear to be determined, to some extent, independently of higher level influences such as beliefs and expectations. For this reason, there is a need for an initial set of data-driven operations called *universal routines*. Universal routines are mandatory and operate independently of higher level influences. In addition, they provide some initial, crude, shape and relational information about objects in the visual field. According to Ullman, these universal routines gather sufficiently useful information to allow a preliminary classification of objects. This crude scene analysis and the computational goal of the perceiver (e.g., to determine whether two Xs are inside the same circle) can then be used to select more specialized visual routines to carry out the relevant task.

Ullman (1984) has proposed that when universal routines and more specialized visual routines operate on a representation, the results of this processing are maintained in an updated version of that representation. These *incremental representations* can then serve as base representations for succeeding visual routines. In other words, intermediate-level visual processing may involve a series of representations of increasing elaborateness that contain information from preceding representations plus the results of visual routines that were invoked to produce them. Low-level vision also is thought to involve a series of representations of increasing elaborateness (e.g., Marr's raw primal sketch, full primal sketch, and 2.5-D sketch). However, low-level processing is data-driven and the nature of the output representation is, in a sense, always the same (e.g., a surface orientation map of the distal stimulus). In contrast, the output representation of intermediate-level processing varies to a much greater degree because the computational goals that lead to the selection of visual routines vary in accordance with the visual property that the perceiver attempts to determine at any given moment. Thus, incremental representations produced by intermediate-level vision can be thought of as a diverse series of data structures of increasing elaborateness containing the results of many possible specialized processes.

The results of a study by Rock and Gutman (1981) are consistent with the notion that incremental representations are produced by visual routines. These researchers presented subjects with a series of interleaved, red and green figures. Subjects were told to attend to figures of just one colour on each trial (e.g., the red one) and to make a "goodness" rating of it. Rock and Gutman found that on a subsequent test, subjects were *only* able to accurately recognize the figures of the colour attended to despite being exposed to figures of the other colour in each display (i.e., the nonattended figures). This finding has been replicated by Butler and McKelvie (1985). In addition, Goldstein and Fink (1981) and Tipper (1985) have found similar results using two overlapping line drawings of different colours (as opposed to the nonsense shapes used by Rock &

Gutman). Ullman's (1984) interpretation of Rock and Gutman's findings is that a series of incremental representations was produced that contained information about the shape of the figures that were attended to. The only information these representations contained about the shape of the unattended figures was that resulting from the crude scene analysis carried out by universal routines. Hence, the information about the shapes of the unattended figures was not adequate for shape recognition.

In a more recent study (Wright & Katz, 1989), subjects performed a task very similar to that performed in the Rock and Gutman experiments. However, in each display the unattended figure had three gaps in its contour that were 0.6° long. On a subsequent, forced-choice recognition test, subjects responded at the chance level in identifying the shape of the background figure. This finding is consistent with Rock and Gutman's (1981) results. In addition, subjects responded significantly below the chance level in a second test condition by consistently choosing a new figure without gaps rather than the original, gapped, unattended figure. In other words, subjects were unaware of the presence of gaps in the unattended figures. This suggests that the incremental representations of the overlapping figure displays did not include information about gaps occurring in the unattended figures, despite the finding (Treisman & Souther, 1985) that gaps are preattentively indexed in some situations. Thus, it appears that the representations stored in long-term visual memory contain only the information about objects in the display that have been indexed and referred to by higher level, visual routines.

Summary. The visual routines hypothesis is a compelling description of the organization of intermediate-level visual operations and the steps involved in the spatial analysis of a scene by visual routines can be summarized as follows:

1. Low-level vision produces a base representation such as the 2.5-D sketch.
2. Mandatory, data-driven, universal routines are automatically applied to this representation to extract crude shape and relational information.
3. The results of the universal routines are maintained in an incremental representation of the visual field.

4. The information yielded by universal routines and the perceiver's computational goal are used in conjunction to select more specific visual routines.

5. These more specific routines are assembled and begin to operate on the incremental representation.

6. A series of incremental representations of increasing elaborateness are produced by subsequent visual routines.

1.32 Attentional Processes

As mentioned previously, one of the basic operations of visual routines is thought to be shifting the processing focus. These shifts are similar in some ways to shifts of attentional focus (also referred to as movement of the attentional spotlight). This section is an outline of research on shifts of attentional focus and a comparison of these shifts with shifts of processing focus and contour tracing during visual routines.

One important aspect of intermediate-level vision is the distinction between preattentive and attentive processing. Treisman (1986, 1988) has proposed a *feature integration theory* to account for how these two types of processes interact. According to this theory, features of objects initially are processed preattentively and in parallel throughout the visual field. Information about these features is then stored in separate maps for object colour, size, orientation, etc. While preattentive processing of features appears to be extensive, Treisman has claimed that the integration of features of a single object requires focused attention. That is, an object will not be perceived until attention is focused at its location and the information at that location in each of the feature maps is integrated. Treisman and her colleagues (e.g., Treisman & Gelade, 1980; Treisman & Gormican, 1988; Treisman & Schmidt, 1982) have collected a substantial amount of empirical evidence to support the theory. These findings indicate that feature analysis takes place at a very early stage of visual processing such as low-level vision, whereas feature integration occurs at a later stage of processing that involves shifts of attentional focus.

The movement of attentional focus to different locations in the visual field is thought by some researchers to be analogous to the movement of a spotlight. The results of experiments by Posner et al. (e.g., Posner & Cohen, 1984; Posner, Davidson, & Snyder, 1980) suggest that the spotlight can be moved to different locations by spatial precuing. Other researchers (e.g., Sperling & Reeves, 1980; Tsal, 1983) have found that the spotlight appears to move independently of eye movements at speeds ranging from 30 to 117 degrees/second. Also, the minimum size of the spotlight appears to be between 0.5 and 1° (Eriksen & Eriksen, 1974; Eriksen & Hoffman, 1972; Humphreys, 1981). According to feature integration theory, features are conjoined and an object is perceived at the location of this attentional spotlight.

Several different ideas have been proposed about the properties of attentional spotlights. For example, Jonides (1983) has described them in terms of a two-state model of attentional distribution. In one state, attention is distributed uniformly throughout the visual field and processing is carried out in parallel. In a second state, however, attention can be focused at a particular location (spotlight mode) and the concentration of attentional resources leads to faster visual processing within this spotlight. Eriksen and Yeh (1985) have modified Jonides' model so that it resembles an attentional zoom lens. Rather than proposing two discrete attentional states, they have argued that focused and distributed modes are poles on a continuum of attentional distribution in the visual field. In particular, lower resolution settings of the zoom lens should lead to widely distributed attention, parallel processing, and slow extraction of information. In contrast, higher resolution settings should lead to focused attention at a particular location and faster extraction of information. According to Eriksen and Yeh, wide range of settings between high and low resolution are also possible. Thus, it has been proposed that the concentration of attentional resources within the spotlight can vary.

Other questions that have been raised about attentional spotlights concern their shape and number. Treisman (1988, p.226) has stated that the belief that attention is unitary derives from the findings of Posner et al. (1980) while conducting tests of the detection of a single light flashing in an otherwise empty visual field. Other research has indicated that more than one spotlight may exist at any given time. For example, Egly and Homa (1984) presented subjects with a series of displays in which two letters were shown for 30 to 60 ms followed by a pattern mask. One letter was always located at the centre of three concentric rings with radii of 1, 2, and 3°. The second letter was always at one of eight positions on one of the rings. They found that the centre letter was identified with 95% accuracy and that valid spatial precuing facilitated the naming latency of the second letter. Moreover, when an invalid spatial precue of the middle ring occurred, the cost of this miscuing was just as great when the second letter was on the inner ring as when it was on the outer ring. This suggests that precuing the middle ring does not cause the attentional spotlight to include both the middle and inner rings. Instead, attention appears to be divided between the centre of the rings and the middle ring as though more than one spotlight was involved. Furthermore, it suggests that the spotlight on the middle ring may have a "donut" shape. This finding is consistent with Pylyshyn's (1988) claim that several locations in the visual field can be indexed simultaneously.

Ward (1982) has proposed a more refined version of the spotlight model called an *attentional frame*. In particular, he has claimed that attention can be thought of as a frame with parameters such as locus and degree of focus that could require adjustments with each new fixation. However, the parameters of the frame stay in their current state until adjustments are required because this conserves processing resources. Using bilevel letter stimuli and a global/local precedence task developed by Navon (1977), Ward has demonstrated that identification of the global stimulus is faster than that of the local stimulus if a global identification was made on the previous trial. However, if a local identification was made on the previous trial, identification of the local stimulus is faster than that of the global stimulus. By hypothesis, making the same type of identification on consecutive trials is faster than making different types of identifications because the attentional frame does not require adjustment across trials. Ward's model is able to account for many of the findings of the global/local precedence literature by taking the degree of attentional focus into consideration.

The notion that attentional spotlights can have different degrees of focal resolution is consistent with a recent finding concerning the resolution of processing focus. In particular, Jolicoeur (1988) found that the proximity of other curves to the curve in question affected the time required to determine whether two Xs lie on the same or on different curves. That is, as the proximity of other curves increased, response times were slower. As mentioned previously, a rapid contour tracing operation appears to be carried out independently of eye movements when this task is performed. One interpretation of Jolicoeur's finding is that increases in the proximity of other curves leads to increases in the resolution of processing focus. If this is the case, then the degree of resolution of processing focus affects the speed of operations of visual routines as well as shifts of attention. Thus, investigations of the effects of focal resolution on attentional spotlight movement may provide some insight into the effects of processing focus resolution on the operations of visual routines.

1.4 DEVELOPING MODELS OF INTERMEDIATE-LEVEL VISUAL PROCESSES

Cognitive modeling is carried out to determine the extent to which a computer simulation performs a task in the same way as a human. According to Pylyshyn (1980, 1984), one of the primary objectives of cognitive science should be to develop computational models of cognitive and brain processes that are *strongly equivalent* to the human processes being studied. In particular, models should be developed that go from the same input to the same output state as the organism in question, and to achieve this mapping from one state to the other with the same processes that the human uses. Thus, a model and an organism carry out the same process in a strongly equivalent manner if they use the same representations and the same sequence of computational operations (the algorithm).

If a computer program is to be viewed as a strongly equivalent model of visual cognition, it must correspond to the process being performed at a sufficiently fine level of resolution. Pylyshyn has argued that this kind of equivalence between a model and an organism only occurs if their processes are indistinguishable, among other things, in terms of computational resource use (e.g., time and memory) as input properties are varied. For this reason the psychophysical measurement of reaction times is an important procedure for discriminating between different computational models of the same organism. In particular, the time required by a strongly equivalent model to go from an input to an output state must correlate highly with the response times of the observers in the task at hand. Thus, information about psychophysical reaction times can serve as a valuable constraint for the development of computational models.

1.41 Modeling the Perception of Inside/Outside Relations

Over the past 20 years, several attempts have been made to model the perception of inside/outside spatial relations. Modeling complements the psychophysical study of human performance of inside/outside tasks. That is, continued refinement of a model often leads to the discovery of operations that are useful for determining this relation. These operations, in turn, yield predictions that can be tested empirically. The results of research motivated by the model's predictions might also provide constraining information for further refinements of the model. Two operations that have been proposed as subprocesses of the perception of inside/outside relations are *ray intersection* and *region colouring*. This section contains descriptions of previous attempts to develop inside/outside models based on these operations.

Ray Intersection Algorithms. One of the first algorithms to be developed that explicitly determines the spatial relation between a bounding figure and a target is the ray intersection algorithm (see e.g., Evans, 1968; Ullman, 1984; Winston, 1977, ch. 2). In brief, this algorithm begins by extending a ray outwards from the target towards "spatial infinity", where spatial infinity is a region that is clearly outside the bounding figure.

After the ray has been extended, the number of intersections of the ray with the contour of the bounding figure are counted. If there are an odd number of intersections then the target is inside. If there are an even number of intersections then the target is outside (see Figure 1a). A tangential meeting counts as two intersections. In the case of displays containing only closed curves, ray intersection algorithms usually produce the correct solution to inside/outside problems.

When compared to human performance of inside/outside tasks, the limitations of this type of algorithm are clear. For example, the presence of other contours in the visual field can disrupt the ray intersection count (see Figure 1b). In this situation the X is inside the bounding figure but the presence of the other contour increases the intersection count to two, thereby causing the algorithm to erroneously come to an "outside" solution. Moreover, this type of algorithm cannot be used to determine whether two Xs are inside the same figure (see Figure 1c) or whether an X is inside more than one figure (see Figure 1d). Thus, ray intersection algorithms can be used to determine inside/outside relations between two figures in simple displays but they are not adequate models in general for inside/outside tasks.

Region Colouring Models. Region colouring is an image segmentation technique also referred to as region merging or region growing. When this operation is carried out, the image is divided into arbitrary elementary regions which are then merged on the basis of a rule (e.g., merging similar regions) until a stopping criterion is reached (see Ballard & Brown, 1982, ch. 5; Brice & Fenema, 1970; and Haralick & Shapiro, 1985). Ullman (1984) has described a method for determining inside/outside relations that is based on region colouring (referred to in the previous section as bounded activation). The algorithm begins with the division of the image into a set of elementary regions. A spread of activation then flows outward from the internal/external point in question as neighbouring regions are merged. This activation is called region colouring because its spread is like a flow of paint and, as mentioned previously, the stopping criterion for

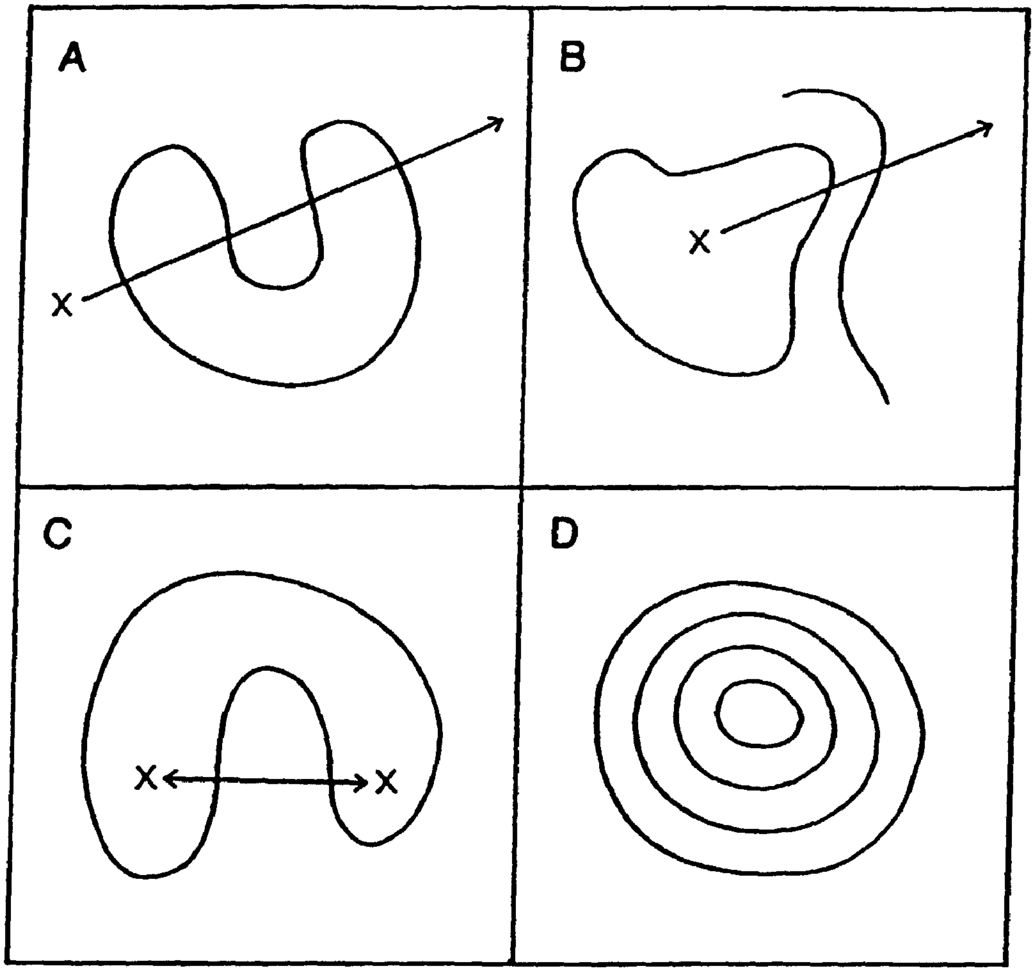


Figure 1. (A) Ray intersection algorithm correctly determines that the X is outside the figure because there is an even number of intersections. (B) Algorithm incorrectly determines that the X is outside because an even number of intersections has occurred due to the presence of an additional contour in the visual field. (C) Algorithm cannot determine whether two Xs are inside the same figure. (D) Algorithm cannot determine whether a target is inside more than one figure. These examples have been adapted from Ullman (1984).

colouring is the presence of a bounding contour. Thus, if the spread of activation is checked by these contours, the point is inside. On the other hand, if the spread of activation flows to "spatial infinity", the point is outside (see Figure 2). Romanycia (1987) has developed a computational model based on region colouring from a seed point that produces correct solutions to a number of different inside/outside problems. This type of model is superior to ray intersection models in two ways. First, its performance is not disrupted by the presence of other contours in the visual field, and it can also be used to determine whether two points are inside the same figure.

Region colouring also has limitations as a model of human performance of inside/outside tasks. For example, an X may be clearly perceived as inside the circle despite the presence of gaps in the circle's contour. If the colouring metaphor is taken literally, then the flow of activation should "leak" through the gaps to spatial infinity and cause the model to produce an outside solution. Despite the problems that arise if the colouring metaphor is taken literally, some researchers have used it as a working model of human perception of inside/outside relations. For example, Ullman (1984, p. 105) reports research in which the responses times of subjects performing inside/outside tasks are affected only weakly by changes in the size of the bounding figure. On the basis of this finding, Ullman and his colleagues (see Mahoney & Ullman, 1988) developed a modified region colouring procedure called *multiscale analysis*. This technique involves simultaneous activation of neighbouring regions of different sizes. As a result, the time required to activate a region is virtually independent of the region's size. Shafrir (cited in Mahoney & Ullman, 1988, pp. 204-207) has investigated the time performance of multiscale colouring algorithms that activate different shapes of neighbour regions. He found that elongated and compact shapes could be activated with equal speed regardless of figural size if the neighbour regions were rectangular. This was not the case if the neighbour regions were square. Thus, a rectangular multiscale colouring model is consistent with Ullman's preliminary finding that inside/outside responses are not affected strongly by changes in figural size.

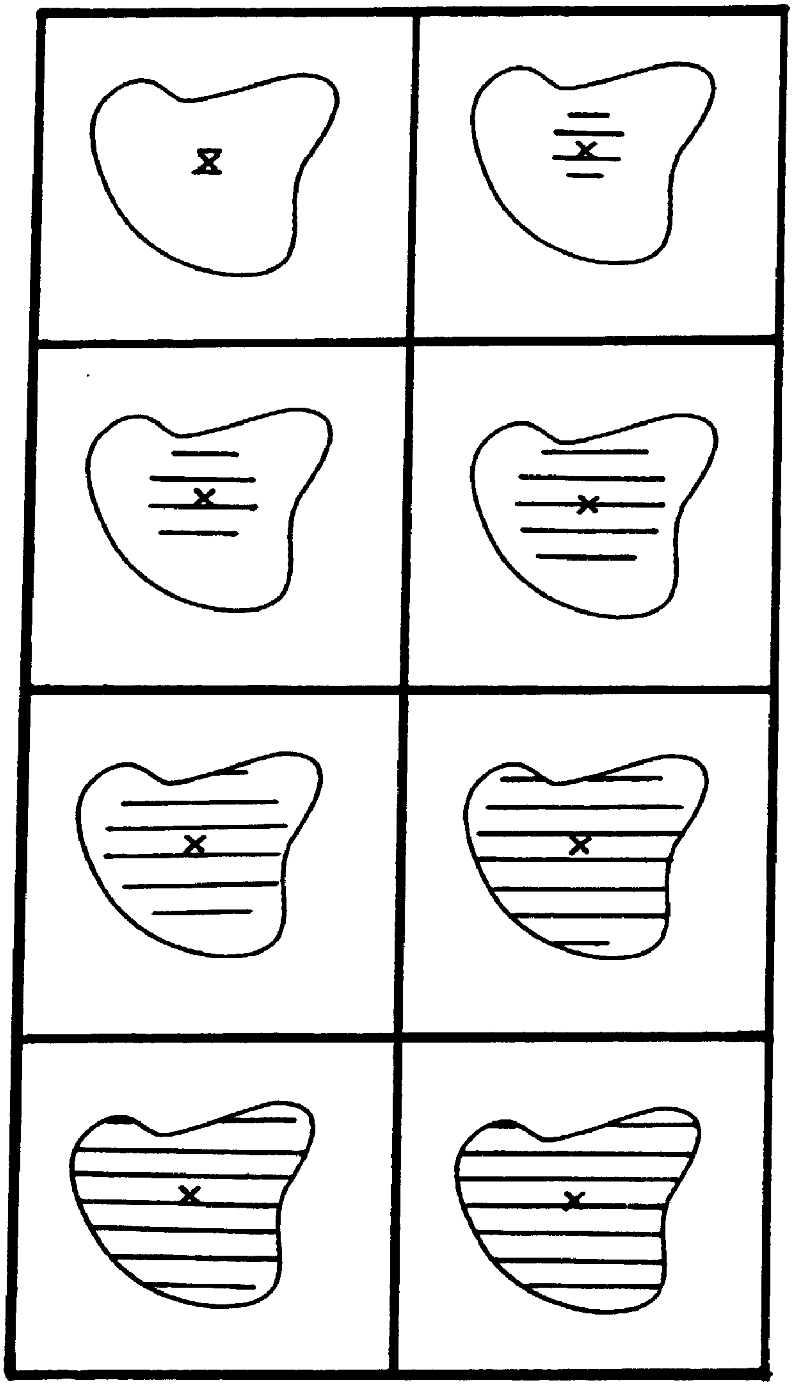


Figure 2. A uniscale colouring algorithm. The size of the region that is coloured increases with each iteration. Eight iterations were required to completely colour the region before activation was checked by bounding contours. An increase in region size would increase the number of iterations required to colour it.

One source of concern about the validity of region colouring models is the notion of spatial infinity. What and where spatial infinity is has not been defined precisely. This ambiguity must be addressed if the speed at which these models find outside solutions is to be compared to human response times because spatial infinity plays an important role as a criterion for stopping activation.

Boltzmann Machines. A more refined region colouring model has been developed by Kienker, Sejnowski, Hinton, and Schumacher (1986) and Sejnowski and Hinton (1987). These researchers implemented a parallel network called a *Boltzmann machine* that separates figure from ground in a number of different displays. This process is essentially the same as determining the inside and outside regions of the figure. In fact, Kienker et al. (1986, p. 198) have suggested that separating figure from ground may be the first step in determining inside/outside relations.

It should be noted, however, that there is some debate about the validity of Boltzmann machines and other connectionist networks as models of higher level visual processing. Feldman and Ballard (1982) and Sabbah (1985) have presented arguments that connectionist networks are indeed suitable as models of both low- and high-level processing. On the other hand, Fodor and Pylyshyn (1988) have outlined a number of points which suggest that connectionist networks are inadequate as models of any type of higher level processing including vision. However, the latter theorists admit that this type of network may be appropriate for describing the "hard-wired", low-level processing thought to be part of the functional architecture of the visual system. Because many of the operations involved in determining spatial relations are carried out at an early stage of visual processing, the examination of Boltzmann machines as models of the perception of inside/outside relations is warranted.

The figure/ground segregation model developed by Kienker et al. (1986) is a parallel network of processing units that receives two types of information. One type is bottom-up input from the image containing information about the outline of a figure. The other type is top-down input containing information about the location of attentional focus. The attentional focus is in the form of a Gaussian attentional spotlight that biases one part of the image to be the inside of the figure. This effect decreases in strength with distance from the centre of the spotlight. In addition, two different types of processing units are involved -- edge units and figure units. Each type is restricted to a separate plane in register with the other. In Figure 3 it can be seen that information about the position of the figure's outline provides bottom-up input to the edge units in one plane, and information about the location of attentional focus provides top-down input to the figure units in the other plane. Also, edge units excite figure units that they point towards and inhibit figure units that they point away from. Over a series of iterations, a set of edge units pointing towards a figure unit can cause it to assume an "inside" state. With this architecture, the network can hold information about the figural outline and the focus of attention fixed, and can find the state of all units that is maximally consistent with the top-down and bottom-up inputs. Hence, visual interpretation of figure/ground problems can be carried out by the network with a combination of top-down and bottom-up information. This is consistent with the nature of operations thought to be carried out at the intermediate-level of visual processing. Furthermore, this model is superior to the standard region colouring models in the sense that processing can be carried out on figures with incomplete outlines. That is, it copes with gapped contours by carrying out a "filling in" operation that involves extending the interaction between the edge and figure units to the next nearest neighbour along the contour's orientation (see Kienker et al., 1986, p. 203).

Each of the aforementioned models represents a significant advance in our understanding of the perception of inside/outside spatial relations. The ray intersection algorithm was an important development because it made explicit a way in which a "connection" could be made between figures when determining their spatial relations. Region colouring models proved to be superior to ray intersection algorithms in many

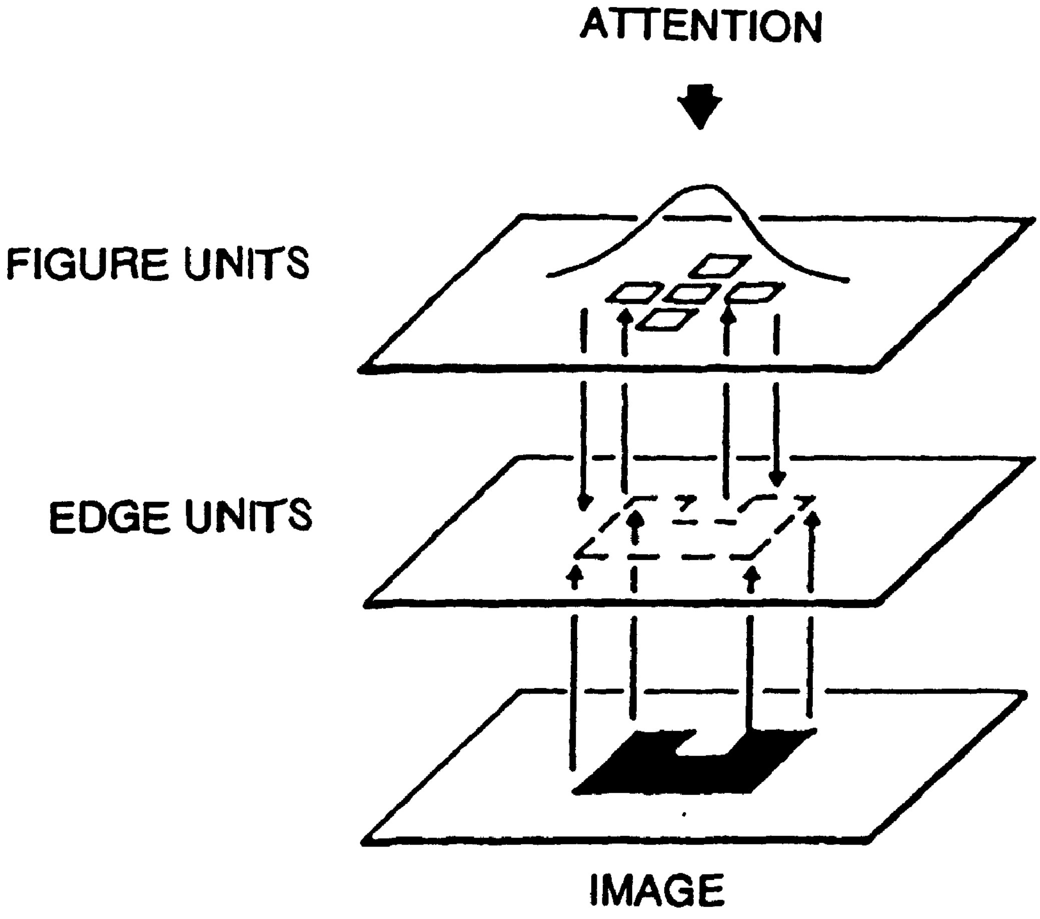


Figure 3. An illustration of the effects of top-down and bottom-up information on the edge and figure unit maps, and the mutual interaction between these maps. This figure was adapted from Kienker et al. (1986).

ways because of their versatility. This is particularly true because the performance of region colouring models is not disrupted to the same extent as ray intersection algorithms by the presence of other contours in the visual field. Kienker et al.'s Boltzmann machine is a refined region colouring model because it has mechanisms for receiving top-down and bottom-up information as input, and is able to find solutions to problems involving gapped figures. Hence, there has been a steady progression of refinements in the development of models of the perception of this relation. However, none approach what Pylyshyn (1984) has called strong equivalence. To develop a model with this property, it appears that more psychophysical data about human performance is required as constraining information.

