Acta Psychologica 83 (1993) 93–154
North-Holland

Visual selective attention: A theoretical analysis

Jan Theeuwes *

TNO Institute for Perception, Soesterberg, The Netherlands

Accepted June 1992

The present paper outlines a framework which allows a consistent interpretation of data regarding visual selection in visual search tasks. It organizes and reviews visual search tasks in which the target is defined by primitive features, by conjunctions of features and when the target is categorically different from non-targets. The special role of spatial attention is reviewed and different theoretical accounts are discussed. Because *visual* selection depends principally on the outcome of the early parallel preattentive stage of processing, the main focus will be on this stage. It is concluded that visual selection is to a large extent determined by the physical characteristics of the stimuli present in the visual field. The early preattentive parallel process computes how different each object is from each of the other objects within a particular stimulus dimension. Attention is automatically drawn to the location having the highest activation, implying that the object at that location is automatically selected irrespective of the intentions of the subject. The model also assumes some top-down control. It is well known that attention can be voluntarily directed to nonfixated locations in visual space, varying from a uniform distribution over the visual field to a highly focused concentration. The model assumes that the endogenous direction of attention to an area in the visual field is the only top-down manner of affecting visual selection. Within the area of directed attention, no top-down control is possible: selection is completely determined by the physical properties of the stimuli.

1. Introduction

The identification of objects in the visual environment plays an important adaptive role in everyday life activities, in particular for acting in a goal-directed manner. A visual scene may contain many objects, that is, potential targets for action, yet the visual system appears to be limited on the number of objects that can be processed at a time. This limitation implies that at some stage (or stages) in the information flow, some objects are excluded from processing. This

Correspondence to: J. Theeuwes, TNO Institute for Perception, P.O. Box 23, 3769 ZG Soesterberg, The Netherlands. E-mail: janthe@izf.tno.nl
* I would like to thank A.F. Sanders and A.H.C. van der Heijden for their valuable suggestions on an earlier version of the manuscript.

0001-6918/93/$06.00 © 1993 – Elsevier Science Publishers B.V. All rights reserved

process of selecting part of simultaneous sources of information, either by enhancing the processing of some objects and/or by suppressing information of others, is traditionally referred to as '*selective attention*' (Johnston and Dark 1986). Theories of human selective attention are concerned with how people select information to provide the basis for responding and with how information, irrelevant to that response, is dealt with.

The main objective of the present review is to examine the selection processes occurring within a single fixation of the eyes. It uses experimental data as a starting point and when necessary will discuss various theoretical approaches to visual search. The paper is divided into five sections. The first section will distinguish processes which should be considered as manifestations of attentional selectivity from processes which are due to the structural constraints of the human eye. In the next section, the two-stage approach is outlined which suggests that visual processing is characterized by two functionally independent stages: an early stage that operates without limitations in parallel across the entire visual field, followed by a later stage characterized by a limited capacity. Because *visual* selection depends principally on the outcome of the early stage of processing, the main focus will be on this early stage. This does imply that selection at later stages of the system is not possible (e.g., categorization of selected items and/or choices among various responses); rather, the present approach confines its concern to *visual selection*, that is, to how items from the first stage of parallel processing are passed on to the second stage of processing. An item entering this second stage of processing is assumed to be *selected*. The third section discusses the extent of parallel processing in different types of search tasks. The fourth section examines the status of research on spatial attention, reviews the results of various cuing experiments, and discusses different theoretical accounts for these data. In the final section, a summary of the conclusions will be provided with respect to three major controversies in visual attention research.

2. Structural constraints

2.1. Selectivity and eye movements

Selective attention in visual perception becomes manifest when orienting in the visual field. One obvious way of visual orienting is overt, involving the movement of eyes and head, or both, so as to overtly focus on an spatial location. This may be considered as a global way of selecting and analyzing certain parts of the visual environment (e.g., Sanders 1963, 1970; Sanders and Donk 1991). Under conditions of free search, observers search a particular field for a possible target while making both eye and head movements. In these circumstances, as in deciding to look first to the left and then to the right before crossing the street, top-down strategies control the macrostructure of the visual scanpath (see, e.g., Lévy-Schoen 1981). Yet, top-down control seems to disappear when a highly salient object is present in the visual field. In the now classic studies of Engel (1977), it was claimed that, as soon as the eye is close to a conspicuous object, the object exerts control over the visual selection system. Engel (1971, 1974, 1977) defined visual conspicuity in terms of a lobe representing the peripheral area around the central fixation point within which an object can be detected with a certain probability within a single glance. In this view, scanning certain parts of the visual environment may depend only on top-down strategies until an eye fixation falls within the conspicuity lobe of a particular object. Then, top-down control disappears and the object exogenously elicits an eye movement toward the object (see Engel 1977).

It may appear that visual selection is equivalent to where the eye fixates in the visual field; yet, such a view is incorrect. It is known that when a saccade is made toward a particular location, *attention* moves to that location *before* the onset of the saccade (Posner and Cohen 1984). In everyday life, a shift of attention is usually followed by an overt movement of the eyes indicating that the eye fixation and attentional locus are usually highly correlated. Yet, in many experiments, it has been shown that movements of attention can occur without making eye movements (e.g., Eriksen and Hoffman 1972; Posner et al. 1980). As a metaphor, properties of this covert orienting have been described as a spotlight (for a detailed discussion see section 5.2.1; see also, e.g., Broadbent 1982) or a zoom lens (Eriksen and Yeh 1985). In these metaphors, the locus of directed attention in visual space is thought of as having greater illumination than the areas to which attention is not directed, or areas from which attention has been removed. In line with the metaphors, many studies have demonstrated that responses to stimuli falling inside the attentional beam are faster and more accurate than responses to stimuli falling outside the beam of attention (e.g., Posner 1980; Posner et al. 1980; Van der Heijden et al. 1987). Again, similar to overt orienting, covert attention may be controlled endogenously by directing attention to a location in visual space or may be captured exogenously by a peripheral sensory signal (Jonides 1980).

The important observation is that, *within a single ocular fixation*, there can be enhanced processing of some objects and reduced processing of other objects in the visual field. The distribution of attention in the visual field can be used to account for the observed selectivity of processing. The point of fixation does not necessarily represent the location from which information is acquired, indicating that visual selection is basically independent of the line of sight. Although the study of eye-movement patterns constitutes an important tool for studying visual selection in tasks requiring a high ecological validity, it should be realized that foveal vision is not necessary for selecting visual information. Therefore, the movements of the eyes should not be considered as the selection process itself, but merely as the outcome of attentional selection processes preceding actual eye shifts. As Broadbent (1982) claimed, eye movements may reinforce selectivity, yet it can occur without them.

2.2. Selectivity and lateral masking

The discussion above suggests that visual selection is basically independent of the line of sight. Yet – and this is often ignored in studies concerning selective attention – processing of spatial information outside the line of sight is rather limited because of a reduced retinal sensitivity in the periphery. As the eccentricity of a target increases, the efficiency of information processing reduces, simply on the basis of the anatomical observation that the density of cone receptors dramatically drops when moving away from the center of the fovea (Yelott et al. 1984). The density of cones 5° from the center of the fovea is about 1/10 that of the foveola (a circular area $1\frac{1}{3}°$ in diameter, Polyak 1957). Therefore, when investigating *postretinal* processes such as selective attention, it is crucial that effects of retinal sensitivity are not confounded with effects of the operation of selective attention.

This drop in accuracy is even more dramatic when other objects are present in the visual field. Bouma (1970) showed that the recognition scores in eccentric vision for randomly chosen target letters between two distractor letters x (i.e., /xax/) dropped sharply as compared to the nonembedded situation (i.e., / a /). The finding that the probability of correctly identifying a target is significantly reduced when there are other items close to the target in the visual field may suggest a selection problem rather than an acuity problem. This effect, generally referred to as lateral masking, is characterized by three properties: (1) lateral masking is more pronounced in the periphery than in the center of the visual field (see e.g., Bouma 1970, 1978), (2) the masking effect of the surrounding item diminishes as the space between target and mask is increased (e.g., Eriksen and Eriksen 1974), (3) a mask placed on the peripheral side is more effective than a mask placed on the foveal side (e.g., Andriessen and Bouma 1973). Since these properties appear to be sensory in nature, most accounts assume that lateral masking results solely from interactions at a sensory level, that is, features of the mask interact with features of the target (for sensory accounts see e.g., Andriessen and Bouma 1973; Estes 1972; Wolford 1975).

In summary, it is crucial to ensure that performance differences observed in various tasks reflect differences in selective attention rather than in structural constraints of the human eye. Note for example, that, independent of the distribution of attention, foveation might facilitate the processing of stimuli, especially in conditions requiring a high visual acuity (e.g., reading). Yet, it is also inappropriate to try to explain performance differences only in terms of sensory interactions without considering possible effects of selective attention at postretinal stages of visual processing. For example, Wolford and Chambers (1983) showed that some effects of lateral masking are inconsistent with sensory interpretations and appear to be related to the distribution of attention in the visual field.

3. Two-stage approach

3.1. Introduction

The present approach is based on the idea that visual information processing consists of two functionally independent, hierarchical stages: An early, *pre-attentive* stage (Neisser 1967) that operates without capacity limitations and in parallel across the entire visual field, followed by a later, *attentive* limited-capacity stage that can deal with only one item (or at best a few items) at a time. When items pass from the first to the second stage of processing, these items are considered to be selected. This central tenet dates back to Broadbent's (1958) classical 'filter' theory and forms the basis of currently influential accounts of visual selection and attention, notably: Treisman's 'feature integration theory' (FIT; e.g., Treisman and Gelade 1980; Treisman 1988; Treisman and Sato 1990), Julesz's 'texton' theory (e.g., Bergen and Julesz 1983; Julesz 1971), Cave and Wolfe's 'guided search' model (1990; Wolfe et al., 1989), Hoffman's two stage model (1978, 1979), and various 'late selection' accounts (e.g., Duncan 1980; Duncan and Humphreys 1989).

The purpose of this section is to describe how recent studies illuminate currently active theoretical issues with respect to the two-stage approach of visual search. First, there will be a discussion on parallel and serial search and the properties that characterize these types of processing. Theoretical controversies regarding these properties will be briefly sketched. In the next section, these controversies will be discussed in more detail in relation to data stemming from three widely used visual search tasks: (1) low-level search tasks in which a target is defined by a single primitive feature, (2) search tasks in which target detection requires the integration of information from two or more separable features, (3) search tasks in which the target is categorically different from the non-targets.

3.2. Two types of search

3.2.1. Preattentive parallel search

The most direct evidence that at least some perceptual operations occur in parallel is provided by visual search tasks (Egeth et al. 1972; Neisser et al. 1963) in which a target is detected with little or no change in reaction time (RT) as the number of non-target items is varied. The consistent pattern of results is a flat or almost flat search function (less than 5 or 6 ms per item; Treisman and Souther 1985; less than 10 ms per item: Treisman and Gormican 1988), in which the detection latency is related to the number of non-target items in the display when the target is present. This particular pattern of flat functions in visual search is referred to as the *pop-out effect* (Treisman and Gelade 1980) and indicates that the operations underlying search are performed spatially in parallel. Typically, a target which is defined by a single physical feature not shared by any other items in the display (e.g., a red item between green non-targets) pops out of the display regardless of how many non-targets are present.

Although inferences from linearly increasing search functions have been critized (Townsend 1972), flat RT functions can only be reconciled with models assuming parallel search of the critical feature that defines the target. It should be realized, though, that search functions are seldom completely flat for reasons that have nothing to do with attention (Duncan 1980). As the number of non-targets increases, so does the chance that a non-target is mistaken for the target simply on grounds of probability (Eriksen and Spencer 1969).

Preattentive parallel processing has been characterized by three basic properties (Folk and Egeth 1989; Posner and Snyder 1975):

(1) Preattentive processing is unlimited in capacity. The absence of an effect of the number of non-targets in the display (flat RT function) suggests that preattentive processing is insensitive to perceptual load. This implies that this process satisfies the load-insensitivity criterion of automaticity, stating that automatic processes are not affected when concurrent information load is increased (e.g., Neumann 1984).

(2) Preattentive processing is spatially parallel operating simultaneously at various locations across the visual field. The absence of a display effect can only be understood when it is assumed that preattentive search takes place across all locations in the visual field at the same time (Townsend 1972). Although pre-attentive means '*before* attention operates', it has been claimed that preattentive search is search in which attention is widely distributed over the whole display rather than narrowly focused and serially directed to one object at a time (Treisman and Gormican 1988)

(3) Preattentive processing operates independent of strategic control. Evidence for the property that preattentive search operates independently of strategic control is rather ambiguous. It refers to the unintentionality criterion of automatic processes, which states that 'automatic processes are under the control of stimulation rather than under the control of the intentions (strategies, expectancies, plans) of the person' (Neumann 1984: 258). Note that if this claim were correct, a subject's knowledge of which target should be found, would not have an effect on the outcome of the preattentive process. Recent studies

have claimed that the preattentive stage does not operate solely data-driven: knowledge about the target to be found may give additional top-down activations in the relevant feature map (Cave and Wolfe 1990), may bias and alter the initial weights given to the items in the visual field (Duncan and Humphreys 1989), or may inhibit feature maps which represent a non-target value (revised FIT, Treisman and Sato 1990).

3.2.2. Attentive serial search

Search functions reflecting parallel search can be contrasted to functions showing a linear increase in RT as the number of non-target items in the display is increased. This pattern of results has been taken as indicative of spatially serial search. The finding that the slope of target absent trials is twice as steep as the slope of target present trials is taken as evidence that serial search is self-terminating (Sternberg 1966). On positive (target present) trials, subjects stop searching as soon as the target is found, whereas in negative trials search continues until the whole display has been checked. Serial functions are usually found in cases when targets are defined by specific arrangements (e.g., search for a T among L's) or when they are defined by conjunctions of features (e.g., a conjunction of color and shape, a red X between red O's and green X's).

The finding that search time increases with display size does not necessarily prove that the underlying processes are performed serially (Townsend 1972). It may be that a linearly increasing search function reflects 'limited' capacity parallel search rather than serial search. This would imply that differences in the slopes of the search functions reflect the amount of capacity required, rather than a qualitative difference between parallel and serial search. For example, there is a small capacity required for searching an item with unique color, whereas a larger capacity might be required for more complex identifications (e.g., T between L's). Without discussing the issue of 'limited versus unlimited capacity' in visual search, it is argued that an increase in RT as a function of the number of non-targets (display size) reflects a qualitatively different search operation than when RT is independent of display size. More specifically, when search functions show a substantial positive slope, it is assumed that search is performed serially, involving the attentive stage of processing. On the other hand, when search functions are essentially flat, it is assumed that search is performed in parallel, involving the preattentive stage of processing. Note, however, that the dichotomy between preattentive and attentive processing might not be as sharp as presented here (see, e.g., Duncan and Humphreys 1989), especially in conditions in which there is not a large difference between target and non-targets (Treisman and Gormican 1988). Yet, there is usually a large contrast between parallel search involving pop-out's, and patterns of RT characterizing serial search. Therefore, in line with the two-process theory of Schneider and Shiffrin (1977; Shiffrin and Schneider 1977) it is assumed that there is a qualitative difference between parallel 'automatic detection' and serial 'controlled search'.

There is not much consensus about the properties of the attentive stage.

