What You Can See Outside the Focus of Attention

Thesis by Jiajun Wen

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California
2000

(Defended May 21, 1999)

Acknowledgements

- Thanks to my chief collaborators - Jochen Braun and Laurent Itti - whose influence has made its way into every chapter. In addition, Achim (Jochen) has been a generous tutor.

- Special thanks to Prof. Terry Sejnowski and Tom Albright for providing space and equipment. Without this help, I could not have finished this work.

- Thanks to Pietro Perona, Shinsuke Shimojo, Giedrius Buracas, Barbara Zenger and Wyeth Bair for stimulating discussion. I have benefited a lot from these interactions.

- Salute to the **Captain**, Christof Koch. His enthusiastism about science and continuing support have been the driving force of this work. I am deeply indebted to him.

- Gratitude to Kathleen Bartle-Schulweis (director of Caltech Women's Center), Parandeh Kia (director of International Student Programs) and Prof. Nai-Chang Yeh. I am extremely grateful for their help during the darkest period in my life, and for their continuing effort to make Caltech a better place for women and minorities.

Abstract

Visual attention is widely considered to be first and foremost a means of controlling the flow of information between different levels of processing. Stimuli selected by attention are thought to gain access to higher levels of processing, including awareness, recognition, memory, and the generation of voluntary responses, while stimuli rejected by attention are deemed to be denied such access. However, visual attention remains poorly understood in many aspects. In this study, we try to investigate some part of the problem.

Several behavioral methods have been devised for the study of visual attention, one of them - the concurrent task paradigm - has gained more and more recognition and is widely used in different studies. In the first part of the thesis, we use this paradigm to investigate an important question: whether attentional capacity is differentiated or not. In non-human primates, neuronal sensitivities to stimulus dimensions such as form, color, or motion are concentrated in different visual cortical areas (Felleman & Van Essen, 1991; Van Essen & Gallant 1994), and functional imaging studies show similar functional specializations in the visual cortex of humans (Corbetta *et al.*, 1990, 1995; Van Essen & Drury, 1997). It has been proposed that attentional capacity is "differentiated" (Sperling & Dosher, 1986; Pashler, 1997) in that discriminations concerning different dimensions draw at least in part on different attentional capacities and are thus less liable to interfere than discriminations concerning the same stimulus dimension. However, Duncan and colleagues have argued at length for a contrary point of view (Desimone & Duncan, 1995; Duncan *et al.*, 1997). In their view, due to the strong interdependence of the subsystems that process different stimulus dimensions, attentional capacity is "undifferentiated". By conducting experiments on the concurrent discrimination of form, color and motion attributes, we test which hypothesis is true. We quantify interference between concurrent discriminations by establishing the attention-operating characteristic. Interference is indistinguishable for similar and dissimilar task combinations (form-form, color-color, motion-motion, and color-form, color-motion, motion-color, motion-form, respectively) for both demanding and less-demanding tasks. These results suggest strongly that different visual discriminations draw on the same attentional capacity, in other words, that the capacity of visual attention is undifferentiated.

After establishing that a task is equally attentional demanding no matter what the stimulus dimension the second task is based on, we use the concurrent task paradigm to investigate how attention affects spatial vision. We measure thresholds for discriminating the contrast, orientation, and spatial frequency of simple patterns that are either fully or poorly attended. We observe differences of about 20% in contrast detection thresholds, 40-50% in contrast discrimination thresholds (and appearance of the "dipper"), 60-70% in orientation and spatial frequency discrimination thresholds, and up to 50% in contrast masking thresholds. These observations tightly constrain any effect attention may have on the visual filters and/or the interactions amongst filters that are thought to underlie basic spatial vision. Comparison with a computational model due to Laurent Itti shows that the observed effects of attention are consistent with stronger interactions amongst filters, but not with a change in noise parameters, as is sometimes thought (Bonnel & Miller, 1994; Lu and Dosher, 1998. Essentially, the effects of attention on different thresholds are too disparate to be accommodated by a single change in noise parameters. In the framework of our model, the strength of the interaction amongst filters is controlled by the exponents of a power law, γ and δ. The larger exponents activate what is best described as a *winner-take-all competition amongst visual filters*. Attention accentuates existing differences between filter responses, boosting filters that respond relatively well to a given stimulus, while suppressing filters that respond relatively poorly. This explains the perceptual advantage conferred by attention: attention enhances the sensory representation by restricting responses to the filters tuned best to the stimulus at hand. Our experimental and modeling results show that the activation of a winner-take-all competition amongst overlapping visual filters explains many basic perceptual consequences of attention.

Contents

Chapter 1

Introduction

1.1 Perception outside focus of attention

"Everyone knows what attention is", was one of William James' famous statements (James, *1890*). There is certain truth about it. The word "attention" is commonly used in everyday language, and people seem to have no problem understanding each other when they use it. This suggests that they must share some common experiences and thus common notions about it. However, by no means does this indicate we really understand what attention is. The past few decades of psychophysical and neurophysiological study have produced a lot of information about many different phenomena related to attention and many theories trying to account for these phenomena have been proposed. Research of attention has been an active and controvertial field. This thesis will deal with one aspect of attention: visual attention.

Visual attention is one of several mechanisms that modulate neural information processing in the human and non-human primate brain (Heilman, *et al.*, 1990; Posner & Petersen, 1990; Colby, 1991; Kinchla, 1992; Posner & Driver, 1992). Stimuli selected by attention are thought to gain access to higher levels of processing, including awareness, recognition, memory, and the generation of voluntary responses, while stimuli rejected by attention are deemed to be denied by such access. This contrasts with lower levels of processing, which process all stimuli indiscriminately, whether selected by attention or not. These lower levels of processing are sometimes thought to contain "feature maps", that is, topographic representations of elementary stimulus features such as orientation, color, motion. Because of this parallel architecture, these lower levels of processing are thought to process an essentially unlimited capacity and to process many stimuli at the same time, whereas the higher levels of processing are thought to have a limited capacity, and to process only a limited number of stimuli selected by attention at one time. Visual attention allows us to separate currently behaviorally relevant from irrelevant information, enabling us to concentrate on a mere fraction of the total information (Helmholtz, 1850; James, 1890; Broadbent, 1958; Neisser, 1967; Treisman, 1969). Visual attention has been described as a "limited capacity", a "spotlight", a "biased competition", and as something that is "guided by a saliency map" (Koch & Ullman, 1985; Robinson & Petersen, 1992; Desimone & Duncan, 1995). These, at best metaphorical, descriptions reflect the two major facts known about attention: first, visual attention can select only a limited amount of information at any one time; and second, the attentional selection is stimulus-driven ("bottom-up") unless volition intervenes ("top-down") (Helmholtz, 1866; Schneider & Shiffrin, 1977; Haenny, *et al.*, 1988; Corbetta *et al.*, 1990; Yantis & Jonides, 1990; Yantis, 1994; Mack *et al.*, 1992; Rock *et al.*, 1992).

Several psychophysical techniques (such as visual search, concurrent-task paradigm, etc.) have been useful in understanding the extent to which different visual tasks depend on attention. The results suggest that the richness of visual experience derives not only from a narrow focus of attention, but also reflects a simultaneous awareness of the entire visual space. Indeed, psychophysical results show that focusing attention narrowly on one location reduces, but does not eliminate, visual performance with respect to other locations in a visual scene. The well-known phenomenon of "pop-out" in parallel search is an example of this ambient visual perception. Thus it appears that observers enjoy a significant visual awareness of poorly attended stimuli, especially when these stimuli are salient and "pop-out" from the scene. In a series of psychophysical experiments, Braun and Julesz (1998a,b) investigated the role of attention in two types of visual search–search for less salient or more salient object. In their experiments, observers were asked to carry out a letter discrimination and a search task concurrently. In order to discriminate the letters, observers had to direct visual attention to the center of the display and leave the periphery (which contained the target and distractors) unattended. In the concurrent task situation, visual search for the least salient item was seriously impaired while search for the most salient item was only moderately affected.

With close collaboration with Dr. Braun, I have continued to investigate how attention affects ambient perception in several qualitative and quantitative ways. In order to study perception outside focus of attention, one has to induce observers to focus attention on one part of the display and thus at least partially withdraw attention from other parts of the display, for a certain amount of time. We manipulate the observer's distribution of attention with a concurrent-task paradigm. Typically, the visual display consists of two parts–central and peripheral, which pose two independent visual tasks, and the observer is asked to perform either one task, or the other, or both (with different priorities, depending on instructions). To prevent the observer from attending to different parts sequentially or making saccades, the display is presented briefly and visual persistence is prevented by masking (stimulus-onset-asynchrony, or SOA, 60-200 ms). This means that each part of the display is visible just long enough for to permit discrimination on its own. Another crucial aspect of this paradigm is that the eyes remains fixated at the center of the display at all times. This ensures that the physical stimulus is the same, no matter which task or tasks are being performed. Thus any difference we observe under different instructions is due to attention manipulation. The point of these experiments is to determine how performance on one task is affected when attention focuses on the other.

1.2 Undifferentiated attentional capacity

According to monkey physiology studies, neuronal sensitivities to stimulus dimensions such as form, color and motion are concentrated in different specialized visual cortical areas (Fellman & Van Essen, 1991; Van Essen & Gallant, 1994), and funtional imaging studies show similar funtional specializations in human visual cortex (Corbetta, Miezin, Dobmeyer, Shulman & Petersen, 1990; Corbetta, Shulman, Miezin & Petersen, 1995; Van Essen & Drury, 1997). In 1986, Sperling and Dosher proposed that attentional capacity is "differentiated" in that discriminations concerning different dimensions draw at least in part on different attentional capacities and thus interfere less than discriminations concerning the same stimulus dimension. However, Duncan and colleagues have argued at length for an opposite point of view (Desimone & Duncan, 1995; Duncan, Humphreys & Ward, 1997). They argued that the neuronal responses to different visual objects compete and "a gain in activity for one object is accompanied by a loss in activity by others". Due to the strong interdependence of the subsystems that process different stimulus dimensions, discriminating different attribute of two visual objects, respectively, should be equally difficult as discriminating same attribute of both objects. Thus, attentional capacity is "undifferentiated".