1.5 OUTLINE OF THE THESIS

The research described in this thesis is a series of studies in which observers were asked to determine the spatial relation "inside of." Very little is known about the operations involved and there is almost no empirical research on this problem in the visual perception literature. Therefore, many of the decisions about experimental design were somewhat arbitrary and the style of experimentation was *exploratory*. In each study, response times were measured to determine which factors systematically and reliably affect the speed at which this relation was perceived. The purpose of the thesis was to compile this type of information in a database in order to gain some insight into the nature of the operations involved in this perception. Chapter 2 outlines a series of experiments in which stimulus size and location were manipulated, and Chapter 3 reports the findings of a set of experiments in which stimuli were presented asynchronously. Chapter 4 is a discussion of this research and its ramifications for future study of the perception of inside/outside relations.

CHAPTER TWO: EFFECTS OF FIGURAL SIZE AND LOCATION ON THE PERCEPTION OF INSIDE/OUTSIDE RELATIONS

Little is known about what factors, if any, reliably affect the speed at which inside/outside relations are determined by the visual system. Two likely candidates, however, are the size of the bounding figure and the location of the target inside or outside this figure. More specifically, if an observer is asked over a series of trials to determine whether a target is inside or outside a rectangle, changes in the size of the rectangle or the target location may increase or decrease response times. The experiments reported in this chapter were conducted to determine whether these factors would systematically affect inside/outside response times and how consistent the results would be with the predictions yielded by region colouring models.

Mahoney and Ullman (1988, pp. 175-176) have suggested that a typical inside/outside colouring routine might include the following steps: Processing focus is shifted from an initial location (e.g., a pretrial fixation point) to the location of the target (see Figure 4). The target is then marked and the region surrounding the target is coloured. Note that with a multiscale algorithm the time required to colour the region is independent of the region's size. However, with a uniscale algorithm the time required to colour the region increases as size is increased. Once the region has been coloured, processing focus is shifted to a peripheral location that is clearly outside the figure (spatial infinity). If this peripheral location is not coloured, then the target is inside a closed contour -- otherwise it is outside. The results of the studies reported in this chapter were compared with the performance of colouring routines of this type.

2.1 EXPERIMENT 1

Ullman (1984, p. 105) has reported that changes in the size of the bounding figure have only a slight effect on inside/outside response times. This result has been taken as an underlying assumption for the development of multiscale region colouring models that

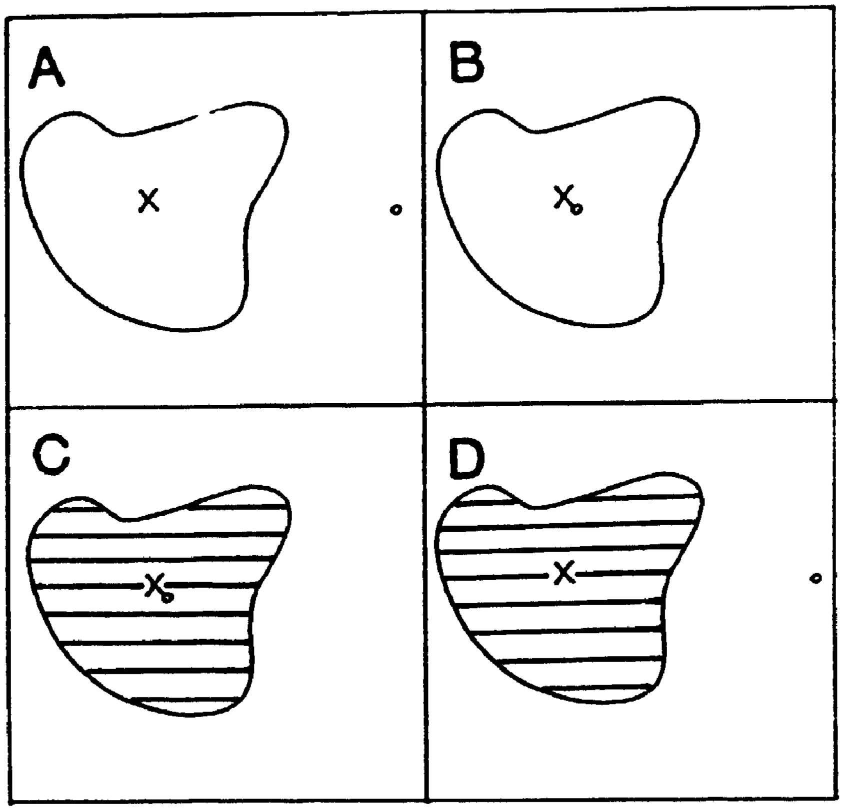


Figure 4. A typical region colouring routine suggested by Mahoney and Ullman (1988). A point outside the figure is marked and processing focus is shifted to the target location. After marking this location, the surrounding region is coloured. Processing focus is then shifted to the marked location outside. This region is not coloured. Therefore, the target is inside a closed contour.

detect inside/outside relations in a time that is independent of the size of the figure involved. However, Ullman's finding was described only briefly in a footnote and it is unclear what the magnitude of this *slight* effect is. In contrast, other research has indicated that size changes do affect inside/outside response times. More specifically, Podgorney and Shepard (1978) conducted a series of experiments in which subjects were shown displays consisting of a block letter (e.g., F) within a 5 x 5 grid (or a 3 x 3 grid in the Podgorney & Shepard, 1983 studies). In each display, a dot was also presented in one of the squares of the grid and was *on* or *off* the letter. Subjects required more time to respond that a dot was *on* the letter as the number of component squares that made up the letter increased. In other words, *on* response times increased as figural size increased. In a two-dimensional display, *on* and *off* are identical to *inside* and *outside*. This suggests that, despite the claims of Ullman, increases in figural size may lead to increases in inside response times in an inside/outside task.

In the following experiment, observers were shown a series of displays on a computer monitor. Their task was to determine whether a target was inside or outside a rectangular figure. The size of this figure varied over trials. If the size manipulation was found to have no significant effect on inside/outside response times, this would support Ullman's claim that inside/outside relations are perceived in a time that is independent of figural size. It would also be consistent with the performance of multiscale colouring models because the solution time involved is independent of size. However, if a significant size effect was found, this would indicate that if region colouring is involved in this perception, a uniscale model would provide a more appropriate account of this perception than a multiscale model. That is, uniscale colouring would require more time to reach a solution as the size of the region to be coloured is increased. Hence, the first experiment was carried out to determine whether size changes would affect inside/outside response times, and to determine which type of colouring model would best account for the results.

2.11 Method

Subjects. Fifteen University of Western Ontario undergraduate students participated in this experiment. All subjects had normal or corrected-to-normal vision and none had any previous experience with this type of study. In addition, all subjects were unaware of the purpose of the study and took part to fulfill a course requirement.

Apparatus and Stimuli. Stimuli were displayed on an NEC JC-1225MA composite colour monitor at a viewing distance at 100 cm. Experimental control and data collection were carried out by a Commodore 64 microcomputer interfaced to two reaction-time keys using a method described by Wright and Dawson (1988). Each trial began with a 1000 ms presentation of a white fixation cross in the centre of a black background. The fixation cross subtended 0.57 x 0.57° of visual angle. This was followed by a 1500 ms blank screen interstimulus interval (ISI) and then the onset of the stimulus display. This display consisted of a white X that subtended 0.57 x 0.57°, and four white lines forming a square figure that varied in size over trials. The figure was either 1.15 x 1.15°, 2.3 x 2.3°, or 3.45 x 3.45°. Positional uncertainty of the figure was introduced so that it appeared in a random order in each of the four quadrants of a central region of the display screen (see Figure 5). On all trials one corner of the square figure was in the centre of the screen. In addition, the position of the target was always the same absolute distance from the fixation point on each trial to control for any effect of fixation cross proximity. As seen in Figure 5, the inside position of the target for a figure in a given quadrant was always the same regardless of the figure's size. Furthermore, the outside position of the target was always in the quadrant diagonally opposite to that containing the square, and was also always the same regardless of the figure's size. For example, the outside position for all figures in the upper left quadrant would be identical to the inside position for all figures in the lower right quadrant. There were 48 trials for each of the six types of conditions per subject (i.e., 3 figure sizes X 2 relations -- inside & outside) and the order of the trials was randomized.

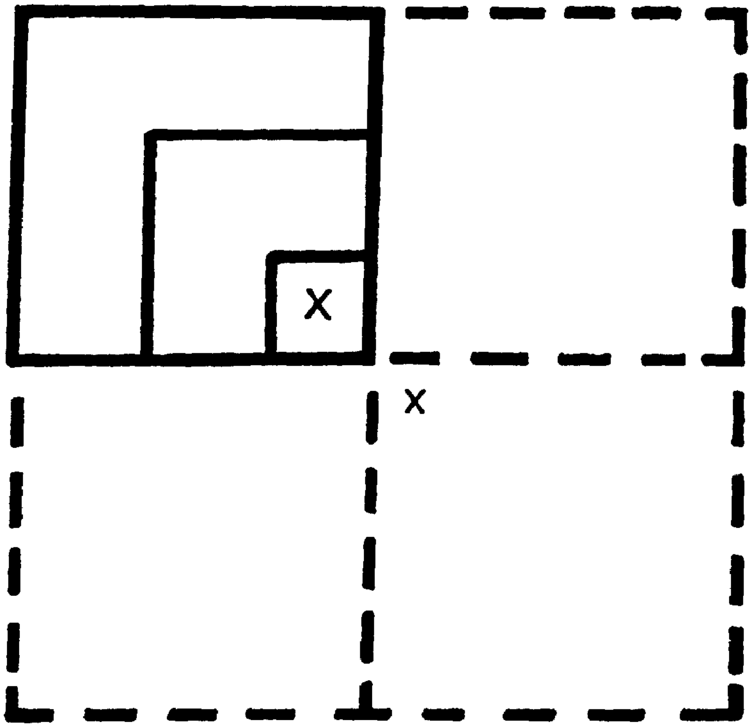


Figure 5. The 3 sizes of bounding figures in Experiment 1. Stimuli were presented individually in a randomly selected quadrant across trials. The dotted lines denote the central display region and its four quadrants.

A rectangle was chosen for the bounding figure in order to conduct experiments with a simple geometric shape. This effectively ruled out other factors such as figural complexity which almost certainly affect inside/outside response times. While rectangular shapes may not have as much ecological validity as elliptical or circular shapes, our proficiency at determining inside/outside relations with schematic diagrams and line drawings is evidence that simple polygons are acceptable as bounding figures in this type of task.

Procedure. Subjects were tested individually in a dark, quiet room. They were also instructed to keep their heads as still as possible and in an upright position throughout the session (no one reported difficulty in doing so). All subjects were asked to judge as quickly and as accurately as possible whether the X presented on the display screen was inside or outside the rectangular figure. In addition, they were told that speed and accuracy were equally important. The display remained visible until a response was made by pressing one of two buttons labeled *inside* and *outside* with the forefingers of the left and right hands. Responses initiated the onset of the next trial. Each subject was given a 20 trial practise run followed by a 288 trial data run. The data runs consisted of eight blocks of 36 trials with a 45 second rest period between each block.

The presentation of response-terminated displays did not rule out the possibility that eye movements would be involved in performing the task. However, eye movements did not appear to affect response times in previous research on the perception of spatial relations. Jolicoeur et al. (1986) presented subjects with stimulus displays for 250 ms followed by a visual mask. The brevity of the stimulus duration eliminated the possibility that eye movements were involved in determining the relation. However, in another experiment in which the stimulus duration was 2500 ms and the same task was performed, they found the same pattern of results. This was the case despite the fact that eye movements could be made during the display interval. Moreover, Jolicoeur (1988, experiment 3) found the same pattern of results when a similar task was performed with a stimulus display that was terminated by a response and a stimulus display presented for 180 ms. Therefore, while the effect of eye movements cannot be ruled out for displays with durations longer than 250 ms, previous research has shown that eye movements do not affect the pattern of response times in some spatial relations tasks.

2.12 Results

Only correct responses within two standard deviations of the mean correct response time for a particular condition (e.g., large figure/inside) were used in the analyses. In the current experiment the outlier analysis led to the removal of 4.14% of the correct responses. After incorrect responses and outliers were removed, 81.42% of the responses remained. The mean response times and accuracy rates are shown in Figure 6.

A two-way ANOVA with repeated measures on both factors (*target inside/outside* X *figural size*) was carried out with the mean response times. These means were 500.19, 511.07, and 514.12 ms for small, intermediate, and large figures respectively and the effect of size was significant ($F_{(2,28)} = 5.25, p < .02$). As in evident in Figure 6, the mean inside response times increased monotonically with increases in the size of the bounding figure. This resulted in a significant linear trend over size for *inside* response times ($F_{(1,28)} = 23.34, p < .01$). However, despite a decrease in the mean outside response times with increases in figural size, a significant linear trend over size for *outside* response times did not result.

A separate two-way ANOVA of response accuracy rates with repeated measures on both factors was also carried out. A main effect of figural size occurred ($F_{(2,28)} = 21.56, p < .0001$) such that subjects were correct on 85.48% of trials with large figures, 80.83% of trials with intermediate figures, and 77.92% of trials with small figures. The mean inside response accuracy rates increased monotonically with increases in figural size, which resulted in a significant linear trend ($F_{(1,28)} = 84.85, p < .01$). A significant linear trend for outside response accuracy rates was not found. However, there was a main effect of target inside/outside ($F_{(1,14)} = 49.16, p < .0001$). Subjects responded

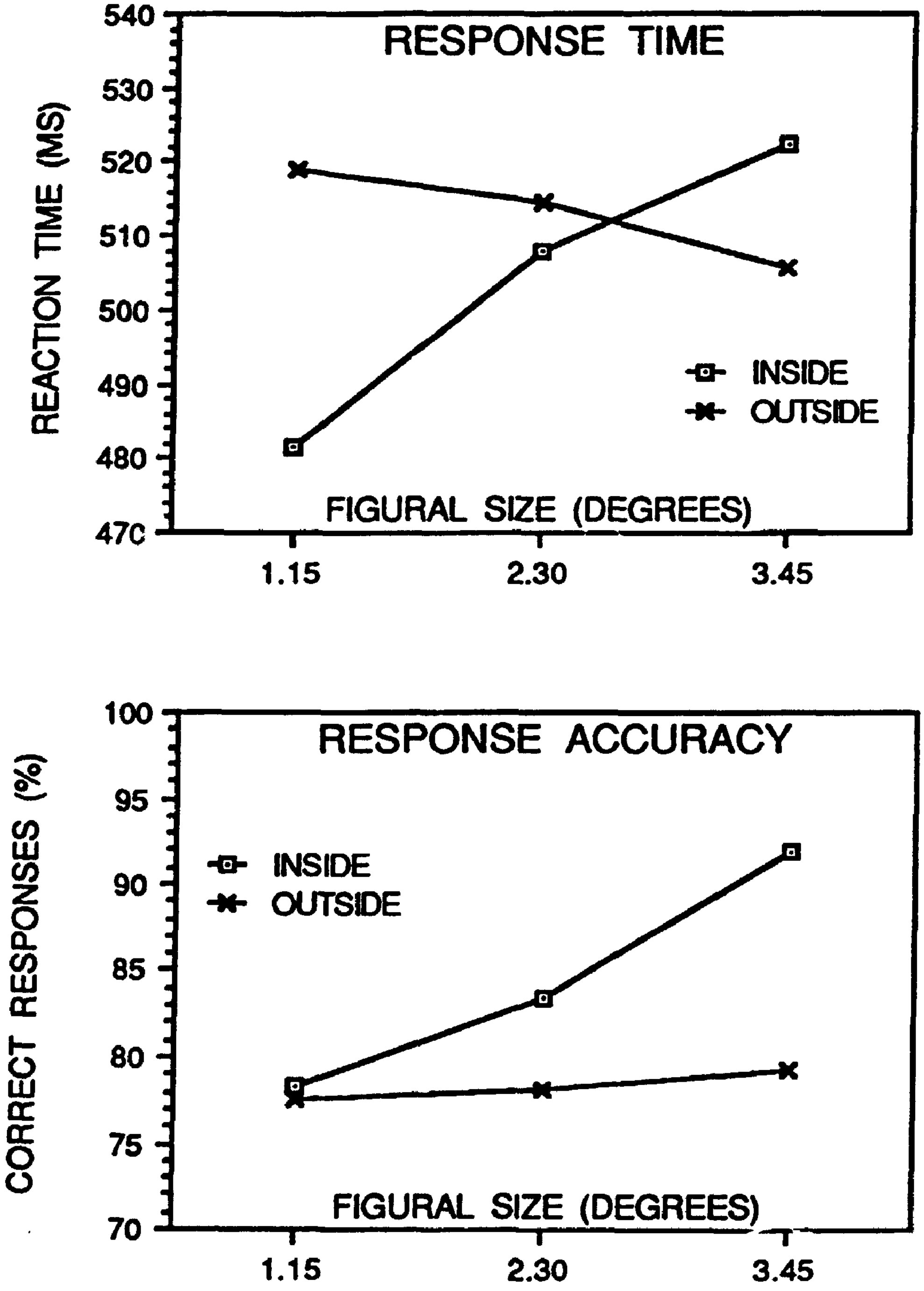


Figure 6. Mean response times and accuracy rates for all subjects in Experiment 1 as a function of figural size.

correctly on 84.54% of inside trials and 78.29% of outside trials. Furthermore, there was a significant interaction of Figural Size x Target Inside/Outside ($F_{(2,28)} = 16.63, p < .0001$). Response accuracy decreased as figural size decreased on inside trials but not on outside trials.

2.13 Discussion

The size of a bounding figure in an inside/outside task was manipulated and these changes affected response times. To elaborate, increases in figural size led to corresponding *increases* in inside response times and *decreases* in outside response times. Furthermore, the effect of changing figural size had the same effect on inside response times in this experiment as it did on *on* response times in the experiment carried out by Podgorney and Shepard (1978). Thus, for inside responses, this finding is consistent with the performance of uniscale colouring algorithms. That is, the time required by these algorithms to colour internal regions increases as a function of the size of these regions.

The results of this experiment suggest that if colouring was carried out, it did *not* involve the simultaneous activation of neighbour regions of different sizes (multiscale colouring). Instead, colouring would have involved the activation of neighbour regions of only one size at a time because colouring time and figural size are related. It should be noted that there is a widely held view that visual attention can only be focused at one resolution at any given time. In particular, previous research (e.g., Navon, 1977; Ward, 1982) on the perception of bilevel letter stimuli indicates that attention can be focused on the global or local level but not on both simultaneously. If region colouring is also uniscalar in this sense, questions arise about the size of the neighbour regions that are activated and how this size is determined. As mentioned previously, Jolicoeur (1988) found that the speed of contour tracing decreased as the proximity of other contours increased. Presumably, increases in this proximity led to a finer resolution of processing focus for carrying out the tracing operation. The size of the regions involved in colouring could also be influenced by figural complexity, proximity to other contours, and other factors that appear to affect the speed of contour tracing. Thus, the time required to colour a region may depend on the size of the neighbour regions that are activated.

It could be proposed that observers made judgements in this task on the basis of some other spatial relation rather than inside/outside. For example, if the figure and the target appeared on the same side of the fixation point, observers may have made *inside* responses. Conversely, if the figure and the target appeared on opposite sides, observers may have made *outside* responses. However, determining a "side-of-fixation-point" relation cannot account for the data of the current study. More specifically, the speed at which this relation is determined should not be affected by changes in figural size because only the part of the figure closest to the fixation point is involved in the task. Thus, changes in figural size should be irrelevant because the part of the figure closest to the fixation point does not change.

Another possibility is that observers may have made inside responses when the vertex closest to the fixation point pointed *away* from the target. Conversely, observers may have made outside responses when this vertex pointed *towards* the target. However, the performance of this task does not account for the effect of figural size changes either.

Still another proposal is that inside responses may have been made if the target was on the same side of the fixation point as the vertex farthest away from this point. Outside responses may have been made when the target was on the opposite side of the fixation point. The determination of this relation should be affected by changes in figural size. That is, size changes directly affect the distance between the vertex involved and the fixation point. Therefore, the time required to shift the processing focus between the two locations should be affected. As a result, response times should also be affected. However, previous research has been conducted with displays in which the target was on the same side of the fixation point as this vertex but was outside (Wright, 1988). A systematic pattern of inside/outside response times resulted from manipulations of target location. This suggests that the relations determined in this type of task did not involve a judgement about the location of stimuli on the *same* or on *different* sides of the fixation point.

Finally, it could be argued that observers were simply determining the relation *"between two points"* rather than inside/outside. In particular, observers may have made inside responses if the target was between two vertices, and outside responses if the target was not. This proposal is consistent with the occurrence of size effects because these changes directly affect the distance between the vertices of rectangular figures. However, in the study described above (Wright, 1988), there was also a systematic pattern of inside response times resulting from target location manipulations. That is, inside response times were slowest when the target was near a vertex as opposed to further away from one. If the relation involved was simply a judgement about whether the target was between two vertices or not, this target location effect would not be expected. More specifically, the operation involved should be a scan between the two vertices in question. There is no reason to believe that the location of a target lying directly between these two points should affect the time required to do so. Therefore, "between two points" may be an oversimplification of the relation determined in this type of task.

In summary, the flexibility involved in assembling different types of visual routines and the simplicity of the figural shape involved in the current study raise questions about whether inside/outside was determined as opposed to some other relation. However, the occurrence of figural size effects in this experiment as well as target location effects in a previous experiment indicate that observers were most likely to be performing inside/outside tasks. In order to be completely sure that this was the case, an additional study (not part of this thesis) could be conducted involving a direct comparison of *inside/outside* response time patterns and *between-two-points* response time patterns. To elaborate, the same distances would be involved between the points or vertices of the bounding figure that the target is between. If, however, response time patterns differ in the two conditions, this would provide evidence that observers were not determining a between-two-points relation in the current study.

Although very different tasks were involved, it should be noted that the pattern of inside response times are similar to those obtained by Bundesen and Larsen (1975). In these studies observers were shown two shapes and were asked to judge whether the shapes were identical except for size and orientation. It was found that reaction times for *same* responses increased approximately linearly as the size ratio between the shapes increased. Similar results have been obtained by Larsen and Bundesen (1978) and Sekuler and Nash (1972). Thus, other findings in the literature indicate that for some visual tasks, increases in stimulus size lead to increased response times.

The corresponding increase in inside response accuracy with increases in size suggests that a speed-accuracy tradeoff for inside responses may have occurred as a result of the size manipulation. This would suggest that on "inside" trials observers were more likely to "fast guess" or somehow reduce the accuracy criterion for displays with small figures than for those with larger figures (see e.g., Pachella, 1974, pp. 71-80; Swensson, 1972; Yellott, 1971). Thus, the effect of size changes would not be a change in the amount of visual processing required before inside responses could be made. Instead, the effect would be to influence a response criterion ranging from "fast guess" with small figures to a more careful, controlled, and accurate response to displays with large figures. However, subjects did not know what type of figure to expect before each trial. Furthermore, the mean response time (508.46 ms) indicates that they did not have much time to set an accuracy criterion before a response was made.

Inside response accuracy also could have been lower on small figure trials because the target was only 0.29° from each side of the figure. In particular, Uttal (1975) has conducted research indicating that for certain visual tasks, lateral retinal interactions will occur if the elements involved are less than 0.5° apart. In addition, Gardner (1973) found that target detection was only independent of target numerosity if the elements in question were more than 0.5° apart. Therefore, inside response accuracy may have been lower in small figure trials than in other types of trials because of the interelement distances involved.

There is also another reason to question the occurrence of a speed-accuracy tradeoff for inside responses. That is, changes in figural size were confounded with target location. When the smallest figure was displayed, the target appeared in the centre of it. However, when the larger figures were displayed, the targets did not appear at a central position within the figures but instead were displayed at the same absolute distance from the fixation point. Hence, it may be the case that inside response accuracy is also affected by target location. If so, then what appears to be a speed-accuracy tradeoff for inside responses may instead be a consequence of target location. One of the purposes of the second study was to test this question.

2.2 EXPERIMENT 2

In this experiment, subjects were shown inside/outside displays involving two sizes of bounding figures. There were also five target locations inside and outside the figures. It was expected that if changes in target location affected response accuracy, then this may indicate that size changes don't lead to a tradeoff between inside response speed and accuracy. Instead, speed and accuracy would be correlated with changes in target location as well as figural size.

There are several reasons for believing that inside/outside response times are affected by target location. As mentioned previously, Podgorney and Shepard (1978) carried out a series of experiments in which subjects were shown a block letter within a 5 x 5 grid. A dot was also presented in the grid and the task was to determine whether the dot was *on* or *off* the letter. Shepard and Podgorney (1986) replotted the data from these studies to assess the dependence of response times on dot location. They found that *off* responses decreased as the distance between the dot and the letter increased. In addition, Wright and Dawson (1987) conducted an experiment in which target location was varied in an inside/outside task. They found that as the distance between the target and the figure increased, outside response times decreased. Furthermore, the same target location effect on outside response times occurred in a study described previously (Wright, 1988). Also, in this study it was found that inside response times were significantly slower when the target was near a vertex rather than further from one. Thus, changes in target location appear to affect the time required to determine inside/outside relations.

According to Ullman (1984), region colouring models reach an *outside* solution if the spread of activation flows to a point referred to as spatial infinity (i.e., a point clearly outside the bounding figure in question). The time required for multiscale colouring to spread across the region from the target to spatial infinity should not depend on the size of the region between the two locations because multiscale colouring activates a region in the same time regardless of its size. Therefore, the occurrence of a target location effect on outside response times would not be consistent with a multiscale colouring account of this perception. Furthermore, changes in target location inside the figure should not affect the solution time of a multiscale algorithm because its performance is size-independent. However, uniscale colouring *should* be affected by target location changes. In particular, the number of iterations of activation spread required to colour a region should increase as the target location becomes closer to the perimeter of the figure (see Figure 7). Therefore, if uniscale colouring is involved in determining inside relations, manipulations of target location should affect the pattern of response times. More specifically, larger mean response times should result when the target is in positions 1 and 5 (near the perimeter of the figure), and smaller mean response times should result when the target is in positions 2 and 4. The smallest mean response time should result when the target is in position 3. That is, position 3 is in the centre of the figure and, therefore, should involve the fewest colouring iterations.

In the current study the target locations were horizontal, collinear, and distributed equally in both hemifields. Thus, a straightforward analysis could be carried out on the

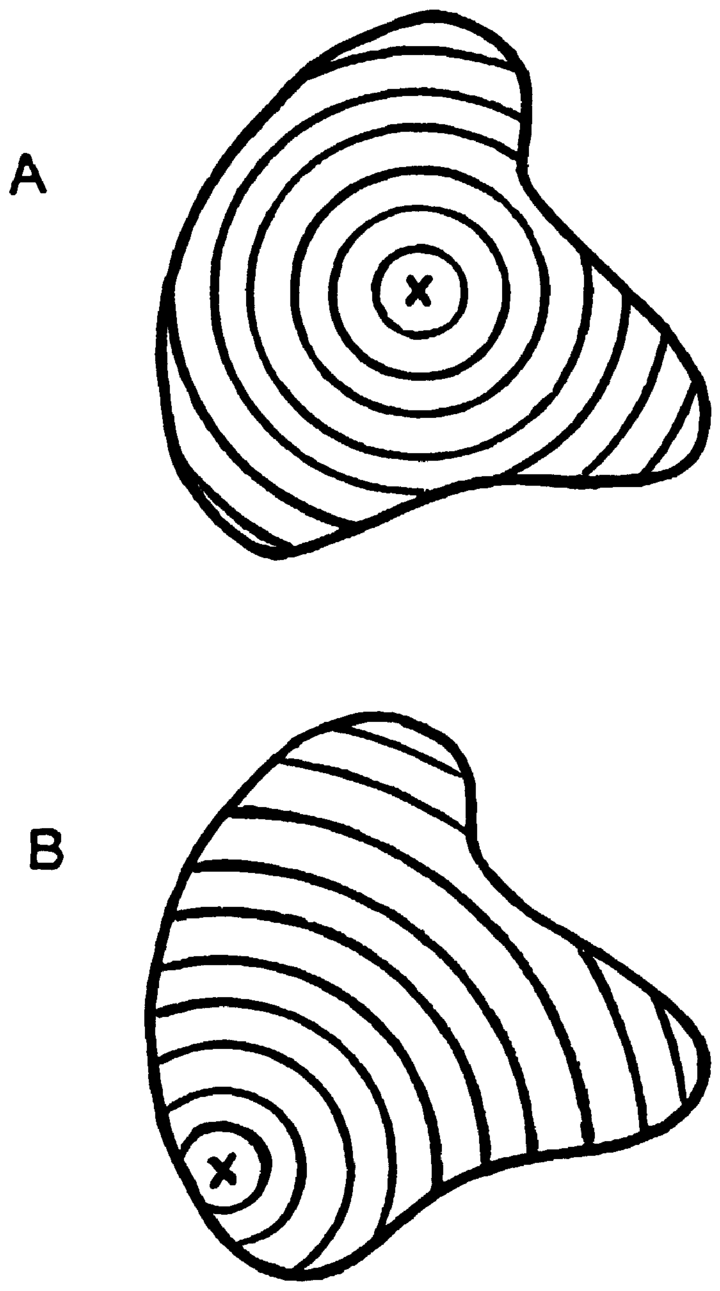


Figure 7. The effect of changes in target location on uniscale region colouring. Eight iterations of the algorithm were required to colour the region in A and eleven iterations were required to colour the region in B.

effects of the hemifield that the figure was presented in on response times. The discovery of hemifield effects could support Kosslyn's (1987) claim that the left hemisphere of the brain is superior to the right for determining spatial relations of a categorical nature (e.g., inside/outside, on/off).