(1) Attentive processing is limited in capacity. This claim is simply the result of the observation that search time increases with display size. The question is: where is the capacity limited? There are various theoretical viewpoints, but they all agree more or less that a second stage is necessary in order to respond to the target (Treisman 1988; Duncan 1980). There is also agreement that some type of *selection* takes place during this stage. Yet, there are divergences on the type of operations performed by the second stage. On the one hand 'late selection' theories assume that the attentive stage does not perform any perceptual operations, but only selects between competing response tendencies arising from multiple stimuli. It is assumed that all stimuli in the visual field are already fully identified at the parallel stage, and that only the target will enter the second stage (Allport 1980; Duncan 1980). On the other hand, early selection views claim that the second stage performs perceptual operations which cannot be performed by the first parallel stage. For example, the second stage is thought to perform cross-dimensional integration of primitive features (Treisman and Gelade 1980); is necessary for conducting finer discriminations on potential targets (Hoffman 1978, 1979); and is required for the localization of objects in the visual field (Treisman and Gelade 1980; Cave and Wolfe 1990). Note that the assumption that capacity is limited is based on the finding that in many visual search experiments, RT increases with display size. Therefore, it is not immediately critical to consider reasons where capacity is limited. It may very well be that capacity is limited because of possible percep-

tual overload (Kahneman and Treisman 1984) and/or conflict in response systems (Neumann 1987, 1990; Van der Heijden 1992).

(2) The attentive stage operates on a limited spatial location. Because the parallel stage operates across all locations in the visual field, it may be argued that the increase of RT with display size reflects the time it takes the attentive stage to serially inspect single items. In this sense, the second stage of focal attention is, as a metaphor, equated to an attentional spotlight (Posner 1980; Treisman 1988) or possibly a zoom-lens (Treisman and Gormican 1988) which serially searches smaller areas within the visual field, causing a linear increase in RT with display size. Therefore, focusing attention on a location in the visual field implies that the item appearing at that location is *selected*, that is, this item enters the second stage of processing. Note that 'late-selection' theories do not assume a special role for spatial attention (see section 5.2.2). Alternatively, it has been claimed that the obtained steeper slopes in typical conjunction search tasks are not due to limitations in the second stage of processing, but merely reflect a reduced power to accurately discriminate between target and non-targets (Duncan and Humphreys 1989). In a late selection view, it is assumed that selection and recognition are aspects of the same process rather than two different stages of processing (Bundesen 1990), suggesting that there is not a qualitative difference between parallel and serial search.

(3) Operations of the attentive stage are controlled strategically. When attentive processing is conceived as a zoom-lens, (see section 5.2.1) which can vary along a continuum from completely divided attention spread out over the display as a whole to sharply focused attention to one item at a time (Treisman and Gormican 1988), it can be claimed that the size of the beam can be set strategically (Humphreys 1981). In this sense, there is a clear top-down control from which location and how much information will *enter* the second stage of processing. Yet, items entering the second stage are already selected, suggesting that, purely on logical ground, it is impossible that top-down effects operating at the second stage of processing affect *visual selection* (see Introduction for a definition of *visual* selection). Top-down effects at this stage (e.g., knowledge regarding properties of the target) merely reflect post-perceptual decision processes (i.e., changes in beta), suggesting that subjects require less evidence to decide that the selected item is in fact the target (e.g., Duncan 1980).

4. Different types of search tasks

4.1. Search for targets defined by primitive features

There is a general consensus that visual scenes are encoded along a set of primitive 'feature' dimensions – such as orientation of edges, width (or spatial frequency), color, brightness, etc. – at the early, parallel and preattentive stage of processing. This early coding parses the scene into separate regions defined by differences in these primitive features, thereby establishing candidate objects for later identification. According to Treisman's FIT and related accounts, the basic features of an object can only be combined into complex object representations at the later stage of processing requiring serial ('focal') attention.

Preattentive parallel search for primitive visual features is demonstrated by 'present–absent' visual search tasks in which subjects are asked to make a speeded decision whether a target defined by a specific simple visual feature is present or not (e.g., Bergen and Julesz 1983; Treisman and Gelade 1980). Thus, Treisman and Gelade (1980: exp. 1) showed that search for a blue letter or an 'S' among brown 'T's' and green 'X's' gave search functions which were essentially flat (for the color condition, 3.8 ms/item; for the shape condition, 2.5 ms/item). For target-absent responses, reaction time linearly increased with display size, probably due to some 'recheck' strategy after the parallel process. The typical flat search pattern suggests that a target 'pops out' of a display, and this observation is treated as evidence that its defining property forms part of the preattentive representation (Treisman 1986). In search for the existence of early separable features, the occurrence of a pop-out is used as a diagnostic tool (Treisman 1988). In a recent review, Enns (1990) provided a list of 2-dimensional features that pop out in visual search. In addition to two-dimensional features, it was demonstrated that three-dimensional orientation of objects and direction of lighting in a scene pop out as well (Enns 1990).

Besides pop-out effects in visual search, the notion of an early perceptual analysis of a particular set of primitive features is supported by a variety of other findings, including physiological recordings from specialized populations of neurons (e.g., Livingstone and Hubel 1987); perceptual aftereffects (Houck and Hoffman 1986);

texture segregation determining the properties that parse a scene into figure and ground (e.g., Beck 1967; Pashler 1988); and illusory conjunctions in which the features from one object are conjoined with the features from another simultaneously present object (Treisman and Schmidt 1982).

Given these findings, Treisman has proposed a feature integration theory of preattentive and attentive processing (FIT; Treisman and Gelade 1980; revised FIT; Treisman and Sato 1990). According to this theory, parallel search occurs over the whole visual field when the target item is defined by a distinctive, preattentively available feature, which non-target items do not share. In Treisman's FIT, it is hypothesized that parallel search occurs because the primitive features are registered separately in different feature maps. A related set of features or a continuum of maps represents a perceptual *dimension* (Treisman and Souther 1985). Thus, for example in the color dimension, there is the color map red, the color map blue, the color map yellow etc. In line with the FIT, is the observation that a target defined by the absence of a critical feature does not pop out from the background. Thus, a target circle with an intersecting line segment (similar to a 'Q') pops out from a background of circles ('O's'). Yet, the opposite does not hold: a circle without an intersecting line segment does not pop out between non-target items having a target line segment (an 'O' between 'Q's') (Treisman and Souther 1985). In addition, serial search is required when the target and non-targets only differ quantitatively on the relevant dimension (i.e., search for target line segment which is somewhat longer than the non-target line segments; Treisman and Gormican 1988). More generally, attention is required when features have to be located and conjoined to specify objects. Attention can select information from the 'master map of locations' which shows *where* primitive feature boundaries are located.

4.1.1. The type of operations performed when searching for targets defined by primitive features

The parallel stage provides information about the presence of a distinctive feature in the visual field. In line with the FIT, the presence of an object with a unique feature, not shared by any other object in the visual field, will result in activity in a 'possibly prespecified' feature map. Thus, deciding whether a red target between green non-targets is present in the display, can be based solely on preattentively available activity in the red color map. Note that in a typical present–absent search task, this information is enough to determine the appropriate response. Without actually *identifying* the target, responses can simply be based on detection of a single difference among features (Folk and Egeth 1989; Pashler and Badgio 1985). Yet, according to the FIT, a high activity in a feature map results in a 'pop-out', indicating that the highly distinctive feature calls attention to itself (Treisman 1988). According to Treisman (1988: 226), this calling of attention is the basis for the 'pop-out' phenomenon. This would imply that a pop-out is always mediated by an automatic shift of spatial attention to the location containing the unique feature (see also Hoffman et al. 1983; Neisser et al. 1963). The item at that location is passed from the preattentive parallel stage to the attentive serial stage of processing, implying that the item at that location is *selected*. Therefore according to the FIT, the presence of a highly distinctive feature results in an automatic (unintentional) selection of the 'popping-out' object. Note that when the item is selected, (i.e., passed on to the second stage) information regarding the location of the distinctive feature becomes available as well (e.g., where is the red target). If features can be identified without the need for spatial attention, it can be inferred that an overt response can be based solely on the processing occurring at the preattentive stage (Folk and Egeth 1989). In other words, information from the first stage does not need to be passed on to the second stage before a response can be given, implying when merely searching for primitive features, there are no attentional limitations at all.

Experiments seeking to clarify whether spatial attention is required for responding provide rather ambiguous results. In a same–different discrimination task, subjects were instructed to indicate whether all elements of a set of target stimuli had the same orientation, or whether one of them had a different orientation. In Sagi and Julesz's (1985a) study, target elements (1 through 4) were horizontal or vertical line segments embedded in a background consisting of diagonals that were all oriented in the same direction. Because the horizontal and vertical line elements requiring a same–different discrimination were embedded among textured elements, the task could not be performed by simply detecting a difference among features, as in a typical present–absent search task. Detection accuracy as a function of SOA revealed that the decision whether 1 through 4 target elements had all the same orientation or whether one was different could not be performed spatially in parallel. Sagi and Julesz concluded that the local detection of differences among features might have been performed in parallel; yet, the identification of each of the potential targets required the serial allocation of attention in the visual field. This conclusion is in line with the claim that spatial attention is necessary for target identification and responding. Using a similar same–different paradigm with RT as dependent measure, Folk and Egeth (1989) challenged this claim and suggested that feature identities are processed and are simultaneously available for responding at the preattentive level. Yet, this conclusion seems not quite substantiated by their data. In fact, their experiment 1 (and exp. 4 of Egeth et al. 1989) replicated Sagi and Julesz's results. In experiment 2, Folk and Egeth varied stimulus quality and showed that the effect of visual quality did not have an overadditive interaction with that of target number as was expected when search was serial (same logic as Pashler and Badgio 1985). Based on these findings, Folk and Egeth suggest that identification of simple features does not require a serial scan as claimed by Sagi and Julesz (1985a; Julesz's texton theory, Bergen and Julesz 1983). Yet, this claim can be questioned: as suggested by Folk and Egeth, it seems likely that the preattentive process cleans up the degradation followed by a subsequent serial processing of feature identities.

A study conducted by Quinlan and Humphreys (1987) suggests that detection (respond target present–absent) of targets defined by primitive features also requires focal attention. In some conditions, subjects searched for two targets simultaneously (i.e., one with a unique shape, one with a unique color) and had to respond 'present' when both targets were present and 'absent' when one or both targets were absent. Results suggested spatially parallel search for two targets defined by two different feature dimensions. Yet, the detection of the targets made available by the preattentive process could *not* be performed in parallel: focal attention had to be switched serially to each of the two target locations.

Thus, it is not immediately clear what type of operations are performed when searching for a target defined by a primitive feature. At one extreme, there is the position of Folk and Egeth (1989) who claim that all feature identities are processed preattentively and are simultaneously available for responding. Along similar lines is the position of Duncan (1980) who claims that the first parallel stage completes the whole input analysis, and that targets only enter the second stage so as to become available for the response mechanism. Rather than claiming that the parallel stage performs a full identification of all items, it is also feasible that the parallel stage only computes *differences* among features (see e.g., guided search model, Cave and Wolfe 1990; Theeuwes 1991c, in press). After this stage, the identity of the elements in the display is still unknown; the only outcome of the parallel stage is some kind of activation map representing how different each element is from each of the other elements within a particular feature dimension (e.g. color dimension, form dimension etc). Thus, after the parallel stage there is no information about the source of the activation (i.e., whether it is due to a difference in color, form, etc.). A single red item between several green items will generate a high activity at its location because it differs from all of the other items. The activity at the location of other (green) items is not very high because the green items differ only from the single red item but not from each other.

Note that the data of Quinlan and Humphreys (1987) are compatible with this last suggestion. In conditions in which two targets, each defined by a separable feature were present, it was claimed that the preattentive stage made the targets available, followed by a focal serial stage required for responding. Quinlan and Humphreys (1987) interpreted their results in line with a 'late' selection account: the targets are supposed to be fully identified at the first stage and a serial process is necessary only because a limited capacity 'central decision mechanism' is required in order to respond to the targets (see Duncan 1980). Such a claim is very odd since the targets did not require separate responses, but only a single response, that is, 'present' when both targets were present. Therefore, the claim that serial processing was necessary because of the capacity limitation of the central decision mechanism seems at least dubious (see also Theeuwes 1991a)

It seems plausible to assume that identification occurs at the second stage of focal attention, and that the first stage only calculates differences between features within a particular dimension. If the target is defined by a primitive feature that differs from each of the other items, then it will generate the highest activity, and will therefore be selected. It will, then, appear as if full identification took place preattentively and that only targets entered the second stage of processing; yet, this is only a consequence of the fact that the target generated the highest activity.

Evidence for the claim that only the item with the highest activity is selected, comes from a study from Theeuwes (1991c, in press) in which it was shown that, during preattentive parallel search, the selection priority depended upon the differences between features within the color and form dimensions. When searching for a unique color (i.e., searching for a green item between red items), a unique form (i.e., a square between circles) did not interfere. On the other hand, when searching for the unique form, the unique color – known to be irrelevant – greatly interfered. When the color difference was reduced (e.g., searching for yellow between orange), this relationship reversed; search for form was not hindered by the presence of the irrelevant color; yet, search for color was hindered by the presence of the irrelevant form. These findings suggest that the differences between the elements with respect to the color and form affect the selection priority. If *all* items had been identified preattentively, then there would be no reason why a unique color or form, known to be irrelevant, would interfere.

This suggest that the parallel stage, which takes care of the detection of differences in features, guides the serial stage in its selection process. Thus, when searching for several red items in a display with green non-target items, search is serial for the red items suggesting a parallel rejection of all non-target items. The basic viewpoint advocated here, is that the first preattentive stage can only carry out some rough discriminations in the sense of computing differences between the items in the display. Before perceptual processing is complete, attentional selection by means of the second stage is necessary. This 'early selection' point of view can be contrasted with a 'late selection' approach (e.g., Deutsch and Deutsch 1963) which claims that all stimuli receive complete perceptual processing before selection takes place.

4.1.2. Properties of processing operations

As indicated earlier, it is unclear whether preattentive parallel processing operates independently of strategic control. The implication of no strategic control is that the preattentive process starts calculating differences among features for each of the feature dimensions (color, form, brightness), as soon as a display is presented to the subject. This computation is assumed to be independent of strategic control and to occur irrespective of whether an item is a target or a non-target. The computation does not depend on target value expectations, will be the same even when nothing is known about the target, and is the basis of what is called the visual 'pop-out' effect (Cave and Wolfe 1990). It indicates that the item generating the highest activity within a particular dimension pops out and calls attention to itself, thereby entering the second stage of processing. In other words, the item with the highest bottom-up activity (i.e., the 'oddest' or most salient item in the display) is selected irrespective of the intentions of the subject.

Some experiments confirm this view. For example, Treisman (1988) showed that knowing what the target was (whether it is blue, red or white between green non-targets), was hardly faster (19 ms) than the condition in which subjects did not know the color of the target. Obviously, expectations regarding the target did not help much. Yet, in a condition in which subjects did not know in which dimension the target would be presented (whether it would be search for a unique color, a unique orientation or a unique size), a rather large increase was found (about 90 ms). Treisman claims that the 'odd one' pops out only within a single, pre-specified dimensional module. These latter findings indicate that there is no top-down selectivity within dimensions. Yet, in experiments regarding selectivity between dimensions, Treisman (1988) showed that knowing the *dimension* of the target gave faster search than when nothing was known about the target. These findings were not confirmed by texture segregation experiments of Pashler (1988). In his experiments subjects searched for a target that was unique in a given dimension (e.g. color or form). For example, subjects located a specific form (e.g., an 'O' between '/') or a specific color (red element between 'green' non-targets) within a display of 90 elements. His experiment 5 showed that specification of the target dimension in advance had negligible effects on RT. Experiment 6 and 7 showed that the presence of a single colored target, known to be irrelevant, interfered with search for a unique form. Pashler concludes that the detection of a unique singleton is mediated by 'discontinuity detectors', which operate irrespective of any attempts by the subject to suppress such an operation.