The question whether visual attention is differentiated or not can be approached in several ways. Our approach is to study the discrimination of two separate objects. When attention is divided between two objects concerning two independent visual tasks, a trained observer can "trade off" performance on one task against performance on the other. And the "trade off" function is a sensitive measure of the extent to which two tasks interfere with each other. We can use this measure to determine whether visual attention is differentiated or not. If attention is differentiated, two tasks concerning two different attributes (e.g., motion and color) should interfere less compared to two tasks concerning the same attribute (e.g., color and color). If attention is undifferentiated, the interference should be the same for the two conditions. In chapter 2 and chapter 3, we report experimental results which test these two alternative hypotheses.

In the early nineties, Braun reported that form discrimination task is highly attentional demanding and two concurrent form discrimination tasks interfere significantly. The experiment combines a central and a periphery task, both of which involve discriminating between

T- and L-shaped elements. In the central task, observers inspect the 5 central letters and report whether they are the "same" or "different". In the periphery task, observers report whether the periphery element, which appears at random locations at 4 deg eccentricity, is a "L" or "T". Braun reported essentially linear trade off curve for the two tasks. In order to determine if attention capacity is differentiated, we conducted analogous experiments with other kinds of stimulus elements, which require the discrimination of motion or color intead of form. In motion task, "dumbbell" shapes rotating either clockwise or anti-clockwise are used; in color task, bisected disks which are either red-green or green-red are used.

The question we asked was whether the "trade off" curve is the same or not between similar task combinations (i.e., form-form, motion-motion, color-color) and dissimilar task combinations (i.e., color-form, color-motion, motion-color, etc.). The critical finding is that in all combinations, double-task performance is characterized by a linear trade-off between tasks: increased performance on one task results same amount decreased performance on the other task. When either task is performed at its best (same performance level as the single task), performance of the other drops to chance level. The fact that the outcome of these experiments is the same no matter which discrimination task is used to engage attention shows that different visual discrimination engage visual attention capacity to the same extent. Thus, there is only one attentional capacity.

One question was raised often when we presented this study: we use highly demanding tasks and find maximum linear trade-off under all conditions. Thus, is it possible that there is some subtle difference between similar and dissimilar task combinations, which is masked by the maximum ceiling trade-off? In order to answer this question, we carried out another set of experiments with less demanding discrimination tasks.

The basic discrimination task is the same: T vs. L for form discrimination, red-green vs. green-red disk for color discrimination, clockwise or anti-clockwise rotation for motion task. Two single elements appear at left and right hemisphere and represent two independent discrimination tasks. The right stimulus appears at four possible locations, and the left element appears at the same elevation and same eccentricity in the contralateral side. Because the two tasks are symmetrical and both only involve one single element, they are much less attentional demanding compared to the central task in the last set of experiments, which involves five elements.

Again, we investigated whether the characteristics of the trade-off curve is the same or different for similar and dissimilar task combinations. All double-task performance is characterized by an intermediate trade-off between tasks: when either task is performed at its best, the other task is performed above chance but less than the single-task performance, and any improvement in the performance of one task is at the expense of a partial reduction in the performance of the other. Quantitative analysis shows no significant difference for characteristics of AOC curves for similar and dissimilar task combinations.

Thus, even though different attributes of a visual scene (e.g., color, motion, etc.) are processed in different parts of the brain as shown in physiological and functional imaging studies, our result shows visual attention can be considered a unitary process, suggesting that visual cortical areas processing different stimulus dimensions must be highly integrated.

1.3 Attentional effect on spatial vision

Having established that attention is a unitary process, we can now safely use concurrent task paradigm to investigate how attention affects many fundamental properties of visual perception. Since attentional capacity is undifferentiated, it does not matter which specific task we use to engage attention as long as it is attention demanding. Thus we can focus on the other task and investigate how its performance is affected by different attentional state.

In chapter 4, we will take a closer look at the discrimination of elementary stimulus attributes such as contrast, orientation, and spatial frequency. The behavioral thresholds for these stimulus attributes collectively characterize "spatial vision" and have been studied for several decades (Nachmias & Sansbury, 1974; Wilson, 1980; Legge & Foley, 1980). Quantitative accounts of these thresholds have become increasingly refined and usually involve a population of "noisy filters" tuned to different orientations and spatial frequencies. Our goal is to compare behavioral thresholds for "poorly attended" stimulus, on the one hand, and thresholds for "fully attended" stimulus, on the other hand. This comparison should reveal in detail how attention alters the early levels of processing which underlie "spatial vision".

We use the concurrent task paradigm described before to establish thresholds when stimulus is poorly attended. One of highly attention demanding tasks (discriminating 5 central letters as the same or different) is used to engage attention. In the double-task condition, observers are instructed that the central task is the primary task and they are not supposed to compromise its performance. Thus, little or no attention remains for peripheral stimulus. Nevertheless, trained observers reliably perform the secondary task with respect to the periphery stimulus, that we can use staircase method to measure the threshold of "poorly attended" stimulus. In the single-task condition, observers view the same display with the same fixation, but ignore the central stimuli. In this condition, attenion is free to focus on the peripheral task, with which we can establish "single-task thresholds". The comparison of single- and double-task thresholds reveals if and how attention affects spatial vision.

Several experiments are conducted to compare threhsolds under single- and double-task conditions. When periphery stimulus is fully attended, contrast detection thresholds are about 20% lower and contrast discrimination thresholds are about 40-50% lower than corresponding thresholds when the periphery stimulus is poorly attended. In addition, the facilitation of detection as mask contrast increases from zero to subthreshold level (the well-known "dipper") is only clearly observed when stimulus is fully attended. Please notice that the location of pheripheral stimulus varies from trial to trial (in order to prevent saccades) and that positional uncertainty is known to reduce the dipper. Therefore, our data may well underestimate the true reduction in the dipper.

Attentional effect on spatial frequency and orientation discrimination are even more dramatic. Spatial frequency discrimination thresholds are about 60% lower and orientation discrimination thresholds are about 70% lower when the stimulus is fully attended, compared to thresholds when it is poorly attended. Both single- and double-task thresholds remain constant for contrast values above 20%.

In order to investigate how different channels of spatial vision interact with each other, we also study how interactions between superimposed stimuli of different orientation or spatial frequency are altered by attention. When target and mask have similar orientation or spatial frequency, they interact most strongly and the attentional effect is biggest as well. Attention lowers the maximal threshold by about 50%. As target and mask become progressively more different, fully and poorly attended thresholds decrease towards the similar baseline level. The converging baseline threshold is comparable to detection threshold without mask, indicating minimal interactions between targets and masks of very different orientation or spatial frequency.

Over all, these observations reveal how attention affects the mechanisms underlying spatial vision. In all five experiments, we are able to establish threshold measurement for both fully and poorly attended stimulus, suggesting that lack of attention only modulates spatial perception instead of eliminating it.

With a closer look, we can draw some qualitative conclusions about the way in which attention alters the neural representation of contrast, orientation and spatial frequency. For example, attention evidently does not act primarily by reducing background noise as the data show that the smallest attentional effects are obtained for stimuli with the lowest contrast. A mere reduction of background noise would produce largest attentional effects for stimuli with lowest contrast level, the exact opposite of experimental data. Other aspects of qualitative conclusions are discussed in detail in chapter 4.

In collaboration with Laurent Itti and Dr. Jochen Braun, we developed a model to account for our observation quantitatively. The modeling work was mainly done by my colleague Laurent Itti and will be presented in detail in his forthcoming thesis. The critical finding is that a simple model of response normalization implemented by divisive inhibition accounts for all the attentional changes to contrast, orientation, and spatial frequency thresholds we have observed. A brief summary of the model is included at the end of chapter 4.

In chapter 5, we briefly discuss some prelinminary data about how attention affects motion perception. This work is done with collaboration of Dr. Giedrius Buracas at Salk Institue. Dr. Buracas is now continuing this work.

The results presented in this thesis have been reported at several meetings including ARVO and NIPS (Wen *et al.* ARVO95; Wen *et al.* ARVO96; Lee *et al.* ARVO97; Lee *et al.* ARVO98; Lee *et al.* ARVO99; Buracas *et al.* ARVO98; Itti *et al.* NIPS97; Itti *et al.* NIPS98). Some results have been published in peer reviewed journals (Lee *et al.* Vision Research, 1997; Lee *et al.* Perception & Psychophysics, 1999; Lee *et al.* Nature Neuroscience, 1999). In most cases, the thesis author has been publishing using the name Dale Kathleen Lee.

Chapter 2

Attentional capacity for demanding tasks

2.1 Background

When two visual objects are presented briefly and simultaneously, observers often readily discriminate one object, but find it next to impossible to discriminate both (Bonnel & Miller, 1994; Bonnel & Prinzmetal, 1998; Bonnel, Stein, & Bertucci, 1992; Braun & Julesz, 1998; Braun & Sagi, 1991; Duncan, 1984; Duncan, Ward, & Shapiro, 1994; Fisher, 1984; Kahneman & Treisman, 1984; Kleiss & Lane, 1986; Sperling & Melchner, 1978; Ward, Duncan, Shapiro, 1996). The difficulty of simultaneously discriminating two visual objects is thought to reflect the "limited capacity" of visual attention (*e.g.*, Broadbent, 1971; Kahnemann, 1973; Neisser, 1967). Here we ask if attentional capacity limitations depend on the similarity of the discriminations in question, specifically, if they depend on whether similar or dissimilar stimulus dimensions are involved (*e.g.*, form, color, or motion).