2.21 Method

Subjects. Five University of Western Ontario graduate students (one of whom was the author) participated in this experiment as paid observers (except the author). All subjects had normal or corrected-to-normal vision, and had previous experience in studies based on reaction-time measures.

Stimuli and Procedure. The same apparatus, testing conditions, and instructions used in the first experiment were employed in this study. The stimuli were a white dot that subtended 0.57 x 0.57° of visual angle, and four white lines forming a rectangle that subtended either 2.56 x 3.15° or 4.88 x 6.01° of visual angle. Positional uncertainty of the rectangle was introduced so that over trials it appeared in a random order either to the left or the right of the fixation cross location. There were 10 collinear positions of the target relative to the rectangular figure (5 inside & 5 outside) and two sizes of figures (see Figure 8). The large figure was almost twice the size of the small one, and the relative positions of the target and the figure remained constant according to this scale. As a result of the positional uncertainty of the figure across trials, no location on the display screen was always inside or always outside. In addition, there were 160 trials for each of the 20 conditions (i.e., 2 figure sizes x 5 target positions x 2 relations -- inside or outside) per subject and the order of trials was randomized within test sessions.

The procedure was the same as that of the previous experiment. Each subject was tested in 10 sessions consisting of a 20 trial practise run and a 320 trial data run for a total of 3200 data trials. Subjects were required to complete the sessions within 10 days but to participate in no more than two sessions per day. The data run of a session consisted of eight blocks of 40 trials with a 45 second rest period between each block.

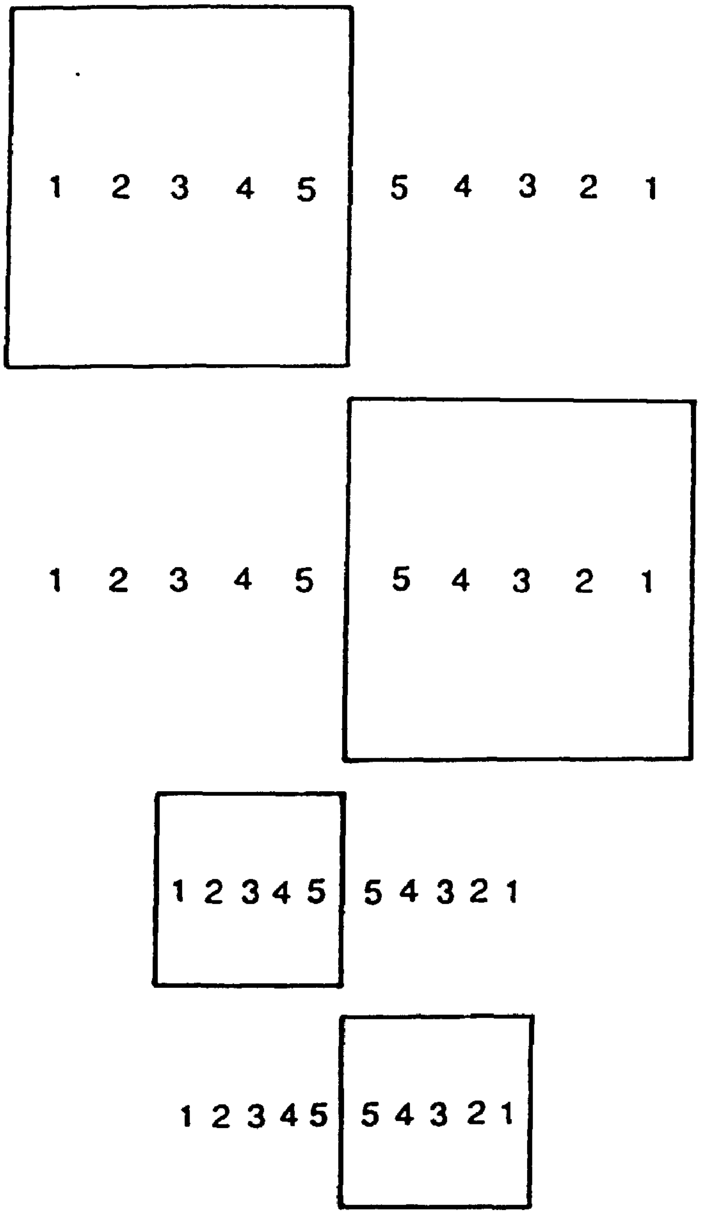


Figure 8. The different locations of the target inside and outside of the rectangular figures in Experiment 2. On each trial either a large figure or a small figure was presented to the left or to the right of the fixation cross. The outside target locations were always in the opposite visual hemifield to the inside target locations. The relative locations of the target and the figure remained constant for both sizes of figures but changed in scale.

2.22 Results

As in the previous experiment, only correct responses within two standard deviations of the mean correct response time for a particular condition were used in the following analyses. This led to the removal of 4.75% of the correct responses. After incorrect responses and outliers were removed, 90.79% of the responses remained.

A three-way ANOVA with repeated measures on all factors (*target position inside figure* X *figural size* X *figure in left/right visual hemifield*) was carried out with the mean *inside* response times for each condition. Unlike the previous experiment, inside and outside response times were analyzed separately. Inside and outside could no longer be considered two levels of the same factor because the set of five target locations inside the figure could not be collapsed with those outside the figure. However, target positions were collapsed across small and large figures. It was found that the difference between the mean inside response times for large (373.34 ms) and small figure (365.28 ms) trials was marginally significant ($F_{(1,4)} = 7.51, p = .051$). In addition, the effect of target position inside the figure on response times resulted in a significant linear trend ($F_{(1,16)} = 35.32, p < .01$). That is, the mean inside response times decreased from 378.65 to 362.62 ms for target positions 1 (peripheral) through 5 (central) respectively. As seen in Figure 9, the mean inside response times on *large* figure trials decreased monotonically with decreases in the distance between the target and the fixation point. This resulted in a significant trend with a linear ($F_{(1,16)} = 67.30, p < .01$) and a quadratic ($F_{(1,16)} = 53.89, p < .01$) component. There was also a Figural Size x Target Position interaction effect on inside response times ($F_{(4,16)} = 6.10, p < .0036$). That is, inside response times for large figures were affected to a greater extent by the target position manipulation than those for small figures. However, this effect is based on relative distances. When absolute distances between the target and the fixation point were considered, this effect did not occur. In addition, the mean inside response time for displays presented in the left visual hemifield (358.94 ms) was significantly faster than that for displays presented in the right visual hemifield (379.69 ms), $F_{(1,4)} = 26.10, p < .007$.

A similar three-way ANOVA (*target position outside figure* X *figural size* X *figure in left/right visual hemifield*) was carried out with the mean *outside* response times in each condition. It was found that the mean response time for large figures (348.11 ms) differed significantly from that for small figures (354.55 ms), $F_{(1,4)} = 12.89, p < .024$. In addition, the effect of target position on outside response times resulted in a significant trend with a linear ($F_{(1,16)} = 35.02, p < .01$) and a quadratic ($F_{(1,16)} = 12.48, p < .01$) component. That is, the mean outside response times increased from 348.78 to 356.99 ms for target positions 1 (peripheral) through 5 (central) respectively. Furthermore, the mean outside response times on *large* figure trials decreased monotonically with increases in the distance between the target and the bounding figure, which resulted in a significant linear trend ($F_{(1,16)} = 5.84, p < .05$). Also, the mean outside response time for displays with bounding figures in the left visual hemifield (341.24 ms) was significantly faster than that for displays with the figure in the right hemifield (361.42 ms), $F_{(1,4)} = 22.03, p < .01$.

A three-way ANOVA was also carried out with the mean *inside* response accuracy rates for each condition. It was found that inside response accuracy was greater for small figures (91.98%) than for large figures (88.68%), $F_{(1,4)} = 10.01, p < .035$. In Figure 10 it can be seen that this effect was due to a reduction in accuracy on trials in which targets were presented in positions 1 and 2 with large figures. There was also a Target Position x Figural Size interaction ($F_{(4,16)} = 3$ [illegible], $p < .035$). This appeared to be the case because target position 1 was a greater absolute distance from the fixation point on trials with large figures than on trials with small figures. Also, the mean inside response accuracy rates for *large* figure trials increased as the distance between the target and the fixation point was decreased. This resulted in a significant trend with linear ($F_{(1,16)} = 37.42, p < .01$) and quadratic ($F_{(1,16)} = 20.05, p < .01$) components. The occurrence of the latter

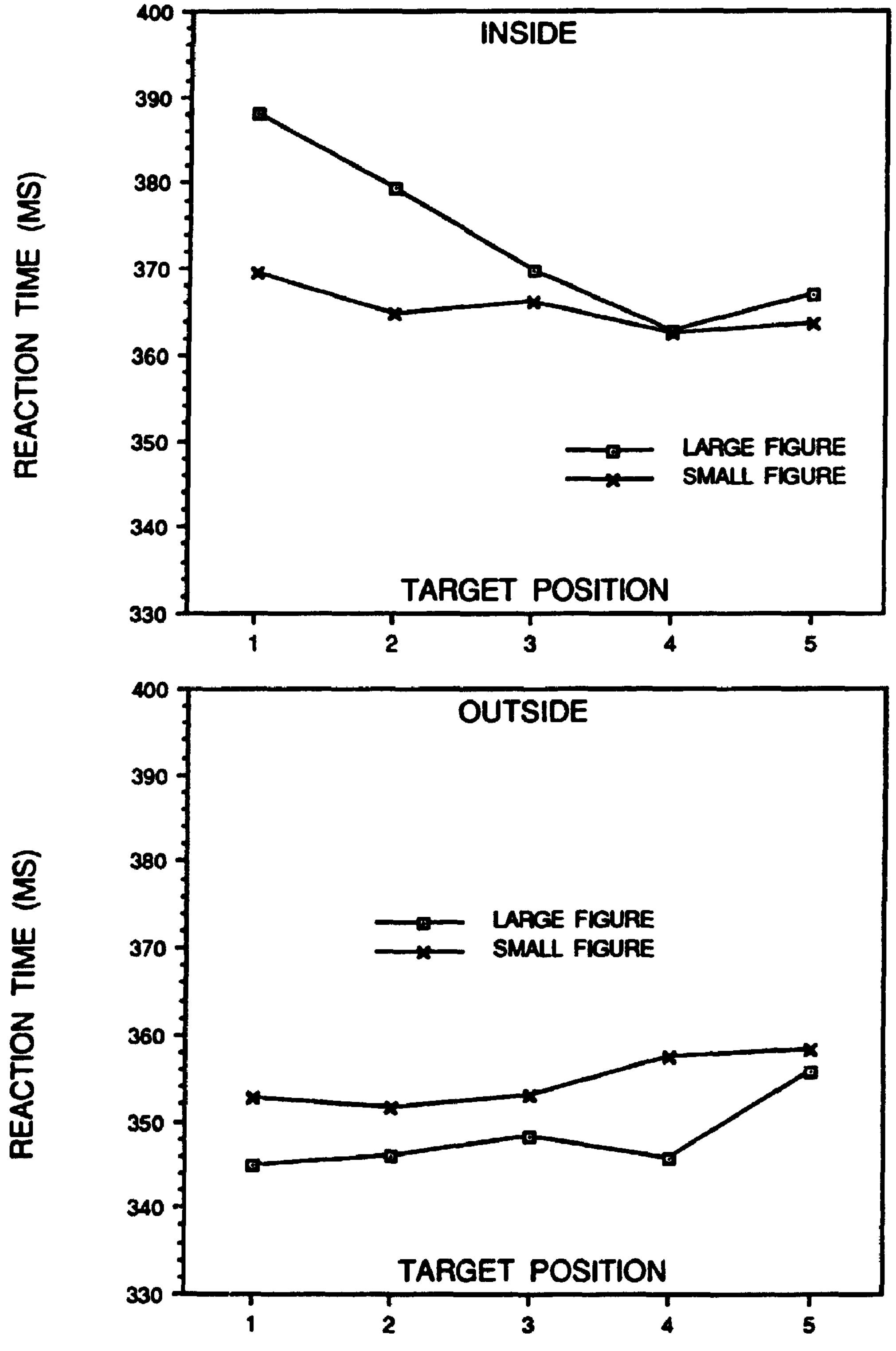


Figure 9. Mean response times in Experiment 2 as a function of target position and figural size. Position 5 was central and position 1 was peripheral.

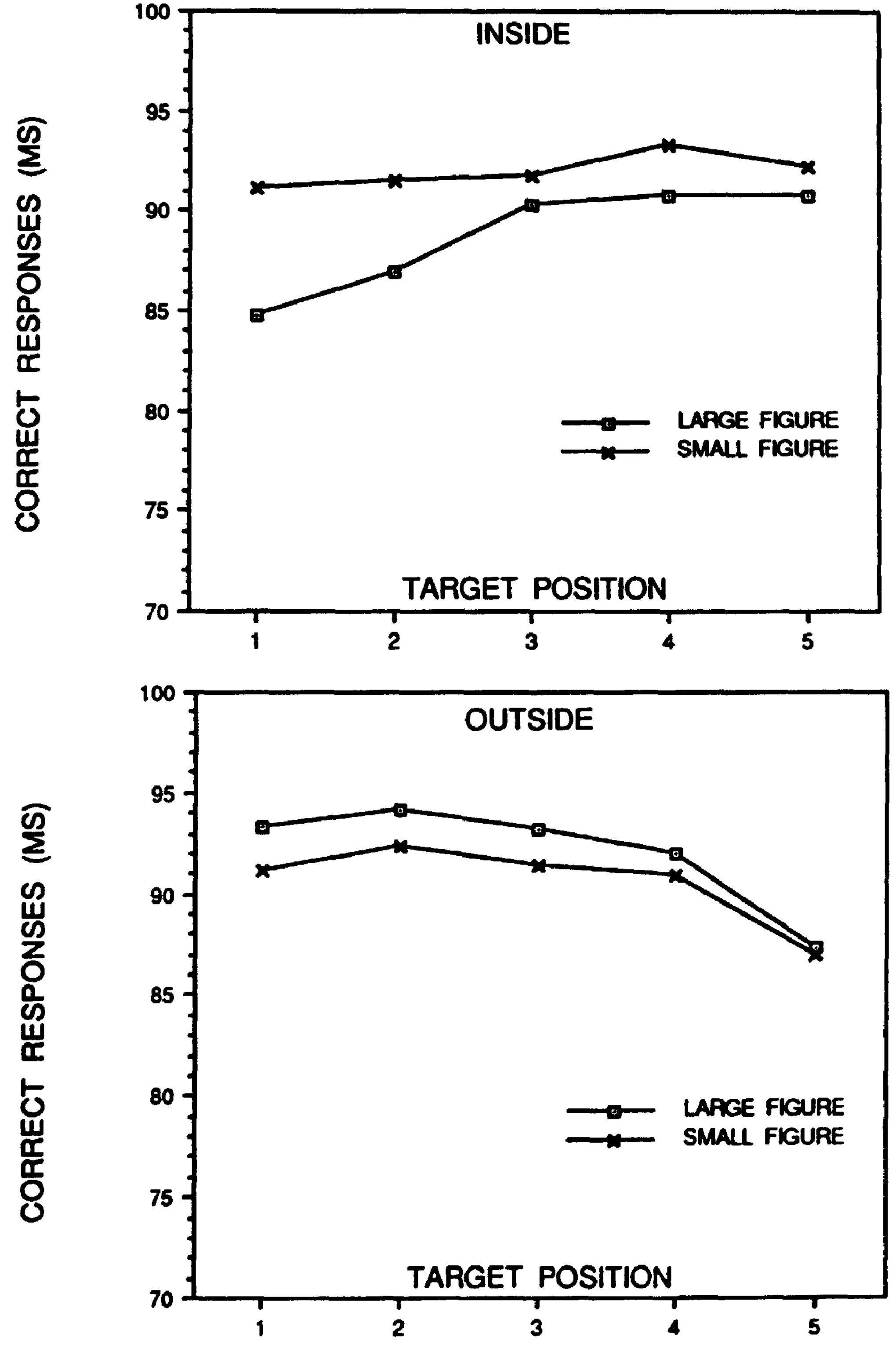


Figure 10. Mean response accuracy as a function of target position and figural size. Position 5 was central and position 1 was peripheral.

suggests that accuracy increases with increasing proximity of the target and the fixation point until the target is adjacent to the fixation point. At this distance, accuracy decreases slightly, perhaps because proximity increases eventually decrease spatial discriminability.

Finally, a three-way ANOVA was carried out with the mean *outside* response accuracy rates for each condition. The only significant effect was of target position ($F_{(4,16)} = 3.11, p < .05$). The mean accuracy rates decreased from 92.19% to 87.19% for target positions 1 through 5 respectively. More specifically, the mean response accuracy rates decreased monotonically as the distance between the target and the figure was decreased for *large* and *small* figure trials. Moreover, significant linear trends were found for large figure ($F_{(1,16)} = 18.25, p < .01$) and small figure ($F_{(1,16)} = 9.01, p < .01$) responses.

2.23 Discussion

As in the first experiment, inside response times increased with increases in figural size. However, inside response accuracy *decreased* with the size increases. This suggests that inside response speed and accuracy are correlated with target location as well as figural size. Therefore, both factors should be taken into consideration when measuring speed and accuracy. Also, the effect of size on inside response times was similar to its effect on *on* response times in the Podgorney and Shepard (1978) studies. If these findings are interpreted in terms of region colouring, a uniscale activation algorithm would provide a more suitable explanation of the data than a multiscale algorithm. That is, the former reaches a solution in a time dependent on figural size, and inside judgements were also made in a time dependent on figural size.

One of the results of this experiment was the occurrence of a visual hemifield effect. In particular, response times were roughly 20 ms faster for displays presented in the left hemifield than for those presented in the right hemifield. This finding suggests that the right hemisphere is more effective than the left at determining inside/outside spatial relations. This is consistent with results indicating that the right hemisphere is more specialized than the left for spatial-visual tasks (e.g., Kimura, 1973). However, it differs somewhat from the predictions of Kosslyn (1987). Kosslyn has proposed that the brain computes two different kinds of spatial-relations representations. More specifically, he has claimed that the left hemisphere is more effective than the right at assigning a spatial relation such as "inside of" to a category. The right hemisphere is said to be more effective at representing one or more locations using a metric coordinate scheme that specifies distances. His claim is based on the results of an experiment (Kosslyn, 1987, pp. 164-165) with stimuli very similar to those used in the current study. In particular, he found that on/off responses (i.e., categorical responses) were faster for displays in the right hemifield than the left and near/far responses were faster for displays in the left hemifield than the right.

Details about the number of different target distances in Kosslyn's (1987) study are not clear. However, there were at least three -- i.e., targets 0 mm from the figure, targets less than 2 mm from the figure, and targets more than 2 mm from the figure. In each of the two conditions the stimulus displays were the same but the instructions differed. In one condition, observers made judgements about whether a target was *on* the contour of the figure or *off* it. This is a categorical judgement in the same sense that inside/outside is. In a second condition, observers made judgements about whether the target was *near* the figure (either on the figural contour or within 2 mm of it) or *far* from the figure (greater than 2 mm from it). This is a metric judgment that involves, according to Kosslyn, computing a coordinate representation of the spatial relation.

As mentioned previously, Kosslyn found that on/off (categorical) judgements were faster when stimulus displays were presented in the right visual hemifield and near/far (metric) judgements were faster when stimulus displays were presented in the left hemifield. This result was obtained even though on/off responses were made with displays involving at least three different metric target distances from the figure, and possibly more. In the current study, there were five distances between the figure and an outside target. Also, subjects were asked to make categorical judgements (inside/outside). Therefore, the two studies are very similar. However, in the current study, responses were faster when the figure was presented in the left hemifield as opposed to the right. According to Kosslyn's proposal, this suggests that judgements in the current study were based on a metric coordinate scheme despite their categorical nature of these judgements. On the other hand, it is not clear from the description of Kosslyn's study how the current study differs from the task in the categorical condition of Kosslyn's study. If subsequent information does not indicate that critical differences exist between the two, then the findings of Experiment 2 raise questions about Kosslyn's claim.

The effect of changes in target location on inside response times appears to be attributable to changes in the absolute distance between the target and the fixation point rather than to colouring. In particular, inside response times were fastest when the target was closest to the fixation point (position 5) and slowest when the target was farthest away from the fixation point (position 1). Presumably, each trial began with a shift of processing focus from the fixation point to the target. Increases in the absolute distance between these two locations presumably increased the time required to shift the processing focus between them. A U-shaped inside response-time function was expected if uniscale colouring was involved in determining the relation. However, the linearity of the function suggests that this was not the case. Instead, the pattern of inside response times may reflect differences in the amount of time required for an initial shift of processing focus to the target. Thus, while the effect of size changes is consistent with a uniscale colouring algorithm, the effects of the target location manipulation suggest that colouring may not play a role in the determination of inside relations.

According to Ullman (1984), the stopping criterion for region colouring algorithms determining "outside of" is the spread of activation to spatial infinity. However, the location of this point is not clear beyond the fact that it is outside the figure.

Furthermore, there is some uncertainty about whether the point initially indexed as spatial infinity is the same distance from the bounding figure on each of a series of trials, or whether this distance can change over trials. It does seem reasonable to assume that as the distance between the figure and a target outside this figure is increased, the distance between the target and the spatial infinity point will decrease. Because of the ambiguities involved with the notion of spatial infinity, region colouring does not seem to be an appropriate account of how outside relations are determined. If colouring does play a role in this perception, a uniscale activation algorithm accounts for the findings of this study better than a multiscale algorithm. That is, increases in the distance between the target and the figure should cause decreases in the distance between the target and spatial infinity. Therefore, the spread of uniscale colouring from the target to spatial infinity should occur in less time as the distance between them is decreased. On the other hand, multiscale colouring should occur in the same time regardless of the size of the region between the target and spatial infinity. Therefore, of the two colouring algorithms, uniscale colouring is more consistent with the effects of target location on outside response times in the current study.

In general, the pattern of outside response times are difficult to account for, and particularly those produced by size changes. One possibility is that these responses may have been affected to some extent by the perimeter of the CRT. To elaborate, if subjects perceived the target as outside the figure but inside the screen perimeter, increases in figural size would lead to decreases in this area outside the figure but inside the screen perimeter. Because figural size increases make this outside region smaller, the time to activate this outside region, perhaps by colouring, and then to make an outside response may be faster as figural size is increased. Experiment 3 was carried out to test this possibility.

2.3 EXPERIMENT 3

If screen perimeter affected outside response times in the previous experiment, then a condition in which this perimeter is no longer detectable should lead to a different pattern of response times than a normal screen perimeter condition. In particular, outside response times should not decrease with increases in figural size because there would no longer be a clearly defined area outside the figure but inside the screen perimeter. Hence, there would not be a clearly defined region that would decrease in size as figural size increased. Similarly, if screen perimeter affected response times in the previous study, then a condition in which this perimeter was visible but reduced in size should lead to different response times than a normal screen perimeter condition. That is, the magnitude of outside response times should be less in the reduced perimeter condition, but the overall response time pattern should be similar to that of the normal perimeter condition. However, if screen perimeter does not affect these responses, the magnitudes and patterns of response times in all three conditions should be similar.

2.31 Method

Subjects. Forty-five University of Western Ontario undergraduate students participated in the experiment, 15 in each of three conditions. All subjects had normal or corrected-to-normal vision and none had any previous experience with this type of task. In addition, all subjects were unaware of the purpose of the study and all took part to fulfill a course requirement.

Apparatus and Procedure. The apparatus was the same as that of the previous experiments except for the following: In Condition A (the No Perimeter Condition) a viewing tube made of black posterboard and lined with black felt was fastened to the monitor. The viewing distance remained the same but the intensity of the monitor was reduced so that the background of the display screen could not be seen. In this condition only the stimuli were visible and the screen perimeter could not be detected. It should be noted that the monitor intensity in Condition A was lower than that of the other conditions. This was the case because each condition was originally carried out as a separate experiment. Thus, holding monitor intensity constant across conditions was not considered when the experiments were conducted. In Condition B (the Reduced Perimeter Condition) a flat black posterboard cover with a 14 x 15 cm opening for the screen was fastened to the monitor. With the addition of white tape around this opening, its perimeter was clearly visible and served as a somewhat smaller screen perimeter (8.01 x 8.58°) than that of the monitor (10.26 x 13.69°). The purpose of Condition B was to determine whether a reduction of this perimeter would lead to a corresponding decrease in outside response times. Nothing was fastened to the monitor in Condition C (the Normal Perimeter Condition) and subjects in this condition served as a control group. Thus, Condition C was identical to the first experiment. The stimulus displays and procedure were also the same as those of the first experiment. A 0.57 x 0.57° dot was used as a target rather than an X because it was more visible at the lower stimulus intensity required to avoid illuminating the sides of the viewing tube in Condition A. Each subject was given a 20 trial practise run followed by a 288 trial data run. The data run consisted of eight blocks of 36 trials with a 45 second rest period between each block.

2.32 Results

As in the previous studies, only correct responses within two standard deviations of the mean correct response time for a particular condition were used in the following analyses. This led to the removal of 4.84% of the correct responses in Condition A, 4.83% of the correct responses in Condition B, and 3.39% of the correct responses in Condition C. After incorrect responses and outliers were removed, 81.53% of the Condition A responses remained, 81.69% of the Condition B responses remained, and 81.16% of the Condition C responses remained. The mean response times for subjects in all conditions are shown in Figure 11, and the mean accuracy rates are shown in Figure 12.

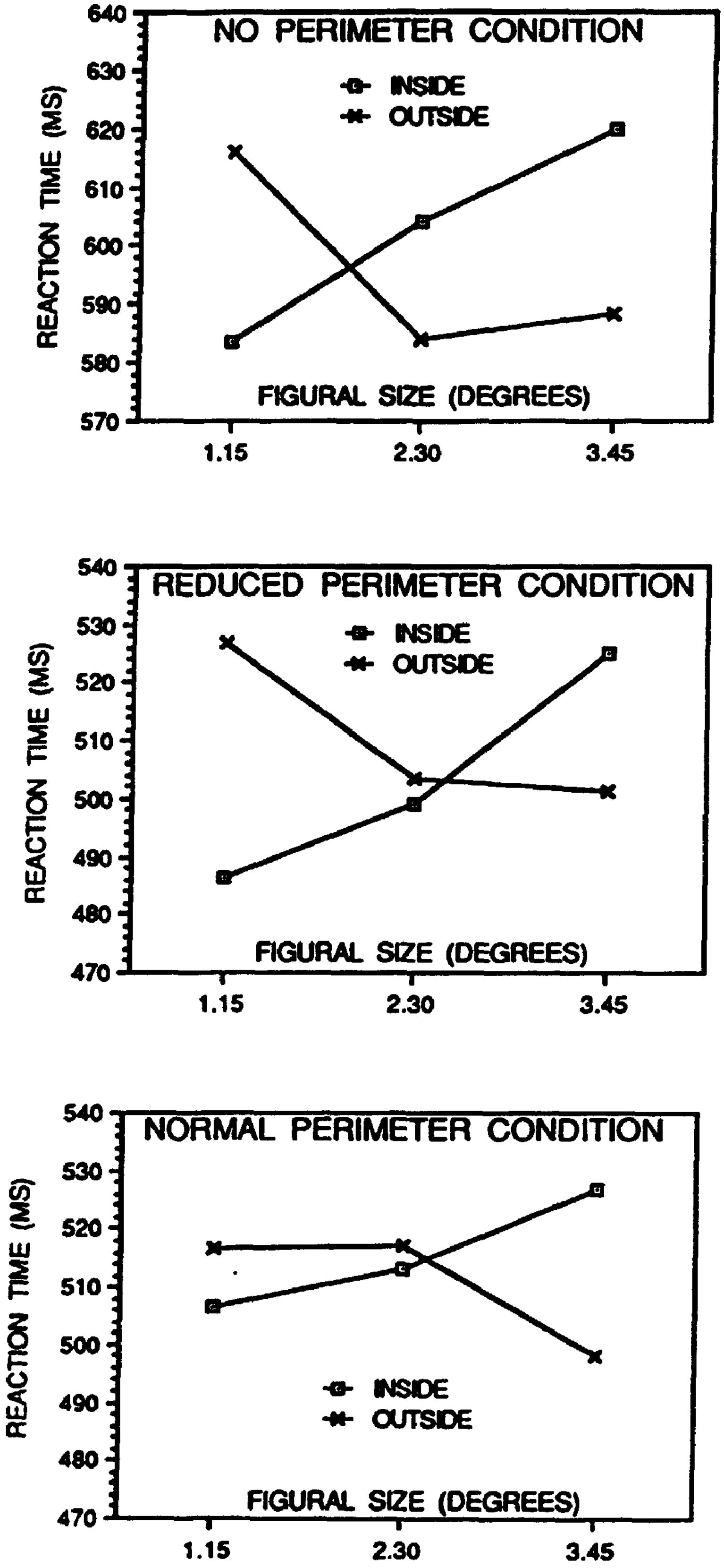


Figure 11. Mean response times for all subjects in Condition A (no perimeter), Condition B (reduced perimeter), and Condition C of Experiment 3 as a function of figural size and the nature of the screen perimeter.