Theeuwes (1991c, in press) provided clear evidence that top-down selectivity during preattentive parallel search was not possible. These studies showed that if processing information along one dimension proceeds in parallel, information from another dimension cannot be intentionally ignored. The stimuli consisted of 5, 7 or 9 items (i.e., circles or squares) that were arranged in a ring. All items contained a slightly tilted (22.5°) line segment, except the target, which was either horizontal or vertical. Subjects were asked to determine the orientation of the target line segment. Note that the task used in these experiments is concerned with what Duncan (1985) has called 'compound search' in which the stimulus information, separating target from non-target, tells nothing about which of the possible responses to choose. In this way it is possible to distinguish perceptual selection factors from factors involved in responding, such as response competition (e.g., Van der Heijden 1992) and post-perceptual response bias (e.g., Duncan 1980). The target line segment was always located in an item which was unique within a certain dimension (e.g., the target line segment was located in an item with a unique form for one group of subjects, and in a unique color for another group of subjects). In the neutral condition, there was only one unique item which contained the target line segment (e.g. a red item between green non-target items). In the different dimension condition, there were two unique items: one along the task relevant dimension (e.g., item with a unique color) and the other along a task-irrelevant dimension (e.g., an item with a unique form). The results showed that *all* search functions were essentially flat, that is, independent of the number of items on display, indicating that search for the unique dimension was performed in parallel. The results also showed that during preattentive parallel search, top-down selectivity towards the task-relevant stimulus dimension was not possible: the presence of the item unique in the task-irrelevant dimension interfered with search for the item unique along the task-relevant dimension. Note that this interference did *not* show up as a display size effect: in the conditions in which a distractor was present the search function remained flat but was higher in comparison to the condition in which no distractor was present. Important was the finding that the interference depended on the relative discriminabilities of the color and form dimensions. For example, when searching for an easy-to-be-discriminated color (green between red) a unique form (a circle between squares) hardly interfered; yet, when the color difference was reduced (searching for orange between orange/yellow) the presence of a unique form greatly interfered.

The study of Theeuwes (1991c) shows that there is no top-down selectivity toward stimulus dimensions. Yet, this does not necessarily imply that knowledge regarding the exact target properties could not have resulted in selective search. In Theeuwes' (1991c) study, subjects only knew the target was unique within a certain dimension; yet they did not know the exact properties of the target (i.e., whether the target was green or red, squared or circular). In an additional study (Theeuwes in press), subjects searched during the whole experiment for a green circle among green squares (form condition) or red circles (color condition). Thus, subjects knew exactly the form and color of the target; and target was never changed during the whole experiment. Again, as found in the 1991c study, the presence of a distractor in the other dimension showed interference which depended on the relative discriminability of color and form. In addition, this study shows that even after extensive and consistent practice, top-down selectivity cannot be obtained.

The finding that the relative discriminability of the dimensions could account for the obtained selectivity during preattentive parallel search suggested that the preattentive process computes for each location in the visual field *differences* in features, separate for each feature dimension. Thus, at each location in the visual field there is a separate activation for color, for form, for brightness etc. Similar to the guided search model of Cave and Wolfe (1990), the present model assumes that the activation for a particular dimension (e.g., color) is the mean of the differences between the stimulus at that location and the stimuli at all other locations. Given the notion of difference signals at each location, access to the second stage of attentive processing may occur when a particular criterion is reached. For example, when access depends on an absolute activation level criterion, the item occurring at the location having the highest activity obtains unintentional access to the second stage of focal attention. Thus, focal attention is attracted to the location producing the strongest pop-out. Alternatively, when access depends on a threshold criterium, the item occurring at the location having an activity exceeding a particular threshold first obtains first access to the second stage of processing. Thus, focal attention is attracted to the location that pops out first. In either case, if the item entering first is not the target, search proceeds serially to the nextmost active location (absolute activation criterion), or to the next item popping out (threshold criterion). When the target is found, search stops. It is most likely that the activation within each dimension builds up according to, for example, an e-power function, implying that both criteria would give the same selection order.

The notion above can easily explain Theeuwes' findings. For example, when searching for a unique color in the presence of a form distractor, the relative discriminability determines the outcome of the selection process. Thus, when searching for an easy-to-be-discriminated color (green between red), the color activation at the location of the green item will be higher than the form activation generated by the unique form, so the uniquely colored item will be selected automatically without interference from the unique form. Yet, when the color difference is reduced, the color activation at the location of the uniquely colored item is not as high as the activation generated by the location of the form distractor, causing an unintentional selection of the distractor. The next most active location is the location of the uniquely colored target, implying that focal attention is then unintentionally switched to the location of the target.

The model assumes that there is no top-down influence on the operation of the preattentive parallel process; at least not when searching for primitive features. The visual system simply calculates differences between features resulting in a pattern of activations at different locations. The preattentive process has no access to the origins of these activation levels (i.e., whether activations are caused by differences in form, color, brightness etc.). Obviously, knowing the properties of the target (whether it has a unique color or form) cannot affect the operations of the preattentive process because this information is not yet available at the preattentive level. Only after entering the second stage of focal attention (i.e., after being selected) this information becomes available. Therefore, knowing the properties of the target can only affect processes occurring in this second stage of processing; that is, after the item has been *selected* (e.g., it may speed up target identification). This view is very similar to what Sagi and Julesz (1985a) and Ullman (1984) claim: parallel processes are limited to *local mismatch detection*, followed by a serial stage in which the most mismatching areas are selected for further analysis.

Yet, most other theories claim this conjecture is too extreme: if selection would be completely bottom-up, it would be very inefficient. Thus, most theories assume some variable internal description of which type of information is currently needed by the task.

In the original FIT, it is assumed that knowing the properties of the target only helps when the distractor is an item on another dimension. Treisman (1988) claims that subjects simply check for activity signalling a contrasting item in the relevant target-defining module, and ignore the others. Yet, if the distractor is in the same feature dimension as the target, it will produce activity in the same 'relevant' module, resulting in slow serial search. In contrast to what is claimed above, the original FIT suggests that the preattentive process does not simply calculate local mismatches but reveals the origins of the activation levels (caused by differences in form, color, etc.) as well, so that only the item in the task-relevant dimension (e.g., the target dimension) pops out. Theeuwes (1991c, in press) showed that this conjecture is questionable: Interference between dimensions occurs irrespective of knowledge about which dimension contains the relevant item.

Most models of visual search assume a top-down component in order to account for conflicting data obtained with conjunction search (see section 4.2); yet, they should be compatible with findings involving search for primitive features as well. In the Cave and Wolfe guided search model (1990), top-down activity depends on the knowledge of the target to be found. The model assumes that the bottom-up activations caused by the differences in feature dimensions can be altered by top-down activations. For example, knowing that the target has a unique color may raise, in a top-down way, the activation in the color map, so that the target receives a higher activation than the distractors that are unique in other dimensions. In the revised FIT (Treisman and Sato 1990) a top-down effect is proposed which is similar to Cave and Wolfe's model: it is assumed that a particular feature map representing a non-target value can inhibit the associated locations in the master map, reducing the activity in all non-target locations.

Although most recent models of visual search assume some top-down control on the preattentive stage in selectively guiding items to the serial stage, the findings of Theeuwes (1991c, in press) can best be explained by assuming no top-down control on the operations of the preattentive parallel stage. In terms of FIT, revised FIT, and the guided search model, one could account for this inconsistency by claiming that the activation caused by the distractor was so large that additional top-down activation in the (target) relevant map could not prevent the distractor from being selected. Yet, in experiment 3 of Theeuwes (1991c), the color difference was about the same as the form difference, and even then, complete selectivity was not possible. When the bottom-up components give similar activation levels, one would have expected that additional top-down activation in the relevant map would result in perfect selectivity.

In summary, during preattentive parallel search top-down selectivity towards a particular stimulus dimension (Theeuwes 1991c) or particular stimulus properties (Theeuwes in press) is not possible. In terms of properties that define automaticity, it appears that the preattentive process is strongly automatic because the absence of a display size effect indicates that the process is insensitive to perceptual load thereby satisfying the load-insensitivity criterium. In addition, it appears that the process cannot be controlled strategically thereby satisfying the unintentionality criterium (see section 3.2.1)

If the preattentive process unintentionally calculates differences between features and if the item at the location having the largest difference signal would automatically enter the second stage of processing, one might expect that salient features would unintentionally pop-out. This would imply that the visual system selects objects that stand out from the environment, irrespective of the intentions of the subject. In fact, such a view is adhered by theories concerned with visual search involving eye movements in large displays. These theories claim that as soon as the eye is close to a conspicuous object, top-down control disappears and selection is completely controlled by the object (e.g., Engel 1977).

Yet, recent studies show that this view is incorrect: pop-out effects in visual search can be strategically controlled (Jonides and Yantis 1988; Theeuwes 1990). It was shown that an item with a unique brightness, color, or shape would not attract attention when these attributes were irrelevant to the task. In Theeuwes' (1990) study, subjects viewed multi-item displays in which one item had either a unique form (experiment 1) or a unique color (experiment 2). In the control condition, the unique item always contained the target line segment: uniqueness was a reliable cue for target line segment search. In the experimental condition, the target line segment was located equally often in all display elements: item uniqueness was unrelated to the position of the target line segment. In the control condition, there was parallel search for the unique item which revealed the position of the target line segment. This implies that the unique item was salient enough to pop-out from the background. Yet, in the experimental condition in which the item's uniqueness was unrelated to the position of the target line segment, the unique item did not receive a priority treatment over any other item in the display. In other words, the unique item did not pop out when it was irrelevant to the search task. Note, that the search function in the experimental condition indicated that search for the target line segment was performed serially.

The results of this study appear to be inconsistent with the earlier claim that the preattentive process operates automatically and selects items unintentionally that stand out from the background. If no top-down control is possible over the preattentive process then, irrespective of the task demands, unique items should pop out from the background. The results show that the pop-out, in fact, did depend on the task demands.

To solve this apparent puzzle, it is assumed that subjects can strategically control the distribution of spatial attention in the visual field. Spatial attention conceived as a processing resource can be strategically varied from a uniform distribution over the entire field to a highly focused concentration (Eriksen and Yeh 1985; Eriksen and St. James 1986). As a metaphor, the size of the area of concentration is supposed to vary like a zoom lens (Eriksen and Yeh 1985). In the last described experiment, dependent on the condition, subjects may have intentionally varied the distribution of attention in the visual field. In the control condition, subjects may have distributed their attention over the entire visual field thereby letting the odd item attract attention automatically. Yet, in the experimental condition, in which the target line segment can only be found by means of focal attention, subjects may have concentrated all attentional resources to a circumscribed area in the visual field. Because attention is focused, the odd item may have been located outside the attentional beam so that it could not attract attention. In terms of the two-stage model, the focusing of attention implies that the preattentive stage is *omitted* so that search for the target line segment proceeds serially.

According to this line of reasoning, strategic control over visual selection is only possible through selectively varying the span of spatial attention. It is assumed that within the attended area, no strategic control is possible. The size of the attended area may be adopted according to the task demands. In this way, spatial attention operates like a filter selecting areas for further processing. The role of spatial attention as a filtering mechanism has been demonstrated by a study of Theeuwes (1991b; also Yantis and Jonides 1990) showing that endogenous focusing of attention to a location in visual space precluded attention attraction by abrupt visual onsets (experiment 1) and offsets (experiment 2) presented elsewhere in the visual field. Yet, when attention was not focused in advance, abrupt onsets and offsets did attract attention automatically. This study demonstrates the role of spatial attention as a filtering mechanism: when attention is unfocused, covering the whole visual field, items appearing at the location of the abrupt transient are automatically selected. When subjects strategically zoom in on a particular area, transients well outside the circumscribed area do not affect performance. Yantis and Jonides (1990) showed that when attention is not rigidly directed to a particular location (e.g., using a low cue validity), onsets elsewhere in the visual field may occasionally attract attention. This may be interpreted as a partial failure of the spatial filter, suggesting that attention may have spilled over to locations containing the transient.

The notion above is in line with a whole series of experiments by Eriksen and his collaborators (e.g., Eriksen and Eriksen 1974; Eriksen and Hoffman 1972; Eriksen and Schultz 1979) in which it was shown that non-target stimuli may slow down responses if they are spatially close to the target, and yet have no effect when they are further away. In effect, this view suggests that spatial location is a very special factor in the control of attention (e.g., Broadbent 1982; Hoffman 1986; Theeuwes 1989), operating as a strategic filter capable of blocking stimuli from further processing. Note that this filter is far from perfect and that suboptimal conditions (see Yantis and Johnston 1990) easily lead to what Broadbent (1982) called 'the breakthrough of the unattended'. It should be mentioned that the effects reported by Eriksen and colleagues (Eriksen and Eriksen 1974; Eriksen and Hoffman 1972) also have been explained without assuming attentional processes; non-targets closer to the target produce more interference because they fall on positions of higher retinal acuity than non-targets further away (Van der Heijden 1992).

The earlier discussed findings that subjects cannot selectively attend to either the color or form dimension (Theeuwes 1991c, in press) is in line with the notion of spatial filtering. If subjects want to find the task-relevant dimension by means of parallel search, attention has to be spread out over the visual field. As a consequence of this strategy, all events occurring within the spread out beam of attention are admitted for further processing and must be admitted irrespective of whether the person intends it or not (Broadbent 1982). Thus, within the beam of attention, top-down control is lost and preattentive processing occurs unintentionally; allowing the item with the highest bottom-up activation to enter the second stage of attentive processing.

The notion that top-down control in visual selection is only accomplished by varying the size of the attended area is in line with the so-called 'group scanning hypothesis' suggested by Treisman and Gormican (1988). When fine discriminations between targets and non-targets have to be made, subjects may reduce the size of the attentional spotlight, resulting in search through subgroups, checking items within groups in parallel. Within the focus of attention, activation is pooled for each feature map, giving an assessment of the likelihood that a particular feature coded by the map is present in the attended area. It is unclear whether the size of the attended area can be varied strategically on the basis of target–non-target discriminability. Yet, with a different paradigm such a strategic adjustment has been demonstrated. Eye movement research shows that subjects adapted the size of a saccade to the *expected* conspicuity of the searched-for-target (Jacobs 1986, 1987). Thus, when searching for a conspicuous object, subjects made larger saccades than when searching for a less conspicuous object. In fact, the saccade size matched exactly the *expected* conspicuity lobe of the searched-for target.

4.2. Search for targets defined by conjunctions of features

A large number of studies have shown that search patterns obtained when searching for a target defined by a conjunction of features are very different from those obtained with search for a single feature (i.e., Treisman and Gelade 1980; Treisman and Schmidt 1982). Conjunction search gave linear functions relating search time to the number of items in the display, suggesting a serial check of each non-target. An additional claim for serial search can be derived from the finding that RTs are faster for trials in which a target is present than for trails in which a target is not present with a ratio of approximately 1 : 2 for target present : absent slopes. The ratio 1 : 2 is indicative for serial self-terminating search. Thus if a target is defined as a red square and the non-targets are a mix of an equal number of red circles and green squares, the red square does not pop out and can only be found when focal attention is serially directed to each item. Focal attention is thought to be required in order to correctly conjoin features producing the percept of a whole object. The serial search by means of spatial attention offered the best account of a whole variety of data (see Treisman 1986, 1988; Quinlan and Humphreys 1987). Yet, in recent years, a number of investigations have reported exceptions to the claim that search for conjunction targets must be serial.