In non-human primates, neuronal sensitivities to stimulus dimensions such as form, color, or motion are concentrated in different visual cortical areas (Felleman & Van Essen, 1991; Van Essen & Gallant, 1994), and functional imaging studies show similar functional specializations in the visual cortex of humans (Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1990; Corbetta, Shulman, Miezin, & Petersen, 1995; Van Essen & Drury, 1997). More than 20 years ago, it was suggested that dimensions such as form, color, and motion are processed by separate subsystems ("perceptual analyzers") and that simultaneous visual discriminations should be problematic only if they engage the same subsystem, but not if they involve different subsystems (Allport, 1971; Treisman, 1969). In other words, two form discriminations or two color discriminations should interfere severely, but one form and one color discrimination should not. A more abstract formulation of this point of view is that attentional capacity is "differentiated" (Sperling & Dosher, 1986) in that discriminations concerning different dimensions draw at least in part on different attentional capacities and are thus less likely to interfere than discriminations concerning the same stimulus dimension. Many current theories of attention are sympathetic to this position (*e.g.*, Kinsbourne, 1981; Pashler, 1997).

Of course, the fact that visual cortex is to some degree functionally specialized does not mean that attentional capacity is necessarily differentiated. It remains to be seen whether the subsystems that process different stimulus dimensions are sufficiently independent to allow simultaneous discrimination of, say, the form of one visual object and the color of another. In fact, Duncan and colleagues have argued at length for a contrary point of view (Desimone & Duncan, 1995; Duncan, Humphreys & Ward, 1997). In their view, the neural responses to different visual objects compete and "a gain in activity for one object is accompanied by a loss in activity for others". Although the response to each object involves multiple subsystems, competition is integrated "as a winning object emerges in one [sub]system, it tends also to become dominant in others" (Duncan, Humphreys, Ward, 1997). Due to the strong interdependence of the subsystems that process different stimulus dimensions, discriminating the form and color of two visual objects, respectively, should be just as difficult as discriminating the forms of both. In this view, therefore, attentional capacity is "undifferentiated".

The question as to whether visual attention is differentiated or undifferentiated can be approached in a number of ways. One approach is to study the discrimination of different attributes of the same object (e.g., color and form), and to take advantage of spontaneous fluctuations in the allocation of attention to the target attributes. If success on one attribute and success on the other are positively correlated, one may conclude that both attributes are subject to the same fluctuations and, therefore, that attention is undifferentiated (Bonnel &

Prinzmetal, 1998; Monheit & Johnston, 1994). If attention were differentiated, one might expect the two attributes to be subject to independent fluctuations and therefore to exhibit no such correlation (Isenberg, Nissen, & Marchak, 1990; Nissen, 1985). The drawback of this approach is that the observed correlations are small and their existence difficult to establish (Johnston, Ruthruff, & Monheit, 1997; van der Velde & van der Heijden, 1997).

A more promising approach is to study the discrimination of two separate objects (Duncan, 1993; Sperling & Dosher, 1986). When attention is divided between two objects, a practised observer is generally able to "trade off" performance on one object against performance on the other. In other words, the observer can alter the division of attention and fare better with one task and worse with the other. For some tasks, such voluntary changes in attentional allocation produce large changes in performance, in some cases from chance to ceiling. As a result, the dependence of the performance of one task on performance of the other ("trade-off function") is a sensitive measure of the extent to which two tasks compete for attention. Armed with this measure, we can determine whether visual attention is differentiated or undifferentiated. If attention is differentiated, two tasks should interfere less with different attributes (e.g., color and form) than when the same attributes are discriminated (e.g., color and color, or form and form). If attention is undifferentiated, however, the discrimination of same attributes and different attributes should lead to the same degree of interference.

A complicating factor is that the discrimination of some stimulus attributes requires more attentional capacity than that of others. Using concurrent tasks, we previously found that discrimination of letter shape requires essentially full attention, while discrimination of triangular or circular form and the discrimination of orientation are significantly less demanding of attention, and the discrimination of color is essentially free of attentional cost (Braun, 1994, 1998; Braun & Julesz, 1998). A similar ranking of attentional cost emerges from studies of visual search: search for a unique letter shape exhibits steeper search slopes than search for a unique orientation or color. In general, the attributes whose discrimination results in greater interference between concurrent tasks are also those for which visual search exhibits steeper "search slopes" (*e.g.*, Enns & Rensink, 1990; Julesz, 1981, 1984; Treisman & Gelade, 1980; Treisman & Souther, 1985; Treisman & Gormican, 1988; Wolfe, 1994). Whether attentional costs of different discriminations are distributed on a continuum, or whether the distribution is bimodal with some discriminations being indeed free of attentional cost, is a matter of continuing debate (*e.g.*, Braun, 1998, Braun & Julesz, 1998; Joseph, Chun, & Nakayama, 1997). What matters here is that the undisputed and sizeable differences in attentional cost need to be taken into account in any *quantitative* comparison of similar and dissimilar task combinations.

A number of previous studies have compared similar and dissimilar task combinations, but with less than conclusive results. Duncan and colleagues have studied the discrimination of form attributes (letter form, size, orientation, length), surface attributes (color, brightness, texture), and motion attributes (left- or rightward motion) in various combinations (Duncan, 1993; Duncan & Nimmo-Smith, 1996; Ward, Duncan, & Shapiro, 1997). For each task combination, interference was assessed by comparing performance of each task alone with performance of both tasks together. In general, similar and dissimilar task combinations exhibited comparable degrees of interference, that is, comparable differences between performance alone and performance together, although some dissimilar combinations — especially those involving color discrimination — exhibited little or no interference. Duncan and colleagues concluded that all discriminations, with the possible exception of color, draw on the same attentional capacity. However, the degree of interference encountered even with similar task combinations was small, and varied substantially between different task combinations. (The difference between performance together and performance alone ranged from $\approx$5% to $\approx$15%, where chance performance is 50%).

The present study modifies this paradigm in order to obtain more conclusive results. First, we maximize task interference by using discriminations with high attentional cost. We increase attentional cost by increasing positional uncertainty, and by requiring the *joint* discrimination of form, color, or motion and *position*. (As a result, the difference between performance together and performance alone ranges from $\approx$28% to $\approx$42%). Second, we establish a complete trade-off function for each task combination. This allows us to quantify the degree of interference in each case, and makes it easier to relate our findings to the substantial theoretical literature on "divided" attention (*e.g.*, Navon & Gopher, 1979; Norman & Bobrow, 1975; Sperling & Dosher, 1986; Sperling & Melchner & 1978).

In the next chapter, we continue to study attention capacity using less demanding tasks. We decrease attentional cost by decreasing positional uncertainty and reducing difficulty of the discrimination task (less element involved). Similarly, we establish a complete trade-off curve for each task combination. This allows us to test whether the conclusion we draw from highly demanding tasks presented in this chapter is specific to high attentional cost tasks, or it is a more general finding about attentional capacity.

2.2 Methods

2.2.1 Equipment and procedure

Stimuli were generated by a Silicon Graphics Indigo^2 workstation and displayed on a high-resolution color monitor (1280 × 1024 pixels, 3 × 8 bit RGB, 60 Hz frame rate). Viewing was binocular, from a distance of about 120 cm, resulting in a display of approximately 12.5 × 16 deg of visual angle, with 1 deg corresponded to 80 pixels. Background luminance was 26.6 cd/m^2, luminance of display elements was $28.4cd/m^2$, $26.3cd/m^2$ (subject DKL) or $29.5/m^2$ (subject SH), and $91.5cd/m^2$, for red, green, and white elements, respectively, and ambient illumination level was $\approx$3 cd/m^2. Subjects fixated a mark at the center of the screen and initiated trials by pressing the "space" key. After viewing a sequence of stimulus and mask displays (see below), they responded by pressing assigned keys on the computer keyboard. Incorrect responses were marked by auditory feedback.

Two subjects participated in the experiment. One was the first author and the other was a volunteer paid for her participation, who was not informed about the purpose of the study. Subjects were well trained and performed 9 (subject DKL) and 15 (subject SH) training sessions before data collection began. Both subjects were highly practised and performed 234 (DKL) and 240 (SH) blocks of 50 trials in $\approx$ 30 sessions over the course of 4 months. 14 (DKL) and 8 (SH) blocks of trials were rejected because of a significant response correlation and are not included in these totals (see below). Each session was dedicated to one particular experiment and included blocks of trials with all three instructions (central task only, peripheral task only, both central and peripheral task, see below for details). At least two sessions were dedicated to each experiment, allowing us to distinguish between the first session and subsequent sessions. In total, we collected between 26 and 52 blocks of trials per experiment and subject (on average $\approx$ 34 blocks). These numbers include both single- and double-task blocks.

Performance of individual tasks was relatively stable and showed little or no improvement over time (see Results). However, the performance of task combinations tended to improve somewhat between the first and subsequent sessions dedicated to a particular combination, as will be discussed.

2.2.2 Stimuli and tasks

All displays comprised a central and a peripheral component, which posed two independent visual tasks. Depending on instructions, subjects performed either the "central task" or the "peripheral task" or both. The central component consisted of 5 elements which appeared at 5 of 7 possible central locations (polar coordinates $(r, \theta) = (0°, 0°)$, $(0.875°, 0°)$, $(0.875°, 60°)$, ..., $(0.875°, 300°)$ where r is in degrees of visual eccentricity and θ in degrees of polar angle), while the peripheral component was a single element at one of eight locations (polar coordinates (r, θ) with $r = 4.375°$ and $\theta = 0°, 45°, \ldots, 315°$). Although different experiments and tasks involved different types of elements (form, color, and motion), only two alternative forms of any element type were used. Central tasks involved reporting whether one of the five elements differed from the other four or whether all five elements were the same. Peripheral tasks involved reporting which of the two alternative elements had been present. The particular elements used here were chosen in pilot experiments because they yielded comparable psychometric functions (similar performance at any given presentation time for form, color, and motion discrimination).