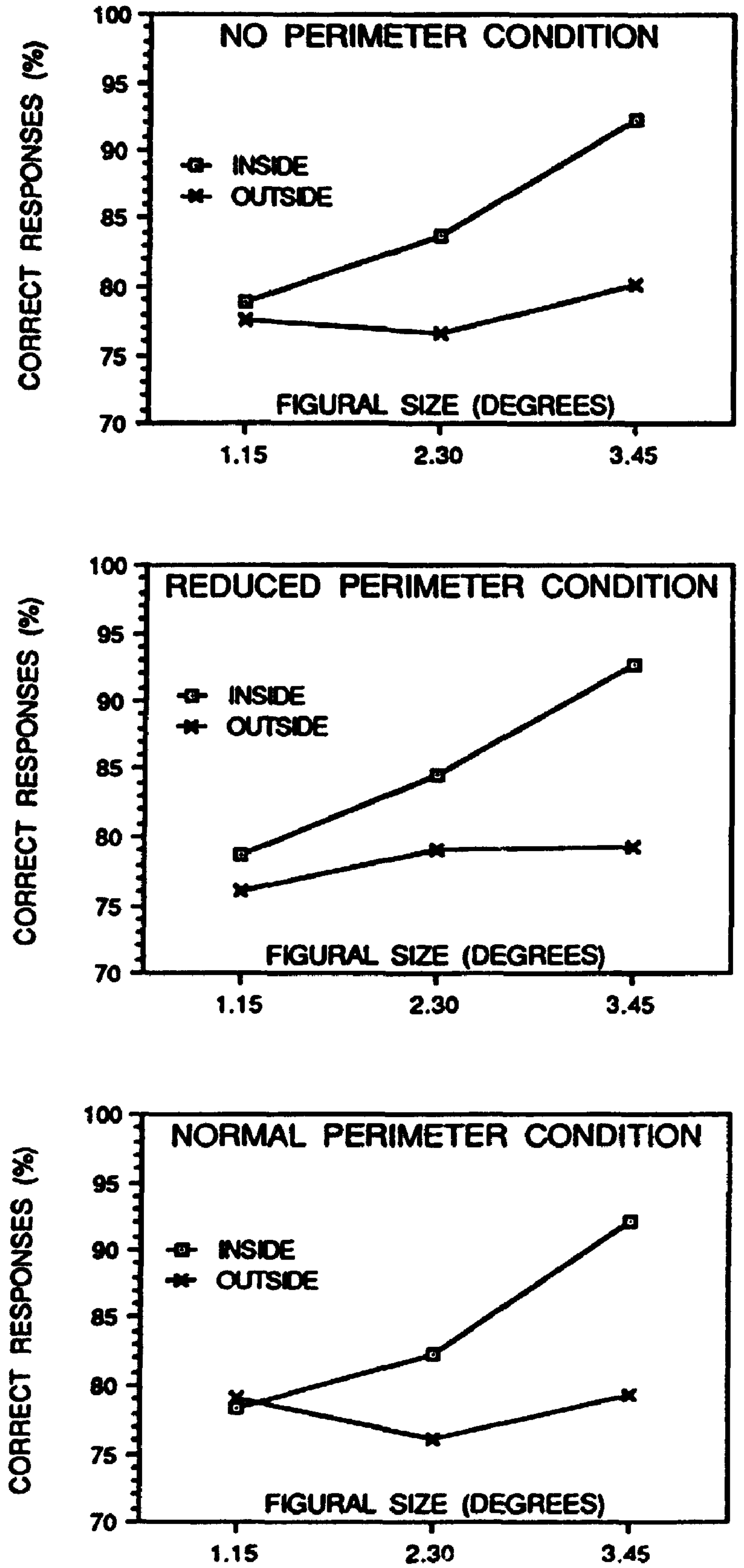


Figure 12. Mean response accuracy rates for all subjects in Condition A (no perimeter), Condition B (reduced perimeter), and Condition C (normal perimeter) of Experiment 3 as a function of figural size and the nature of the screen perimeter.

Separate two-way ANOVAs with repeated measures on both factors (*target inside/outside* X *figural size*) were carried out on the mean response times for Conditions A, B, and C. A main effect of figural size occurred in the Reduced Perimeter Condition ($F_{(2,28)} = 3.63, p < .04$) but not in the other conditions. However, mean *inside* response times increased monotonically with increases in figural size in all three conditions. This resulted in significant linear trends in Condition A ($F_{(1,28)} = 19.71, p < .01$), B ($F_{(1,28)} = 31.74, p < .01$), and C ($F_{(1,28)} = 4.32, p < .05$). The mean *outside* response times decreased with increases in figural size in two of the three conditions. That is, a significant trend with a linear ($F_{(1,28)} = 11.34, p < .01$) and quadratic ($F_{91,28)} = 6.67, p < .05$) component resulted in Condition A, and a significant linear trend ($F_{(1,28)} = 13.89, p < .01$) resulted in Condition B. Also, there was no effect of the target being inside vs. outside in any condition. However, there was a significant interaction between this factor and figural size in all conditions (i.e., $F_{(2,28)} = 17.35, p < .0001$ in Condition A; $F_{(2,28)} = 22.01, p < .0001$ in Condition B; and $F_{(2,28)} = 4.62, p < .02$ in Condition C).

Response times in Condition A (the No Perimeter Condition) were roughly 100 ms longer than those in the other conditions. This may have occurred because the stimulus intensity was reduced relative to that of the other conditions so that the inner walls of the viewing tube would not be illuminated, thereby making the screen perimeter detectable. However, the stimulus intensity was the same in Conditions B and C and the mean outside response times for each condition were not significantly different (510.53 ms in Condition B and 510.63 ms in Condition C).

Separate two-way ANOVAs of response accuracy with repeated measures on both factors were also carried out for Conditions A, B, and C. As in Experiment 1, a main effect of figural size occurred (i.e., $F_{(2,28)} = 37.57, p < .0001$ in Condition A; $F_{(2,28)} = 47.18, p < .0001$ in Condition B; and $F_{(2,28)} = 35.12, p < .0001$ in Condition C). Subjects' response accuracy was 86.18%, 80.21%, and 78.19% for trials with large, intermediate, and small figures respectively in Condition A; 85.79%, 81.74%, and

77.36% for trials with large, intermediate, small figures respectively in Condition B; and 85.69%, 79.10%, and 78.69% for trials with large, intermediate, and small figures respectively in Condition C. As seen in Figure 12, mean *inside* response accuracy rates increased monotonically with increases in figural size in all three conditions. This resulted in significant linear trends in Condition A ($F_{(1,28)} = 108.94, p < .01$), B ($F_{(1,28)} = 172.95, p < .01$), and C ($F_{(1,28)} = 202.92, p < .01$). In two of the three conditions, mean *outside* response accuracy rates also increased monotonically with increases in figural size. Significant linear trends were found in Condition A ($F_{(1,28)} = 4.29, p < .05$) and B ($F_{(1,28)} = 9.03, p < .01$). Increases in figural size resulted in a quadratic trend ($F_{(1,28)} = 14.54, p < .01$) in Condition C.

There was also a main effect of target inside/outside in each of the conditions (i.e., $F_{(1,14)} = 107.75, p < .0001$ in Condition A; $F_{(1,14)} = 122.47, p < .0001$ in Condition B; and $F_{(1,14)} = 54.87, p < .0001$ in Condition C). Subjects' response accuracy was 84.95% and 78.10% for inside and outside trials respectively in Condition A; 85.23% and 78.15% for inside and outside trials respectively in Condition B; and 84.21% and 78.10% for inside and outside trials respectively in Condition C. Furthermore, there was a Figural Size x Target Inside/Outside interaction in all conditions (i.e., $F_{(2,28)} = 17.53, p < .0001$ in Condition A; $F_{(2,28)} = 27.61, p < .0001$ in Condition B; and $F_{(2,28)} = 48.66, p < .0001$ in Condition C). Response accuracy decreased as figural size decreased, and the rate of decrease was greater for inside trials than for outside trials (see Figure 12). The pattern of response times that constituted this interaction effect was very similar to that of the first experiment.

The response time results of all three conditions were also analyzed in a single three-way ANOVA (*condition* X *target inside/outside* X *figural size*) with repeated measures on the second and third factors. There was a main effect of condition ($F_{(2,28)} = 6.74, p < .004$) in which the mean response times were 599.27, 506.97, and 513.02 for Conditions A, B, and C respectively. There was also a significant Target Inside/Outside

X Figural Size interaction ($F_{(2,84)} = 33.53, p < .0001$) as was the case in each of the conditions when analyzed separately. As is evident in Figure 13, mean inside response times increased monotonically with size increases, resulting in a significant linear trend ($F_{(1,84)} = 43.38, p < .01$). Conversely, mean outside response times *decreased* monotonically with size increases, which also resulted in a significant linear trend ($F_{(1,84)} = 24.62, p < .01$).

A similar three-way ANOVA was carried out with the mean response accuracy rates of the three conditions. No effect of condition was found, but there was a significant effect of target inside/outside ($F_{(2,42)} = 269.36, p < .0001$). Subjects responded with 84.79% accuracy on inside trials and 78.13% accuracy on outside trials. In addition, there was a significant effect of figural size ($F_{(1,84)} = 107.58, p < .0001$) and a significant Target Inside/Outside X Figural Size interaction ($F_{(2,84)} = 82.29, p < .0001$). As seen in Figure 13, mean inside and outside response accuracy rates increased with increases in figural size. This resulted in a significant trend for inside responses with linear ($F_{(1,84)} = 453.10, p < .01$) and quadratic ($F_{(1,84)} = 12.59, p < .01$) components. This was also the case for outside responses (i.e., ($F_{(1,84)} = 10.47, p < .01$) for the linear component and ($F_{(1,84)} = 5.54, p < .05$) for the quadratic component).

2.33 Discussion

The results of this study replicated those of the previous experiments. Increases in figural size led to increased inside response times and decreased outside response times. As in the first experiment, increases in figural size also led to increases in the accuracy of inside responses. In addition, the presence of the screen perimeter had no effect on the pattern of outside responses times. Even when the screen perimeter was not visible, outside response times decreased with increases in figural size. Furthermore, a clear reduction in screen perimeter had no effect on the pattern of outside response times relative to that of a condition with a normal screen perimeter. Thus, the effect of changes in figural size on inside/outside response times cannot be attributed to the perimeter of the CRT.

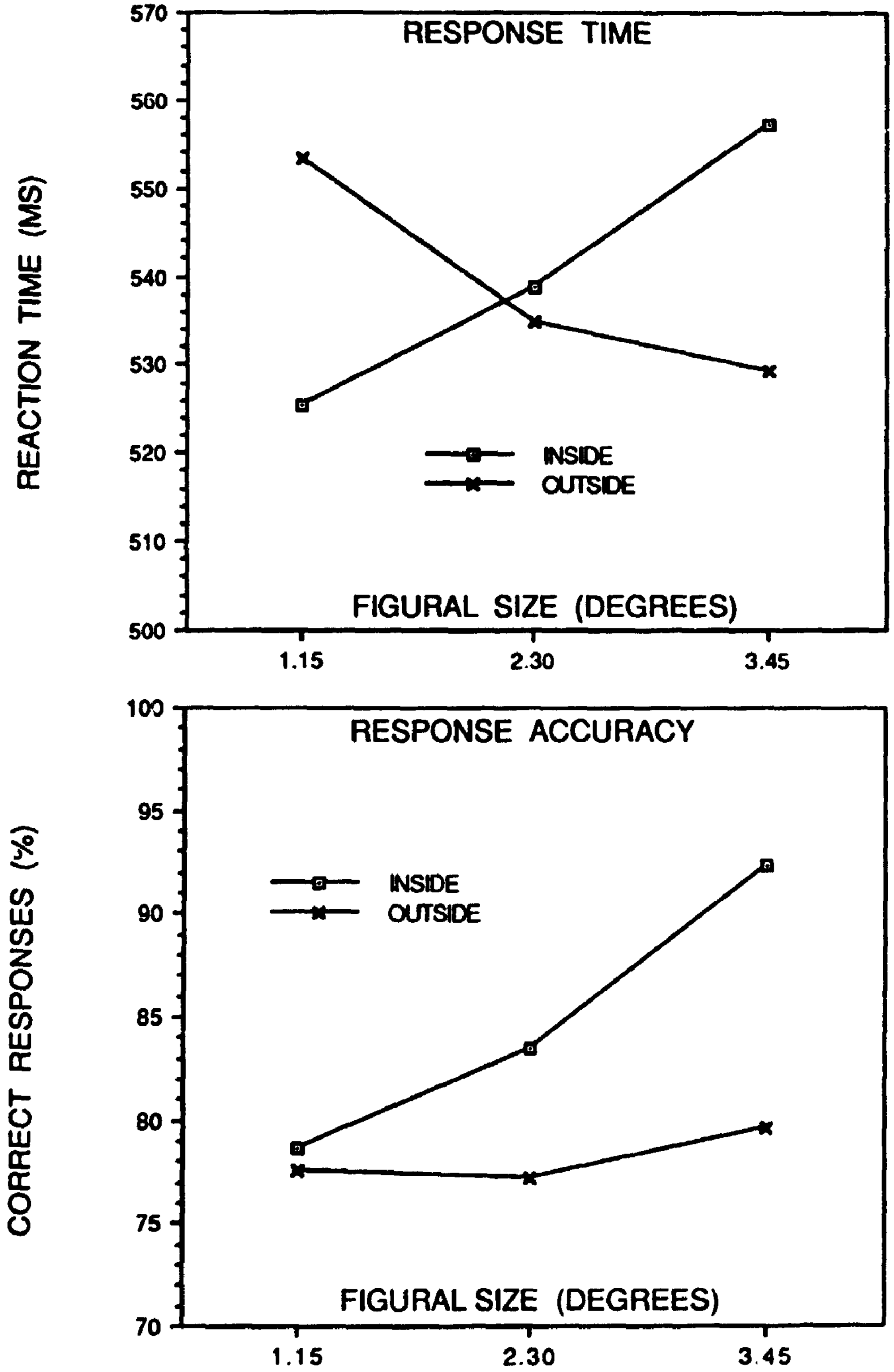


Figure 13. Mean response times and accuracy rates collapsed across the three perimeter conditions of Experiment 3 as a function of figural size.

The pattern of outside response times produced by size changes was very robust, having occurred in all conditions in the first three experiments. However, it is difficult to account for this effect in terms of region colouring. If colouring begins at the target location and flows toward spatial infinity, changes in the size of the region inside the figure should not affect the time required to carry out this activation. Moreover, it is unclear why decreases in the size of the figure's internal region should increase the time required to colour the region between the target and spatial infinity. Therefore, region colouring does not appear to be an appropriate account of how "outside of" is determined.

2.4 GENERAL DISCUSSION

The experiments described in this chapter involved manipulations of target location and the size of the bounding figure in inside/outside tasks. It was found that inside response times increased as a function of figural size. This suggests that if region colouring was involved in this perception, colouring models should produce a solution in a time that increases as a function of figural size. Therefore, the performance of models with uniscale algorithms is more consistent with this finding than the performance of multiscale algorithms. As seen in Figure 2, uniscale colouring time increases as a function of region size.

Unlike the effect of figural size changes, the effect of target location changes on inside response times did not indicate that colouring was involved in these studies. In particular, the pattern of inside response times in Experiment 2 indicated that judgements were slower as the distance between the target and the fixation point was increased. This suggests that increases in this distance caused a corresponding increase in the time required to shift the processing focus from the fixation point to the target. Thus, the inside response time differences produced by changes in target location may be due to processing focus shifts.

The effects of changes in target location and figural size on outside response times are more difficult to account for with a region colouring algorithm. It may be the case that increases in the distance between the target and the figure decrease the distance between the target and spatial infinity. However, the notion of spatial infinity is abstract. Although it plays a critical role as a stopping criterion for determining "outside of" with colouring algorithms, its location cannot be defined concisely. Moreover, it is not clear from the data that is available why region colouring models should require less time to determine "outside of" as the size of the figure's internal region is increased. In conclusion, the performance of uniscale colouring algorithms are consistent with the pattern of inside response times in these studies but the involvement of region colouring in the perception of outside relations is questionable.

The occurrence of inside response time increases with size increases raises questions about how large an effect on human performance must be before those modeling that performance feel that it should be incorporated into the algorithm. In the four conditions in Experiments 1 and 3, size (based on the length of one side of the square figure) increased from 1.15 to 3.45°. Inside response times increased by an average of 34.18 ms with this increases, or 14.86 ms/degree (see e.g., Figure 11). The effect was statistically reliable and robust as indicated in Experiment 2 in which it was obtained with displays of a different type. For this reason, it is my contention that this effect should be taken into account by those attempting to develop strongly equivalent models of the processes involved. This may include, for example, the development of uniscale colouring algorithms that depend to some extent on figural size.

In addition to region colouring, there may be other explanations of the effect of size changes on inside response times. For example, perhaps inside response times were affected by the time required to adjust the attentional frame. As mentioned in the previous chapter, Ward (1982) has proposed that attention can be thought of as a frame (similar to an attentional spotlight) with parameters such as locus and degree of focus that could require adjustments with each new fixation. According to Ward, the parameters of the frame stay in their current state until adjustments are required in order to conserve processing resources. Presumably, the focus of the attentional frame in the current studies was at a small resolution (0.57 x 0.57°) when the fixation cross was displayed. When inside responses were made, the magnitude of adjustment of this focal resolution would be greater for larger figures than for smaller ones. Hence, the prediction that follows from Ward's proposal is that inside responses should require more time with larger figures than with smaller ones because more extensive adjustments of focal resolution would be required at the time of the onset of the stimulus display. Note that this prediction is based on the assumption that the size of the attentional frame is changed to a relatively small focal resolution to process the fixation point or target at the beginning of each trial. However, it may be the case that the size of the attentional frame remains constant throughout a given block of trials (e.g., frame size is set to the average size of the figures in that block). Therefore, the role that attentional frame adjustments play in determining inside relations raises several empirical questions (e.g., does target size affect inside response times as well?).

The attentional frame proposal still cannot account for the pattern of outside response times produced by changes in figural size. Moreover, the results of previous research (e.g., Eriksen & Yeh, 1985) indicate that increases in the size of the attentional frame or spotlight lead to decreases in accuracy in the performance of some visual tasks. In Experiments 1 and 3, response accuracy increases as a function of figural size. Therefore, the operations may be more complex than the adjustment of an attentional frame.

CHAPTER THREE: PERCEIVING INSIDE/OUTSIDE RELATIONS BETWEEN ASYNCHRONOUSLY PRESENTED STIMULI

One important aspect of Ullman's (1984) visual routines hypothesis is the notion of incremental representations. These are created after low-level vision produces the base representations. The content of incremental representations is thought to be modified in the course of applying visual routines. These representations were described in the first chapter as a series of data structures of increasing elaborateness that contain information from preceding representations plus the results of the visual routines that were carried out to produce them. Therefore, an incremental representation contains more information about objects and their relations in the visual field than a base representation. For this reason, the determination of spatial relations between stimuli should be faster in many cases if it is based on an incremental representation as opposed to a base representation. Ullman (1984, p. 113) made the following proposal about how incremental representations may facilitate inside/outside response times. In a hypothetical experiment an observer is asked to respond to a target in an inside/outside display that contains a fairly complex bounding figure. Immediately after a response is made, a second target is presented and another inside/outside judgement must be made. Ullman proposed that the second judgement would be faster than the first because the results of previous computations (e.g., region colouring) would already be summarized in an incremental representation of the figure in question.

In general, if two stimuli are presented in succession, then the first stimulus will probably undergo some degree of processing before the onset of the second stimulus. The results of this preliminary processing of the first stimulus may be stored in an incremental representation of the visual field. If this is the case, then determining inside/outside relations at the time of onset of the second stimulus should involve operations that are carried out on an incremental representation containing some information about the first stimulus. If this is the case, then the time required to make an inside/outside judgement should be less than that required in a condition with simultaneous stimulus onsets.

Asynchronous stimulus presentation may also enable researchers to examine the operations or subprocesses involved in the determination of inside/outside relations. In particular, if part of the stimulus display is presented initially, perhaps one of the operations involved (e.g., region colouring, shifting of processing focus, etc.) could begin. Then when the rest of the display is presented, the response time may be reduced by the amount of time that the operation required. For example, if a bounding figure is presented before a target, the internal region may be coloured before the target's onset. When the target is presented, the inside/outside response time may be reduced by the amount of time that the colouring operation required because this has already been carried out. Moreover, examining the facilitation produced by varying the stimulus onset asynchrony (SOA) may provide some insights into the time course of the operations involved in this perception. That is, some operations may not require much processing time (e.g., shift of processing focus) and the facilitation produced by increasing SOA may saturate at some point. On the other hand, other operations (e.g., colouring a large region with a uniscale algorithm) may require more processing time. In this case, the point at which increases in SOA no longer produce corresponding increases in response-time facilitation may not occur until much longer SOAs are involved.

In the experiments described in this chapter, the relative onset times of asynchronously presented stimuli were manipulated. In some cases the target was presented before the bounding figure, and in some cases the figure (or some part of it) was presented before the target. It was expected that there would be a response-time savings in the asynchronous conditions relative to a condition with simultaneous onsets because spatial relations should be perceived faster if some of the operations involved in determining the relation have already been carried out.

3.1 EXPERIMENT 4

The task in this study was identical to that used in the previous experiments. However, in the current studies the stimuli were sometimes presented in temporal succession. On some trials the target was presented before the bounding figure and on other trials the figure was presented before the target. It was expected that if some of the subprocesses involved in determining inside/outside relations could begin with the onset of the first stimulus, then response time facilitation would occur. More specifically, if one of the subprocesses is region colouring, then the presentation of the figure before the target was expected to produce greater facilitation than the presentation of the target before the figure. In particular, the internal region of the figure can be coloured before the target is presented in the former case, but when the target is presented first, colouring must still be carried out when the figure is presented. In addition, it was expected that as SOA was increased, response-time facilitation would also increase. To elaborate, increased SOAs may allow the preliminary processing to be carried out to a greater extent before the onset of the rest of the display.

3.11 Method

Subjects. Thirty University of Western Ontario undergraduate students, 15 in each of two conditions, participated in the experiment as part of a course requirement. All subjects had normal or corrected-to-normal vision and none had any previous experience with this type of study. In addition, all subjects were unaware of the purpose of the study.

Stimuli and Procedure. The same apparatus, testing conditions, and instructions used in Experiment 1 were employed in this study. The stimulus display consisted of a white X that subtended 0.57 x 0.57° of visual angle and four white lines forming a rectangle that subtended 2.86 x 5.73°. Positional uncertainty of the rectangular figure across trials was achieved by presenting the figure horizontally in either the top half or bottom half of a central region of the display screen, or vertically in either the left or the right half of this region. Figural position was determined randomly across trials and no location within the central display region was always *inside* or always *outside*. In addition, there were four positions of the X relative to that of the rectangular figure (see Figure 14).

In Condition A, the target and the figure were presented either simultaneously or in succession with a 100, 300, 500, or 700 ms delay between their presentation onsets. On half of the asynchronous trials the target was presented before the figure, while on the other half the figure was presented before the target. Condition B was identical with the exception that on asynchronous trials the delay was either 50, 100, 150, or 200 ms. This condition was originally a separate experiment conducted after Condition A to determine whether effects of the manipulation would be robust across different ranges of delay durations. The temporal order of stimulus presentation is shown in Figure 15. There were 16 trials for each of the 18 conditions for each subject (4 SOAs X target/figure first X inside/outside AND a simultaneous onset condition X inside/outside). The order of trials was randomized. Also, the procedure was the same as that of the previous experiments. Each subject was given a 20 trial practise run and a 288 trial data run. The data run consisted of eight blocks of 36 trials with a 45 second rest period between each block.

3.12 Results

As in the previous experiments, only correct responses within two standard deviations of the mean correct response time for a particular condition were used in these analyses. This led to the removal of 4.33% of the correct responses in Condition A and 3.51% of the correct responses in Condition B. The mean correct response times for the two conditions are shown in Figure 16 and 17, and the mean response accuracy rates are shown in Figure 18 and 19. After incorrect responses and outliers were removed, 90.07% of the Condition A responses remained and 90.03% of the Condition B responses remained.

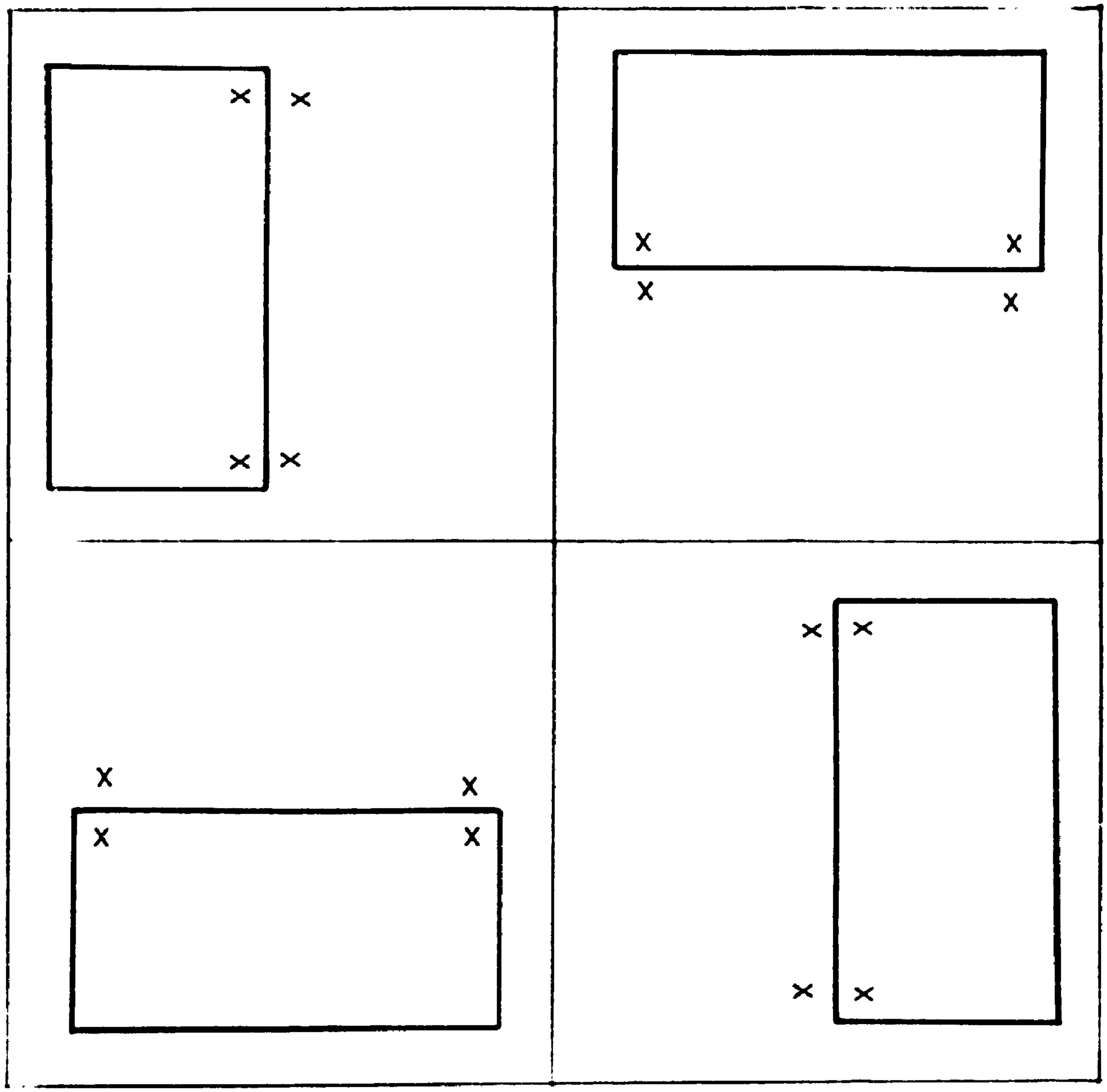


Figure 14 Positions of the target relative to the rectangular figure in each of the experiments described in this chapter.