In an experiment of Egeth et al. (1984), subjects searched for a red 'O' in between 4, 14, or 24 non-targets, consisting of a mix of an equal number of red N's and black O's. In line with the FIT, search time for a conjunction target defined by a combination of color and form, increased linearly with display size. Yet, in conditions in which the ratio between the number of different categories of non-targets was changed (e.g., 2 red N's and 2, 12, or 22 black O's) subjects limited search to the red items. Because the number of red items did not vary with display size, search functions remained flat, indicating that the FIT's claim that conjunction search requires serial examination of all items in the display was incorrect. The results suggested that subjects serially search through a subset of stimuli (e.g., all the red items) and ignore the other subset. Note that the findings of Egeth et al. elegantly fit the earlier outlined model suggesting that attention is switched to the location having the highest difference signal. Obviously, in case of an equal number of red N's and black O's, all display elements (including the target) in both color and shape dimension will have more or less the same activity. By changing the ratio between the number of red and black items or between N's and O's some elements obtain a higher activation. If the target is among the elements with a higher activation, as in Egeth et al.'s experiments, the target is found by search which may appear to be completely parallel. Note that also in Theeuwes' (1991c, in press) studies, the search function remained flat even in conditions in which a distractor was present. Egeth et al. (1984) suggest that subjects may *voluntarily* restrict search to the more salient of the two subsets. As argued earlier, whether this is truly voluntary is disputable; it may be merely a bottom-up effect caused by a change in the differential activity level.

Experiments by Nakayama and Silverman (1986) are more problematic for the FIT. They found parallel (or close to parallel) search functions for color-motion conjunctions and parallel search for different combinations of conjunction of features involving stereoscopic depth (see also, Steinman 1987). These findings can be treated as exceptions to the general finding that conjunctions require serial search. Thus, Nakayama and Silverman (1986) suggest that depth and motion may behave as special features. If subjects can direct their attention to particular planes in depth or to directions of movements, targets which appear to be defined by a conjunction of features become, within a particular plane, a target defined by a single primitive feature. For example, if the target is a red bar moving up and down among green bars moving up and down and red bars moving left to right, it appears that the target is defined by a conjunction of movement and color. Yet, if subjects can direct selectively their attention to the movement in the vertical plane, the target will pop out because, within this plane, it is the only red item (comparable to feature search). Similar arguments were raised by McLeod et al. (1988; 1991) who showed that grouping by a common motion has different effects on visual search than grouping by a common color. In other words, movement and possibly also depth are in some sense 'special', therefore challenging the common finding of serial search for conjunction targets.

Findings of Pashler (1987) also dispute the claim of serial and self-terminating search for conjunction targets in displays with fewer than eight items. His search latencies increased linearly with display size, so that the findings do not challenge the serial claim. Yet, slopes for target present and target absent search were parallel rather than showing the 1 : 2 ratio normally found for conjunction search in larger displays (see Houck and Hoffman, 1986, for similar results). These findings suggest that search is not self-terminating, implying that in target-present trials subjects keep searching even after the target is found. Yet, when redundant targets were added to the display, search became faster, a finding which argues against exhaustive search (e.g., Holmgren et al. 1974). As an alternative account, Pashler (1987) suggests that for clumps up to eight items at a time, search is a limited capacity *parallel* self-terminating process giving small display size effects and parallel slopes for target-present and target-absent trials. It should be realized that the whole argument against FIT is based on the 1 : 1 slope ratio for positive and negative trials at small displays. This finding of a 1 : 1 ratio at small display is not general: experiments of Quinlan and Humphreys (1987) and Treisman and Gelade (1980) did show the typical 1 : 2 present–absent ratio also for small displays.

Experiments conducted by Duncan and Humphreys (1989) and Humphreys et al. (1989) suggested the parallel encoding of so-called 'within-object conjunctions', that is, 'conjunctions' which deal with the spatial arrangement of strokes within a letter (e.g. an L among non-target T's). Humphreys et al. (1989) showed that there was hardly a display size effect when searching for an inverted T among upright T's. Duncan and Humphreys (1989) claimed that the slopes of the search functions for feature search (a tilted 'T' among upright 'T's) were more or less the same as for within-object conjunction search (search for an 'L' among upright non-target 'T's'); although their data indicate that feature search always provided flatter slopes than conjunction search. Duncan and Humphreys (1989) showed that large letters give flatter slopes than small letters, and that the heterogeneousness of the non-targets set (e.g., search for an L among non-target T's rotated 180° and 270°) had a large effect on the slopes. Because these experiments show that under certain circumstances, stroke 'conjunctions' are little affected by display size, it is claimed that these findings are problematic for the FIT. Duncan and Humphreys (1989) refute the distinction between parallel and serial search, and claim that the effect continuously varies when irrelevant (non-targets) information is added to the display (next section outlines their model).

Finally, experiments conducted by Wolfe et al. (1989) seem to pose difficulties to the claim that conjunction search is serially performed. The main argument is that there is a wide variation across subjects with respect to the search slope when searching for a conjunction of form and color. Between subjects, search slopes varied from 2.0 ms/item to 20.2 ms/item (Wolfe et al. 1989: exp. 1) suggesting that from individual to individual, search varies from serial to almost parallel. The finding that there are large differences between subjects is usually taken as an indication that some curious manipulations took place. It seems odd that basic processes such as visual search give so much inter-individual noise. In search for peculiarities, it occurred that Wolfe et al. (1989) used rather large displays (up to 36 items) which were arranged 'at any of 36 locations in slightly irregular 6 by 6 displays' (p. 420). Note that Treisman (Treisman and Gelade 1980) used displays in which items 'were scattered over the card in positions which appeared random' (Treisman and Gelade 1980: 102). The fact that Wolfe et al. used 'slightly' irregular displays might have caused additional distraction which has nothing to do with typical conjunction search. Thus, when less than 36 items are present in the display (the low display sizes), regular clumps of items with empty spaces in between may have caused a search pattern in which all locations (even the empty locations) are checked causing relatively long search times at lower display sizes. In addition, clumps of items scattered in more or less regular groups might have caused peculiar jumps of attention through the visual field. Because the search times at these lower display sizes are longer than expected on the bases of 'normal' conjunction search, it appears that the slopes are flatter, yet possibly with higher intercepts (unfortunately, Wolfe et al. (1989: exp. 1) do not provide intercepts; only slopes of the individual search functions).

In order to account for the conflicting data regarding conjunction search alternatives to the FIT have been formulated.

Cave and Wolfe guided search model: Cave and Wolfe (1990; Wolfe et al. 1989; Wolfe et al. 1990) have adapted the FIT to what is called the 'guided search model' which can account for fast conjunction search. The model is very similar to the FIT. Guided search has a parallel stage that processes primitive features spatially in parallel followed by a serial stage that performs complex operations on a limited part of the visual field. Yet, contrary to the FIT, the guided search model assumes that there is some interaction between the parallel and serial stage. Rather than assuming that the second stage has to serially check each item in the display, the guided search model assumes that such a check only has to be made for elements that are closest to the target value for a particular dimension. The parallel stage is supposed to be capable of dividing items into distractors and candidate targets. Thus, for example, when the target is a red X among red O's and green X's, a parallel feature map excites all locations with a red item (color dimension) and all locations having an X (form dimension). Obviously, on the activation map, the location having the red X (the target) will be double activated. The serial scan with the attentional spotlight is then directed to locations in accordance to their current level of activation. If this mechanism would work perfectly, one would expect parallel search for the conjunction target. Yet, it assumed that there is noise so that distractors may obtain a high activation directing attention to the wrong location. As pointed out earlier, the bottom-up component is determined by calculating the differences in features for each dimension (e.g. a single green item that differs from all other red items will give a high activation at the location of the green item in the color dimension). The top-down component depends on knowledge about the target. If the feature at a particular location is close to the target value this will result in an extra excitation at that location (or a reduced excitation for features that do not resemble the target). Given the baseline bottom-up activity, the top-down component modulates these activations.

Treisman and Sato's revised FIT: In the original FIT, it was assumed that conjunction search was performed by a sequential spatial scan of one location at a time. Because it was shown that knowing what the target was, gave faster search functions (both in slope and intercept) than when the target was unknown, it was hypothesized that subjects can organize their search by restructuring the display, thereby inhibiting different pairs of features and scanning the remaining active locations. Similar to the Cave and Wolfe model, the revised FIT assumes some top-down control in searching for conjunction targets. Rather than removing just one set of distractors from the search process and searching the other set in parallel (as suggested by Egeth et al.'s (1984) data), the revised FIT suggests that 'feature inhibition' takes place implying that inhibition is generated in several feature maps coding non-target features, resulting in a reduced activity at the non-target locations.

Duncan and Humphreys (1989) similarity theory: A rather different account of visual search is provided by Duncan and Humphreys (1989) in which the similarity between targets and non-targets plays a crucial role. First, search difficulty (the slope of the function relating performance to the number of non-targets) increases with increasing similarity between target and non-targets. Second, difficulty increases with decreasing non-target similarity (Duncan and Humphreys 1989). Similar to the Cave and Wolfe and revised FIT models, Duncan and

Humphreys (1989) assume a parallel stage of perceptual segmentation followed by selective access of chosen material into visual short-term memory (VSTM), although it is claimed that there is no clear dichotomy between serial and parallel search modes. Similar to the serial stage of the models above, stimuli must make it into the VSTM before a response can be given; that is, before it gains control of behavior. In fact, the theory of Duncan and Humphreys (1989) is designed to model top-down effects on visual search. It is assumed that the access to the VSTM is strictly limited. At the parallel stage, there is a limited capacity resource which can be divided in varying proportions among structural units in the input description. Increasing the assignment to one unit, decreases the assignment to the others. The total available resources are distributed across inputs in proportion to relative weights. Initially, all weights are set to some constant; the selection system biases and alters the initial weights so that it matches an internal 'template' of information currently needed in behavior. In search tasks, this template will be some description of the target (i.e., its color, shape, location). If non-targets resemble targets, non-targets receive rather high weights because they closely match the target template. Obviously, from this follows the assertion that search difficulty depends on the similarity between target and non-targets. The second factor influencing the selection weights is a process that is called 'weight linkage', which refers to the distribution of weights to inputs which are perceptually linked. Linkage refers here to perceptual grouping controlled by such considerations as shared brightness, color, motion, shape or spatial proximity. Weight linkage is especially effective in rejecting strongly grouped non-targets, a process which is called 'spreading suppression'. Thus if all non-target items are green, reducing the weight to a single green non-target item reduces also the weight of all other green non-target items. In this way (perceptual) groups of non-targets can be rejected very effectively. Given this mechanism, it is clear that search is also affected by the similarity between all non-targets.

In summary, there are exceptions to the general claim of the FIT that search for conjunction targets is performed serially. Under certain circumstances, e.g., relatively large display sizes as in Pashler (1987) and Wolfe et al. (1989), or when searching for conjunctions involves a special type of feature, e.g., movement and depth as in

Nakayama and Silverman's (1986) study, search functions become relatively flat. In addition, Duncan and Humphreys' (1989) results suggest relatively 'parallel' search (e.g., slopes between 6 to 11 ms/item) for elementary conjunctions of letter features. The alternative theories for the FIT all incorporate some top-down mechanism on the parallel preattentive stage so that non-targets which are very dissimilar to the target do not have the same probability of entering the attentive stage, as non-targets do which are more similar to the target. Although this notion appears to be viable for search for conjunction targets (see Theeuwes in press), it is not in accordance with the findings on search for targets defined by a primitive feature (e.g., Theeuwes 1991c). In these experiments, subjects could not ignore the distractor in the irrelevant dimension although the distractor was maximally dissimilar from the target.

It should be realized that conjunction search may be explained without supposing top-down control at the preattentive stage. If it is assumed that subjects adjust the size of the attentional spotlight (i.e., the group scanning hypothesis; Treisman and Gormican 1988) when searching for conjunction targets, then, possibly dependent on the lay-out of the display, the target might occasionally be surrounded by non-targets which only differ in primitive features, giving a pop-out of the conjunction target within the attended area. This strategy gives better results with large displays, especially when items are presented in a square-like array around the fixation point. Therefore, in relation to small displays, the search times for large displays are underestimated, giving relatively flat slopes. The efficiency of the strategy depends on the – from trial-to-trial changing–relative position of target and non-targets, possibly explaining the observed large variability of results within and between experiments and subjects.

4.3. Search for categorically defined targets

In the preceding sections it was shown that discriminations at the level of primitive features can easily be made parallel, and that parallel processing becomes somewhat problematic when discriminations require the conjunctions of different features. Yet, there is even evidence that suggests that search for targets which are *semantically* different from non-targets may proceed in parallel. Such findings have been taken as support for late selection theories (e.g., Allport 1977;

Duncan 1980, 1981; Duncan and Humphreys 1989; Shiffrin and Schneider 1977) which claim that all perceptual encoding, including identification, proceeds in parallel across the visual field. Selection is supposed to occur 'late' in processing, primarily to select between competing response tendencies arising from different stimuli (Allport 1980). Early selection theories, such as the FIT and the guided search model (e.g., Broadbent 1971, 1982; Treisman and Gelade 1980; Cave and Wolfe 1990) claim that only the discrimination of basic features can be conducted in parallel, and that for full identification, a second stage of limited capacity is needed.

Although there is not much evidence for parallel processing of semantically defined targets, a few sources of evidence can be mentioned.

In various visual search experiments it has been shown that the slope of the search function depends on the categorical relationship between target and non-target (Egeth et al 1972; Gleitman and Jonides 1976, 1978; Jonides and Gleitman 1972, 1976). This finding is referred to as the category effect, and denotes that it is much easier to find a letter among digits or a digit among letters, than it is to find a letter among letters or a digit among digits. Thus, almost flat search functions were obtained for detecting a categorically different target, whereas relatively steep slopes were found when searching for a within-category target (i.e., Jonides and Gleitman 1976). The category effect has often been taken to imply that all items in the display are identified and categorized in parallel, and that the target can be distinguished from its background by a category code; just as a red item can be segregated by its color from a field of blue distractors. Yet, the category effect can also be explained without the radical conjecture that all stimulus characteristics are encoded in parallel to a semantic level. Instead, one may assume parallel processing of only those features needed for categorizing a character as letter or digit. In a second stage, full identification might take place (Broadbent 1982; Kahneman and Treisman 1984; Neumann 1984). According to this line of reasoning, the parallel stage separates the target from non-targets on the basis of some simple feature, rather than on the basis of some categorical difference. Krueger (1984) provided direct evidence for this claim: when letters and digits were matched with respect to featural difference, the category effect disappeared. Similar arguments have been raised by Duncan (1983) and Rabbitt (1967). Also, Theeuwes (1991a) provided evidence against the late selection account. In one of his experiments subjects searched for a letter among a variable number of digits. The control experiment showed a relatively flat search function (about 8.5 ms/item) replicating the typical category effect. Yet, when two target letters, which had to be matched, were presented simultaneously among a variable number of non-target digits, parallel search was disrupted, and subjects searched the whole display serially (about 25 ms/item). Contrary to the late selection theory (e.g., Duncan 1980, 1981; Garner 1973), it was impossible to reject all non-targets in parallel.

A similar line of evidence that has been used to support the idea of parallel processing of semantically defined targets is provided by experiments using consistently mapped (CM) visual search tasks (Schneider and Shiffrin 1977; Schiffrin and Schneider 1977) in which targets and distractors never exchange roles. After extended practice there is a reduction (and sometimes elimination) of the effects of the number of items in both the display and target sets. It has been claimed that the consistent pairing of output processes with particular shapes results in an automatic process that operates independent of attention, is not capacity limited and cannot be modified by the subject (Shiffrin and Schneider 1977; but see Fisher 1982, 1984). Automatic search is contrasted with controlled search which requires attention, is capacity limited, and can be modified by the subject. Similar to the category effect, the results of automatic search tasks may be explained by claiming that all stimuli are identified in parallel. Yet, similar as in the case of the category effect, due to consistent training, it is possible that subjects learn some general, possibly feature-like, code that enables them to distinguish targets from non-targets (e.g., Cheng 1985; Hoffman 1986). The claim of Shiffrin and Schneider (1977) that attention does not play a role in automatic search tasks is challenged by their own data (exp. 4d): subjects could not ignore previously valid CM targets when performing a controlled search task. If attention is not involved why do previously valid CM targets interfere? A two-stage model could possibly account for these types of effects: due to extensive training, some physical code is learned which enable the subjects to discriminate the target at the preattentive level; followed by an automatic shift of spatial attention necessary for target identification. In Shiffrin and Schneider (1977: exp. 4d), a previously valid CM target interfered because attention was unintentionally attracted to its location. As Hoffman (1986) suggested, the previously valid CM targets might have interfered because they were located relatively close to the areas which had to be attended, thus precluding the possibility to filter on a spatial basis (same account is used for explaining the Stroop word-color effect; e.g., Kahneman and Henik 1981; Kahneman and Treisman 1984). The hypothesis that an automatic target causes a shift of spatial attention is confirmed by an experiment of Hoffman et al. (1983: exp. 3), in which it was shown that the detection of a digit among letters was accompanied by an increased ability to perceive shapes in the vicinity of the digits.