For form discrimination, we used randomly rotated T- and L-shaped elements of white color **(Fig. 2.1a)**. The central task was to inspect the five central forms and to report "same" (5 Ts or 5 Ls) or "different" (4 Ts and 1 L, or 4 Ls and 1 T). The peripheral task was to inspect the single peripheral element and to report "T" or "L".

For color discrimination, we used vertically bisected discs with red and green halves **(Fig. 2.1b)**. The colors were isoluminant (flicker photometry) and the color order was either green left and red right (GR), or red left and green right (RG). The central task was to inspect the five central disks and to report "same" (5 RGs or 5 GRs) or "different" (4 RGs and 1 GR, or 4 GRs and 1 RG). The peripheral task was to inspect the single peripheral disk and to report "GR" or "RG".

For motion discrimination, we used "dumbbell" forms of white color rotating around their center **(Fig. 2.1c)**. The angular velocity of each dumbbell was chosen randomly in the range from 1.5rps (revolutions per second) to 3.5rps, and the rotation was either clockwise (CW) or counterclockwise (CCW). The central task was to inspect the five central dumbbells and to report "same" (5 CWs or 5 CCWs) or "different" (4 CWs and 1 CCW, or 4 CCWs and 1 CW). The peripheral task was to inspect the single peripheral dumbbell and to report "CW" or "CCW".

For form and color discrimination, the visual persistence of stimulus elements was curtailed by masking, and *stimulus onset asynchrony*, (SOA) was defined as the time between stimulus and mask onset. For motion discrimination, no masking was used, as there was no discernible visual persistence of dumbbell rotation. Here, SOA was defined as the time between stimulus onset and offset. For each task and subject, the SOA was adjusted individually to obtain a performance level of $\approx$ 80% correct when the task in question was carried out alone. Since elements had been chosen such as to yield comparable psychometric functions, SOAs for form, color, and motion discrimination were comparable. Central task SOAs ranged from 183ms to 200ms and peripheral task SOAs from 67ms to 100ms. Note that due to the difference in SOA, peripheral display elements were masked earlier than central display elements.

All tasks in the present study were designed to place strong demands on attention. Although we have yet to study the issue systematically, we tentatively attribute the attention-demanding nature of the tasks in question to the fact that they involve judging the relative position of *two* features (of vertical and horizontal bars for the T- and L-shaped elements, of red and green for the bisected disks, and of the respective directions of motion of terminal points for the rotating dumbbells). In the terminology of Treisman, the tasks in question involve "conjoining" two features with two locations.

In spite of their somewhat complex nature, the discrimination tasks would have been "easy" had they concerned a single element at a known location: The necessary SOA to achieve 80% correct performance in discriminating a single T- or L-form, GR- or RG-disk, or CW- or CCW-rotating dumbbell at fixation can be as short as 40ms. For the peripheral task, the necessary SOA was increased by visual eccentricity and positional uncertainty of the discriminated element. For the central task, this value was increased by the fact that the task involved more than one element and that these elements appeared at variable positions.

Three elements would have been sufficient to obtain SOAs above $100ms$, but five elements were used in order to allow direct comparison with earlier work (Braun, 1994; Braun & Julesz, 1998).

2.2.3 Concurrent-task procedure

Subjects were asked, in separate blocks of trials, to carry out either one task, or the other, or both. The display remained the same and always contained both central and peripheral stimuli. Subjects were asked to remain fixation all the time. In some blocks of trials, subjects were instructed to concentrate on one task and ignore the other task, responding only once after each trial. This situation, which allows subjects to focus attention on one task, yields "single-task" performance. In other blocks of trials, subjects were instructed to perform both tasks and to respond twice after each trial. This situation, which forces subjects to divide attention between tasks, produces "double-task" performance. The order of responses has no noticeable effect on double-task performance (Braun & Julesz, 1998). Subjects can voluntarily affect the way in which they divide attention between tasks, so that there is an entire family of possible double-task outcomes (one for every possible division of attention). To obtain the full range of outcomes, we used three variations of the basic double-task instruction: (i) "perform both tasks but give priority to the central task", (ii) "perform both tasks but give priority to the peripheral task", (iii) "perform both tasks and give equal priority to both tasks". The concurrent task results are reported in the form of attention-operating characteristics (AOC), which describes how the performance of both tasks varies with the division of attention.

2.2.4 Differentiated and undifferentiated capacity

To formalize the distinction between differentiated and undifferentiated capacity, we treat attentional capacity as a divisible resource. The empirical justification is that when observers attend to two tasks they can trade performance on one task off against performance on another. Thus we can nominally speak of the "fraction of attention" allocated to a particular task, and can postulate a function which describes visual performance as this fraction

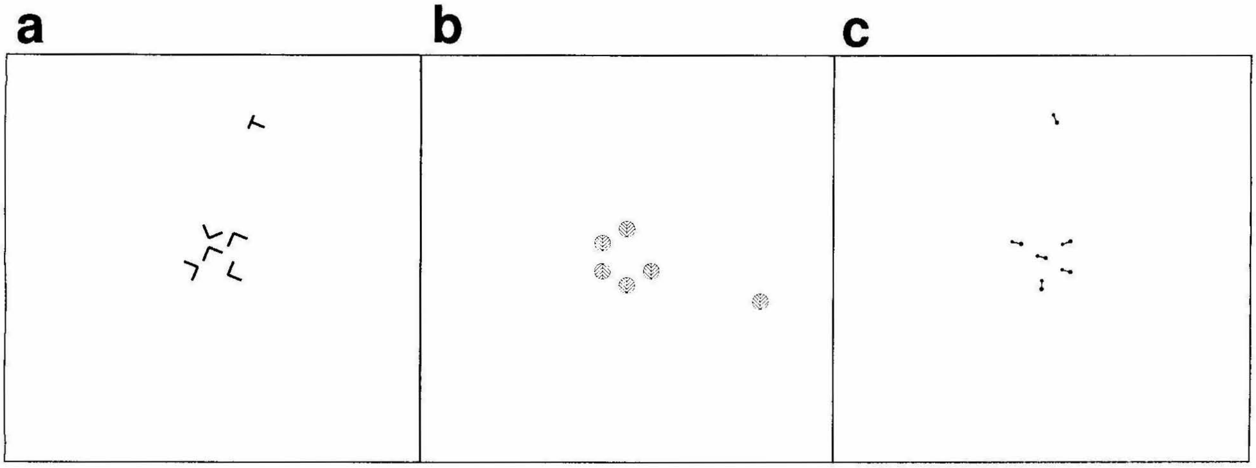


Fig. 2.1 Examples of stimulus displays (schematic, drawn to scale). Central and peripheral form **(a)**, color **(b)** and motion **(c)** discrimination. Form elements are T- or L-shaped and randomly rotated. Color elements are vertically bisected discs with red and green halves (indicated by hatching). Motion elements are dumbbell forms rotating either clock- or anti-clockwise around their center of gravity. In each case, there are five central elements and one peripheral element. With respect to the central elements, subjects report whether all five elements exhibit the same form, coloring, or motion, or whether one element differs from the other four in form, coloring, or motion. With respect to the peripheral element, subject identify the form, color, or motion. Central and peripheral elements were masked at appropriate (and usually different) times to obtain performance levels for individual tasks between ≈80% and ≈90% correct.

increases from zero to unity ("performance-resource function", PRF).

Although we cannot observe PRFs directly, we can infer them from the performance of concurrent tasks. In such an experiment, observers vary the relative priority of two concurrently performed tasks. If observers perform optimally (which is not always the case!), the results show how performance of one task varies as a function of performance of the other. This curve is known as the "attention-operating characteristic" (AOC) or, more informally, as the trade-off curve.

The AOC of a task combination almost determines the PRFs of the two tasks in question (one degree of freedom remains, see **Fig. 2.2**). If the PRF of one of the two tasks is known from some other source, the AOC completely determines the PRF of the other task. In the latter case, the chain of inference is that performance of the first task reveals the fraction of attention available for the second, which fraction is then related to performance on the second. This method evidently works best when performance of the first task depends strongly on attention, that is, when performance of this task ranges from chance to ceiling as the fraction of attention increases from zero to unity. This is the reason why we use tasks of high attentional cost.

If attentional capacity is differentiated, the PRF of a given task will depend on the task with which it is paired. Dissimilar task pairs will result in more relaxed trade-off curves and more rapidly saturating PRFs than similar task pairs. If attentional capacity is undifferentiated, however, a given task will have the same PRF, no matter with which other task it is paired. In this case, similar and dissimilar task pairs will result in the same PRFs.

2.2.5 Theoretical PRF and AOC curves

With input and help from my colleague Jochen Braun, we generated theoretical AOC curves for a task combination by assuming that attention is always fully allocated, to one or both tasks, so that no attention is ever wasted. Given this assumption (which does not always hold in practice) we can compute the AOC curve from the PRFs of the two tasks. For simplicity, we assume that PRFs have the following functional form:

$$p(r) = \begin{cases} \dfrac{p(1)}{2}\left(\dfrac{2r}{\alpha}\right)^{\beta} & \text{if } 0 \leq r \leq \dfrac{\alpha}{2} \\ p(1) - \dfrac{p(1)}{2}\left(\dfrac{2(\alpha - r)}{\alpha}\right)^{\beta} & \text{if } \dfrac{\alpha}{2} < r \leq \alpha \\ p(1) & \text{if } \alpha < r \leq 1 \end{cases}$$

where $p(r)$ is performance (in units of fraction correct, $p(r) \in [0,1]$), r is the fraction of attention allocated to the task ($r \in [0,1]$), $p(1)$ is performance with full attention (i.e., single-task performance), and α and β are parameters. The analytical form of this function is not motivated by theoretical considerations but simply by convenience in generating sigmoidal PRFs with an inflexion point at $r = 1/2$. As illustrated in **Fig. 2.2bc**, the parameter α ($0 \leq \alpha \leq 1$) determines what fraction of attention suffices for optimal performance, and the parameter β ($\beta \geq 1$) determines whether the relation between attention and performance is linear ($\beta = 1$), quadratic ($\beta = 2$), and so on.