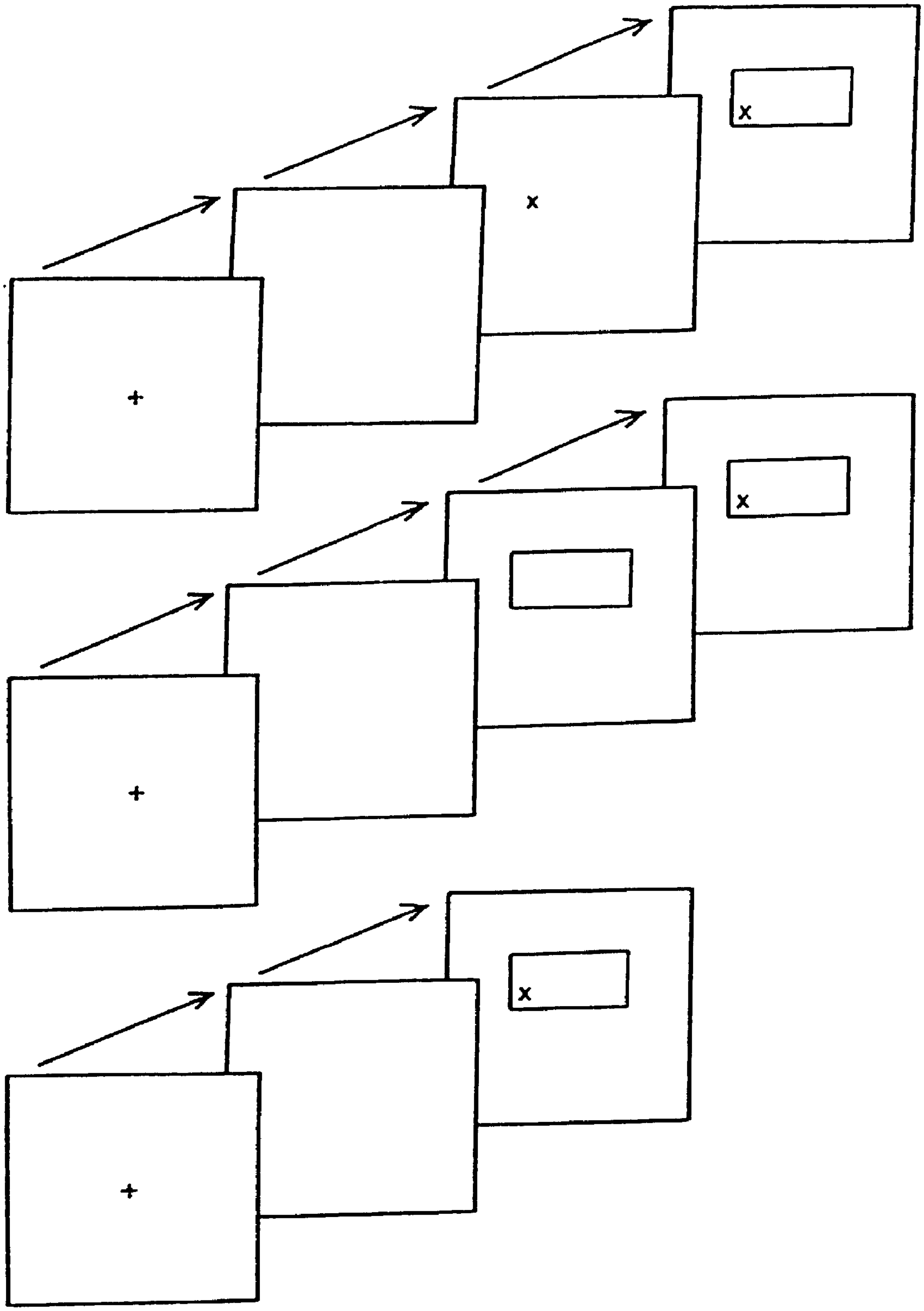


Figure 15. The temporal order of stimulus presentation in the target-first, figure-first, and simultaneous presentation conditions in Experiment 4.

Two separate three-way ANOVAs (one for Condition A and one for Condition B) were carried out with the mean correct response times. The factors were target inside/outside, target/figure first, and SOA (e.g., 100, 300, 500, or 700 ms). A main effect of target/figure first occurred in Condition A ($F_{(1,14)} = 51.91, p < .001$) and in Condition B ($F_{(1,14)} = 17.56, p < .001$). In particular, response times were significantly faster if the figure was presented first as opposed to the target (503.1 vs. 526.5 ms in Condition A and 490.4 vs. 511.2 ms in Condition B). Also, the mean outside response time in Condition B (492.70 ms) was significantly faster than the mean inside response time (509.00 ms), $F_{(1,14)} = 5.70, p < .03$. The effect of SOA on response times resulted in a significant trend with linear ($F_{(1,42)} = 55.73, p < .01$) and quadratic ($F_{(1,42)} = 10.15$, $p < .05$) components in Condition A, and a significant linear trend ($F_{(1,42)} = 21.92, p < .01$) in Condition B. In Condition A the mean response times were 548.93, 511.85, 501.83, and 496.62 ms for displays with SOAs of 100, 300, 500, and 700 ms respectively. The mean response time for displays with a 0 ms SOA was 594.32 ms. This mean was not included in the analysis because it was used as a comparison for both the figure first and target first conditions. Therefore, it could not be included in an ANOVA with a balanced design. On target-first trials in Condition A, significant linear trends resulted from inside ($F_{(1,42)} = 6.43, p < .05$) and outside ($F_{(1,42)} = 4.29, p < .05$) response times. On figure-first trials in this condition, a significant linear trend also resulted from outside response times ($F_{(1,42)} = 12.78, p < .01$). In addition, a significant trend with linear ($F_{(1,42)} = 14.76, p < .01$) and quadratic ($F_{(1,42)} = 16.99, p < .01$) components resulted from inside response times. In Condition B the mean response times were 520.86, 498.14, 496.41, and 487.91 ms for displays with SOAs of 50, 100, 150, and 200 ms respectively. The mean response time for displays with a 0 ms SOA was 548.27 ms. It was not included in the analysis for the reasons given above. On figure-first trials in Condition B, a significant linear trend resulted from inside response times ($F_{(1,42)} = 9.62, p < .01$). In addition, a significant trend with linear ($F_{(1,42)} =$

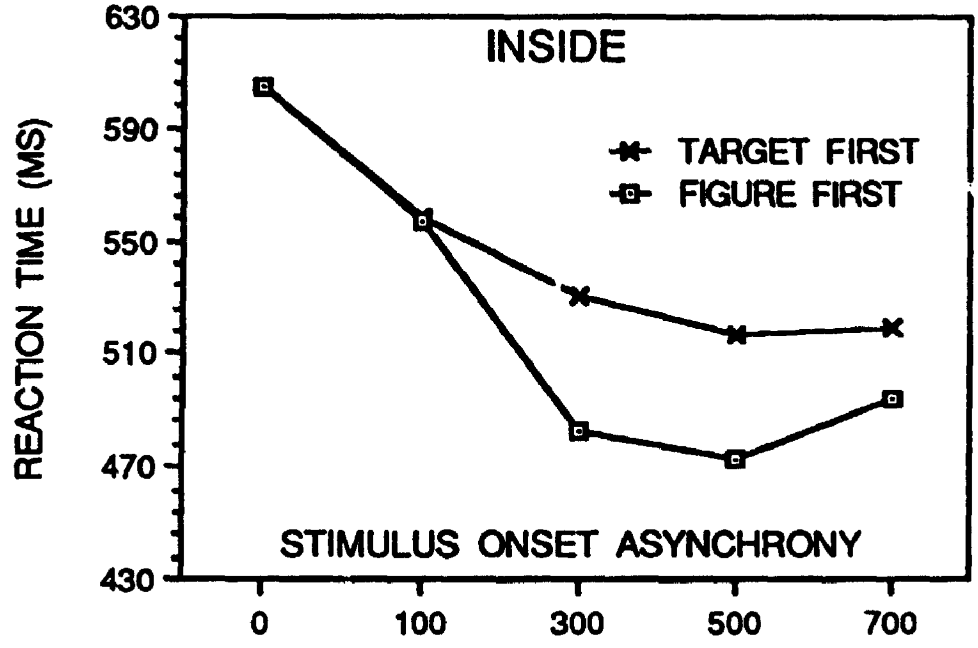
11.07, $p < .01$) and cubic ($F_{(1,42)} = 8.79, p < .01$) components resulted from the pattern of outside response times.

Seperate three-way ANOVAs with repeated measures on all factors were also carried out with the mean response accuracy rates. In Condition A subjects responded with 92.81% accuracy for inside responses and 86.72% accuracy for outside responses ($F_{(1,14)} = 101.08, p < .0001$). In Condition B subjects responded with 91.88% accuracy for inside responses and 86.88% accuracy for outside responses ($F_{(1,14)} = 28.15, p < .0001$). Also, subjects were more accurate when the target was presented first as opposed to the figure in Condition A (92.69% vs. 86.81%) and in Condition B (92.69% vs. 86.13%); $F_{(1,14)} = 166.14, p < .0001$ in Condition A; and $F_{(1,14)} = 121.86, p < .0001$ in Condition B.

As seen in Figure 18 and 19, there was an interaction between target inside/outside and target/figure first for response accuracy in both conditions ($F_{(1,14)} = 119.84, p < .0001$ in Condition A and $F_{(1,14)} = 61.86, p < .0001$ in Condition B). Moreover, on figure-first trials in Condition B, a significant cubic trend ($F_{(1,42)} = 8.02, p < .01$) resulted from inside response accuracy rates and a significant quadratic trend ($F_{(1,42)} = 8.87, p < .01$) resulted from outside response accuracy rates. Thus, subjects were faster at making outside responses than inside responses on trials in which the figure was presented first, but the accuracy of outside responses was significantly lower on these trials.

3.13 Discussion

In both conditions the response-time savings were greater on asynchronous trials in which the figure was presented first rather than the target. In addition, increases in SOA decreased response times on all asynchronous trials. These findings indicate that if the first stimulus undergoes some preliminary processing before the onset of the rest of the display, responses will be faster than when stimuli are presented simultaneously. Also, the effect of SOA of response-time facilitation suggests that if preliminary processing of



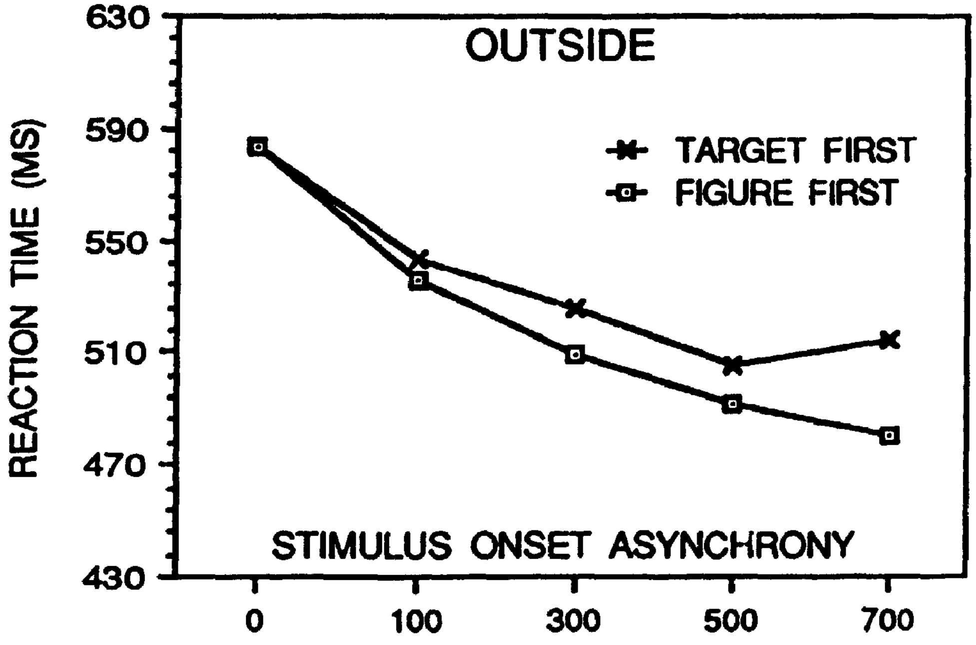
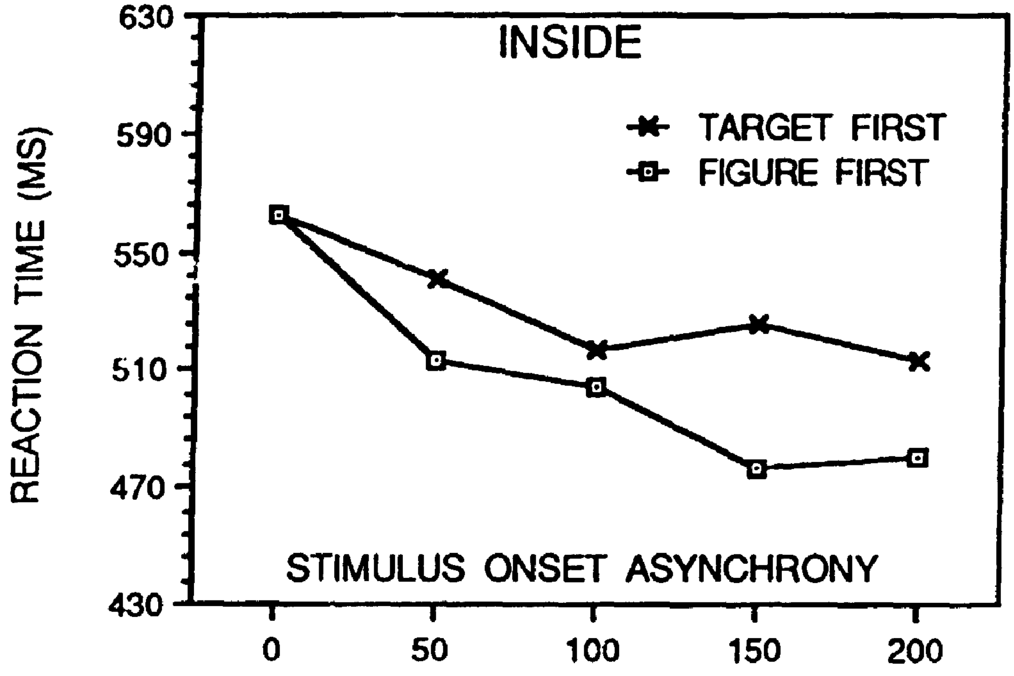


Figure 16. Mean response times for all subjects in Condition A of Experiment 4 as a function of SOA.



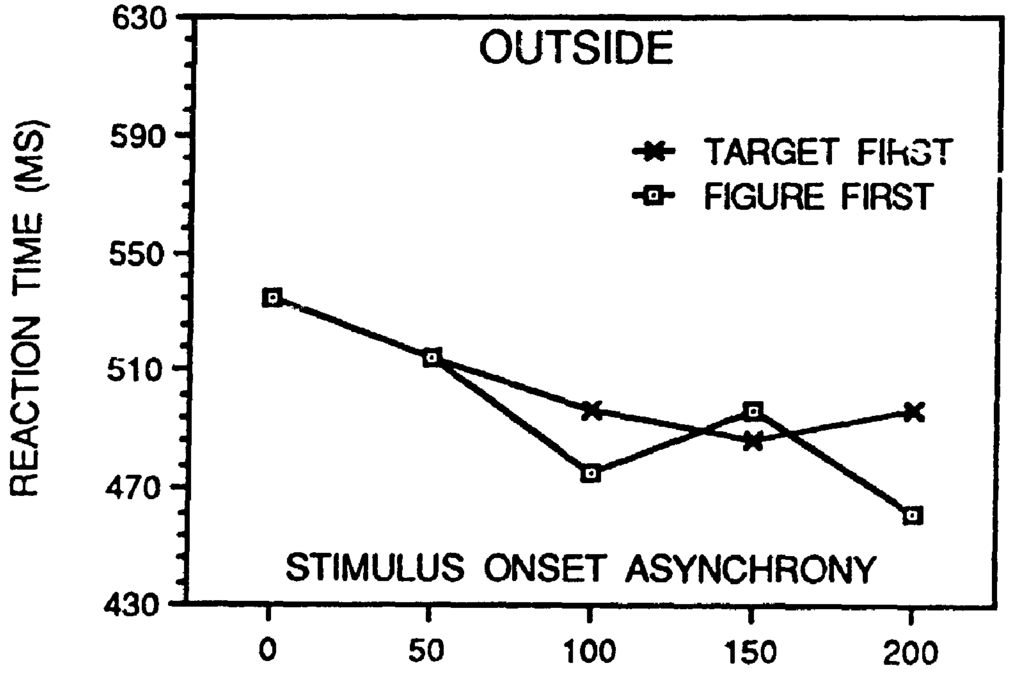


Figure 17. Mean response times for all subjects in Condition B of Experiment 4 as a function of SOA.

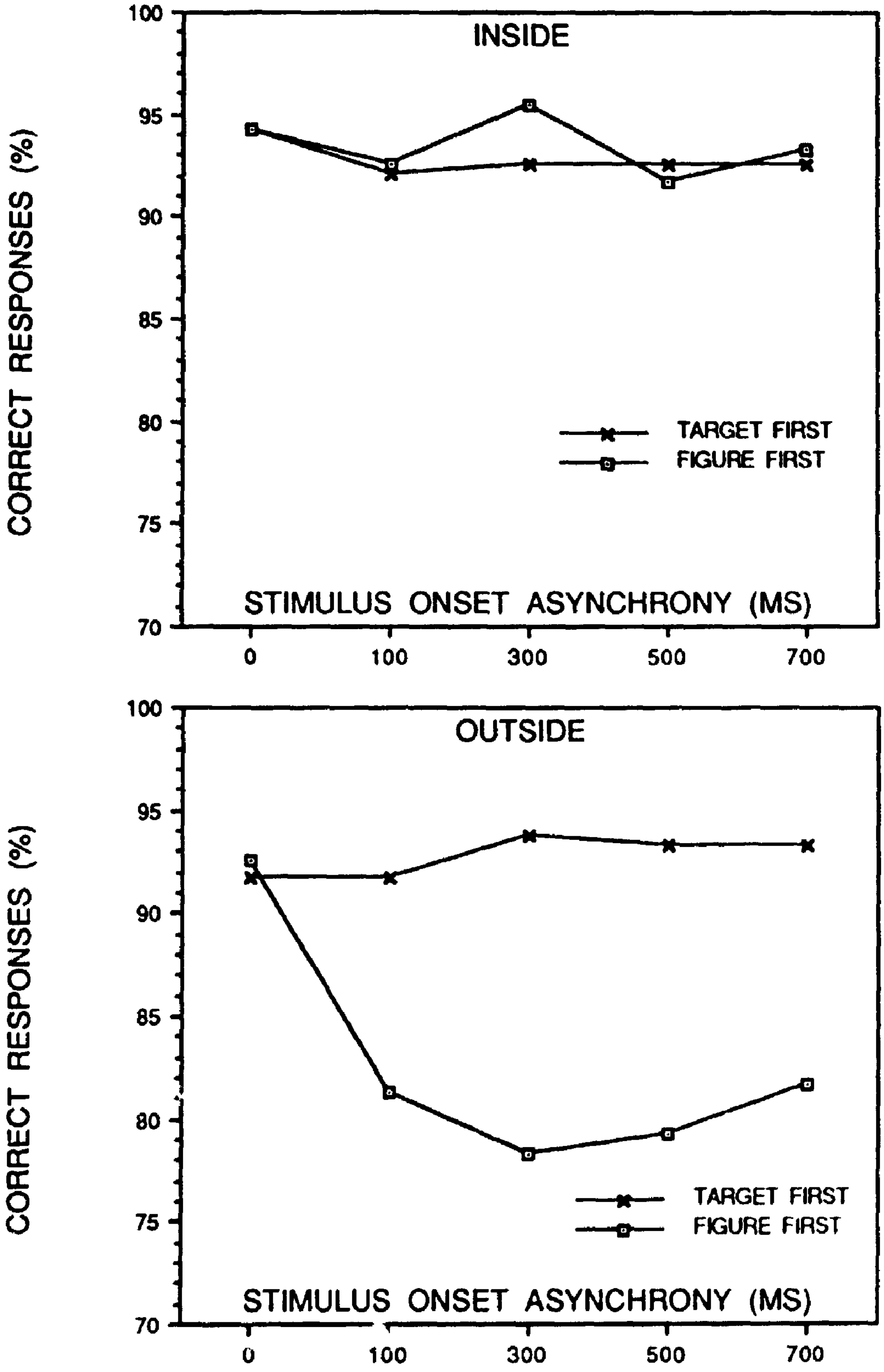


Figure 18. Mean response accuracy rates for all subjects in Condition A of Experiment 4 as a function of SOA.

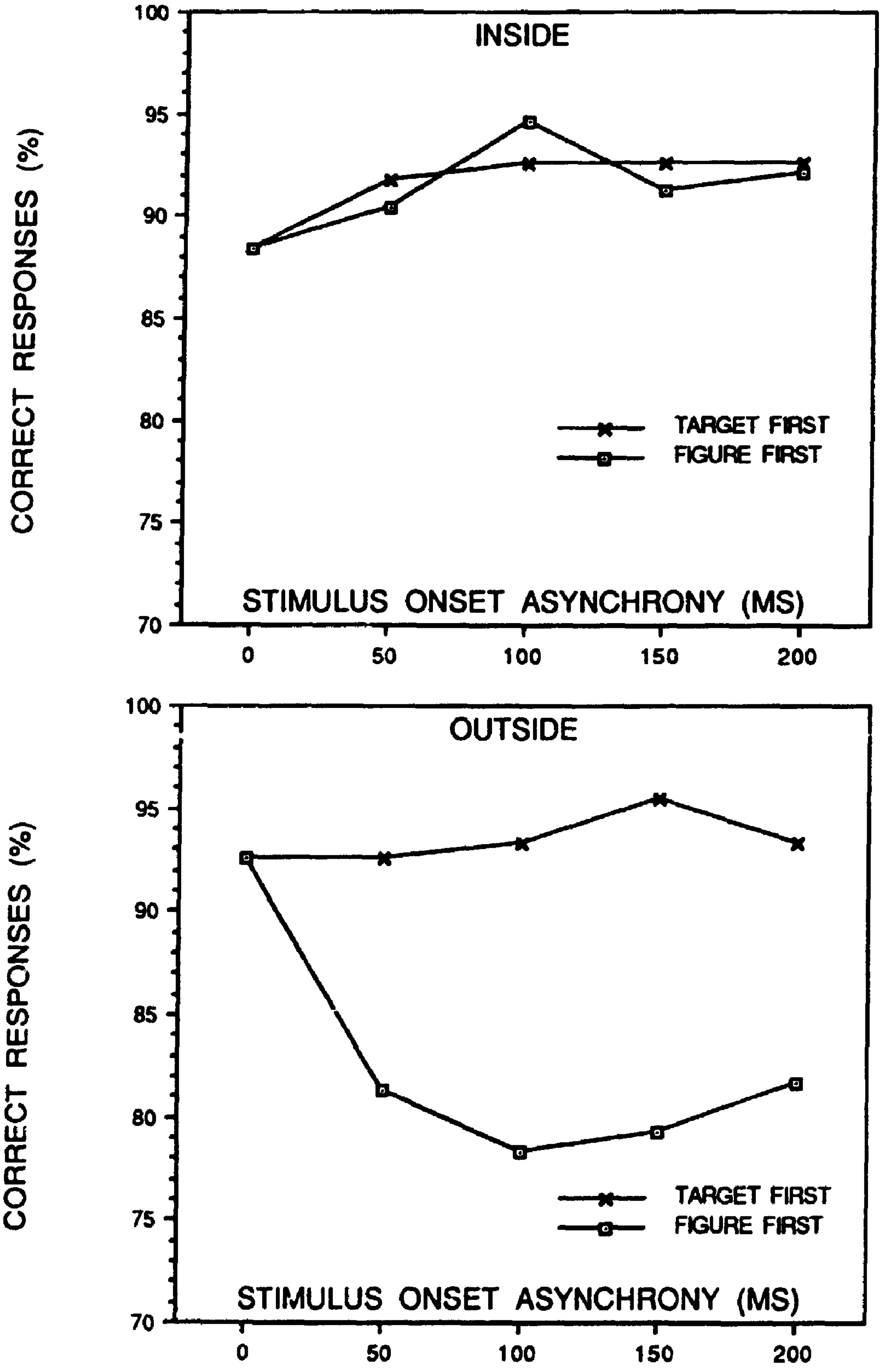


Figure 19. Mean response accuracy rates for all subjects in Condition B of Experiment 4 as a function of SOA.

the first stimulus can be carried out over a longer period of time before the onset of the rest of the display, less processing is required once the display is present. One account of the greater response-time facilitation on figure-first trials rather than target-first trials is that the figure has a clearly defined internal region. Thus, on target-first trials colouring cannot begin before the onset of the figure. Hence, if the routine to determine inside/outside relations consists primarily of a colouring operation, more of the routine can be carried out on the first stimulus before the onset of the second stimulus on figure-first trials.

One aspect of asynchronous presentation that must be addressed is spatial precuing of the display's location with the onset of the first stimulus. Because the location of the display varied randomly across trials, some of the response-time facilitation produced by asynchronous presentation has to be the result of spatial precuing. However, response-time facilitation cannot be entirely accounted for by precuing. In particular, it is not clear why facilitation is greater on figure-first trials than on target-first trials if spatial precuing is the sole cause of this response-time savings. Moreover, it is not clear why facilitation should increase as SOA is increased to 500 ms on target-first trials. To elaborate, spatial precuing is thought to be mediated by the shift of an attentional spotlight (e.g., Posner et al., 1980). As mentioned in the first chapter, estimates of the speed of movement of this spotlight range from 30-177 degrees/second. Therefore, based on the most conservative estimate of this speed, the time required to shift processing focus 2.8° from the fixation point to the target would be 94 ms. If response-time facilitation is caused by shifting the processing focus to the precued location, no additional facilitation should result from increasing the SOA past 94 ms. In the current studies, the SOA effect did not saturate on target-first trials until 500 ms. Thus, some figural processing of the target also must have been carried out.

Response accuracy rates were also examined and it was found that on target-first trials, subjects responded with approximately 93% accuracy. This was also the case on figure-first trials when the target was inside. When the target was outside, however, response accuracy was approximately 80%. I have no explanation to offer for this interaction beyond a suggestion that it may be easier for the visual system to erroneously shift the processing focus to the wrong target location on figure-first trials when the target is outside than it is in other types of trials.

3.2 EXPERIMENT 5

In the previous experiment, response-time facilitation was greater on figure-first trials than on target-first trials. This finding raises questions about whether the facilitation generalizes to a case in which part of the figure is presented before the rest of the figure and the target. That is, will response-time facilitation occur to the same extent with the initial presentation of a figure part as opposed to the whole figure? The current study was conducted to test this question. The stimulus displays were identical to those of the previous experiment with the exception that the target-first condition was replaced by a *partial-figure* condition. On these trials, one of the long sides of the figure (i.e., the side bisecting the central square region of the display screen) was presented before the rest of the figure and the target. The figure-first condition of the previous experiment was retained and referred to as the *whole-figure* condition. It was expected that response-time facilitation would be greater in the whole-figure condition than in the partial-figure condition. To elaborate, colouring could be carried out in the former condition with the onset of the first stimulus. However, in the partial-figure condition, colouring could not begin until a clearly defined internal region was present. Otherwise, colouring would have to spread in both directions from the figure part, and may spread to spatial infinity before the rest of the display is presented (see Figure 20). Therefore, greater facilitation was expected in the whole-figure condition than in the partial-figure condition.