A study conducted by Pashler and Badgio (1985) also suggests parallel identification of multiple familiar stimuli. In their study, subjects had to name the highest digit of an array of 2, 4, 6 digits, a task which requires exhaustive identification. Manipulations of display size and visual quality were used to test the serial versus parallel processing hypotheses. The results showed that visual quality was additive with display size, a result which cannot be explained with a model which assumes that the degraded quality of each item is 'cleaned up' serially. Therefore, it is concluded that all stimuli are processed in parallel; although it might be argued that only the 'clean-up' of the whole display was performed in parallel, followed by a serial search through the 'cleaned up' digits. Yet, according to Pashler and Badgio (1985) such an interpretation is implausible because visual quality interacted with response factors.

Finally, experiments performed by Van der Heijden and his colleagues (e.g., Van der Heijden 1975; Van der Heijden et al. 1983) suggest parallel processing of simultaneous stimuli, conceivably of letter identities. In a redundant-target paradigm, it was shown that more than one target can contribute simultaneously to target identification. For example, in Van der Heijden et al. (1983) subjects had to discriminate E's from F's. Two types of tasks were used, in a 'go–no go' task subjects pressed the button when one, two or three E's were presented and refrained from responding when only F's were presented. In a 'yes–no' task, subjects pressed 'yes' in case of one or more E's and 'no' in case of F's. In both tasks, RTs became faster as more targets were present in the display. This redundancy gain can only be explained by assuming (fairly) unlimited capacity, parallel, self-terminating processing. Van der Heijden et al. (1983) considered the possibility that the results are due to a 'favored position' artifact. If subjects selectively attend to a single favored position in the display, even when this position randomly varies across trials, than the three target item displays will always give faster RTs than the displays with less targets. By means of an analysis of RT distributions, it was shown that the findings in the 'yes–no' tasks could, in fact, be explained by the favored position artefact. Yet, the RT distributions of the 'go–no go' task clearly indicated parallel processing of redundant targets. Recently Egeth et al. (1989) also demonstrated the redundancy gain effect. In a 'go–no go' task, subjects responded to letters (e.g., 'go') and digits (e.g., 'no go'). Either one, two or three identical digits or letters were presented simultaneously. Again, an RT decrease with increasing number of targets was found, a result which could not be attributed to random favored position artifacts. In another experiment, Egeth et al. (1989) used the go–no go paradigm in a semantic categorization task in which subjects responded to the meaning of one or two identical words (e.g., 'go' when the word is an animal category, 'no go' otherwise). No redundancy gain was found in this semantic categorization task. Yet, using similar stimulus material, in a lexical decision task in which subjects had to respond to words and refrain from responding when nonwords were presented, the redundancy gain was observed again. By using the redundancy gain paradigm Mordkoff et al. (1990) showed that detecting conjunctions of color and form are processed in parallel. In a 'go–no go' task in which subjects responded to a red X, and refrained from responding when a red O or a green red X were presented, a redundancy gain was found suggesting that red X's are processed in parallel.

The results regarding the redundancy gain are difficult to interpret. A basic issue is whether theoretical models for target recognition using small displays in which the target has to be *detected* (e.g., the tasks used for studying the redundancy gain) apply to visual search tasks in which a target has to be *selected* among non-targets. Also, it is unclear whether the redundancy gain studies really indicate that letters and words are processed in parallel with unlimited capacity or whether only simple features are processed in parallel, as for example suggested by the FIT. For example, in Van der Heijden et al.'s (1983) experiments, subjects only responded to the letter 'E' and refrained from responding to the letter 'F'. The letter 'E' can obviously be discriminated from the letter 'F' purely on the bases of a single feature '_'. If it is assumed that the preattentive process encodes for each location spatially in parallel, the presence of the discriminating feature, the redundancy gain is simply due to a horse race between low-level perceptual processes (see Egeth et al., 1989, for a similar argument). Then, the findings do not implicate the parallel processing of letter identities, but the parallel processing of some basic feature, completely in line with the FIT and related accounts. Purely on theoretical grounds, Bundesen (1990) claims that Van der Heijden et al.'s (1983) redundancy gain is due to position uncertainty for the displays with less than three targets. Bundesen predicts that if the position uncertainty would be eliminated, the RTs for the one- and two-target conditions would be the same as for the three-target condition.

In summary, in line with the late selection view, there is some evidence that processing of semantic identities can occur in parallel. Yet, this evidence is not overwhelming, and it seems that many findings that suggest parallel identification of semantic identities can also be explained by assuming that only some basic feature distinguishing target from non-targets is processed in parallel. In a study in which selection by color was directly compared with selection by category, Navon (1989) concluded that selection by color and category are not as similar as the late selection view claims. Selection by color was fast, efficient, and load free; whereas selection by category was slow, prone to intrusions and load-sensitive. As Broadbent claims 'the popularity of late selection does not stem from any empirical evidence. Rather it seems to be one of separation between different academic communities' (1982: 281).

5. The special role of spatial attention

5.1. Introduction

There is a considerable controversy regarding the role of spatial attention. Space-based theories assume that spatial location plays a unique role in the selection of information for further processing (e.g., Posner et al. 1980; Eriksen and Hoffman 1973), whereas others claim that location is just one selection dimension (although an extremely efficient one) that is, in principle, not different from selection dimensions such as color and shape (Duncan 1981). In this view, spatial information enables efficient selection not because it is 'special', but simply because it has a lot of discriminative power. Visual selection can be facilitated by any type of cue as long as it provides discriminative power to separate target from non-targets.

Studies suggesting a special role for spatial attention show that advance knowledge of the target location improves processing of that target (e.g. Posner at al. 1980). In many of these covert orienting experiments, a cue prior to stimulus presentation serves as a signal to expect a target in a specific location. Selective allocation of visual attention (without the aid of eye movements) has been examined by means of a cost–benefit analysis (Posner 1978, 1980) in which the performance of detecting signals at expected locations is compared to the performance of detecting signals at unexpected locations. Various studies have demonstrated better performance for valid than for invalid cued targets across a number of tasks, including simple luminance detection (Bashinski and Bacharach 1980; Posner et al. 1978; Posner et al. 1980), identification (e.g., Eriksen and Hoffman 1972; Van der Heijden et al. 1987), and discrimination (e.g. Downing 1988), as well as across two dependent measures, response latency (e.g., Jonides 1981; Posner et al. 1980) and accuracy (e.g., Downing 1988; Van der Heijden et al. 1987).

Two different types of cues may be used to inform the subject about the likely target location. Before display onset, a cue appearing at the center of the display (e.g., an arrowhead) can direct attention to the likely target location. In response to this cue attention is supposed to be voluntarily directed to the cued location in visual space. This can be contrasted with 'peripheral cues', which usually appear at a position near the actual target location. This type of cue, usually a stimulus with an abrupt visual onset (e.g., a 'barmarker', Jonides 1981; Theeuwes 1991b), is supposed to *reflexively* draw attention to the location of the onsetting event. Posner (1980; Posner and Cohen 1984) suggested that dependent on the type of cues, control of visual attention can be either *endogenous* (e.g., intentionally directed to a location) or *exogenous* (e.g., unintentionally drawn to a location; Todd and Van Gelder 1979). It has been demonstrated that the endogenous and exogenous shift of attention are often not comparable (Briand and Klein 1987; Jonides 1981; Müller and Rabbitt 1989). With exogenous orienting, attention is reflexively drawn to a location in space and is actively engaged in processing the cue (or target) that attracted it there; with endogenous cuing, there is a passive expectancy that the target is likely to occur at a position in space, without active processing of visual information taking place (Klein and Hansen 1990). The next section will discuss different theoretical explanations which can account for the findings on spatial cuing.

5.2. Different theoretical accounts

5.2.1. Space-based accounts

It is assumed that a spatial cue enables attention to focus on a particular region in visual space in which the quality of the perceptual representation is enhanced. As a metaphor, it is suggested that spatial attention is analogous to shining a spotlight on the visual field. Objects falling within the spotlight can be identified or detected more rapidly than objects outside the spotlight. This analogy has been advocated by many authors (e.g., Broadbent 1982; Downing and Pinker 1985; Eriksen and Hoffman 1973; Posner 1980: Shulman et al. 1979; Tsal 1983). Two properties are implicit in the metaphor. First, attention is limited in spatial extent (Yantis 1988), implying that focused attention cannot be directed to two or more locations at the same time (Posner et al. 1980). Second, directing attention in visual space is necessary before responding to the target is possible (Posner 1980). Recall that most theories on visual search assume the involvement of attention in order to identify the stimulus (Bergen and Julesz 1983; Hoffman 1978, 1979; Treisman and Gelade 1980; Wolfe et al. 1990), or to select among different responses (Allport 1980; Duncan 1981; Duncan and Humphreys 1989; but note that the latter theories do not claim that focused attention has a spatial locus). Thus if attention is directed to an invalid location, it has to be shifted to the valid location in order to respond to the target. Since the spotlight of attention as well as overt eye movements can both be summoned by a peripheral event (e.g., exogenous cue) and both can be directed voluntarily to a location in visual space (e.g., the endogenous cue) it has been suggested that there is a close functional (Posner 1980) and physiological (Wurth and Mohler 1976) relationship between these systems. Thus, attention is like an 'internal eyeball' (Skelton and Eriksen 1976) or, attention is the 'mind's eye' (Jonides 1980, 1981, 1983). Yet, the simple spotlight theory, analogous to the movement of the eyes, had to be adapted in order to account for conflicting data.

Many studies have been directed to reveal the mechanism and structures that subserve the relocation of attention (e.g., Downing and Pinker 1985; Eriksen and St. James 1986; Posner 1980; Shulman et al. 1979). Two contrasting claims have emerged: (1) The spotlight moves with a constant velocity through visual space, taking more time as it moves a greater distance (e.g., the eye movement analogy; Posner 1980; Shulman et al. 1979; Tsal 1983). (2) Shifting attention occurs in a discrete abrupt manner so that it does not take more time to move longer distances (e.g., Eriksen and Murphy 1987; Murphy and Eriksen 1987; Sagi and Julesz 1985a, 1985b; Skelton and Eriksen 1976). Recently, Kwak et al. (1991) provided strong evidence for the second claim in a task in which focal attention was required for identification (matching rotated T's and L's). It was shown that the distance between the targets had no effect, suggesting that there is no movement of attention at all but that attention is simply reallocated from one location to another (see also Eriksen and Webb 1989). Note that this finding reduces the viability of the spotlight metaphor and also flaws the claim that spatial location is a very special factor in the control of attention.

Eriksen and Yeh (1985; Eriksen and St. James 1986) suggested a 'zoom lens' model that emphasized the variable size of the spotlight, indicating that attention, conceived as a processing resource, can vary from a uniform distribution over the entire field to a highly focused concentration. When reliable spatial information is provided, the system switches to its focused mode, concentrating all resources on a circumscribed area, leaving little or no resources for simultaneous processing of other display locations. When it is not necessary to focus attention on a specific location, attention is distributed uniformly over the visual field. Downing (1988) suggested that processing perceptual information is most efficient at attended areas, and becomes increasingly less efficient as the distance from the attended area is increased. In addition, Murphy and Eriksen (1987) proposed that a precue facilitates processing of stimuli within the attended area, with a gradient of decreasing facilitation along the borders of focused attention. According to the gradient model of attention (LaBerge and Brown 1989) attending to a small, compared to a large region in visual space gives a narrow peak in the resource gradient, resulting in a better performance for targets falling within such a peak than for targets falling within the more distributed peak. Again, attention is considered as a limited processing resource implying that a narrow peak removes resources from other locations. The narrower the peak, the more strongly resources fall off with the distance from the peak.

Pre-knowledge of information has been found to be a very effective cue that enables the separation of targets from non-targets. In the original experiments of Colgate et al. (1973) and Eriksen and Hoffman (1974) subjects responded to one out of 12 letters which were positioned in a circular display around a fixation point. It was shown that a response to a target was faster when its location was indicated by a barmarker close to the target position. These results were originally interpreted as indicating that the precue facilitated the selection of the correct letter out of the twelve non-target letters. Yet, in a control experiment in which only the target letter was presented, the precue speeded up the response as well, even though nothing had to be selected visually. The finding that prior knowledge of the position of a target in an 'empty' visual field has a beneficial effect on recognition has been demonstrated many times (e.g. latency, Eriksen and Hoffman 1973, 1974; Hoffman 1975; accuracy, Van der Heijden et al. 1987). The finding that foreknowledge of location affects the identification of stimuli even in an empty field in which no visual selection is necessary suggests that spatial location operates early in vision facilitating perceptual processing (e.g., Van der Heijden et al. 1987). In a late selection view, location information can only be helpful to select among several already fully identified stimuli, a claim which is hard to understand when nothing has to be selected. Yet, Duncan (1981) contended that even blank parts of the field are themselves real stimuli which may be admitted to, or rejected from, further processing by the attentional system, a claim recently substantiated by Müller and Humphreys (1991).

Related to this issue is the discussion regarding the locus of the effect of spatial cuing. The space-based account of attention assumes that a spatial cue facilitates processing because all attentional resources are concentrated on a relatively small portion of the visual field increasing the speed of processing for objects appearing at that location (e.g. Eriksen and Yeh 1985). Because in this view, attention modulates the efficacy of sensory processing it is expected that cued and uncued location differ in sensitivity (d') (e.g., Bashinski and

Bacharach 1980; Downing 1988; Hawkins et al. 1990). Alternatively, it has been claimed that attention has no effect on sensory processing at all. Instead, reaction time differences due to spatial cuing are attributed to differences in decision strategies at cued and uncued locations (Shaw 1978, 1984). Because targets are more likely to appear at cued locations, observers may simply require less sensory evidence to decide that a target has appeared there. Consequently, one would expect that cued and uncued locations differ in decision bias (e.g., beta). In Broadbent's (1970, 1982) terminologies these effects are also known as stimulus set or filtering (changes in sensitivity), and response set or pigeon holding (changes in response bias). Experiments trying to pinpoint the effects of spatial cuing seem to suggest that spatial cuing affects sensitivity (d'). For example, Bashinski and Bacharach (1980) showed in a luminance detection task that spatial cuing affected detection sensitivity but not decision processes. Similar results were provided by Downing (1988) and Hawkins et al. (1991). On the other hand, Müller and Findlay (1987) claimed that spatial cuing only affected decision biases and not sensitivity. Recently, Müller and Humphreys (1990) conducted a study in which subjects performed a localization-and-detection task showing that spatial cuing enhanced the accuracy of detecting simple luminance increments at a cued location, even when there was only a single target in the display. In addition to the effects on sensitivity, central cues also affected beta, whereas peripheral cues had no such effect.