Given two tasks A and B, we obtain the AOC curve as the set of all $(p_A(r), p_B(1-r))$ with $r \in [0,1]$ **(Fig. 2.2ad)**. By varying the PRF parameters α_A, β_A, α_B, and β_B, a wide range of different AOCs is obtained. If the tasks draw on pools of attentional resources that overlap only partially, optimal performance will be reached at some fraction of attention α_A and α_B which is smaller than unity **(Fig. 2.2a)**. As a result, there will be a certain performance range (smaller than the full range) in which the two tasks compete for attentional resources **(Fig. 2.2d)**. Specifically, task A will compete for attention between $p_A(1-\alpha_B)$ and $p_A(\alpha_A)$, and task B will compete between $p_B(1-\alpha_A)$ and $p_B(\alpha_B)$. The resulting AOC curve will lie above and to the right of the linear trade-off line, but its precise form will depend on β_A and β_B.

If the tasks draw on the same pool of attentional resources, optimal performance will require full attention. In this case, $\alpha_A = \alpha_B = 1$ and the resulting AOC curve will cross the linear trade-off line. In the special case that $\beta_A = \beta_B$ the AOC curve will be identical to the linear trade-off line. Note that in this latter case the absolute value of β is indeterminate.

For the purpose of this study, the critical question for any given pair of tasks is whether α_A and α_B are equal to unity so that the tasks compete for the same pool of resources, or whether α_A and/or α_B are smaller than unity so that the tasks draw on somewhat different pools of resources.

2.2.6 Maximum likelihood analysis

Given a family of theoretical AOC curves, we would like to know which particular curve is the most likely to have produced the observed experimental result. To this end, we compute first the probability of a given AOC curve producing a particular experimental result, and then use Bayes' rule to obtain the probability of a given experimental result being produced by a particular AOC curve.

Given two tasks A and B, AOC curves are specified by the four parameters $\alpha_A, \beta_A, \alpha_B, \beta_B$, and experimental results are given in terms of the number of correct responses n_A and n_B in N concurrent task trials. The probability $P(\alpha_A, \beta_A, \alpha_B, \beta_B | n_A^1, n_B^1, n_A^2, n_B^2, \ldots)$ that a particular set of parameters $\alpha_A, \beta_A, \alpha_B, \beta_B$ will produce a given set of observations $n_A^1, n_B^1, n_A^2, n_B^2, \ldots$ (in multiple blocks of trials) is, according to Bayes' rule, related to the probability $P(n_A^1, n_B^1, n_A^2, n_B^2, \ldots | \alpha_A, \beta$ of obtaining a particular set of observations $n_A^1, n_B^1, n_A^2, n_B^2, \ldots$, given a set of parameters $\alpha_A, \beta_A, \alpha_B, \beta_B$:

$$P(\alpha_A, \beta_A, \alpha_B, \beta_B | n_A^1, n_B^1, n_A^2, n_B^2, \ldots) = \frac{P(n_A^1, n_B^1, n_A^2, n_B^2, \ldots | \alpha_A, \beta_A, \alpha_B, \beta_B)}{P(n_A^1, n_B^1, n_A^2, n_B^2, \ldots)} \quad P(\alpha_A, \beta_A, \alpha_B, \beta_B)$$

$$P(\alpha_A, \beta_A, \alpha_B, \beta_B | n_A^1, n_B^1, n_A^2, n_B^2, \ldots) \propto P(n_A^1, n_B^1, n_A^2, n_B^2, \ldots | \alpha_A, \beta_A, \alpha_B, \beta_B)$$

The proportionality relation holds because the *a priori* probabilities $P(n_A^1, n_B^1, n_A^2, n_B^2, \ldots)$ and $P(\alpha_A, \beta_A, \alpha_B, \beta_B)$ can be assumed to be the same for all $\alpha_A, \beta_A, \alpha_B, \beta_B$.

All computations are carried out numerically in Matlab (Math Works). As a first step, we compute the conditional probability $P(n_A, n_B | r, \alpha_A, \beta_A, \alpha_B, \beta_B)$ of obtaining n_A and n_B correct responses in N trials of tasks A and B, given a particular division of attention r and choice of α_A, β_A, α_B, and β_B:

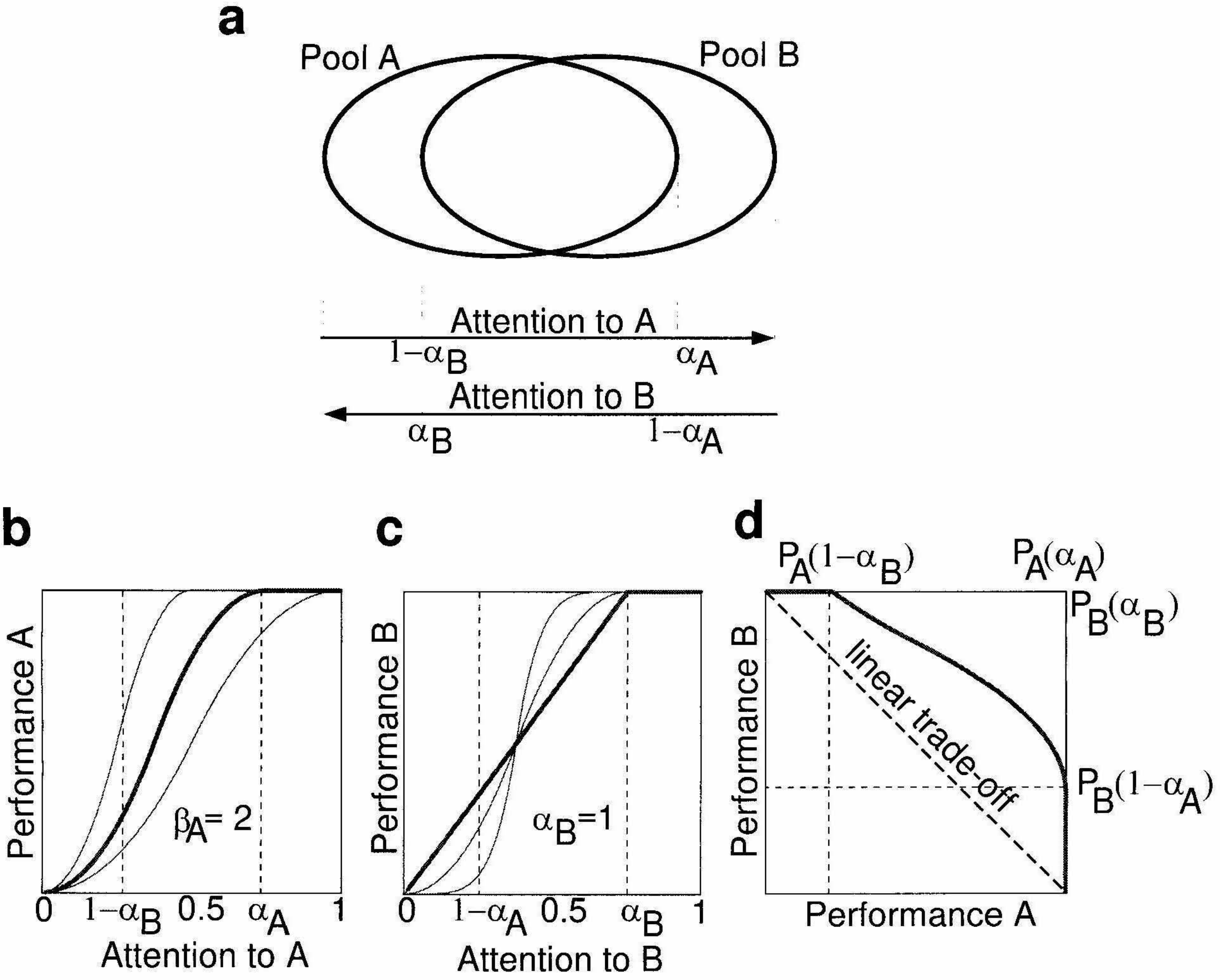


Fig. 2.2 Attention-operating characteristic (AOC) and performance-resource functions (PRFs). **(a)** In the most general case, task A and B draw on somewhat different "pools" of attentional resources. To generate AOCs, we assume that attention can be divided arbitrarily between tasks A and B, but that attention to A and attention to B always sum to unity. **(b)** Performance A as a function of the fraction of attention allocated to task A. Three possible PRFs are shown, with $\alpha_A = 0.5, 0.75, 1.0$ and $\beta_A = 2.00$ (see Methods). **(c)** Performance B as a function of attention allocated to task B. Three possible PRFs are shown, with $\alpha_B = 0.75$ and $\beta_B = 1, 2, 5$. **(d)** AOC resulting from the PRFs drawn with a thicker pen in **(b)** and **(c)** ($\alpha_A = 0.75$, $\beta_A = 2$, $\alpha_B = 0.75$, $\beta_B = 1$). Since $\alpha_A, \alpha_B < 1$, the AOC lies above and to the right of the linear trade-off line. Linear trade-off is obtained if $\alpha_A = \alpha_B = 1$ and $\beta_A = \beta_B$. Note that only **(d)** is experimentally observable.