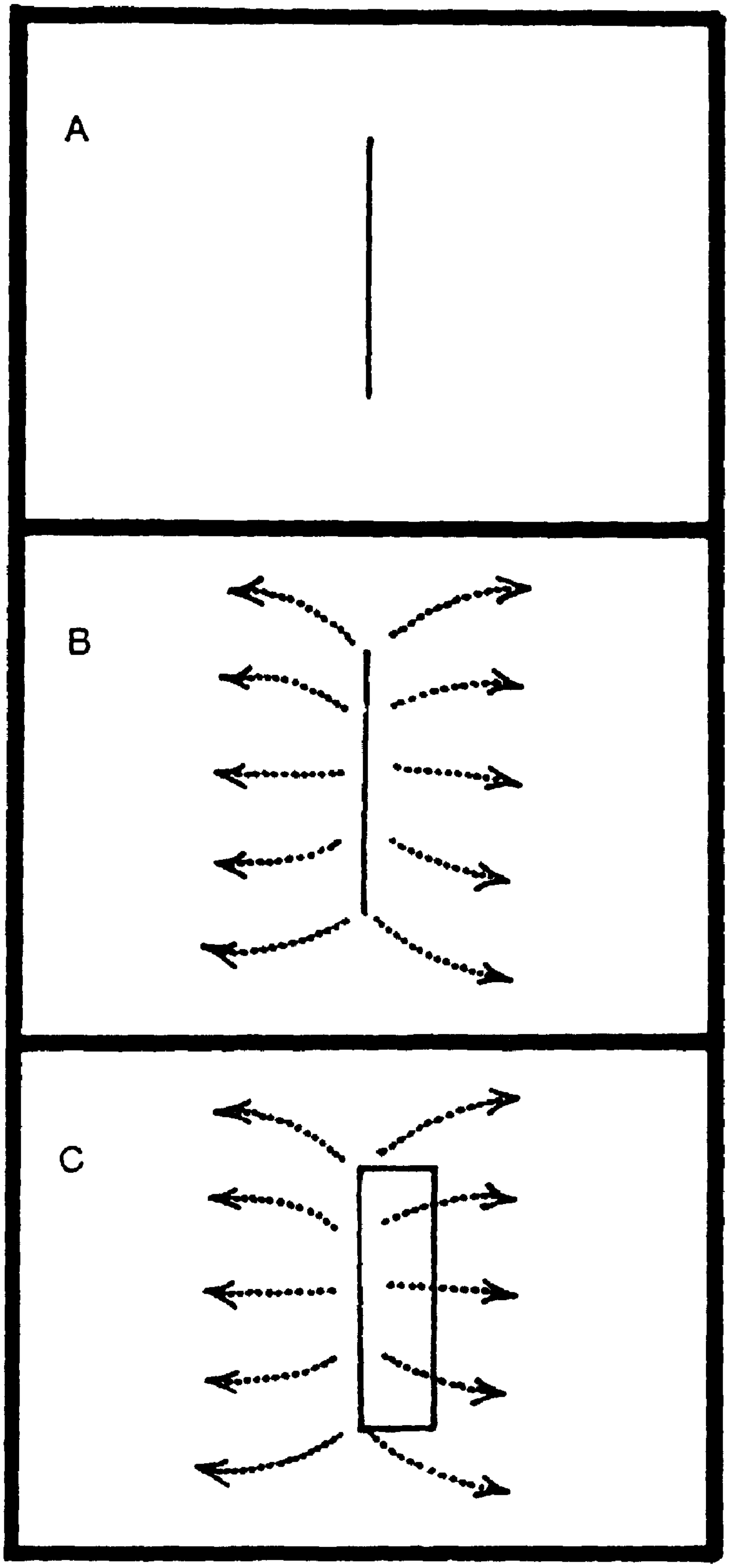


Figure 20. The problem with colouring without a clearly defined internal region. The figure part presented in (A) does not indicate the direction the colouring should start in. In (B) the colouring begins in both directions. In (C) the colouring spreads to spatial infinity before the rest of the figure is presented.

3.21 Method

***Subjects*..** Fifteen University of Western Ontario undergraduate students participated in this experiment as part of a course requirement. All subjects had normal or corrected-to-normal vision and none had any previous experience with this type of study. In addition, all subjects were unaware of the purpose of the study.

***Stimuli and Procedure*.** The same apparatus, testing conditions, instructions, and procedure used in the previous experiment were employed in this study. The stimulus display was also the same as that of the previous experiment, but on some trials it was preceeded by the presentation of a horizontal or vertical white line 5.73° in length (the partial-figure condition). This line constituted one side of the complete rectangular figure shown in the second frame of view, and on each trial it bisected the central square region of the display screen. The locations of the figure and positions of the target relative to that of the figure across trials were identical to those in Experiment 4. In addition, stimuli were presented either simultaneously or with a 100, 200, 300, or 400 ms delay between their presentation onsets. On half of the asynchronous trials the line was presented before the rest of the figure and the target, while on the other half the whole figure was presented before the target. The temporal order of stimulus presentation for each condition is shown in Figure 21. There were 16 trials for each of the 18 conditions for each subject (4 SOAs X partial/whole figure X inside/outside AND a simultaneous onset condition X inside/outside). The order of trials was randomized. In addition, each subject was given a 20 trial practise run and a 288 trial data run. The data run consisted of eight blocks of 36 trials with a 45 second rest period between each block.

3.22 Results

Only correct responses within two standard deviations of the mean correct response time for each condition were used in the analyses. This led to the removal of 4.32% of the correct responses. The mean correct response times are shown in Figure 22, and the mean accuracy rates are shown in Figure 23. After incorrect responses and outliers were removed, 86.18% of the responses remained.

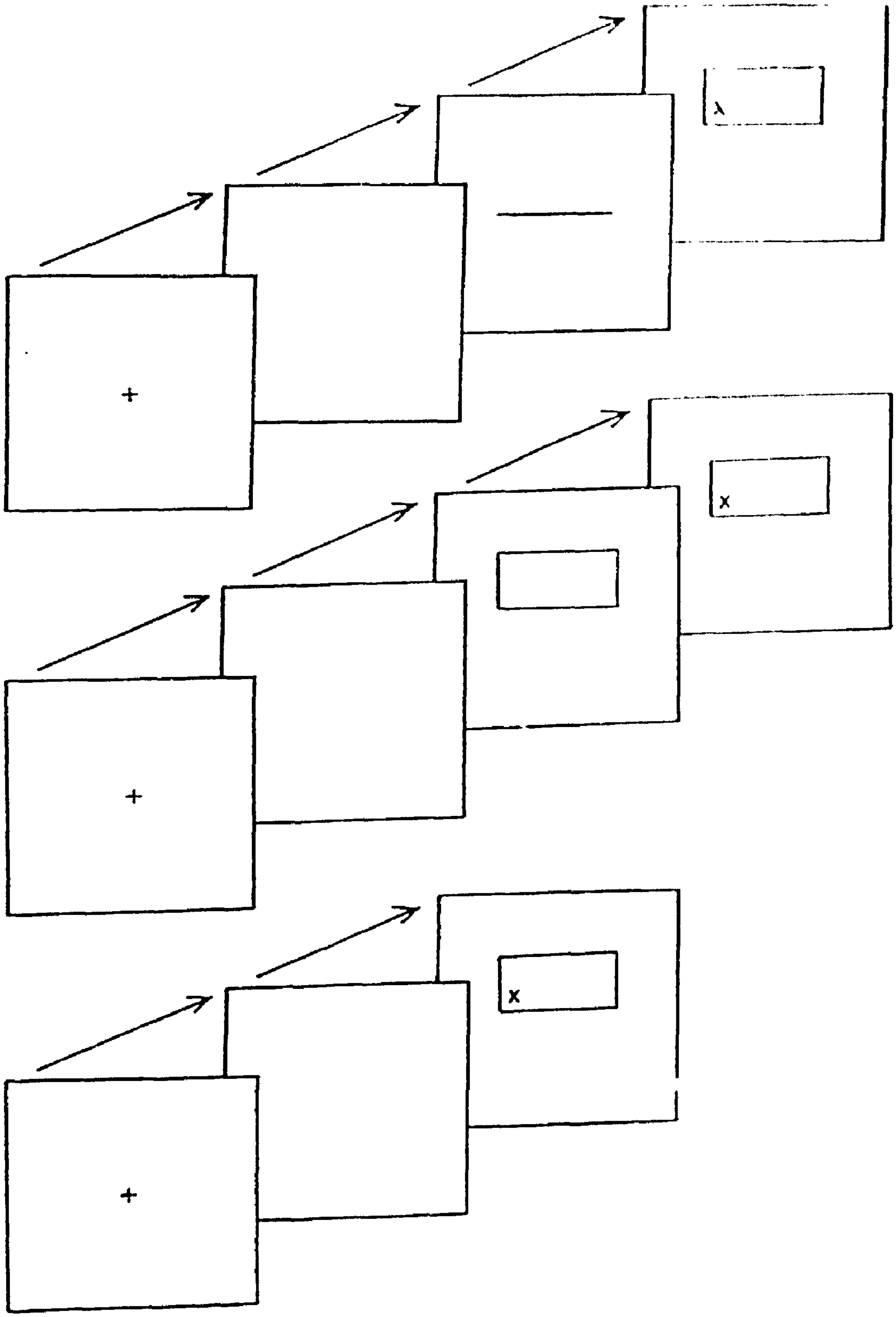


Figure 21. The temporal order of stimulus presentation in the partial-figure, whole-figure, and simultaneous presentation conditions in Experiment 5.

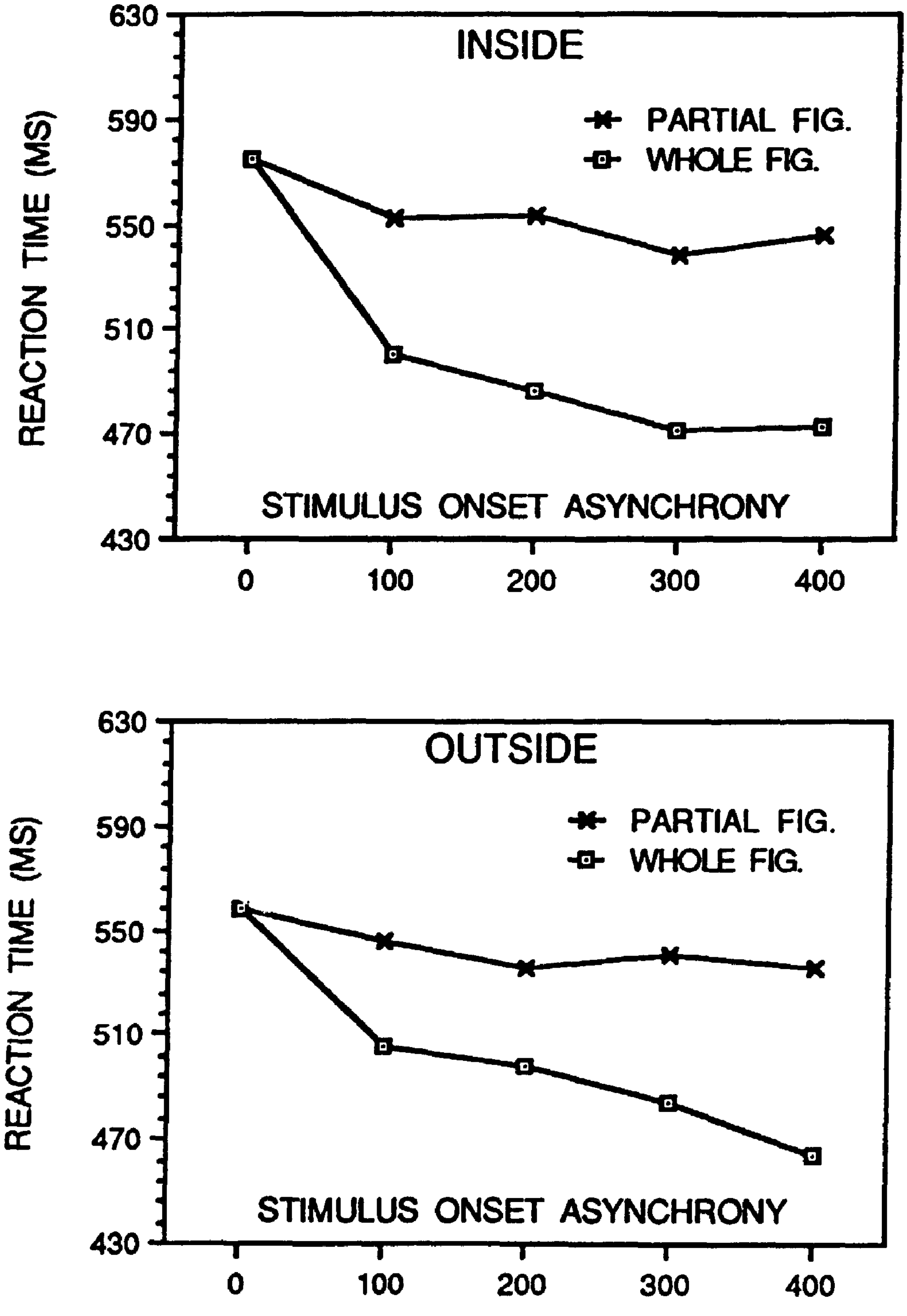


Figure 22. Mean response times for all subjects in Experiment 5 as a function of SOA.

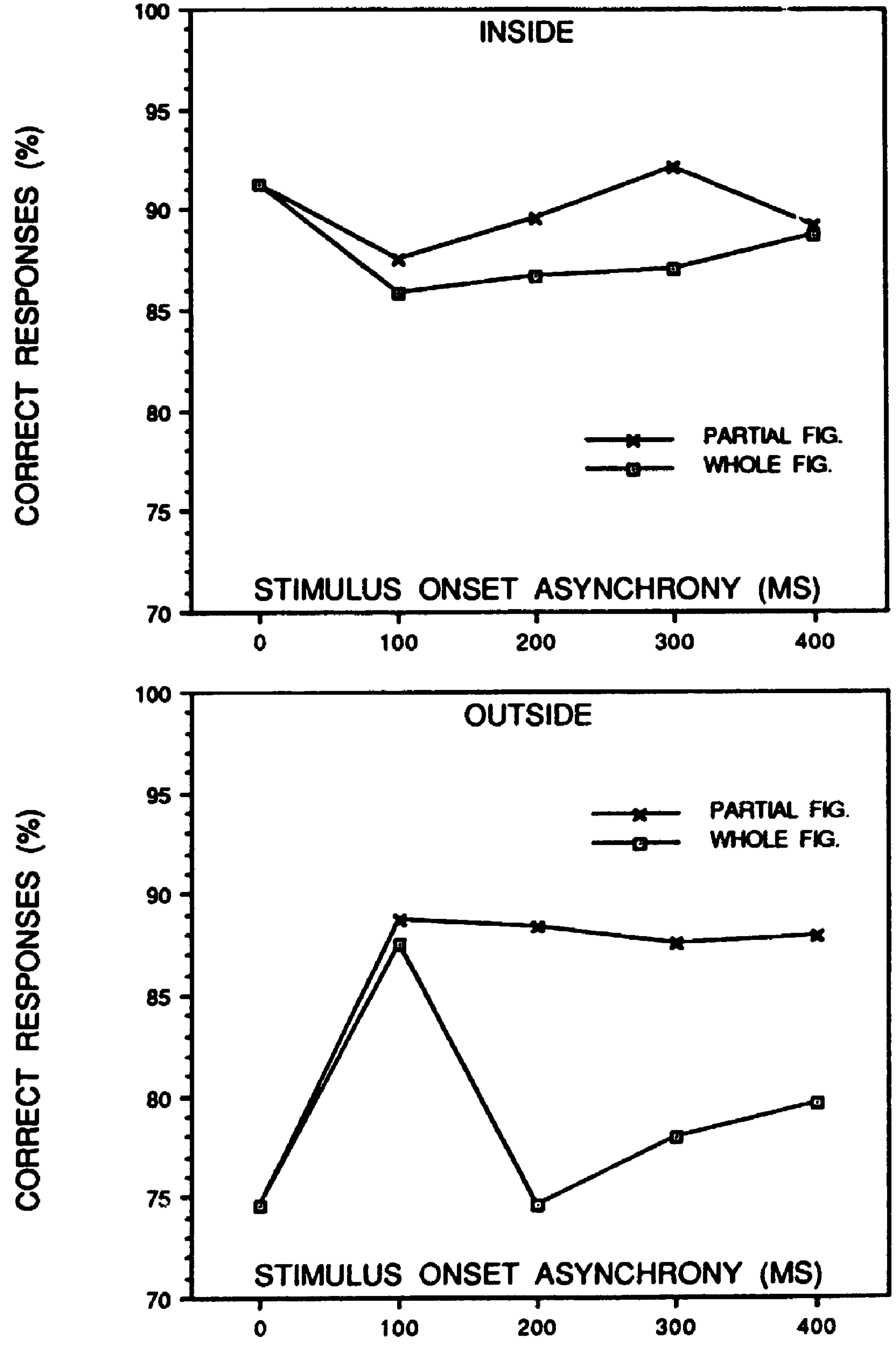


Figure 23. Mean response accuracy rates for all subjects in Experiment 5 as a function of target location.

A three-way ANOVA with repeated measures on all factors was carried out with the response time means. The three factors were target inside/outside, whole/partial figure first, and SOA (i.e., 100, 200, 300, or 400 ms). A main effect of whole/partial figure first occurred ($F_{(1,14)} = 162.96, p < .0001$) as well as an effect of SOA ($F_{(3,42)} = 9.43, p < .0001$). The mean response time for whole-figure trials was significantly lower than that on for partial-figure trials (485.30 vs. 543.28 ms). Responses were also faster as SOA was increased. The mean response times were 525.80, 518.25, 508.65, and 504.45 ms for displays with SOAs of 100, 200, 300, and 400 ms respectively. The mean response time for displays with a 0 ms SOA was 566.75 ms. The latter was not included in the analysis for reasons given in the previous experiment. The pattern of inside response times on whole-figure trials resulted in a significant linear trend ($F_{(1,42)} = 8.71$, $p < .01$) as did the pattern of outside response times ($F_{(1,42)} = 17.08, p < .01$). Moreover, there was a significant interaction between Whole/Partial Figure X SOA ($F_{(3,42)} = 3.05, p < .05$). Response times in the whole-figure condition decreased at a greater rate as SOA was increased than did response times in the partial-figure condition.

A three-way ANOVA of the response accuracy rates was also carried out. As in the previous experiment, subjects were more accurate in making inside responses (88.31% correct) than outside responses (84.00% correct), $F_{(1,14)} = 43.52, p < .0001$. In addition, accuracy was greater on partial-figure trials (88.80% correct) than on whole-figure trials (83.50% correct) ($F_{(1,14)} = 13.23, p < .0027$). There was also a Whole/Partial Figure X Target Inside/Outside interaction effect ($F_{(1,14)} = 13.23, p < .0027$). That is, subjects responded with roughly equal accuracy on all partial-figure trials and on whole-figure trials in which the target was inside the figure. However, on whole-figure trials in which the target was outside the figure, response accuracy was roughly 9% lower. As seen in Figure 23, the effect of SOA on outside response accuracy for whole-figure trials resulted in a significant trend with linear ($F_{(1,42)} = 6.12, p < .05$) and quadratic ($F_{(1,42)} = 15.54, p < .01$) components. The pattern of response accuracy rates is similar to that in the previous experiment and the same comments apply here.

3.23 Discussion

As in the previous experiment, increases in SOA led to increases in response-time facilitation. In addition, this facilitation was greater on whole-figure trials than on partial-figure trials. One reason for this finding may be that region colouring is possible between the onsets of the first and second stimulus in the whole-figure condition. However, the occurrence of facilitation in the partial-figure condition indicates that some type of processing of the figure part was carried out before the onset of the rest of the display. One possibility is that the contour of the figure part was traced, and the results were stored in an incremental representation. With the onset of the rest of the display, the results of processing of the figural part were already available and inside/outside relations were determined on the basis of this representation.

3.3 EXPERIMENT 6

The results of the previous experiments indicate that greater response-time facilitation occurred in a whole-figure condition rather than a partial-figure condition. However, several investigators (e.g., Eriksen & Hoffman, 1974; Posner, Davidson, & Snyder, 1980) have shown that when a warning signal precedes the rest of the display, general arrousal or alertness effects can facilitate response times. This could account for much of the response-time facilitation in the partial-figure condition of the previous experiment. To test this possibility, the size of the figure part in the partial-figure condition was varied in this study. It was expected that facilitation would occur with all sizes of figure parts, but that it would increase as a function of this size. This would suggest that response-time facilitation produced by the bar is due to other factors besides a warning signal.

3.31 Method

Subjects. Four University of Western Ontario graduate students (one of whom was the author) participated in this experiment as paid observers (except the author). All subjects had normal or corrected-to-normal vision, and had previous experience in studies based on reaction-time measures. Two of the subjects also participated in Experiment 2.

***Stimuli and Procedure*.** The same apparatus, testing conditions, and instructions used in the previous experiment were employed in this study. On all trials the display was preceeded by the presentation of a horizontal or vertical white line along the plane bisecting the central square region of the display screen. This line was either 1.43°, 2.87°, 4.3°, or 5.73° in length (i.e., 1/4, 1/2, 3/4, or the complete side of the rectangular figure presented in the second frame of view). The presentation of the 5.73° line was in one vertical or one horizontal position, and the 4.3° line was in one of two horizontal or one of two vertical positions (see Figure 24). Similarly, the 2.87° line was in one of three horizontal or one of three vertical positions, and the 1.43° line was in one of four horizontal or in one of four vertical positions. These positions were determined randomly over trials. Note that displays preceeded by the 5.73° line were the same as those in the partial-figure condition of the previous study. Because the other lines were 1/4, 1/2, and 3/4 of the length of the line segment in the partial-figure condition, the displays are referred to as the 1/4 bar, 1/2 bar, 3/4 bar, and the 1/1 bar displays. Positional uncertainty of the rectangular figure across trials and position of the target relative to that of the figure was identical to that of Experiment 5. All stimuli were presented asynchronously with a 100, 200, or 300 ms delay between the onset of the bar and that of the rest of the figure and the target. The temporal order of stimulus presentation is shown in Figure 25. There were 80 trials for each of the 24 conditions for each subject (3 SOAs X 4 bar lengths X 2 relations -- inside & outside). The order of the trials was randomized within test sessions.

The procedure was the same as that of the previous experiments. Each subject was tested in 10 sessions consisting of a 20 trial practise run and a 192 trial data run for a total

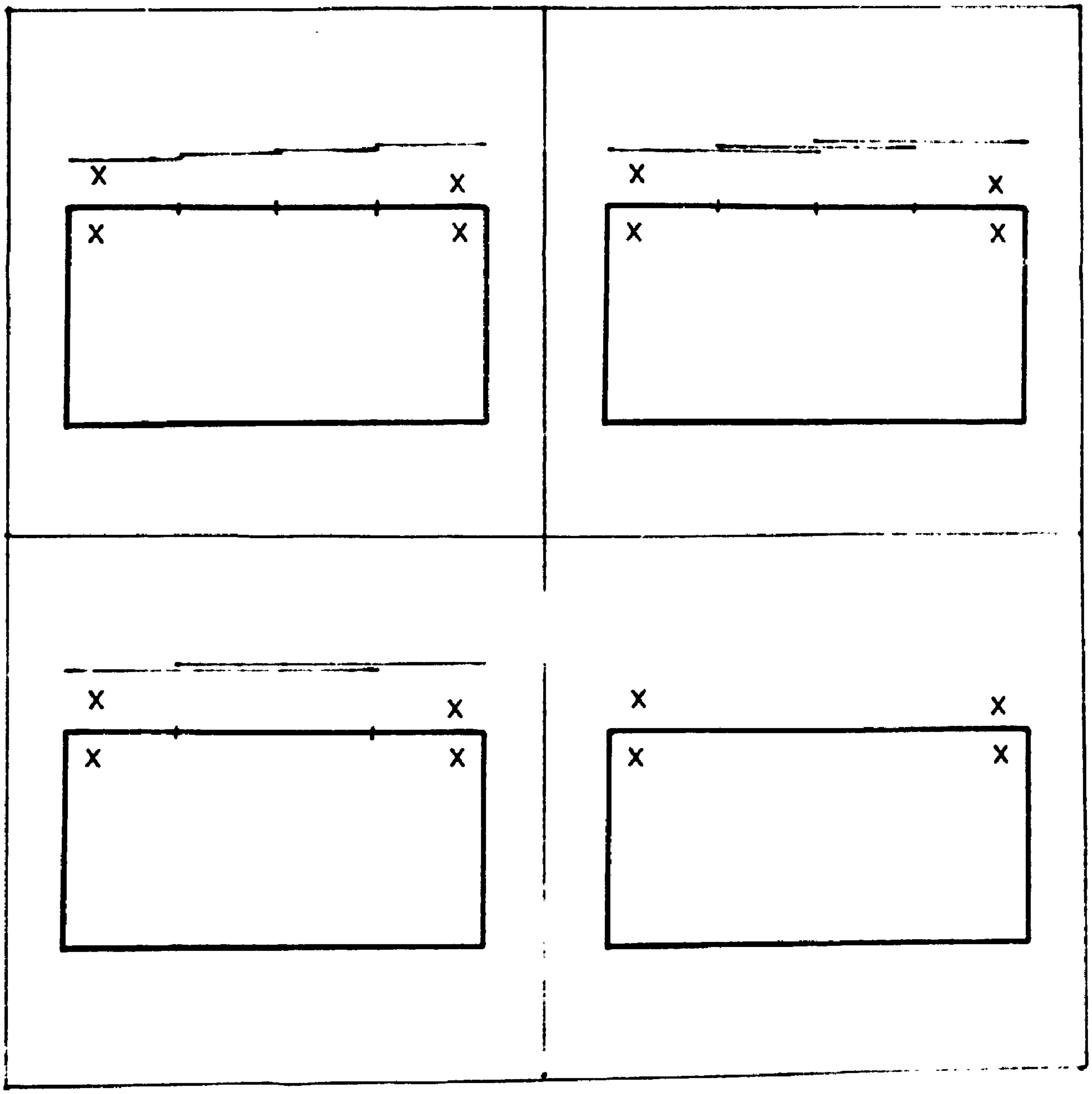
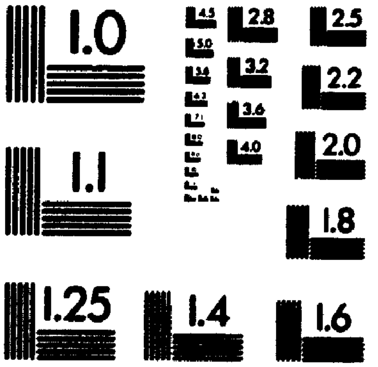


Figure 24. The four bar length conditions of Experiment 6. In the 1/4 bar condition the bar could occur in one of four locations relative to the rest of the figure. In the 1/2 bar condition the bar could occur in one of three locations. In the 3/4 bar condition the bar could occur in one of two locations, and in the 1/1 bar condition the bar could only occur in one location relative to the rest of the figure (as was the case in the partial-figure condition of the previous study).

2 of/de 2



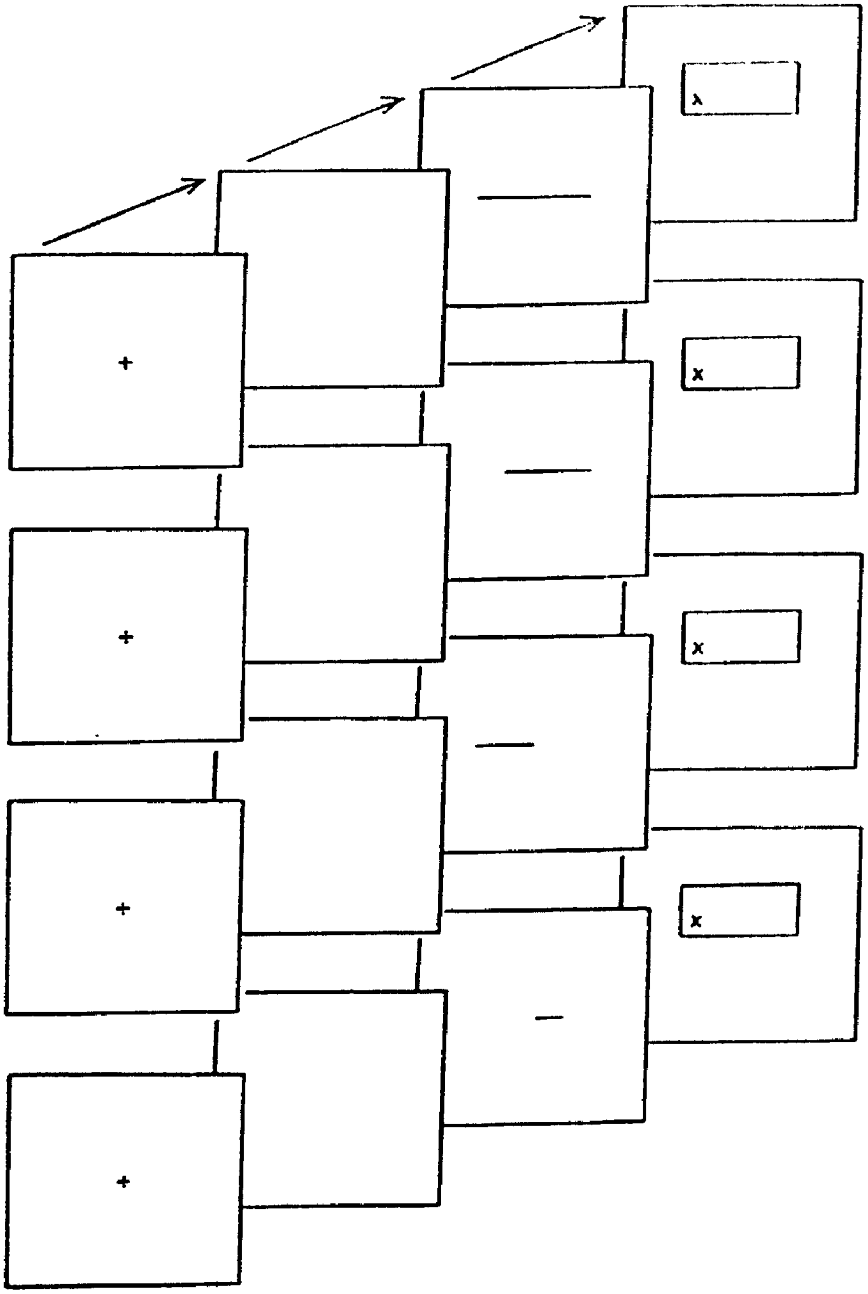


Figure 25. The temporal order of stimulus presentation in Experiment 6 with an example of each bar length condition.