The finding that spatial cues affect sensitivity rather than beta, has been interpreted as evidence against 'late selection' favoring the spotlight metaphor and claiming a special role for spatial attention. Thus, Downing (1988) who found a sensitivity effect of spatial cuing, asserts that spatial cuing 'affects the quality of relatively early levels of representation'. It is found that sensitivity for a particular location is determined by the distance from the attentional focus (e.g., LaBerge 1983; LaBerge and Brown 1989; Downing and Pinker 1985), with a maximum sensitivity at the focus of attention and a gradual decrease for locations further away. Yet, both changes in sensitivity and the finding of an attentional gradient can also be interpreted in terms of a late selection account (Müller and Humphreys 1989). In line with Duncan (1981) and Duncan and Humphreys (1990) spatial cuing is supposed to affect the priority by which information is read out from the first parallel stage, where fully identified representations are subject to decay. The cued location is assigned priority over targets at uncued locations, which compete for selection. Because the attentional priority is hard to interrupt, the cued location is selected even when no target is present at that location. Consequently, at least invalid cues (e.g., a cue pointing to the right while the target appears at the left) will produce costs even for a target in an empty field. The gradient effect is explained by assuming that stimuli at locations closest to the cued location are placed higher on the selection stack than stimuli further away (Müller and Humphreys 1991).

The finding that the time to respond to a target letter becomes faster as the angular distance between target and distractor letter increases (Eriksen and Eriksen 1974) has been treated as strong evidence for the spotlight theory and suggests a special role of spatial attention. In a recent study on the flanker compatibility effect (i.e., subjects respond to a central letter and ignore irrelevant flanking letters; the identity of the flankers produce a response compatibility effect), Miller (1991) provided again strong evidence that spatial separation has a special status for visual selective attention (see also LaBerge et al. 1991). In line with the spotlight theory, near distractors produce interference because they fall within the attentional spotlight focused on the target; and hence are selected for further processing or for potential response. Distant distractors fall outside the spotlight and are therefore ineffective (Broadbent 1982). This reasoning is in line with the earlier developed view which suggested that strategic variation of the span of attention acts as a selective filter.

In recent years, spotlight theories of visual attention have come under attack. In a study conducted by Driver and Baylis (1989) it was shown that spatial separation per se, as in the Eriksen and Eriksen (1974) paradigm, could not account for selectivity. In an experiment in which some items moved in a similar direction, it was shown that distant distractors that moved with the target produced more interference than stationary letters that were near the moving target. It appeared that movement was a strong feature for perceptual grouping. Yet, a whole series of recent experiments by Kramer et al. (in press) demonstrated that the effects reported by Driver and Baylis (1989) could not be replicated.

A study conducted by Juola et al. (1991) also challenges the spotlight type of theories. Subjects responded to a target letter which appeared between non-target letters in three rings around the fixation point. Subjects were instructed to attend to the ring (i.e., the outside, middle or inside ring) which was cued before display onset. In all conditions, responses were faster to valid cues than to invalid cues. Therefore, if subjects were instructed to attend to the outside ring, responses to targets appearing in the inside ring were relatively slow. The results were neither consistent with a zoom-lens nor with a spotlight model, and suggested a model that enables attention to be concentrated in ring-like areas of anticipated stimulus location.

Studies performed by Klein and Hansen (1990, 1987) have also been interpreted as a failure of the spotlight theory. On each trial, central precues were used to indicate the probable position of the imperative signal, which required either a manual or a saccadic response depending on the form of the imperative signal. When the stimulus forms were not equally probable, latencies to the more probable form showed the typical spatial cuing effect, whereas RTs to the improbable form was unaffected by spatial cuing. This latter result suggests that responding to the improbable form is unaffected by the location of spatial attention. It is concluded that dual-response experiments with asymmetric response probabilities fail to show the spotlight effect for the improbable stimulus/response combinations. Although these results are difficult for the spotlight theory, it should be realized that the improbable stimulus/response combination may simply have attenuated the spatial cuing effect through some response inhibition process (note that in the improbable condition, there were still spatial cuing effects, yet they failed to reach significance). Typically, subjects had the tendency to give the probable response to the improbable stimulus at the cued location. In studies by Lambert (1987; Lambert and Hockey 1986) in which subjects expected particular shapes to occur at particular locations, RT data suggested that visual selective attention can be sensitive to combinations of shape (category) and location. Yet, the finding that prior knowledge of the to-be-presented shape (category) affects RT does not necessarily dispute the spotlight theory because expectations regarding shapes (categories) may very well reflect response biases rather than effects on visual selection. Kingstone and Klein (1991) conducted a similar experiment and suggested a hierarchical-processing model of attentional selectivity. In this model shape and location information is processed simultaneously, yet at different processing rates such that the early resolution of one attribute expectancy (e.g., location) affects the processing of the other attribute expectancy (e.g., shape).

In experiments such as those cited above, the effects of expectancy of location are sometimes inappropriately compared to the effects of expectancies for other target attributes. It is assumed by a late selection account that location is just one selection dimension to separate target from non-target that is in principle not different from any other selection dimensions. When testing such a hypothesis, it is crucial to separate expectancy effects on perceptual processing from expectancy effects on responding. In a location cuing experiment, location says nothing about the ultimate response to be given. Hence, in a direct test, expectancies about form, color etc. should say nothing about the response either. Therefore, the required response should be completely independent of the information provided by the precue. Theeuwes (1989) tested the late selection hypothesis in a study in which subjects responded to the orientation of a line segment which was located in either a square or circle, presented either left or right from fixation. Central cues provided either information about the likely location (left/right) or about the likely form (square/circle). Obviously, RT differences due to the precue information could only be attributed to perceptual selection processes because the precue did not provide any information about the response. The results showed clear RT costs–benefits for location cuing, but not such effects for form cuing; implying that, opposite to the late selection account, these cues have quite different effects. As a typical counterargument, late selection theory can claim that the form cue did not provide enough discriminative power to affect selection. For example, Navon (1989: 54) notes that 'late selection theorists might respond that any difference in selective efficiency of different selection cues may be due to differential ease of retrieving them from a post-categorical representation, or searching for them in such a representation'.

5.2.2. Late selection accounts

Late-selection accounts do not assume a special role for location (Duncan 1980; Duncan and Humphreys 1989). For example, the earlier outlined theory of Duncan and Humphreys (1989) assumes that the selection system biases and alters the initial weights assigned to the structural units so that the selection system matches an internal 'template' of the information sought for. This template can specify more or less any attribute of the information needed such as its color, shape, category (letter or digit) or its location. Thus, in search for particular information (e.g., a target with particular color, or a target at a particular location) the selection system matches the template to the current input by changing weights proportional to the degree of match. Bundesen (1990) developed a theory which is similar to Duncan and Humphreys' (1989) account. According to Bundesen (1990), foreknowledge of the spatial position of a target facilitates selection because the subject can increase the attentional weight of any element at the cued location by increasing the 'pertinence' value of the cued location. The higher attentional weight speeds up the processing of items at cued locations at the expense of items at uncued locations. Pertinence value in Bundesen's theory represents a top-down component reflecting the importance of attending to elements that belong to a certain category. For example, if subjects have to search for red between black items, the pertinence value of the perceptual category red will be high. It appears that both late selection accounts (Duncan and Humphreys 1989; Bundesen 1990) are quite similar and assume that location cuing is just one of the many dimensions which help selecting items that compete for selection. The notion is similar to the Müller and Humphreys (1991) interpretation of the cuing effect although they assume that selection specifically takes place from a *spatial* map of multiple display locations.

A basic issue in the whole dispute between early and late selection is the distinction between selection and identification of items. In an early selection view, selection has to take place before full identification can begin, whereas in a late selection view these two aspects are viewed as part of the same process. Because late selection theory assumes that identification of items occurs *before* selection takes place, factors affecting the identification of items (e.g., factors speeding up the categorization of an item) are also considered as factors affecting the selection of items. In other words, a variable influencing the identification process is treated by late selection theory as a factor influencing the selection process as well. In the view as presently advocated, factors affecting the recognition process are considered irrelevant for the selection process because these factors exert their effect on already selected items (i.e., items that have entered the second stage of processing). Therefore, priming effects are for example not discussed because they do not influence the selection of items, but the recognition processes of already selected items (note that Johnston and Dark's (1986) review on selective attention does include priming).

5.2.3. Post-categorical filtering

The notion of post-categorical filtering and selection as advocated by Van der Heijden (1981, 1992) can neither be considered as an 'early' nor a 'late' selection view. In line with a 'late' selection claim, the model suggests that before selection takes place, visual information is automatically and in parallel registered and processed into an *identity* code. Yet, retrieval of this identity code can only occur by means of the *position* of the object, suggesting that attention operates via spatial locations, a claim in line with an early selection view. According to Van der Heijden (1992), selection operates on an 'early' representation using position to select one identity among a number of available identities. Visual selection takes place, not because there is a limited capacity to process information, but because there are many potential targets for action and each time a particular object has to be selected among others.

The basic structure of Van der Heijden's model (1992) consists of an input module which sends information in parallel to two independently operating channels: an identity and a position channel. Identity processing occurs in parallel, yet without position information from the position channel, the identity information cannot become available. Through a feedback loop from the position channel to the input module, a particular location in the visual field will be addressed implying that the identity of the information at the selected location becomes available for responding. The location channel feeds back in parallel the location information of *all* stimuli represented in the input module; yet, the position that receives additional activation is thought to be *selected*. In other words, selective attention is the process of feeding back particular location information to the input module. Important is that the location channel has no access to the identity of the elements in visual field, suggesting that selective attention can in principle only operate through position information. Location information is needed to select identified information.

To this basic (data-driven) structure a higher center is added to account for findings such as endogenous location cuing and selection based on primitive features such as color. The (top-down) higher center operates directly on the location channel or operates through other modules (e.g., color module) which activates the location containing the particular feature.

The model proposed by Van der Heijden is based on findings from partial report and location cuing experiments. Yet, the model is also applicable to visual selection in visual search. The data-driven structure of Van der Heijden's model is to some extent similar to the view developed in this paper. In the present view, primitive features can be processed in parallel without capacity limitations and also location is the entity on which visual *selection* is based (in our view, a shift of focal attention to a location implies that the object at that location is *selected*). Yet, in our view, after selection, additional processing goes on before the identity of the object becomes available whereas Van der Heijden assumes that after selection identify becomes available without additional processing.

With respect to top-down control, both views claim that only through the 'location system', visual selection can be affected in a top-down way (in our view 'endogenously direct spatial attention to a location is visual space'). Also both views assume that if information is coming from feature modules, this information affects the 'location system' before it will affect selection. Van der Heijden assumes that these feature modules can be modulated in a top-down way (changing the criterion), whereas the present view suggests that effect on the 'location system' (in our view 'causing an automatic shift of spatial attention to the location of the unique feature') is automatic and not subject to cognitive control.

Van der Heijden's theory is compatible with the results of the earlier discussed visual search experiments if it is assumed that the position channel automatically activates the location of the largest feature difference present in the input module (possibly calculating the differences among features within each module) implying that the odd item is automatically selected first. If attention is captured by a target, a response will be given; otherwise attention is 'withdrawn' and 'moves' to the target; in Van der Heijden's view by means of top-down control.

The above analyses suggest that both views are very similar expect that the terminology used is quite different (e.g., 'position channel' versus 'spotlight'; 'the higher center can activate a relevant position in the location channel' versus 'subjects can endogenously direct spatial attention to a location in visual space', 'location system' versus spotlight of attention'). Yet the views differ with respect to the extent of parallel processing. Van der Heijden assumes that unlimited capacity processing is not limited to primitive features as the present view assumes, but that all objects (identity channel) and all locations (location channel) are fully identified in parallel without any capacity limitations.

5.2.4. FIT and related accounts

Like the space-based theories discussed above, various models dealing with visual search (e.g., FIT, Treisman and Gelade 1980; guided search model, Cave and Wolfe 1990; Hoffman's two-stage model, Hoffman 1978, 1979; Julesz's texton theory; Bergen and Julesz 1983) also presume a special role of spatial attention. Although features may be registered preattentively without attention, focused spatial attention (e.g., the spotlight of attention) directed serially to locations in visual space, is necessary in order to combine information from different feature maps into complex object representations and is necessary to locate items in the visual field. Treisman (1988) provides evidence that visual attention operates by selecting stimuli in particular locations: identification of targets defined by primitive features was well above chance even though the target was mislocated. Yet, mislocated conjunction targets were not identified at levels above change. In addition, Nissen (1985; see also Isenberg et al. 1990) presented four items varying in location, color, and shape. In one condition, subjects had to report the color and shape of items cued by location. This condition showed that responses for color and shape were basically independent. In another condition, subjects had to respond to the location and shape for items cued by color. In this condition, localization of the cued item was necessary in order to correctly report the correct shape. This study indicates that in order to report the shape of an item which is cued by a color, spatial attention is necessary. Yet, recently Bundesen (1991) showed that the findings of Nissen (1985) can also be interpreted as evidence for the view which assumes that selection based on color or shape is independent of location information. Tsal and Lavie (1988) presented circular displays for 100 msec containing a mixture of three red, three green, and three brown letters. Subjects had to report the first item of a given color, and then any other letter they could identify. If attention was allocated to the location to the first reported letter, letters adjacent to the first letter should be reported more frequently than other letters. Similarly, if attention would be allocated to the color of the first reported letter, letters of the same color (irrespective of their position) should be reported more frequently than other letters. The results showed that subjects reported items that were close in space to the first reported letter, suggesting that the direction of attention to a relevant spatial location is a mandatory process that takes place irrespective of the dimension according to which the stimulus was initially selected for processing (e.g., color and shape).

According to the FIT and related accounts, the attentional spotlight selects objects in their location allowing the conjoining of features. Briand and Klein (1987) addressed the question whether the allocation of attention in the FIT is the same attentional mechanism as the direction of attention in covert orienting tasks like those of Posner (1980). Theoretically, such a question denotes whether the second attentive stage involved in visual search tasks is equivalent to the covert direction of attention in the visual field. If this were correct, then, in line with the two-stage model, covert orienting to a location in the visual field would imply that the first preattentive stage is bypassed so that items at the location focused directly enter the second stage of attentive processing. Briand and Klein's (1987) experiments showed that the effect of a central cue was the same in feature and conjunction conditions (both costs and benefits implying the involvement of attention); whereas a peripheral cue gave costs and benefits in the conjunction condition but not (very little) in the feature condition. Because in the FIT the detection of a simple feature is supposed not to require attention while conjunction targets require focal attention, it is then concluded that the spatial attention involved in the FIT is equivalent to exogenously controlled attention. Although this conclusion has been considered as too tentative (Treisman 1988), the findings of this study fit nicely the model outlined in section 4.1.2 which suggested a special role for the spatial distribution of attention as a filtering device (see also Theeuwes 1991b). If it assumed that, when attention is spread out over the visual field, the unique feature in the feature condition attracts attention automatically, the exogenous cue cannot have much effect on the detection of the feature target (i.e., hardly any costs–benefits). In the conjunction condition, in which the target does not pop out, the exogenous cue will produce large costs and benefits. In case of an endogenous cue, subjects are forced to focus their attention to a location in the visual field before display onset. When attention is directed to a invalid location, the unique feature falls outside the attentional spotlight and cannot attract attention anymore (see Theeuwes, 1991b, for similar effect). Reshifting attention to the valid location will take time and produce costs. For the conjunction condition exact the same reasoning holds, explaining why the same costs and benefits are found in conjunction and feature conditions when using an endogenous cue.

In conclusion, in visual search, the operation of attentive processing involves the serial direction of the 'spotlight of attention' (Posner and Cohen 1984) to locations in the visual field. The direction of the 'spotlight of attention' can be considered as the same attentional phenomenon as occurring in covert orienting tasks (e.g., Posner 1980; Posner et al. 1980) in which, independent of eye movements, attention is exogenously shifted toward a particular location (Briand and Klein 1987). Because a target defined by a single discriminable feature already attracts attention automatically, target detection does not benefit much from the advance direction of attention. Targets defined by properties that do not pop out (i.e., conjunction target, semantically defined targets) benefit from the advance direction of attention.