$$P(n_A, n_B | r, \alpha_A, \beta_A, \alpha_B, \beta_B) = \binom{N}{n_A} p_c^{n_A} (1 - p_c)^{N-n_A} \binom{N}{n_B} p_p^{n_B} (1 - p_p)^{N-n_B}$$

Next, we assume that all divisions of attention r are equally likely, and compute the probability $P(n_A, n_B | \alpha_A, \beta_A, \alpha_B, \beta_B)$ of obtaining n_A and n_B correct responses as

$$P(n_A, n_B | \alpha_A, \beta_A, \alpha_B, \beta_B) = \int_0^1 P(n_A, n_B | r, \alpha_A, \beta_A, \alpha_B, \beta_B) dr$$

The joint probability of obtaining particular results in multiple blocks of trials is

$$P(n_A^1, n_B^1, n_A^2, n_B^2, \ldots | \alpha_A, \beta_A, \alpha_B, \beta_B) = \prod_i P(n_A^i, n_B^i | \alpha_A, \beta_A, \alpha_B, \beta_B)$$

Finally, we normalize this probability by integrating over the particular part of the parameter space that is of interest and setting the integral equal to one. For example, to compare likelihoods for $0 \leq \alpha_A \leq 1$ and fixed $\alpha_B, \beta_A, \beta_B$, we set

$$\int_0^1 P(n_A^1, n_B^1, n_A^2, n_B^2, \ldots | \alpha_A, \beta_A, \alpha_B, \beta_B) d\alpha_A = 1$$

and plot the resulting $P(n_A^1, n_B^1, n_A^2, n_B^2, \ldots | \alpha_A, \beta_A, \alpha_B, \beta_B)$ as a function of α_A.

2.2.7 Contingency analysis

An important aspect of concurrent task experiments is whether the division of attention is consistent throughout a block of trials, or whether it varies from trial to trial (Sperling & Dosher, 1986). This can be decided by analyzing the joint probabilities of responding correctly and/or incorrectly on each task (Braun & Julesz, 1998). If the division of attention is constant, success or failure on one task will be independent of success or failure on the other. If the division varies, there will be a negative correlation.

To estimate the degree of attentional variation implied by an observed (negative) correlation, we can use the theoretical PRFs defined above. Given reasonable parameter values

($\alpha_A = \alpha_B = 1$, $\beta_A = \beta_B \in [1, 3]$), we can compute performance for any particular division of attention and, thus, the correlation resulting from a variable division of attention (e.g., $r = 0.43$ in 50% of the trials, and $r = 0.57$ in the remaining trials). The theoretical curves in **Fig. 2.8** are based on such a calculation.

2.3 Results

2.3.1 Similar task combinations

The first series of experiments concerned concurrent performance of similar task pairs. Three pairs of tasks were investigated: central and peripheral discrimination of form, color, and motion, respectively. These experiments permit us to quantify the attentional cost of the tasks involved.

Results are shown in **Fig. 2.3**, and were comparable for both subjects and all task combinations. Each subject performed between 26 and 52 blocks of 50 trials on each task combination (average $\approx$34 blocks). Generally, double-task performance was inferior to single-task performance. Whenever double-task performance of one task approached the single-task level, double-task performance of the other task approached chance. Indeed, double-task performance was characterized by a roughly linear trade-off between tasks: when either task was performed at its best, the other task was performed at chance, and any improvement in the performance of one task was at the expense of a reduction in the performance of the other. In some blocks of trials, double-task performance was well below the linear-trade line, suggesting that subjects did not efficiently allocate attention (in the sense that some of the available attention was not allocated to either task). This aspect of the results is discussed further below. A contingency analysis was performed after pooling blocks for each subject and task combination to determine whether subjects had maintained a stable division of attention. No significant ($p < .05$) response correlation was found. Pooled results from all three task combinations are shown in **Fig. 2.9** and **Fig. 2.10**.

A maximum likelihood analysis of these results reveals the attentional requirements of the tasks in question (see Methods). As the results are roughly compatible with a linear trade-off in performance, we expect the most likely values of α_c and α_p to be near unity, and the most likely values of β to satisfy $\beta_c = \beta_p$. (The subscript 'c' and 'p' denotes parameters of the central and peripheral task, respectively.) As it is impractical to investigate the entire four-dimensional parameter space which contains all possible values of α_c, β_c, α_p, and β_p, we explored four one-dimensional subspaces by varying the value of one parameter while

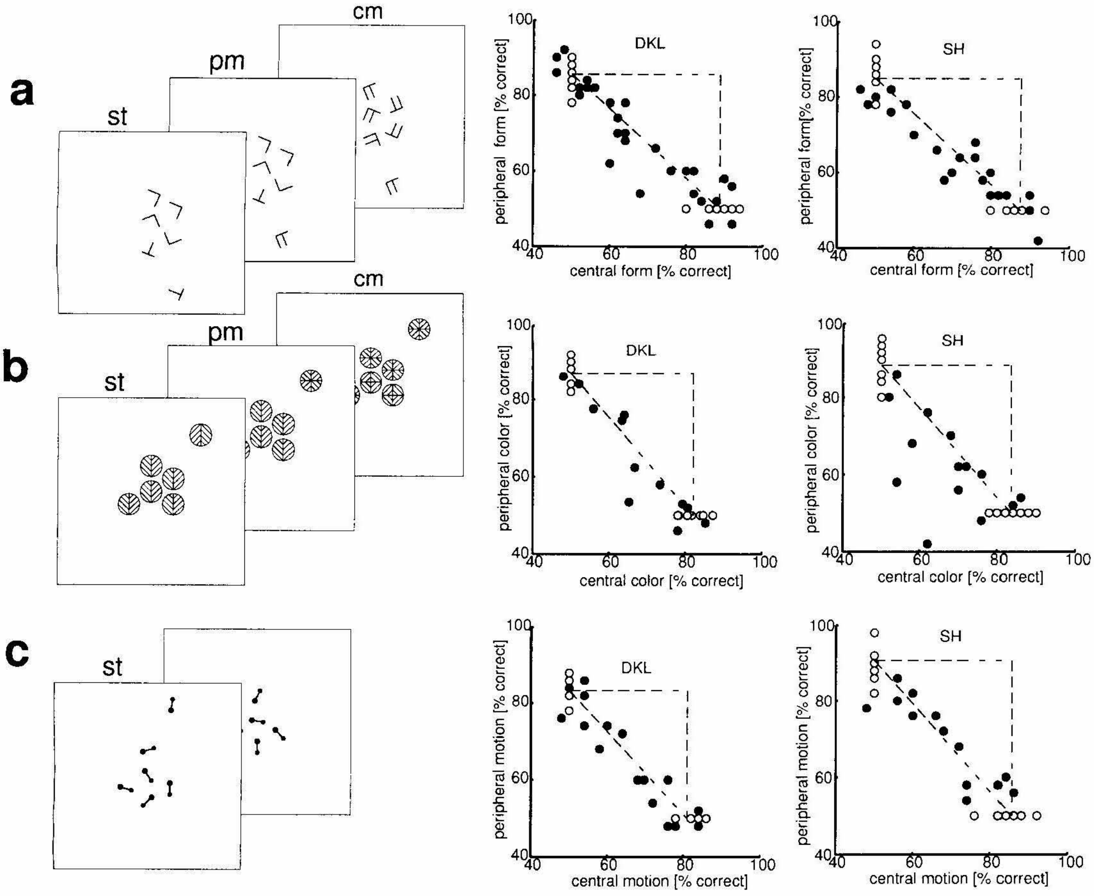


Fig. 2.3 Displays and results for combinations of similar tasks. Stimulus and mask displays are **not** drawn to scale. Red and green color is indicated by hatching. Results are presented in the form of an attention-operating characteristic (AOC), with center task performance on the horizontal and peripheral task performance on the vertical axis. Open circles indicate single-task and filled circles, double-task performance. Each circle represents one block of 50 trials. Dashes lines indicate the theoretical extremes of the AOC. Subjects are identified by their initials (DKL or SH). **(a)** Central and peripheral form discrimination. The stimulus display (st) is followed by a peripheral (pm) and a central mask (cm). **(b)** Central and peripheral color discrimination. The stimulus display (st) is followed by a peripheral (pm) and a central mask (cm). **(c)** Central and peripheral form discrimination. The stimulus display (st) is not masked, but the peripheral element is presented for less time than the central elements. In all three experiments, double-task performance follows the lower extreme of the AOC, indicating that optimal performance of either task requires full attention.

keeping the other three values at unity. For the α's, unity is close to the most likely value. The justification for setting β equal to unity is that the results are close to the special case in which the absolute value of β is arbitrary (*i.e.*, the case of $\alpha_c = \alpha_p = 1$ and $\beta_c = \beta_p$; see Methods). **Fig. 2.4** shows the probability that a given choice of parameters value generates the particular results obtained for every subject and task combination. The most likely values of α_c and α_p are unity in all cases except subject SH and motion-motion combination, where the most likely $\alpha_c = 0.98$. With confidence 0.95, the value of all α_c and α_p is larger than 0.75 (the median value of this lower bound is 0.86). The most likely values of β_c and β_p are generally also near unity, although here there also is a significant likelihood of larger values. With confidence 0.95, we can say only that all β_c and β_p are smaller than 4.8 (median value 3.8). For the form-form combination and observer SH, the color-color combination and both observers, and the motion-motion combination and observer SH, we obtain particularly disparate values of β_c and β_p. The reason is that in these cases the data include blocks of trials well below the linear trade-off line, creating the appearance of an asymmetric AOC curve (which would imply disparate values of β_c and β_p).