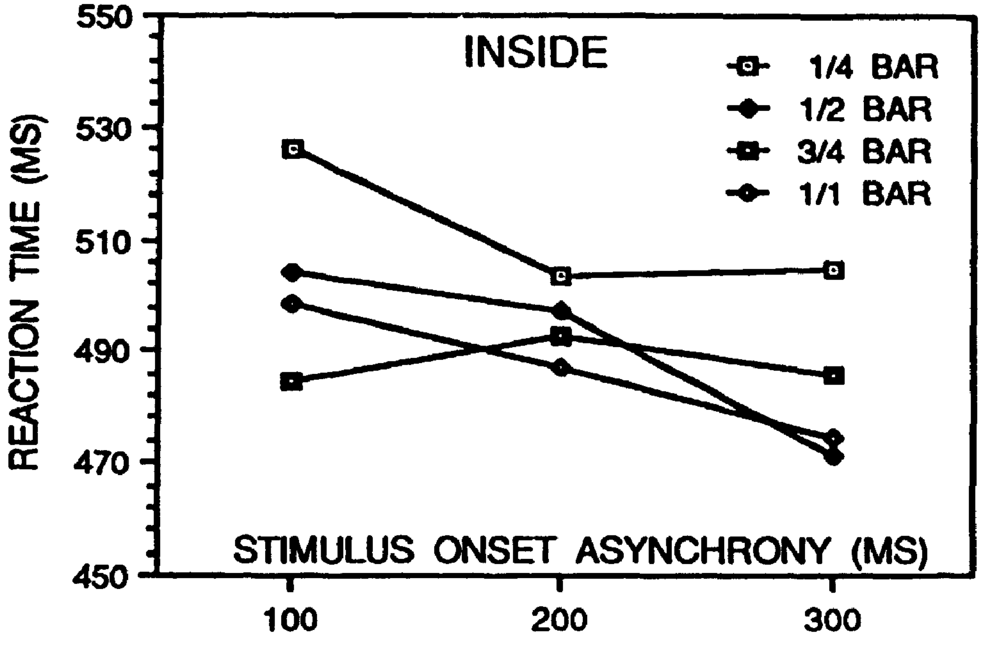
of 1920 data trials. Subjects were required to complete the sessions within 10 days but to participate in no more than two sessions per day. The data run of a session consisted of eight blocks of 24 trials with a 45 second rest period between each block.

3.32 Results

As in the previous experiments, only correct responses within two standard deviations of the mean correct response time for a particular condition were used in these analyses. This led to the removal of 4.75% of the correct responses. The mean correct response times are shown in Figure 26, and the mean response accuracy rates are shown in Figure 27. After the incorrect responses and the outliers were removed, 89.24% of the responses remained.

A three-way ANOVA with repeated measures on all factors was carried out with the response time means. The three factors were target inside/outside, SOA (i.e., 100, 200, or 300 ms), and bar length. As in the previous two experiments, increases in SOA led to decreases in response times ($F_{(2,6)} = 11.14, p < .0096$). The mean response times were 497.16, 484.29, and 475.30 ms for displays with SOAs of 100, 200, and 300 ms respectively. This resulted in a significant linear trend for inside ($F_{(1,6)} = 31.75, p < .01$) and outside ($F_{(1,6)} = 51.43, p < .01$) response times. There was also an effect of bar length on response time that produced a significant linear trend ($F_{(1,9)} = 15.00, p < .01$). The mean response times were 494.97, 485.61, 482.12, and 479.64 ms for displays with bar lengths of 1/4, 1/2, 3/4, and 1/1 respectively. Thus, the SOA and bar length manipulations both affected the magnitude of response-time savings.

A three-way ANOVA was also carried out with the response accuracy means. The only significant effect yielded by this analysis was a Target Inside/Outside X SOA interaction ($F_{(2,6)} = 9.63, p < .02$). In Figure 27 it can be seen that inside response accuracy increased as SOA was increased but outside response accuracy decreased. The pattern of outside response accuracy rates resulted in a significant trend with linear ($F_{(1,6)} = 7.35, p < .05$) and quadratic components ($F_{(1,6)} = 7.75, p < .05$) components. Hence, subjects tended to respond "inside" more often as SOA was increased.



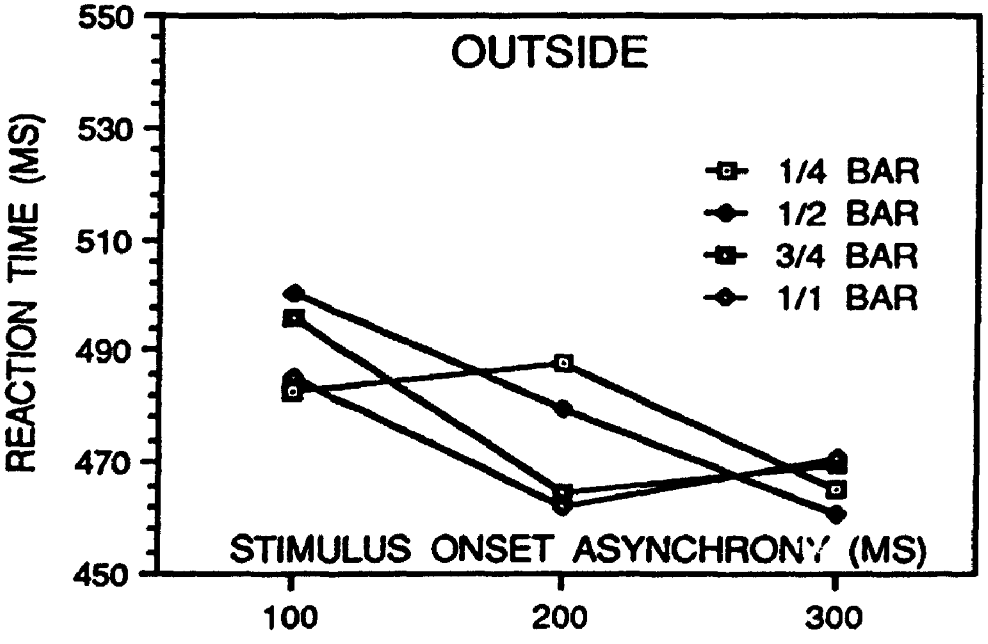


Figure 26. Mean response times for all subjects in Experiment 6 as a function of bar length and SOA.

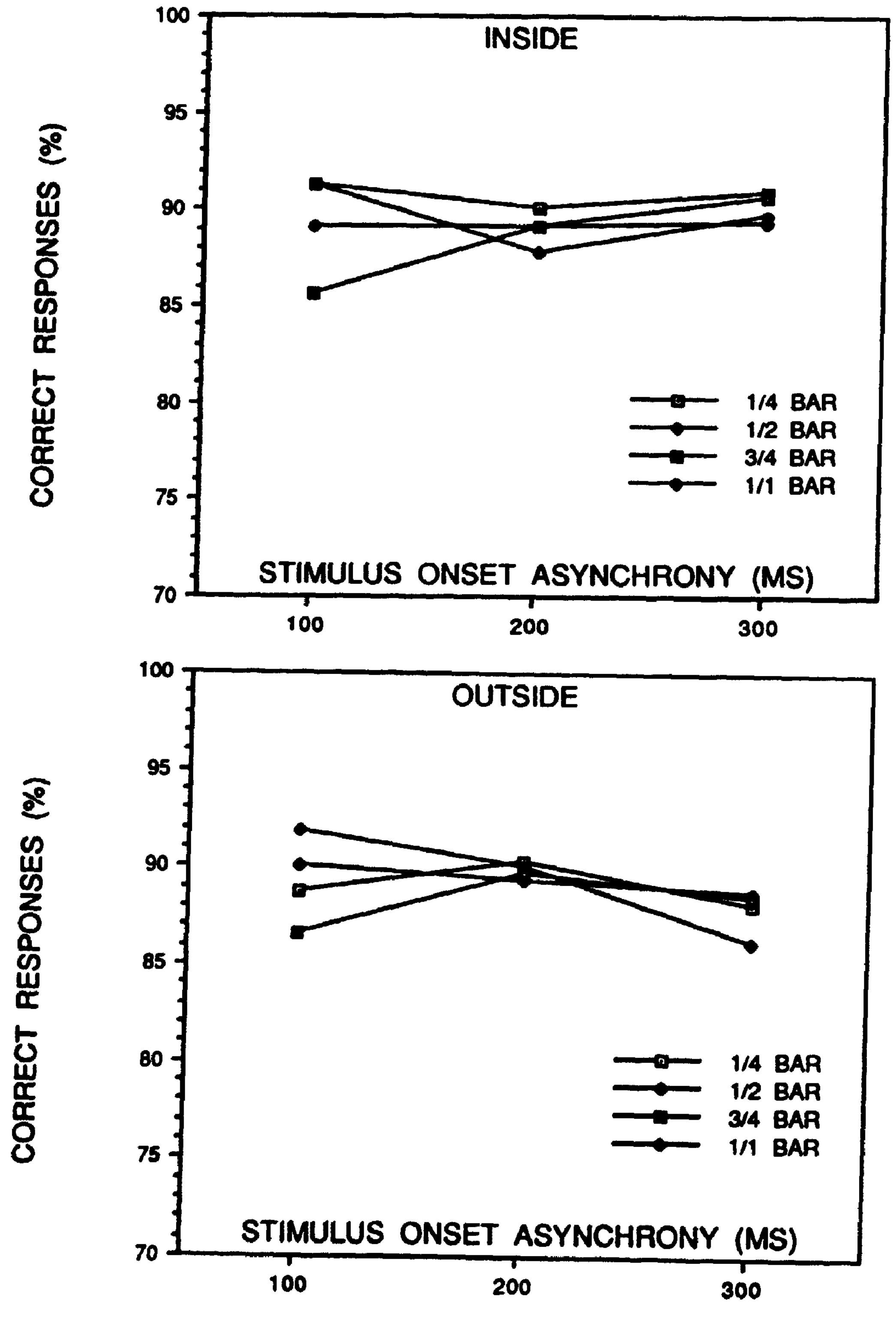


Figure 27. Mean response accuracy rates for all subjects in Experiment 6 as a function of bar length and SOA.

3.33 Discussion

The results of this study replicated and extended those of the previous two experiments. The response time savings produced by presenting the stimuli asynchronously increased as SOA was increased. Furthermore, this savings increased as the size of the figure part (bar length) was increased. One interpretation of this bar length effect is that increases in figure part size led to an increase in the amount of preliminary processing that was carried out before the onset of the rest of the display. However, another interpretation is that the bar provided a more accurate spatial precue about the eventual location of the figure, and shorter bars provided less of this information.

3.4 EXPERIMENT 7

The effect of increases in bar length on response times in the previous study may have been due to a reduction in uncertainty of the eventual location of the bounding figure. This experiment was carried out to test this proposal. On all trials a rectangular figure and a target were preceded by one or more of the figure's vertices. As a result, a spatial precue was provided by the vertices to indicate where the figure would be presented with absolute certainty whenever two or more vertices were presented. If the number of figure parts presented before the rest of the display serves only as a spatial precue, then response times should not have been affected by the number of vertices initially presented provided that at least two were present. However, if the size or number of the figure parts presented first determines how much processing is carried out before the onset of the rest of the display, then increasing the number of vertices should increase response-time facilitation.

3.41 Method

Subjects. Subjects were five University of Western Ontario graduate students (one of whom was the author) who participated in the experiment as paid observers (except the author). All subjects had normal or corrected-to-normal vision, and all had participated in one of the earlier experiments.

Stimuli and Procedure. The stimulus display on each trial was identical to that of the previous experiment with the following exception: Rather than part of one side of the rectangular figure being presented before the rest of the figure and the target, one to four vertices of the figure were presented first (see Figure 28). The combination of vertices was selected randomly before each trial so that all combinations would be tested. After an SOA of 100, 200, or 300 ms, the rest of the figure and the target were also displayed. Each vertex subtended 0.57 x 0.57° of visual angle. There were 80 trials for each of the 24 conditions (4 vertex numbers X 3 SOAs X 2 inside/outside conditions) and the order of trials was randomized within test sessions. The procedure, number of data trials, and number of test sessions per subject were also identical to that of Experiment 6.

4.42 Results

Only correct responses within two standard deviations of the correct response time for a particular condition were used in these analyses. This led to the removal of 4.83% of the correct responses. After incorrect responses and outliers were removed, 92.31% of the responses remained. A three-way ANOVA with repeated measures on all factors (*target inside/outside* x *SOA* x *number of vertices*) was carried out with the response time means. There was a main effect of SOA ($F_{(2,8)} = 11.33, p < .005$) in which the mean response times were 411.51, 398.25, and 392.94 ms for displays with SOAs of 100, 200, and 300 ms respectively. As seen in Figure 29, the pattern of inside response times resulted in a significant trend with linear ($F_{(1,8)} = 120.47, p < .01$) and quadratic ($F_{(1,8)} = 6.45, p < .05$) components. Similarly, the pattern of outside response times resulted in a trend with linear ($F_{(1,8)} = 141.90, p < .01$) and quadratic ($F_{(1,8)} = 9.70, p < .05$) components.

In addition, there was a main effect of the number of vertices initially presented ($F_{(3,12)} = 56.72, p < .0001$). The mean response times were 410.75, 401.65, 397.69, and

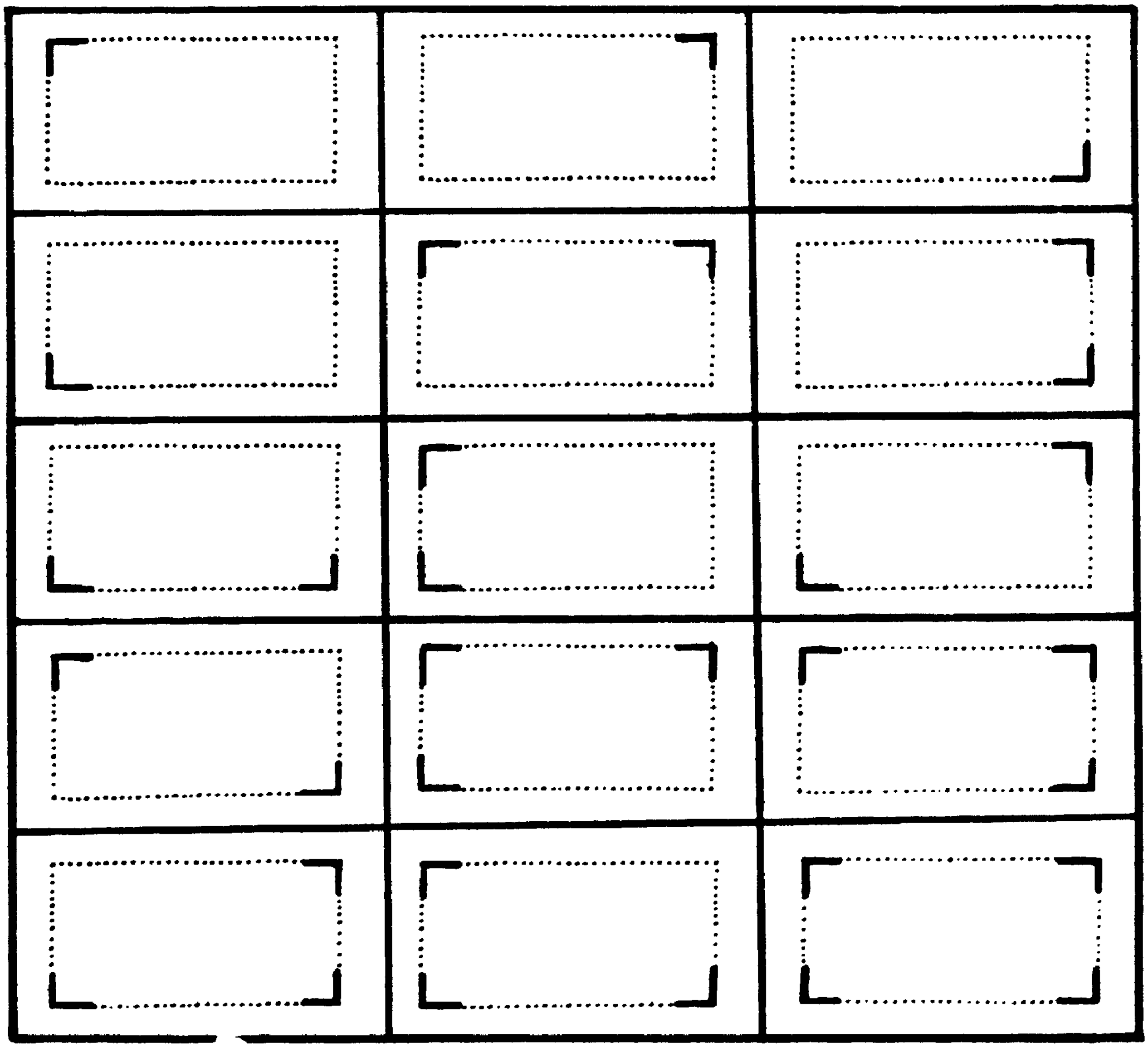


Figure 28. The different combinations of vertices that were presented prior to the onset of the rest of the display in Experiment 7.

393.68 ms for displays preceded by one, two, three, and four vertices respectively. This pattern of response times resulted in a significant trend with linear ($F_{(1,12)} = 162.26$, $p < .01$) and quadratic ($F_{(1,12)} = 6.55$, $p < .05$) components. There was also a Number of Vertices x SOA interaction ($F_{(6,24)} = 2.51$, $p < .05$). Response times for the one vertex condition were affected to a lesser degree by the SOA manipulation than response times in the two, three, and four vertices conditions. Also, a three-way ANOVA was carried out with the response accuracy rates of each condition and all possible effects were tested. None of these tests was significant, indicating that no speed-accuracy tradeoff occurred (see Figure 30).

Finally, the response time results of this experiment were compared to those of the previous experiment using a four-way ANOVA (*bar/vertex first* X *target inside/outside* X *SOA* X *number of figure parts*) with repeated measures on all factors but the first. Increases in the number of figure parts (either vertices or bar length) caused a decrease in response times. More specifically, the pattern of inside response times resulted in a significant trend with linear ($F_{(1,21)} = 48.55$, $p < .01$) and quadratic ($F_{(1,21)} = 9.05$, $p < .01$) components. Similarly, the pattern of outside response times resulted in a significant linear trend ($F_{(1,21)} = 17.70$, $p < .01$). Also, increases in SOA caused a decrease in response times. As seen in Figure 31, the pattern of inside response times resulted in a significant linear trend ($F_{(1,14)} = 111.24$, $p < .01$). Similarly, the pattern of outside response times resulted in a significant trend with linear ($F_{(1,14)} = 152.75$, $p < .01$) and quadratic ($F_{(1,14)} = 9.05$, $p < .01$) components. Despite the large difference between the mean response times of the bar condition (485.88 ms) and the vertex condition (400.90 ms), the effect of condition was not significant ($p = 0.106$). A similar four-way ANOVA was carried out with the response accuracy rates of each condition and no effects were significant (see Figure 32).

3.43 Discussion

The results of this study replicate the SOA effect of the previous experiments.

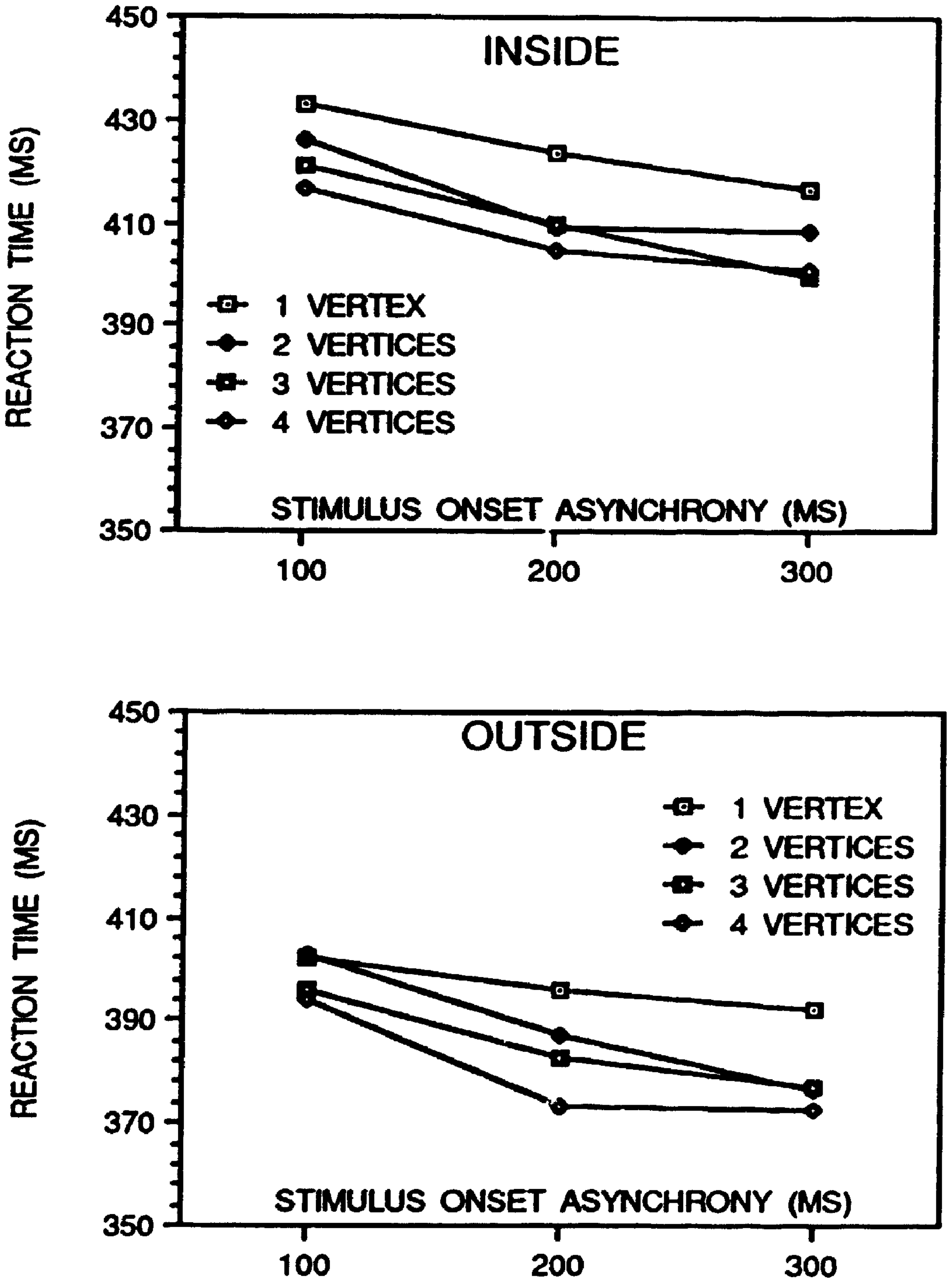


Figure 29. Mean response times for all subjects in Experiment 7 as a function of the number of vertices and SOA.

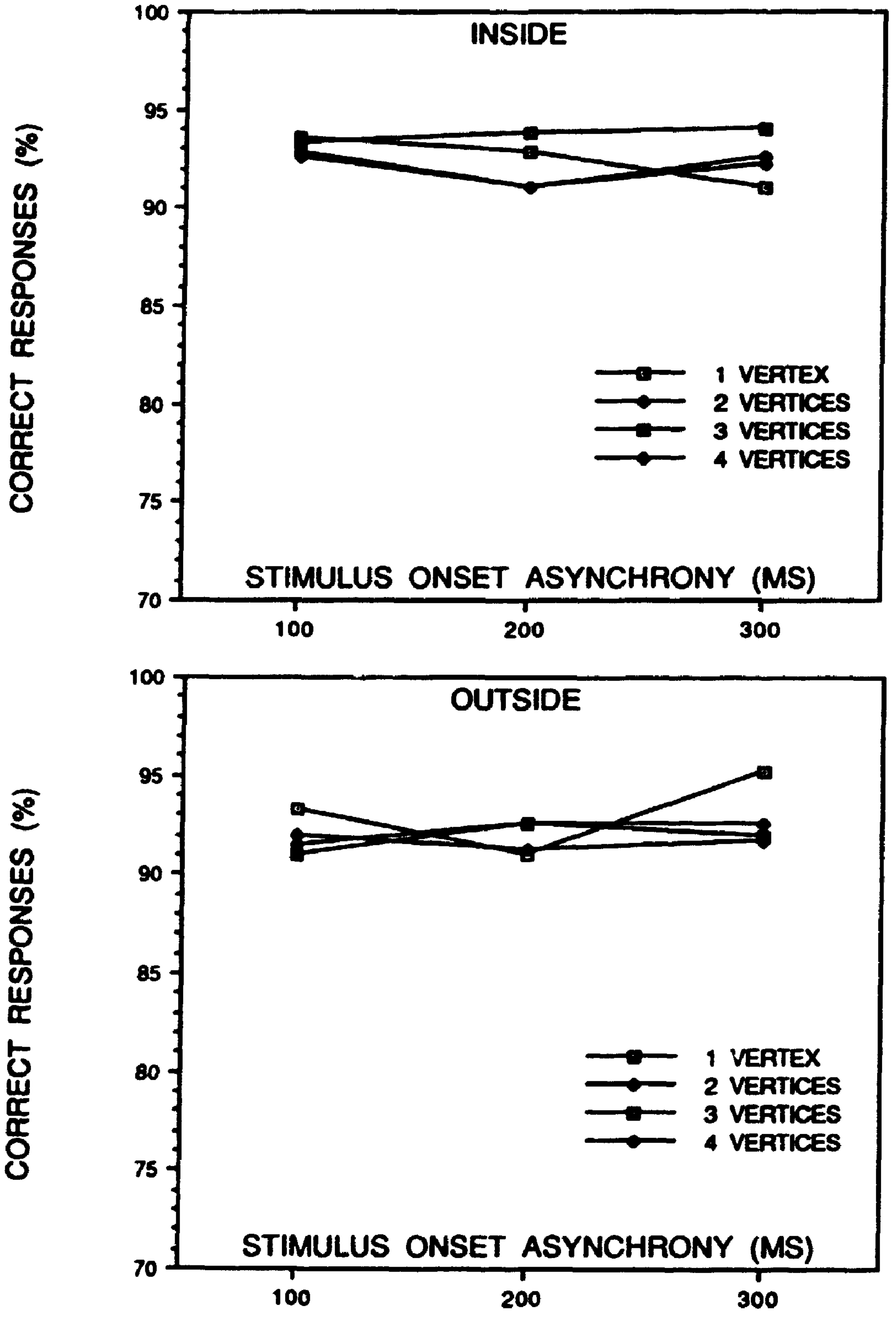


Figure 30. Mean response accuracy rates for all subjects in Experiment 7 as a function of the number of vertices and SOA.