6. An integrative summary

The present paper outlined a framework which allowed a consistent interpretation of a great deal of experimental data. It is realized that some interpretations allow counterinterpretations, yet the view advocated here is tenable and seems to provide the most parsimonious explanation for many findings. This section will summarize the present viewpoints along the lines of three (interrelated) sources of controversies in the field of attention. Finally, the relation between the present approach and various other 'perceptual' approaches will be discussed.

6.1. Top-down versus bottom-up processing

It is assumed that the strategic control over 'what' is selected in the visual field is limited, yet, 'where' (from which location) it is selected from, can be controlled strategically. Subjects can strategically control visual *selection* by varying the span of spatial attention in the visual field. Dependent on the task demands, attention can be spread out over the entire visual field (or a portion of the visual field) or be focused on a specific location. If attention is spread out, preattentive parallel processing will occur; within the beam of spatial attention, top-down selectivity is not possible: the item producing the highest bottom-up difference signal within any stimulus dimension will attract attention, and therefore will enter the second stage of attentive processing implying that the item is *selected*. It is assumed that the bottom-up activity consists of a difference signal for each feature dimension separately depicting how different an item at a certain location is from all other items in the visual field with respect to a certain dimension (cf. Cave and Wolfe 1990). The preattentive process cannot reveal the origins of the difference signal (whether the item differs in color, shape, brightness etc). Hence, the processing focus is directed toward the location with the highest activity. Thus, within the beam of attention, perceptual selectivity (i.e., what enters the second stage of processing) is determined entirely by the input. Koch and Ullman's (1985; Ullman 1984) model is similar to the present view. According to their model, visual information occupying the position of the highest salience or conspicuity, is marked and passed on to the 'central representation' that is responsible for further stimulus analysis.

The present account assumes a genuine filtering by location in suggesting that the filter may be operationally set to encompass multiple objects. The deployment of visual spatial attention may be the *only* mechanism for selection, that is, there are no additional mechanisms for selection within the deployment of attention. Because visual attention is specifically spatial, other kinds of apparent filtering may in fact operate indirectly. For example, if a task requires selection by a unique color, the preattentive process automatically extracts the positions of the items which differ in color, causing a shift of attention to the spatial sites of the uniquely colored items.

Evidence suggesting parallel processing of conjunctions of features and semantic identities is not overwhelming. With respect to parallel conjunction search, it was speculated that the occasionally found almost flat search functions are obtained because subjects adjust the attentional beam, which allows a strategy in which a conjunction target occasionally will pop out (group scanning strategy, Treisman and Gormican 1988). With respect to the parallel processing of so-called semantic identities, it was speculated that in most cases, the target is distinguished from non-targets on the basis of some basic feature (or set of features) which can be extracted preattentively in parallel.

The endogenous or exogenous direction of attention on a location in the visual field acts like a spatial filter, restricting the encoding of items to those appearing at the focused location and blocking out information from all other parts in the visual field. This filter is far from perfect implying that items close to the focus of attention or items having a relatively high saliency (i.e., abrupt visual onsets and offsets; Breitmeyer and Ganz 1976; Theeuwes 1990; Yantis and Jonides 1984) occasionally break through (e.g., Broadbent 1982). Obviously, items breaking through which also require some action will produce larger interferences (Eriksen and Eriksen 1974). Because focusing attention acts like a filter, preattentive processing of other locations cannot occur. Items appearing on the location on which attention is focused are assumed to enter the second stage of focal processing.

6.2. Early versus late selection

The view advocated here is that attentional selectivity operates at an early perceptual level. It is suggested that preattentive processing is limited to the extraction of low-level perceptual characteristics. Shifting focused spatial attention to a location represents the only way by which information in the visual field is selected for further processing. Shifts of focused attention may occur exogenously when, within the beam of attention, a highly salient item is present, or, endogenously, when subjects focus their attention to a location in visual space. The 'late-selection' claim that all stimuli are processed regardless of attention (Duncan 1981) seems to apply to the present account to some extent: low-level perceptual operations on stimuli occurring within the attended area are also processed without attention and as a consequence, cannot be controlled strategically.

6.3. Automatic versus controlled processing

The present view assumes that, *within the attended area*, preattentive processing (i.e., the calculation of the difference signals for each stimulus dimension at each location) and the subsequent shift of focused attention are automatic processes because they fulfil both the load insensitivity and the unintentionality criterium of automaticity (Theeuwes 1991c). Alternatively, since the size of the attended area can be varied strategically, it may be argued that the preattentive process is not completely under the control of stimulation: information outside the attended area is not processed at all. The endogenous direction of attention in the visual field is a controlled process since it requires effort and is not under the control of stimulation.

6.4. Relation to other approaches

The present view which claims that the operations of the first stage are purely bottom-up driven and are limited to the extraction of low-level perceptual characteristics, followed by a second stage which allows detailed perceptual analysis is similar to Marr's (1982) computational approach of vision. Marr recognized two types of early representations in vision: the primal sketch which is a representation of the incoming image, and the $2\frac{1}{2}$-D sketch, which is a representation of surfaces in three dimensional space. These early visual representations are assumed to be viewer-centered, bottom-up driven, and they are essentially local descriptions that represent properties such as depth, orientation, color, and the direction of movement. The extraction of more complicated units and the description of spatial relations among the elements is not achieved at this early level (Ullman 1984). Marr (1982) viewed the organization of the visual system as *modular* (e.g., Barlow 1986), suggesting that for example, color, motion, depth, and shape are processed independently of each other; a view very similar to the FIT (Treisman and Gelade 1980). This view also appears to be supported by physiological evidence (e.g., Livingstone and Hubel 1987; Maunsell and Newsome 1987). As suggested by Fodor (1983) special-purpose modules have many advantages: they allow a fast and mandatory coding of relations within dimensions without possible cross-talk from other dimensions. Note, though, the present approach indicates that the output of these modular systems may cause interferences at the second stage of processing. As addressed earlier, the present approach is related to the texton theory (Bergen and Julesz 1983), the geon theory (Biederman 1987), and various theories on visual search (Cave and Wolfe 1990; Hoffman 1978, 1979; Treisman and Gelade 1980).

References

Allport, D.A., 1977. 'On knowing the meaning of words we are unable to report: The effect of visual masking'. In: S. Dornic (ed.), Attention and performance VI. Hillsdale, NJ: Erlbaum.

Allport, D.A., 1980. 'Attention and performance'. In: G. Claxton (ed.), Cognitive psychology: New directions (pp. 112–153). London: Routledge and Kegan Paul.

Andriessen, J.J. and H. Bouma, 1973. Eccentric vision: Adverse interactions between line segments. Vision Research 16, 71–78.

Barlow, H.B., 1986. Why have multiple cortical area? Vision Research 26, 81–90.

Bashinski, H.S. and V.R. Bacharach, 1980. Enhancement of perceptual sensitivity as the result of selectivity attending to spatial locations. Perception and Psychophysics 28, 241–280.

Beck, J., 1967. Perceptual grouping produced by line figures. Perception and Psychophysics 2, 491–495

Bergen, J.R. and B. Julesz, 1983. Parallel versus serial processing in rapid pattern discrimination. Nature 303, 696–698.

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

Bouma, H., 1970. Interaction effects in parafoveal recognition of initial and final letters of words. Nature 226, 177–178.

Bouma, H., 1978. 'Visual search and reading: Eye movements and functional visual field: A tutorial review'. In: R. Requin (ed.), Attention and performance VII (pp. 115–146). Hillsdale, NJ: Erlbaum.

Breitmeyer, B.C. and L. Ganz, 1976. Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression and information processing. Psychological Review 83, 1–36.

Briand, K.A. and R.M. Klein, 1987. Is Posner's 'beam' the same as Treisman's 'glue'?: On the relation between visual orienting and feature integration theory. Journal of Experimental Psychology: Human Perception and Performance 13, 228–241.

Broadbent, D.E., 1958. Perception and communication. London: Pergamon.

Broadbent, D.E., 1970. 'Stimulus and response set: Two kinds of selective attention'. In: D. Mostofsky (ed.), Attention: Contemporary theories and analysis. New York: Appleton-Century-Crofts

Broadbent, D.E., 1971. Decision and stress. London: Academic Press.

Broadbent, D.E., 1982. Task combination and the selective intake of information. Acta Psychologica 50, 253–290.

Bundesen, C., 1990. A theory of visual attention. Psychological Review 97, 523–547.

Bundesen, C., 1991. Visual selection of features and objects: Is location special? A reinterpretation of Nissen's findings. Perception and Psychophysics 50, 87–89.

Cave, K.R. and J.M. Wolfe, 1990. Modeling the role of parallel processing in visual search. Cognitive Psychology 22, 225–271.

Cheng, P.W., 1985. Restructuring versus automaticity: Alternative accounts of skill acquisition. Psychological Review 93, 414–423.

Colgate, R., J.E. Hoffman and C.W. Eriksen, 1973. Selective encoding from multielement visual displays. Perception and Psychophysics 14, 217–224.

Deutsch, J.A. and D. Deutsch, 1963. Attention: Some theoretical considerations. Psychological Review 70, 80–90.
Downing, C.J., 1988. Expectancy and visual-spatial attention: Effects on perceptual quality. Journal of Experimental Psychology: Human Perception and Performance 14, 188–202.
Downing, C.J. and S. Pinker, 1985. 'The spatial structure visual attention'. In: M. Posner and O. Marin (eds.), Attention and performance XI. Hillsdale, NJ: Erlbaum.
Driver, J. and G.C. Baylis, 1989. Movement and visual attention: The spotlight metaphor breaks down. Journal of Experimental Psychology: Human Perception and Performance 15, 448–456.
Duncan, J., 1980. The locus of interference in the perception of simultaneous stimuli. Psychological Review 87, 272–300.
Duncan, J., 1981. Directing attention in the visual field. Perception and Psychophysics 30, 90–93.
Duncan, J., 1983. Perceptual selection based on alphanumeric class: Evidence from partial reports. Perception and Psychophysics 33, 533–547.
Duncan, J., 1985. 'Visual search and visual attention'. In: M. Posner and O. Marin (eds.), Attention and performance XI. Hillsdale, NJ: Erlbaum.
Duncan, J. and G.W. Humphreys, 1989. Visual search and stimulus similarity. Psychological Review 96, 433–458.
Egeth, H.E., C.L. Folk and P.A. Mullin, 1989. 'Spatial parallelism in the processing of lines, letters and lexicality'. In: B.E. Shepp and S. Ballesteros (eds.), Object perception: Structure and process. Hillsdale, NJ: Erlbaum.
Egeth, H.E., J. Jonides and S. Wall, 1972. Parallel processing of multielement displays. Cognitive Psychology 3, 674–698.
Egeth, H.E., R.A. Virzi and H. Garbart, 1984. Searching for conjunctively defined targets. Journal of Experimental Psychology: Human Perception and Performance 10, 32–39.
Engel, F.L., 1971. Visual conspicuity, directed attention and retinal locus. Vision Research 11, 563–576.
Engel, F.L., 1974. Visual conspicuity and selective background interference in eccentric vision. Vision Research 14, 459–471.
Engel, F.L., 1977. Visual conspicuity, visual search and fixation tendencies of the eye. Vision Research 17, 95–108.
Enns, J.T., 1990. 'Three dimensional features that pop out in visual search'. In: D. Brogan (ed.), Visual search. London: Taylor and Francis.
Eriksen, B.A. and C.W. Eriksen, 1974. Effects of noise letters upon the identification of a target letter in a nonsearch task. Perception and Psychophysics 16, 143–149.
Eriksen, C.W. and J.E. Hoffman, 1972. Temporal and spatial characteristics of selective encoding from visual displays. Perception and Psychophysics 12, 201–204.
Eriksen, C.W. and J.E. Hoffman, 1973. The extent of processing of noise elements during selective encoding from visual displays. Perception and Psychophysics 14, 155–160.
Eriksen, C.W. and J.E. Hoffman, 1974. Selective attention: noise suppression or signal enhancement? Bulletin of the Psychonomic Society 4, 587–589.
Eriksen, C.W. and T.D. Murphy, 1987. Movement of attentional focus across the visual field: A critical look at the evidence. Perception and Psychophysics 42, 299–305.
Eriksen, C.W. and D.W. Schultz, 1979. Information processing in visual search: A continuous flow conception and experimental results. Perception and Psychophysics 25, 249–263.
Eriksen, C.W. and T. Spencer, 1969. Rate of information processing in visual perception: Some results and methodological considerations. Journal of Experimental Monographs 79, 1–16.
Eriksen, C.W. and J.D. St. James, 1986. Visual attention within and around the field of focal attention: A zoom lens model. Perception and Psychophysics 40, 225–240.
Eriksen, C.W. and J.M. Webb, 1989. Shifting of attentional focus within and about a visual display. Perception and Psychophysics 45, 175–183.

Eriksen, C.W. and Y.Y. Yeh, 1985. Allocation of attention in the visual field. Journal of Experimental Psychology: Human Perception and Performance 11, 583–597.

Estes, W.K., 1972. Interactions of signal and background variables in visual information processing. Perception and Psychophysics 12, 278–286.

Fisher, D.L., 1982. Limited-channel models of automatic detection: Capacity and scanning in visual search. Psychological Review 89, 662–692.

Fisher, D.L., 1984. Central capacity limits in consistent mapping, visual search tasks: Four channels or more? Cognitive Psychology 16, 449–484.

Fodor, J.A., 1983. The modularity of mind. Cambridge, MA: The MIT Press.

Folk, C.L. and H. Egeth, 1989. Does the identification of simple features require serial processing? Journal of Experimental Psychology: Human Perception and Performance 15, 97–110.

Garner, G.T., 1973. Evidence for independent parallel channel in tachistoscopic perception. Cognitive Psychology 4, 130–155.

Gleitman, H. and J. Jonides, 1976. The cost of categorization in visual search: Incomplete processing of target and field items. Perception and Psychophysics 20, 281–288.

Gleitman, H. and J. Jonides, 1978. The effect of set on categorization in visual search. Perception and Psychophysics 24, 361–368.

Hawkins, H.L., S.A. Hillyard, S.J. Luck, M. Moulana, C.J. Downing and D.P. Woodward, 1990. Chronometric analysis of apparent spotlight failure in endogenous visual orienting. Journal of Experimental Psychology: Human Perception and Performance 16, 802–811.

Hoffman, J.E., 1978. Search through a sequentially presented display. Perception and Psychophysics 23, 1–11.

Hoffman, J.E., 1979. A two-stage model of visual search. Perception and Psychophysics 25, 319–327.

Hoffman, J.E., 1986. Spatial attention in vision: Evidence for early selection. Psychological Research 48, 221–229.

Hoffman, J.E., B. Nelson and M.R. Houck, 1983. The role of attentional resources in automatic detection. Cognitive Psychology 15, 379–410.

Holmgren, J., J. Juola and R. Atkinson, 1974. Response latency in visual search with redundancy in the visual display. Perception and Psychophysics 16, 123–138.

Houck, M.R. and J.E. Hoffman, 1986. Conjunction of color and form without attention: Evidence from an orientation-contingent color after effect. Journal of Experimental Psychology: Human Perception and Performance 12, 186–199.

Humphreys, G., 1981. On varying the span of visual attention: Evidence for two modes of spatial attention. Quarterly Journal of Experimental Psychology 33A, 17–31.

Humphreys, G.W., P.T. Quinlan and M.J. Riddoch, 1989. Grouping processes in visual search: Effects with single- and combined-feature targets. Journal of Experimental Psychology: General 118, 258–279.