Of course, the individual block results are expected to scatter around the AOC curve somewhat simply because of the finite number of trials per block. In general, the observed scatter was not significantly different from the scatter expected from binomial sampling of a linear trade-off. This was true for all observers and task combinations (χ^2 test, $p < .0001$), except the color-color combination of observer SH (who performed exceptionally poorly on two blocks). However, individual block results tend to fall somewhat more frequently below than above the linear trade-off line. For observers DKL and SH, the average distance of individual block results from the linear trade-off line is -0.91 ± 0.49 and -1.84 ± 0.74 percentage points, respectively (mean ± standard error, negative values indicate positions below the line). This small departure from strict linearity is weakly significant (t-test, $p < 0.05$).

2.3.2 Dissimilar task combinations

A second series of experiments established concurrent performance of dissimilar tasks. Four task combinations were investigated: central discrimination of color combined with periph-

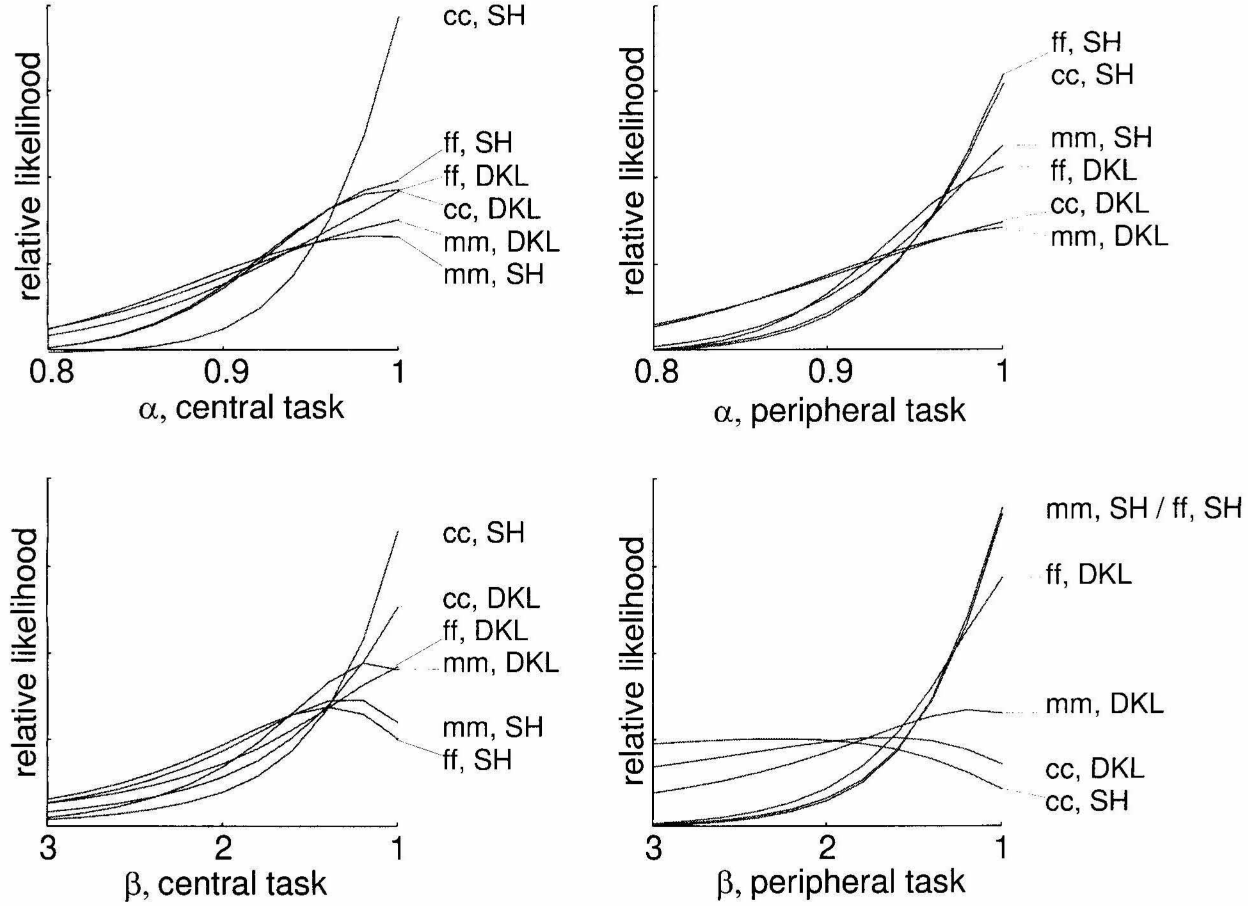


Fig. 2.4 Maximum likelihood analysis of similar task combinations. Relative likelihood is shown as a function of α_c **(a)**, α_p **(b)**, β_c **(c)** and β_p **(d)**. The other three parameters are set equal to unity in each case. Each curve is labeled by task combination ('ff' for form-form, 'cc' for color-color, and 'mm' for motion-motion) and by subject (DKL or SH).

eral discrimination of form or motion, and central discrimination of motion combined with peripheral discrimination of form or color.

Results are shown in **Fig. 2.5** and **Fig. 2.6**. Each subject performed between 29 and 35 blocks of 50 trials on each task combination (average ≈33 blocks). As was the case for similar task combinations, double-task performance was generally inferior to single-task performance. Whenever double-task performance of one task approached the single-task level, double-task performance of the other task approached chance. However, in many blocks of trials double-task performance was well below the linear-trade-off line, suggesting that subjects frequently failed to efficiently allocate attention. A contingency analysis was performed after pooling blocks for each subject and task combination, but no significant ($p < .05$) response correlation was found. Pooled results from all four dissimilar task combinations are shown in **Fig. 2.9**.

Maximum likelihood analysis shows that, for every task and subject, the most likely value of α is unity, except for where it is 0.95 (**Fig. 2.7**). With confidence 0.95, we can say that α exceeds 0.8 for all tasks and subjects. This shows that the tasks in this study require full or nearly full attention, and that this is equally true when they are performed in dissimilar or in similar combinations. The values of β are once again compromised by data points below the linear trade-off line. **Fig. 2.7** shows the probability that a given choice of parameter value generates the particular results obtained for every subject and task combination. The most likely values of α_c and α_p are unity in all cases except subject DKL and the color-form and motion-form combinations, where $\alpha_c = 0.94$ and $\alpha_c = 0.98$, respectively. With confidence 0.95, the value of all α_c and α_p is larger than 0.79 (median 0.89). The values of β_c and β_p are distributed more widely, although the most likely values are generally near unity. With confidence 0.95, the value of all β_c and β_p is smaller than 4.9 (median 3.9). Moreover, the distribution of β_c and β_p are often quite different. This is particularly true for subject DKL and the color-motion (0.95 confidence limits for β_c and β_p are 1.6 and 4.9, respectively) and motion-form combination (confidence limits 4.5 and 2.6). The reason for the disparate values of β is that in these cases an especially large fraction of the results falls well below the linear trade-off line **(see Figs. 2.5, 2.6)**.

The distribution of individual block results is not significantly different from the distri-

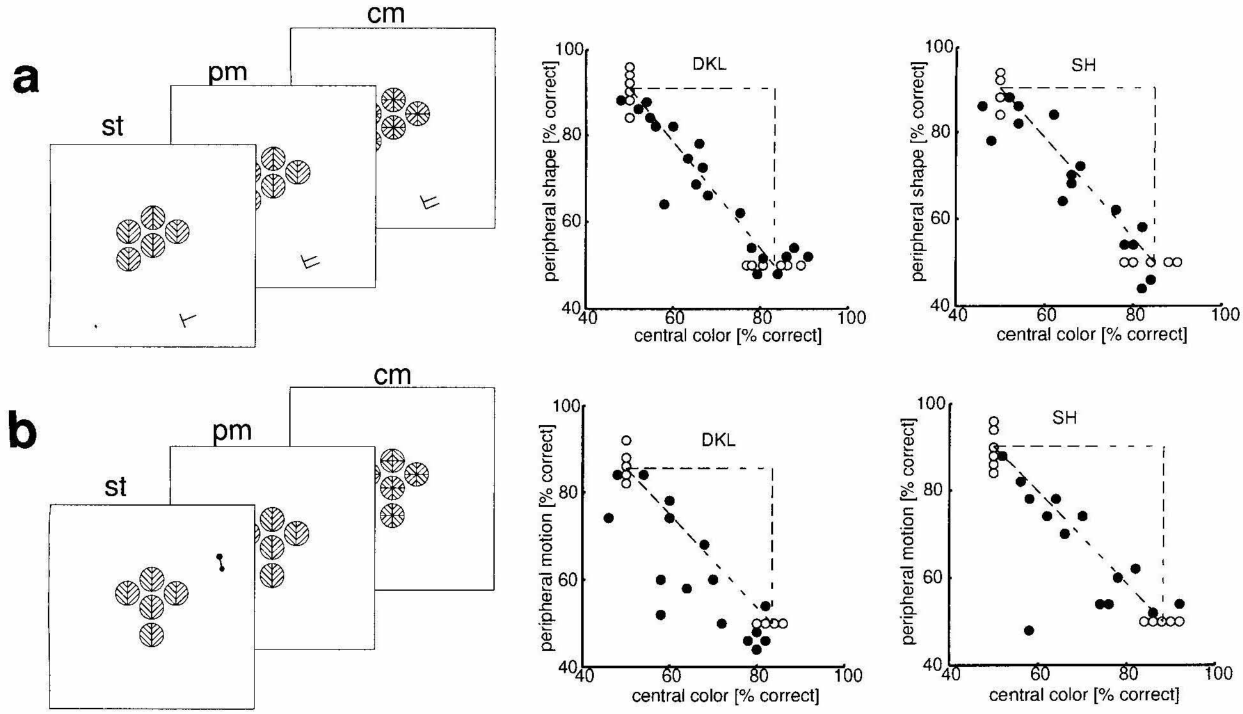


Fig. 2.5 Displays and results for combinations of dissimilar tasks. See caption of **Fig. 2.3**.