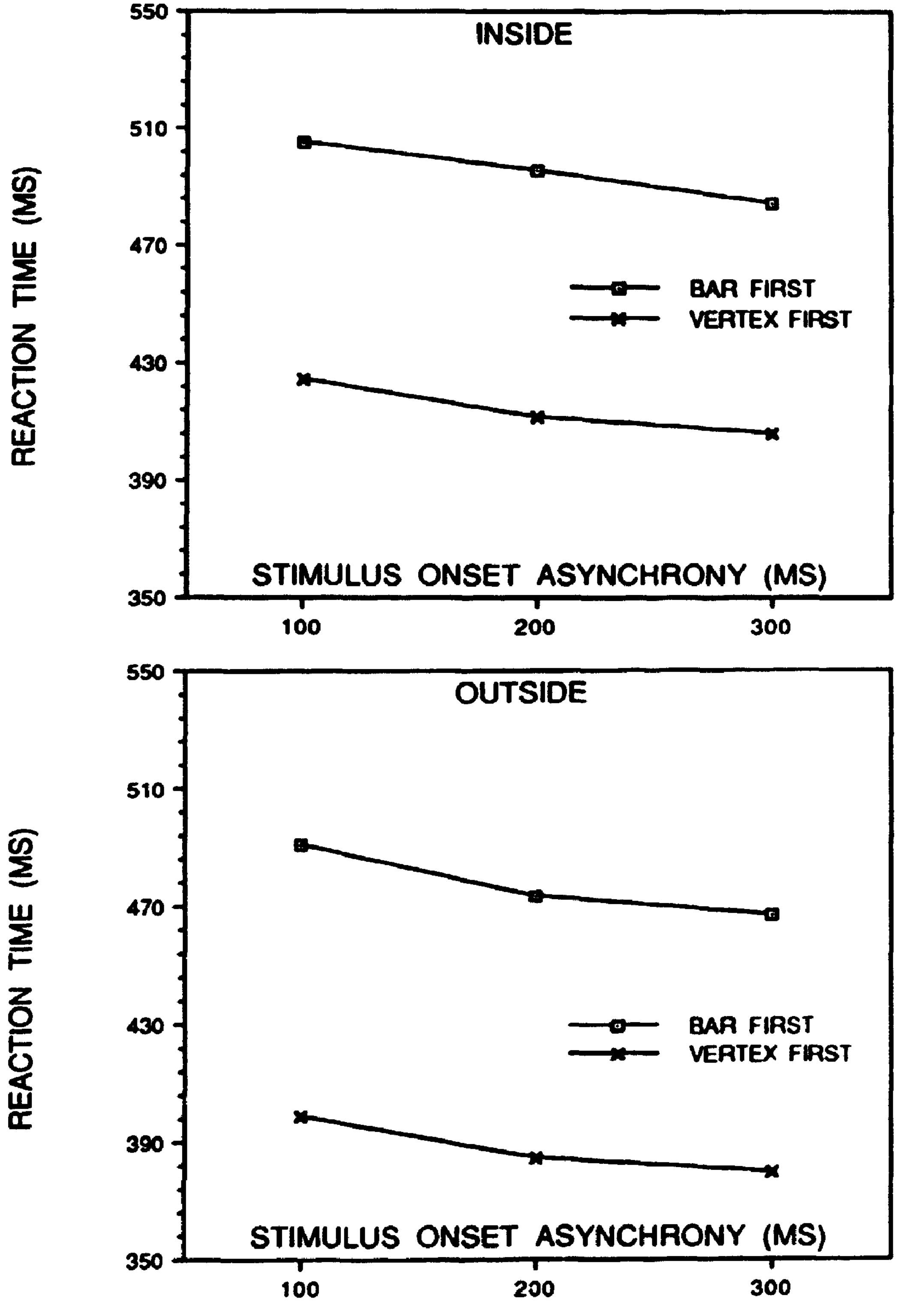


Figure 31. The mean responses times as a function of the number of parts presented first (i.e. vertices or bar length) and SOA in a comparison of the results of Experiment 6 and Experiment 7.

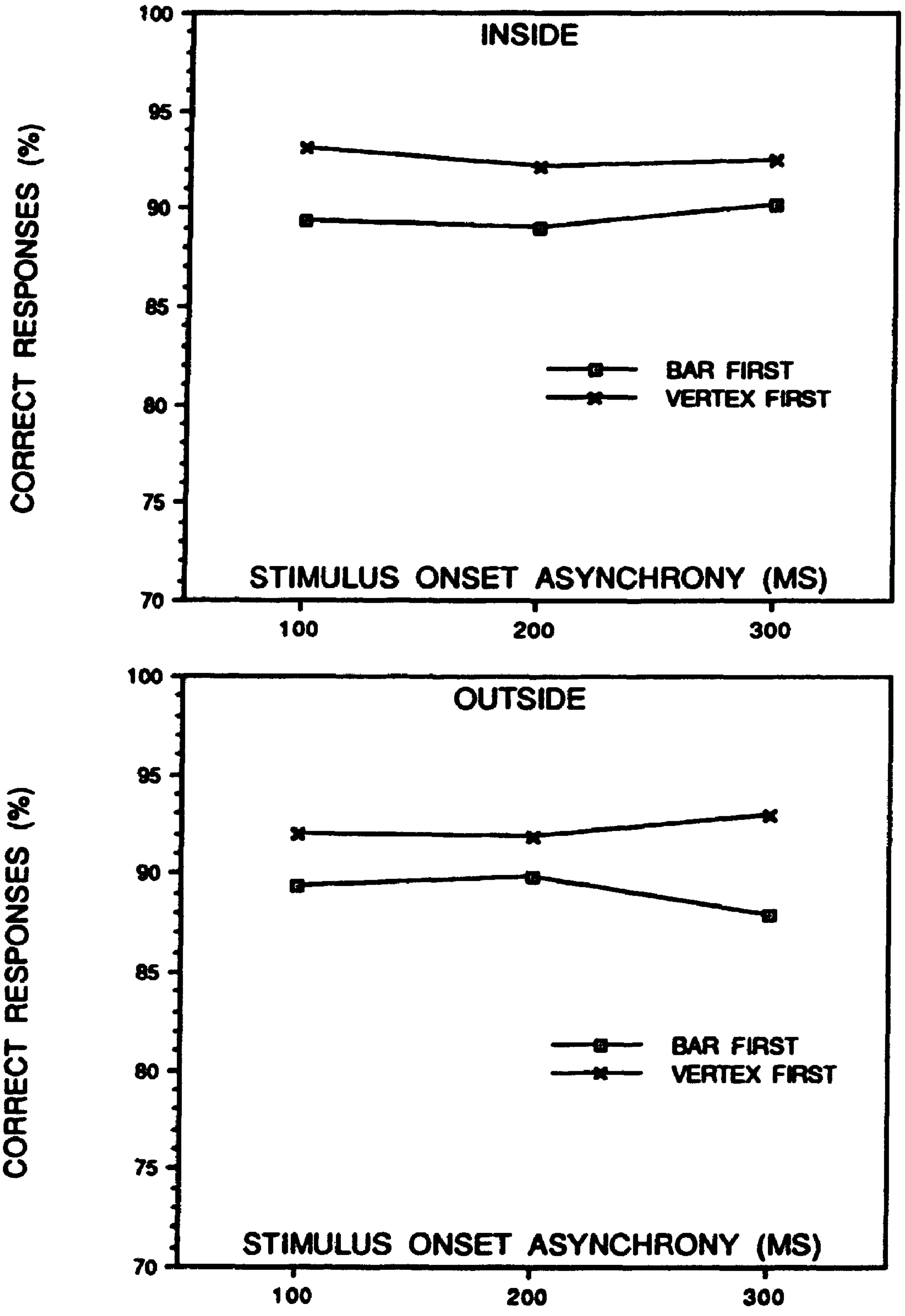


Figure 32. The mean response accuracy rates as a function of the number of parts presented first (i.e. vertices or bar length) and SOA in a comparison of the results of Experiment 6 and Experiment 7.

They also indicate that increases in the number of figure parts (vertices) leads to increases in response-time facilitation. Moreover, it appears that this effect is *not* due to a decrease in uncertainty about the figure's eventual location as the number of vertices were increased. Instead, the increases in the number of figure parts available before the onset of the rest of the display appears to cause increases in the amount of preliminary processing that can be carried out on the first stimulus.

Due to the nature of the stimulus displays, there was some question about whether changes in the number of vertices presented affected precuing of the eventual target location. As seen in Figure 14, the inside/outside target locations were close to the ends of the straight-line figure part bisecting the central display region. When four vertices were initially presented, there was always a vertex close to both target locations. However, when three vertices were initially presented, there was a vertex close to both target locations on only 50% of the trials (see Figure 28). Therefore, on half of the three-vertices trials, only one of the target locations was precued. Moreover, on these trials there was a 50% probability that the single vertex provided an invalid cue. Posner et al. (1980) have shown that invalid cues sometimes *increase* response times relative to a no-cue condition. Thus, on three-vertices trials in which only a single target location was precued, the facilitation of valid precuing should be nullified by the cost of invalid precuing. Furthermore, when two vertices were initially presented, there was a vertex close to both target locations on only 16.67% of trials. Based on these probabilities, the increase in facilitation should be greater when the number of vertices is increased from three to four rather than two to three. To elaborate, both target locations are accurately precued on 100% of four-vertices trials, 50% of three-vertices trials, and 16.67% of two-vertices trials. However, the increase in response-time facilitation was linear as the number of vertices was increased from two to three (a savings of 3.96 ms) and three to four (a savings of 4.01 ms). This indicates that target precuing cannot account for all of the response-time facilitation in this study.

Colouring cannot begin until a clearly-defined internal region is present in the visual field. This is not the case if only one or two vertices are displayed. The presentation of two or more vertices provides an indication of *where* the contours bounding the internal region will be presented on the current trial but only if the shape and size of the figure are the same as those of the previous trials. It should be noted that the linear increase in response-time facilitation with increases in the number of vertices should not be expected if region colouring is possible with four but not with two or three vertices. More specifically, a quadratic increase in facilitation would be expected because facilitation should be greater in conditions (e.g., the four-vertices condition) in which region colouring is possible. A direct test of the effect of vertices on colouring could be conducted by comparing performance in a four-vertices condition with that in a full-figure condition (as in Experiment 4). If vertices provide sufficient information for region colouring to begin, then the patterns of inside response times in the two conditions should be the same. However, if inside response times are faster in the full-figure condition, this could suggest that vertices do not provide enough information for colouring to begin.

3.5 GENERAL DISCUSSION

The experiments described in this chapter involved manipulations of the relative onset times of stimuli in inside/outside tasks. It was found that when stimuli were presented in temporal succession, response times decreased as the SOA between stimuli increased. In addition, greater response-time facilitation occurred if the stimulus presented first was the bounding figure rather than the target. One account of these findings is that when stimuli are presented asynchronously, inside/outside relations are determined on the basis of an incremental representation containing the results of processing carried out on the first stimulus before the onset of the second stimulus. There appears to be two ways to increase the amount of this preliminary processing. One is to increase the SOA between stimuli to allow the first stimulus to be processed in isolation for a longer period of time. The other is to increase the size or number of figure parts presented first. In this way, a greater degree of the processing that will have to be done at some point can be carried out and its results can be stored in an incremental representation before the second stimulus is presented. Thus, the results of the studies indicate that increases in the amount of preliminary processing of the first stimulus will produce greater response time facilitation when the second stimulus is displayed.

It appears that the subprocesses operating on the first stimulus before the onset of the second stimulus are carried out on both contour lines and bounded regions. In particular, in the experiments described in this chapter response-time facilitation occurred in target-first and partial-figure (straight line figure part) conditions even though region colouring was not possible until the rest of the display was presented. This suggests that another subprocess besides colouring is involved in determining inside/outside relations. For example, perhaps there is a subprocess A that colours regions and a subprocess B that somehow paints or activates contours. Mahoney and Ullman (1988, pp. 194-203) have described how this type of contour activation operation might work. If this is the case, then subprocess B could be active before the onset of the figure in a target-first display. In contrast, subprocesses A and B could both be active (perhaps in parallel) before the onset of the target in a figure-first display. Presumably, more preliminary processing will be carried out before the onset of the second stimulus in the latter case because both types of subprocesses will have been invoked. This would account for the greater response time facilitation in the figure-first condition than in the target-first condition. This proposal could be tested by manipulating certain contour properties (e.g., presence or absence of contour gaps; the contour's proximity to other contours) and region properties (e.g., figural size) that are known to affect inside/outside response times. If stimuli are presented asynchronously in these experiments, the findings may provide some indication about whether separate subprocesses of this type can be teased apart.

CHAPTER FOUR: CONCLUDING REMARKS

The studies conducted in this thesis involved inside/outside tasks in which the size of the bounding figure, the location of the target, and the relative onset times of the stimuli were manipulated. Each manipulation had some effect on response times. In particular, increases in the distance between the figure and the target led to faster outside response times. Also, inside response times increased as a function of the distance between the target and the pretrial fixation point. Furthermore, the effect of figural size changes on inside/outside responses was quite robust. That is, inside response times increased with increases in figural size while outside response times decreased. Thus, manipulations of stimulus size and location affect inside/outside response times.

In the third chapter it was found that manipulations of SOA had a systematic effect on inside/outside response times. These results suggest the following processing sequence: When two stimuli are presented asynchronously, the first stimulus is visually processed from the time of its onset. When the second of the two stimuli is presented, some degree of processing of the first stimulus has already been carried out and the results of this processing are stored in a temporary visual representation (e.g., an incremental representation). Therefore, the response time facilitation that occurs as a result of asynchronous presentation suggests that the determination of spatial relations between two stimuli is based on an incremental representation containing the results of preliminary processing carried out on the first stimulus. In other words, the visual system can determine the spatial relation in less time with asynchronous stimuli because some of the required processing of the first stimulus has already been carried out. Moreover, this time appears to decrease as the degree of preliminary figural processing is increased.

Three of the experiments were conducted with trained observers who participated as paid subjects over a period of ten testing sessions. The other four experiments were conducted with naive, first year undergraduate students who participated in a single test session as a psychology course requirement. There is reason to believe that the data was less reliable in the latter case because the subjects' motivation was probably lower, they were not trained to perform the task (with the exception of 20 practise trials), and there were fewer observations per subject. Moreover, the mean error rate of the paid subjects across the three experiments was roughly 5% while that of the undergraduate students was roughly 10%. However, each of the major effects on response times reported in this thesis were obtained with both undergraduate subjects and trained, motivated observers (target location effects were obtained with undergraduate subjects in a study mentioned in Experiment 1). Thus, the response time patterns are robust.

The results of this research suggest that while colouring can account for the effect of figural size changes on the determination of *inside* relations, it is more difficult to explain how *outside* relations are determined with this type of model. More specifically, the increase in inside response times as a function of figural size is consistent with the performance of a colouring model based on a uniscale activation algorithm. However, the decreases in outside response times with figural size increases do not appear to have a straightforward explanation in terms of colouring. This suggests that region colouring may not always be invoked when inside/outside relations are determined, or may only be invoked to determine inside of" but not "outside of."

The results of this research also indicate that several subprocesses or operations besides colouring are involved in the determination of this relation. For example, it was found that inside response times were faster as the proximity of the target to the location of the pretrial fixation point was increased. One account of this finding is that each trial involved an initial shift of processing focus from the fixation point to the target at the time of onset of the stimulus display, and that the time required for this shift decreased as the absolute distance between the two locations was decreased. Thus, it appears that some inside/outside routines are preceded by shift of processing focus operations. In addition, the results of the studies involving asynchronous stimulus presentation indicated that some operations were carried out on the target or on straight line figure parts (e.g., some form of contour activation) before the onset of the bounding figure even though there was no clearly defined region to colour. This suggests that operations other than region colouring are involved in determining inside/outside relations, and that these processes operate together (perhaps in sequence) as a visual routine.

There are many possible extensions of this research. For example, the effect of varying target location could be also be studied at a finer level by increasing the number of possible locations inside or outside of a figure. The type of bounding figure used in such a study could also be quite different. For example, if Experiment 2 was replicated using circular figures, it would be possible to determine whether target proximity to the centre of the figure affects response times without being concerned about the effect of target proximity to vertices. A positive relationship between response time performance and target proximity to the centre of the figure may suggest that observers are carrying out some form of computation like the derivation of a centroid when making inside responses. This would be consistent with evidence that observers tend to scan a figure by beginning at its centre (Noton & Stark, 1971). Thus, extensions of research on manipulations of target location may yield more refined accounts of the processes involved in this perception.

There are also several possible extensions of the studies involving figural size manipulations. As mentioned previously, inside response times in Experiments 1, 2, and 3 may have been affected by the time required to adjust what Ward (1982) has called an attentional frame. In particular, it was suggested that the focal resolution of the attentional frame was set to the size of the fixation cross at the beginning of each trial. As a result, more extensive adjustments of this focal resolution would be required for attending to larger figures than for attending to smaller figures that followed the fixation cross. Thus, more time would be required to make inside responses with larger figures than with smaller figures. This proposal could be tested by varying the size of the fixation cross displayed at the beginning of each trial of a replication of Experiment 1. That is, if the fixation cross is the same size as the largest of the bounding figures, then the focal resolution of the attentional frame will not require adjustment on trials involving the largest figures. However, the magnitude of focal adjustment would increase as the size of the bounding figure is decreased. Thus, if a large fixation cross is used, it may be the case that inside response times would be faster with large figures than with smaller ones. Presumably, this would occur because less adjustment of the attentional frame would be required on trials involving large figures.

There are also many possible extensions of the studies involving asynchronously presented stimuli. For example, Experiment 7 could be replicated with four vertices in each display prior to the onset of the rest of the display so that the two possible target locations would always be precued. In addition, the number of complete sides of the rectangular figure that are presented in the first frame of view could also be manipulated. In this way, the effect of spatial precuing of the figure and target locations could be completely controlled for. Any response-time facilitation that results could be attributed exclusively to the degree of preliminary processing carried out on these figure parts before the onset of the rest of the stimulus display. As mentioned previously, the possible roles of region and contour activation could also be investigated by manipulating the size of asynchronously presented stimuli as well as their contour properties. This may provide some insight into the nature of the activation operations involved when inside/outside relations are determined by the visual system.

In each study described in this thesis, the manipulation had a systematic and robust effect, and the findings provided suggestions for several possible extensions of this research. Therefore, the goal of the project, to put together an initial database of information about factors that affect the perception of inside/outside spatial relations, has been achieved. Continued investigation of this type should provide additional information about the operations involved in this processing, and perhaps converging evidence in support of a particular model of this perception.

REFERENCES

Ballard, D.H., & Brown, C.M. (1982). *Computer Vision*. Englewood Cliffs, N.J.: Prentice-Hall.

Brady, M. (1981). The changing shape of computer vision. *Artificial Intelligence, 17*, 1-15.

Brice, C.R., & Fennema, C.L. (1970). Scene analysis using regions. *Artificial Intelligence, 1*, 205-261.

Bruner, J.S. (1957). On perceptual readiness. *Psychological Review, 64*, 123-151.

Bundesen, C., & Larsen, A. (1975). Visual transformation of size. *Journal of Experimental Psychology: Human Perception & Performance, 1*, 214-220.

Butler, L., & McKelvie, S.J. (1985). Processing of form: Further evidence for the necessity of attention. *Perceptual & Motor Skills, 61*, 215-221.

Dawson, M.R.W., & Pylyshyn, Z.W. (1986). Using relative velocity information to constrain the motion correspondence problem: Psychophysical data and a computational model. *Proceedings of the Sixth Canadian Conference on Artificial Intelligence.*

Dawson, M.R.W., & Pylyshyn, Z.W. (1988). Natural constraints in apparent motion. In Z.W. Pylyshyn (Ed.), *Computational Processes in Vision: An Interdisciplinary Perspective*, 99-120. Norwood, N.J.: Ablex.

Egly, R., & Homa, D. (1984). Sensitization of the visual field. *Journal of Experimental Psychology: Human Perception & Performance, 10*, 778-793.

Eriksen, B.A., & Eriksen, C.W. (1974). Effect of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*, 143-149.

Eriksen, C.W., & Hoffman, J.E. (1972). Temporal and spatial characteristics of selective encoding from visual displays. *Perception & Psychophysics, 12*, 201-204.

Eriksen, C.W., & Hoffman, J.E. (1974). Selective attention: Noise suppression or signal enhancement? *Bulletin of the Psychonomic Society, 4*, 587-589.

Eriksen, C.W., & Yeh, Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception & Performance, 11*, 583-597.

Evans, T.G. (1968). A heuristic program to solve geometric analogy problems. In M. Minsky (Ed.), *Semantic Information Processing*. Cambridge, Mass.: MIT Press.

Feldman, J.A., & Ballard, D.H. (1982). Connectionist models and their properties. *Cognitive Science, 6*, 205-254.

Fodor, J.A. (1975). *The Language of Thought.* Sussex: Harvester Press.

Fodor, J.A., & Pylyshyn, Z.W. (1980). How direct is visual perception? Some reflections on Gibson's "ecological approach." *Cognition, 9*, 139-196.

Fodor, J.A., & Pylyshyn, Z.W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*, 3-71.

Gardner, G.T. (1973). Evidence for independent parallel channels in tachistoscopic perception. *Cognitive Psychology, 4*, 130-155.

Goldstein, E.B., & Fink, S.I. (1981). Selective attention in vision: Recognition memory for superimposed line drawings. *Journal of Experimental Psychology: Human Perception & Performance, 7*, 954-967.

Gregory, R. (1970). *The Intelligent Eye.* London: Weidenfield & Nicolson.

Haralick, R.M., & Shapiro, L.G. (1985). Image segmentation techniques. *Computer Vision, Graphics, & Image Processing, 29*, 100-132.

Hildreth, E. (1983). *The Measurement of Visual Motion.* Cambridge, Mass.: MIT Press.

Hinton, G.E., & Anderson, J.A. (1981). *Parallel Models of Associative Memory.* Hillsdale, N.J.: Erlbaum.

Humphreys, G. (1981). On varying the span of attention: Evidence for two modes of spatial attention. *Quarterly Journal of Experimental Psychology, 334*, 1-15.

Ingleton, M.A. (1987). Rate of curve tracing and the effect of pattern size. Unpublished Master's thesis, University of Waterloo, Waterloo, Ontario, Canada.

Jolicoeur, P. (1988). Curve tracing operations and the perception of spatial relations. In Z.W. Pylyshyn (Ed.), *Computational Processes in Human Vision: An Interdisciplinary Perspective*, 133-168. Norwood, N.J.: Ablex.

Jolicoeur, P., Ullman, S., & MacKay, M. (1986). Curve tracing: A possible basic operation in the perception of spatial relations. *Memory & Cognition*, *14*, 129-140.

Jonides, J. (1983). Further towards a model of the mind's eye's movement. *Bulletin of the Psychonomic Society*, *21*, 247-250.

Kienker, P.K., Sejnowski, T.J., Hinton, G.E., & Schumacher, L.E. (1986). Separating figure from ground with a parallel network. *Perception*, *15*, 197-216.

Kimura, D. (1973). The asymmetry of the human brain. *Scientific American*, *228*, 70-78.

Koch, C., & Ullman, S. (1984). Selecting one among the many: A simple network implementing shifts in selective visual attention. MIT AI Memo No. 770.

Kosslyn, S.M. (1987). Seeing and imagining in the cerebral hemispheres: A computational approach. *Psychological Review*, *94*, 148-175.

Larsen, A., & Bundesen, C. (1978). Size scaling in visual pattern recognition. *Journal of Experimental Psychology: Human Perception & Performance*, *4*, 1-20.

Lowe, D.G. (1985). *Perceptual Organization and Visual Recognition.* Hingham, Mass.: Kluwer Academic Publishers.

Mahoney, J.V., & Ullman, S. (1988). Image chunking: Defining spatial building blocks for scene analysis. In Z.W. Pylyshyn (Ed.), *Computational Processes in Human Vision: An Interdisciplinary Perspective*, 169-209. Norwood, N.J.: Ablex.

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

Marr, D., & Poggio, T. (1976). A cooperative computation of stereo disparity. *Science*, *194*, 283-287.

Marr, D., & Poggio, T. (1979). A computational theory of human stereo vision. *Proceedings of the Royal Society of London*, *B 211*, 151-180.

McClelland, J.L., Rumelhart, D.E., & the PDP Research Group (1986). *Parallel Distributed Processing*. Cambridge, Mass.: MIT Press.

Minsky, M., & Papert, S. (1969). *Perceptrons*. Cambridge, Mass.: MIT Press.

Navon, D. (1977). Forest before the trees: The precedence of local and global features of forms. *Cognitive Psychology, 9*, 353-383.

Noton, D., & Stark, L. (1971). Eye movements and visual perception. *Scientific American, 224* (6), 34-43.

Neisser, U. (1963). Decision-time without reaction-time: Experiments in visual scanning. *American Journal of Psychology, 76*, 376-385.

Pachella, R.G. (1974). The interpretation of reaction time in information processing research. In B.H. Kantowitz (Ed.) *Human Information Processing: Tutorials in Performance and Cognition*, pp. 41-82, Hillsdale, N.J.: Erlbaum.

Podgorney, P., & Shepard, R.N. (1978). Functional representations common to visual perception and imagination. *Journal of Experimental Psychology: Human Perception & Performance, 4*, 21-35.

Podgorney, P., & Shepard, R.N. (1983). Distribution of visual attention over space. *Journal of Experimental Psychology: Human Perception & Performance, 9*, 380-393.

Posner, M.I, & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. Bowhuis (Eds.), *Attention & Performance Vol. 10*. Hillsdale, N.J.: Erlbaum.

Posner, M.I., Snyder, C.R.R., & Davidson, B.J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109*, 160-174.

Pringle, R., & Egeth, H.E. (1988). Mental curve tracing with elementary stimuli. *Journal of Experimental Psychology: Human Perception & Performance, 14*, 716-728.

Pylyshyn, Z.W. (1980). Computation and Cognition: Issues in the foundations of cognitive science. *The Behavioral & Brain Sciences, 3*, 111-169.

Pylyshyn, Z.W. (1984). *Computation and Cognition*. Cambridge, Mass.: MIT Press.

Pylyshyn, Z.W. (1988). Here and there in the visual field. In Z.W. Pylyshyn (Ed.), *Computational Processes in Human Vision: An Interdisciplinary Perspective*, 210-236. Norwood, N.J.: Ablex.

Pylyshyn, Z.W., & Storm, R.W. (1986). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. University of Western Ontario, Centre for Cognitive Science Memo No. 22.

Rock, I., & Gutman, D. (1981). The effect of inattention on form perception. *Journal of Experimental Psychology: Human Perception & Performance*, *7*, 275-285.

Romanycia, M.H.J. (1987). The design and control of visual routines for the computation of simple geometric properties and relations. University of British Columbia, Department of Computer Science Technical Report No. 87-34.

Sabbah, D. (1985). Computing with connections in visual recognition of origami objects. *Cognitive Science*, *9*, 25-50.

Sejnowski, T.J., & Hinton, G.E. (1987). Separating figure from ground with a Boltzmann machine. In M.A. Arbib and A.R. Hanson (Eds.), *Vision, Brain, and Cooperative Computation*, 703-724. Cambridge, Mass.: MIT Press.

Sekuler, R., & Nash, D. (1972). Speed of size scaling in visual pattern recognition. *Psychonomic Science*, *27*, 93-94.

Shepard, R.N., & Podgorney, P. (1986). Spatial factors in visual attention: A reply to Crassini. *Journal of Experimental Psychology: Human Perception & Performance*, *12*, 383-387.

Sperling, G., & Reeves, A. (1980). Measuring the reaction time of a shift of visual attention. In R.S. Nickerson (Ed.), *Attention & Performance Vol. 8*. Hillsdale, N.J.: Erlbaum.

Swensson, R.G. (1972). The elusive tradeoff: Speed vs accuracy in visual discrimination tasks. *Perception & Psychophysics*, *12*, 16-32.

Tipper, S.P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *Quarterly Journal of Experimental Psychology*, *37A*, 571-590.

Treisman, A.M. (1986). Features and objects in visual processing. *Scientific American, 255* (5), 114-125.

Treisman, A.M. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology, 40A*, 201-237.

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

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

Treisman, A.M., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology, 14*, 107-141.

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

Trick, L.M., & Pylyshyn, Z.W. (1988). Spatial attention and subitizing: An investigation of the FINST hypothesis. In *Proceedings of the Tenth Annual Conference of the Cognitive Science Society*, Montreal, Canada.

Tsal, Y. (1983). Movements of attention across the visual field. *Journal of Experimental Psychology: Human Perception & Performance, 9*, 523-530.

Ullman, S. (1979). *The Interpretation of Visual Motion.* Cambridge, Mass.: MIT Press.

Ullman, S. (1984). Visual routines. *Cognition, 18*, 97-159.

Uttal, W.R. (1975). *An Autocorrelation Theory of Form Detection.* New York: John Wiley.

Ward, L.M. (1982). Determinants of attention to local and global features of visual forms. *Journal of Experimental Psychology: Human Perception & Performance, 8*, 562-581.

Winston, P.H. (1977). *Artificial Intelligence.* Reading, Mass.: Addison-Wesley.

Witkin, A.P., & Tenenbaum, J.M. (1983). On the role of structure in vision. In J. Beck,

B. Hope, & A. Rosenfeld (Eds.), *Human and Machine Vision*. New York: Academic Press.

Wright, R.D. (1988). Developing a visual routine for determining inside/outside spatial relations. Paper presented to the University of Alberta Vision Interest Group, Edmonton, Alberta.

Wright, R.D., & Dawson, M.R.W. (1987). Determinants of the speed of perceiving inside/outside spatial relations. Paper presented at the annual meeting of the Canadian Psychological Association, Vancouver, British Columbia.

Wright, R.D., & Dawson, M.R.W. (1988). Using hardware interrupts for timing visual displays and reaction-time key interfacing on the Commodore 64. *Behavior Research Methods, Instruments, & Computers*, *20*, 41-48.

Wright, R.D., & Katz, A.N. (1989). Detecting contour gaps in unattended figures. Manuscript in preparation.

Yellott, J.I. (1971). Correction for guessing and the speed-accuracy tradeoff in choice reaction time. *Journal of Mathematical Psychology*, *8*, 159-199.

Zucker, S.W. (1987). The emerging paradigm of computational vision. *Annual Review of Computer Science*, *Vol. 2*, 69-89.