Isenberg, L., M.J. Nissen and L.C. Marchak, 1990. Attentional processing and the independence of color and orientation. Journal of Experimental Psychology: Human Perception and Performance 16, 869–878.

Jacobs, A.M., 1986. Eye movement control in visual search: How direct is visual span control? Perception and Psychophysics 39, 47–58.

Jacobs, A.M., 1987. On the role of blank spaces for eye movement control in visual search. Perception and Psychophysics 41, 473–479.

Johnston, W.A. and V.J. Dark, 1986. Selective attention. Annual Review of Psychology 37, 43–75.

Jonides, J., 1980. Towards a model of the mind's eye's movements. Canadian Journal of Psychology 34, 103–112.

Jonides, J., 1981. 'Voluntary vs. automatic control over the mind's eye's movement'. In: J.B. Long and A.D. Baddeley (eds.), Attention and performance IX (pp. 187–203). Hillsdale, NJ: Erlbaum.
Jonides, J., 1983. Further toward a model of the mind's eye's movement. Bulletin of the Psychonomic Society 21, 247–250.
Jonides, J. and H. Gleitman, 1972. A conceptual category effect in visual search: 'O' as letter or digit. Perception and Psychophysics 10, 457–460.
Jonides, J. and H. Gleitman, 1976. The benefit of categorization in visual search: Target location without identification. Perception and Psychophysics 20, 289–298.
Jonides, J. and S. Yantis, 1988. Uniqueness of abrupt visual onset in capturing attention. Perception and Psychophysics 43, 346–354.
Julesz, B., 1971. Foundations of cyclopean perception. Chicago, IL: University of Chicago Press.
Juola, J.F., D.G. Bouwhuis, E.E. Cooper and C.B. Warner, 1991. Control of attention around the fovea. Journal of Experimental Psychology: Human Perception and Performance 17, 125–141.
Kahneman, D. and A. Henik, 1981. 'Perceptual organization and attention'. In: M. Kubovy and J.R. Pomerantz (eds.), Perceptual organization. Hillsdale, NJ: Erlbaum.
Kahneman, D. and A.M. Treisman, 1984. 'Changing views of attention and automaticity'. In: R. Parasuraman and D.R. Davies (eds.), Varieties of attention (pp. 29–61). New York: Academic Press.
Kingstone, A. and R. Klein, 1991. Combining shape and position expectancies: Hierarchical processing and selective inhibition. Journal of Experimental Psychology: Human Perception and Performance 17, 512–519.
Klein, R. and E. Hansen, 1987. Spotlight failure in covert visual orienting. Bulletin of the Psychonomic Society 25, 447–450.
Klein, R. and E. Hansen, 1990. Chronometric analysis of apparent spotlight failure in endogenous visual orienting. Journal of Experimental Psychology: Human Perception and Performance 16, 790–801.
Koch, C. and S. Ullman, 1985. Shifts in selective visual attention; towards the underlying neural circuitry. Human Neurobiology 4, 219–227.
Kramer, A.F., M.P. Tahm and Y.Y. Yeh, in press. Movement and focused attention: A failure to replicate. Perception and Psychophysics.
Krueger, L.E., 1984. The category effect in visual search depends on physical rather than conceptual differences. Perception and Psychophysics 35, 558–564.
Kwak, H.W., D. Dagenbach and H.E. Egeth, 1991. Further evidence for a time-independent shift of the focus of attention. Perception and Psychophysics 49, 473–480.
LaBerge, D., 1983. Spatial extent of attention to letters and words. Journal of Experimental Psychology: Human Perception and Performance 9, 371–379.
LaBerge, D. and V. Brown, 1989. Theory of attentional operations in shape identification. Psychological Review 96, 101–124.
LaBerge, D., V. Brown, M. Carter, D. Bash and A. Hartley, 1991. Reducing the effects of adjacent distractors by narrowing attention. Journal of Experimental Psychology; Human Perception and Performance 17, 65–76.
Lambert, A., 1987. Expecting different categories at different locations and spatial selective attention. Quarterly Journal of Experimental Psychology 39, 61–76.
Lambert, A. and R. Hockey, 1986. Selective attention and performance with a multidimensional visual display. Journal of Experimental Psychology: Human Perception and Performance 12, 484–495.
Lévy-Schoen, A., 1981. 'Flexible and/or rigid control of oculomotor scanning behavior'. In: D.F. Fisher, R.A. Monty and J.W. Senders, (eds.), Eyemovements: Cognition and visual perception. Amsterdam: North-Holland.

Livingstone, M. and D. Hubel, 1987. Segregation of form, color, movement and depth: Anatomy, physiology and perception. Science 240, 740–749.

Marr, D., 1982. Vision: A computational investigation into the human representation and processing of visual information. San Francisco, CA: Freeman.

Maunsell, J.H.R. and W.T. Newsome, 1987. Visual processing in monkey extrastriate cortex. Annual Review of Neuroscience 10, 363–401.

McLeod, P., J. Diver and J. Crisp, 1988. Visual search for a conjunction of movement and form is parallel. Nature 332, 154–155.

McLeod, P., J. Diver, Z. Dienes and J. Crisp, 1991. Filtering by movement in visual search. Journal of Experimental Psychology: Human Perception and Performance 17, 55–64.

Miller, J., 1991. The flanker compatibility effect as a function of visual angle, attentional locus, visual transients and perceptual load: A search for boundary conditions. Perception and Psychophysics 49, 270–288.

Mordkoff, T.J., S. Yantis and H.E. Egeth, 1990. Detecting conjunction of color and form in parallel. Perception and Psychophysics 48, 157–168.

Müller, H.J. and J.M. Findley, 1987. Sensitivity and criterion effects in the spatial cuing of visual attention. Perception and Psychophysics 42, 383–399.

Müller, H.J. and G.W. Humphreys, 1991. Luminance increment detection: Capacity-limited or not? Journal of Experimental Psychology: Human Perception and Performance 17, 107–124.

Müller, H.J. and P.M.A. Rabbitt, 1989. Reflexive and voluntary orienting of visual attention: Time course of activation and resistance to interruption. Journal of Experimental Psychology: Human Perception and Performance 15, 315–330.

Murphy, T.D. and C.W. Eriksen, 1987. Temporal changes in the distribution of response to precues. Perception and Psychophysics 42, 576–586.

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

Navon, D., 1989. Attentional selection: Early, late, or neither? European Journal of Cognitive Psychology 1, 47–68.

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

Neisser, U., R. Novick and R. Lazar, 1963. Searching for ten targets simultaneously. Perceptual and Motor Skills 17, 955–962.

Neumann, O., 1984. 'Automatic processing: A review of recent findings and a plea for an old theory'. In: W. Prinz and A.F. Sanders, Cognition and motor processes (pp. 255–290). Berlin: Springer-Verlag.

Neumann, O., 1987. 'Beyond capacity: A functional view of attention'. In: H.Heuer and A.F. Sanders (eds.), Perspectives on perception and action. Hillsdale, NJ: Erlbaum.

Neumann, O., 1990. 'Visual attention and action'. In: O. Neumann and W. Prinz (eds.), Relationships between perception and action: Current approaches. New York: Springer-Verlag.

Nissen, M.J., 1985. 'Accessing features and objects: Is location special?' In: M.I. Posner and O.S.M. Marin (eds.), Attention and performance XI (pp. 205–219). Hillsdale, NJ: Erlbaum.

Pashler, H., 1987. Detecting conjunctions of color and form: Reassessing the serial search hypothesis. Perception and Psychophysics 41, 191–201.

Pashler, H., 1988. Cross-dimensional interaction and texture segregation. Perception and Psychophysics 43, 307–318.

Pashler, H. and P. Badgio, 1985.Visual attention and stimulus identification. Journal of Experimental Psychology: Human Perception and Performance 11, 105–121.

Polyak, S., 1957. The vertebrate visual system. Chicago, IL: University of Chicago Press.

Posner, M.I., 1978. Chronometric explorations of mind. Hillsdale, NJ: Erlbaum.

Posner, M.I., 1980. Orienting of attention. Quarterly Journal of Experimental Psychology 32, 3–25.

Posner, M.I. and Y. Cohen, 1984. 'Components of visual orienting'. In: H. Bouman and D.G. Bouwhuis (eds.), Attention and performance X. Hillsdale, NJ: Erlbaum.

Posner, M.I., M.J. Nissen and W.C. Ogden, 1978. 'Attended and unattended processing models: The role of set for spatial location'. In: H.L. Pick and E. Saltzman (eds.), Modes of perceiving and processing information. Hillsdale, NJ: Erlbaum.

Posner, M.I. and C.R.R. Snyder, 1975. 'Facilitation and inhibition in the processing of signals'. In: P.M.A. Rabbitt and S. Dornic (eds.), Attention and performance V. New York: Academic Press.

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

Quinlan, P.T. and G.W. Humphreys, 1987. Visual search for targets defined by combinations of color, shape and size: An examination of the task constraints on feature and conjunction searches. Perception and Psychophysics 41, 455–472.

Rabbitt, P.M.A., 1967. Learning to ignore irrelevant information. American Journal of Experimental Psychology 80, 1–13.

Sagi, D. and B. Julesz, 1985a. Detection versus discrimination of visual orientation. Perception 14, 619–628.

Sagi, D. and B. Julesz, 1985b. 'Where' and 'what' in vision. Science 228, 1217–1219.

Sanders, A.F., 1963. The selective process in the functional visual field. Assen: v. Gorkum.

Sanders, A.F., 1970. Some aspects of the selective process in the functional visual field. Ergonomics 13, 101–117.

Sanders, A.F. and M. Donk, 1991. 'Visual search'. In: O. Neumann and A.F. Sanders (eds.), Attention. New York: Wiley.

Schneider, W. and R.M. Shiffrin, 1977. Controlled and automatic human information processing. I. Detection, search and attention. Psychological Review 84, 1–66.

Shaw, M.L., 1978. A capacity allocation model for reaction time. Journal of Experimental psychology: Human Perception and Performance 4, 586–598.

Shaw, M.L., 1984. 'Division of attention among spatial locations; A fundamental difference between detection of letters and detection of luminance increments'. In: H. Bouma and D.G. Bouwhuis (eds.), Attention and performance X. Hillsdale, NJ: Erlbaum.

Shiffrin, R.M. and W. Schneider, 1977. Controlled and automatic human information processing. II: Perceptual learning, automatic attending and a general theory. Psychological Review 84, 127–190.

Shulman, G.L., R. Remington and J.P. Mclean, 1979. Moving attention through visual space. Journal of Experimental Psychology: Human Perception and Performance 9, 522–526.

Skelton, J.M. and C.W. Eriksen, 1976. Spatial characteristics of selective attention in letter matching. Bulletin of the Psychonomic Society 7, 136–138.

Steinman, S.B., 1987. Serial and parallel search in pattern vision. Perception 16, 389–399.

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

Theeuwes, J., 1989. Effects of location and form cuing on the allocation of attention in the visual field. Acta Psychologica 72, 177–192.

Theeuwes, J., 1990. Perceptual selectivity is task dependent: Evidence from selective search. Acta Psychologica 74, 81–99.

Theeuwes, J., 1991a. Categorization and identification of simultaneous targets. Acta Psychologica 76, 73–86.

Theeuwes, J., 1991b. Exogenous and endogenous control of attention: the effect of visual onsets and offsets. Perception and Psychophysics 49, 83–90.

Theeuwes, J., 1991c. Cross-dimensional perceptual selectivity. Perception and Psychophysics 50, 184–193.

Theeuwes, J., in press. Perceptual selectivity for color and form. Perception and Psychophysics.

Todd, J.T. and P. van Gelder, 1979, Implications of a sustained-transient dichotomy for the measurement of human performance. Journal of Experimental Psychology: Human Perception and Performance 5, 625–638.

Townsend, J.T., 1972. Some results on the identifiability of parallel and serial processes. British Journal of Psychology 25, 168–199.

Treisman, A.M., 1982. Perceptual grouping and attention in visual search for features and for objects. Journal of Experimental Psychology: Human Perception and Performance 8, 194–214.

Treisman, A.M., 1986. 'Properties, parts and objects'. In: K. Boff, L. Kaufman and J. Thomas (eds.), Handbook of perception and human performance, Vol. 2: Cognitive processes and performance (ch. 35, pp. 1–70). New York: Wiley.

Treisman, A.M., 1988. Feature and objects: The fourteenth Bartlett memorial lecture. The Quarterly Journal of Experimental Psychology 40, 201–237.

Treisman, A.M. and G. Gelade, 1980. A feature integration theory of attention. Cognitive Psychology 12, 97–136.

Treisman, A.M. and S. Gormican, 1988. Feature search in early vision: Evidence from search asymmetries. Psychological Review 95, 15–48.

Treisman, A.M. and S. Sato, 1990. Conjunction search revisited. Journal of Experimental Psychology: Human Perception and Performance 16, 459–478.

Treisman, A.M. and H. Schmidt, 1982. Illusory conjunctions in the perception of objects. Cognitive Psychology 14, 107–141.

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

Tsal, Y., 1983. Movements of attention across the visual field. Journal of Experimental Psychology: Human Perception and Performance 9, 523–530.

Tsal, Y. and N. Lavie, 1988. Attending to color and shape: The special role of location in selective visual processing. Perception and Psychophysics 44, 15–21.

Ullman, S., 1984. Visual routines. Cognition 18, 97–159.

Van der Heijden, A.H.C., 1975. Some evidence for a limited capacity parallel selfterminating process in simple visual search. Acta Psychologica 39, 21–41.

Van der Heijden, A.H.C., 1981. Short-term visual information forgetting. London: Routledge and Kegan Paul.

Van der Heijden, A.H.C., 1992. Selective attention in vision. London: Routledge and Kegan Paul.

Van der Heijden, A.H.C., W. LaHeij and J.P.A. Boer, 1983. Parallel processing of redundant targets in simple visual search tasks. Psychological Research 45, 235–254.

Van der Heijden, A.H.C., G. Wolters, J.C. Groep and R. Hagenaar, 1987. Single-letter recognition accuracy benefits from advance cuing of location. Perception and Psychophysics 42, 503–509.

Wolfe, J.M., K.R. Cave and S.L. Franzel, 1989. Guided search: An alternative to the feature integration model for visual search. Journal of Experimental Psychology: Human Perception and Performance 15, 419–433.

Wolfe, J.M., K.P. Yu, M.I. Stewart, A.D. Shorter, S.R. Friedman-Hill and K.R. Cave, 1990. Limitations on the parallel guidance of visual search: Color × color and orientation × orientation conjunctions. Journal of Experimental Psychology: Human Perception and Performance 16, 879–892.

Wolford, G., 1975. Perturbation model for letter identification. Psychological Review 82, 184–199.

Wolford, G. and L. Chambers, 1983. Lateral masking as a function of spacing. Perception and Psychophysics 33, 129–138.

Wurtz, R.H. and C.W. Mohler, 1976. Organization of monkey superior colliculus: Enhanced visual response of superficial layer cells. Journal of Neurophysiology 39, 745–765.

Yantis, S., 1988. On analog movements of visual attention. Perception and Psychophysics 43, 203–206.

Yantis, S. and J.C. Johnston, 1990. On the locus of visual selection: Evidence from focused attention tasks. Journal of Experimental Psychology: Human Perception and Performance 16, 135–149.

Yantis, S. and J. Jonides, 1984. Abrupt visual onsets and selective attention: Evidence from visual search. Journal of Experimental Psychology: Human Perception and Performance 10, 601–621.

Yantis, S. and J. Jonides, 1990. Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. Journal of Experimental Psychology: Human Perception and Performance 16, 121–134.

Yellott, J.L., B.A. Wandell and T.N. Cornsweet, 1984. 'The beginnings of visual perception. The retinal image and its initial encoding'. In: Handbook of physiology: The nervous system. Berlin: Springer-Verlag.