(a) Central color and peripheral form discrimination. The stimulus display (st) is followed by a peripheral (pm) and a central mask (cm). **(b)** Central color and peripheral motion discrimination. The stimulus display (st) is followed by a peripheral mask (pm) and the peripheral element is presented for less time than the central elements. In both experiments, optimal performance of either task requires full attention.

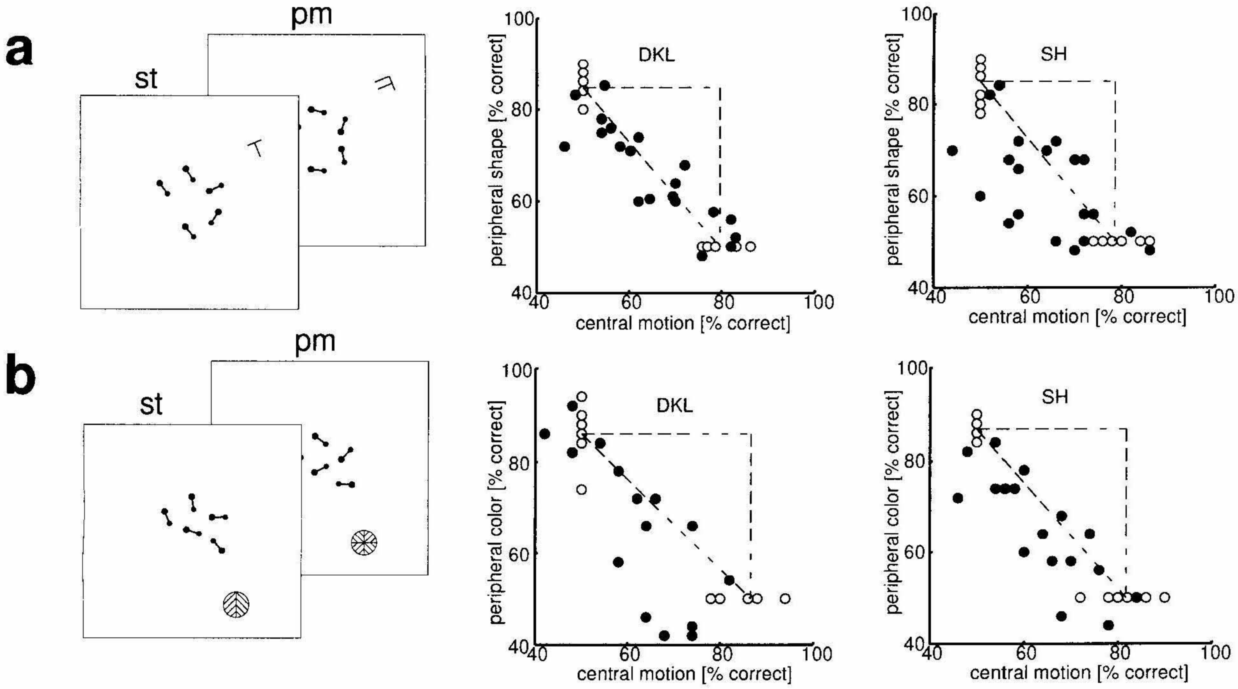


Fig. 2.6 Displays and results for combinations of dissimilar tasks. See caption of **Fig. 2.3**.

(a) Central motion and peripheral form discrimination. The stimulus display (st) is followed by a peripheral mask (pm). **(b)** Central and peripheral form discrimination. The stimulus display (st) is not masked, but the peripheral element is presented for less time than the central elements. In all three experiments, double-task performance follows the lower extreme of the AOC, indicating that optimal performance of either task requires full attention.

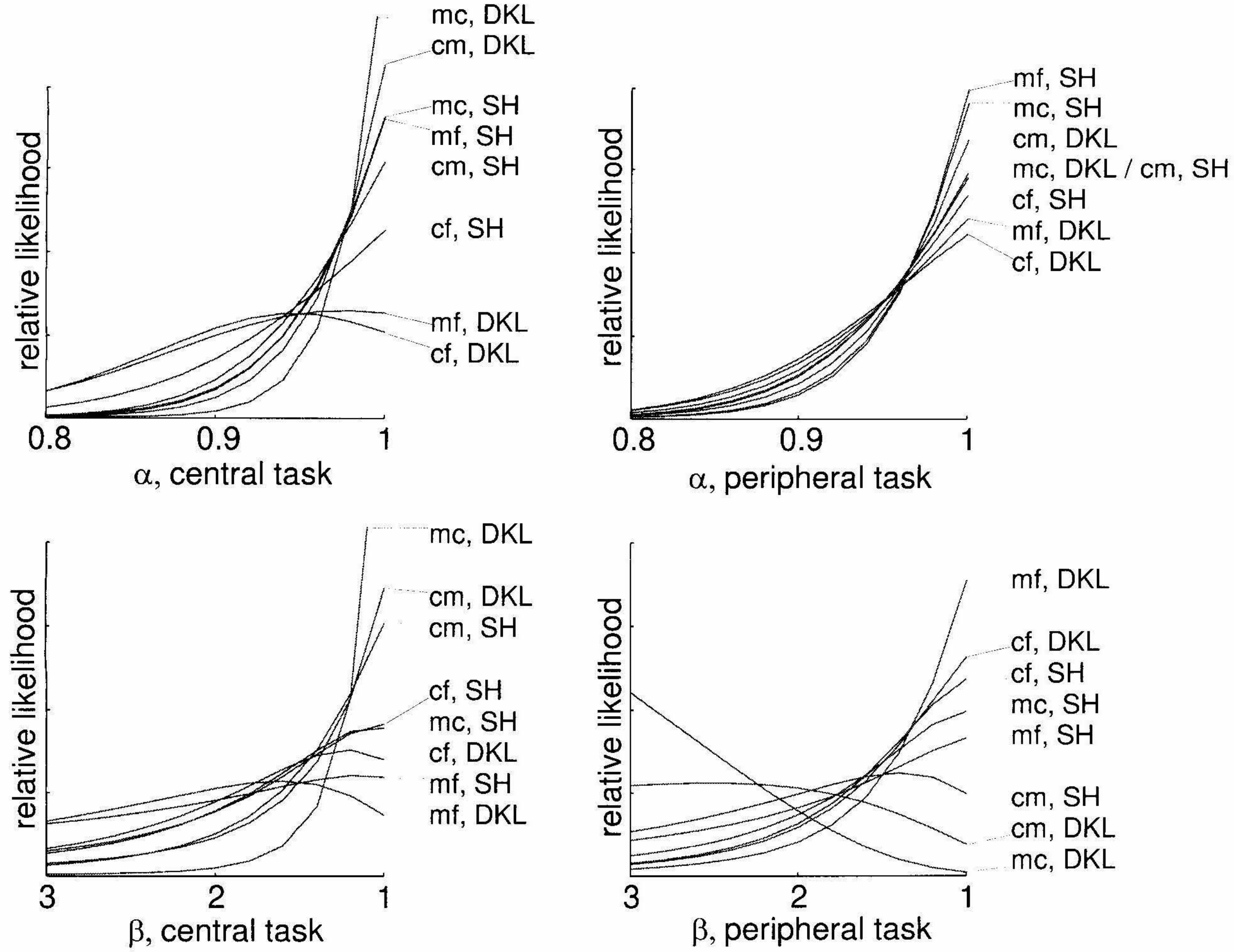


Fig. 2.7 Maximum likelihood analysis of dissimilar task combinations. Relative likelihood is shown as a function of α_c **(a)**, α_p **(b)**, β_c **(c)** and β_p **(d)**. The other three parameters are set equal to unity in each case. Each curve is labeled by task combination ('cf' for color-form, 'cm' for color-motion, 'mf' for motion-form, and 'mc' for motion-color) and by subject (DKL or SH).

bution expected from binomial sampling of a linear trade off (χ^2 test, $p < .0001$), except color-motion combination of observer SH and the motion-color combination of observer DKL. In both cases, the departure from the expected distribution is due to exceptionally poor performance in a few blocks of trials. However, individual block results fall more frequently below than above the linear trade-off line. The average distance of individual block results from the linear trade-off line is significantly less than zero: -2.48 ± 0.79 percentage points for observer DKL and -2.92 ± 0.71 percentage points for observer SH (t-test, $p < 0.01$).

2.3.3 Contingency analysis of pooled results

As mentioned, contingency analysis of data for individual subjects and task combinations revealed no significant response correlation, suggesting that subjects maintained a consistent division of attention during each block of trials. To see whether a residual response correlation would reach significance for an even larger data set, we pooled the results of all experiments on similar and dissimilar task combinations. In doing so, we considered only blocks of trials in which double-task performance of both tasks was comparable (neither performance above 70% of single-task level). The rationale was that near the middle of the AOC curve (when attention is about equally divided between tasks and performance comparable), a response correlation is more likely than at the ends of the AOC curve (when attention is directed mostly at one task and performance is disparate). Thus, by limiting the analysis to the middle of the AOC curve, we maximize the chances of observing a response correlation.

To show how responses to one task depend on responses to the other task, we plot percentage correct when subjects fail on the other task against percentage correct when subjects succeed on the other task **(Fig. 2.8)**. As expected, performance is slightly better when subjects fail on the other task. This is true for both similar and dissimilar task combinations (3950 and 4450 trials, respectively). On average, the difference in performance is less than 3 percentage points, large enough to be significant (χ^2-text, $p < .005$) in the pooled data sets but too small to reach significance in individual data sets.

Model calculations (see Methods) shows that a negative response correlation of this small magnitude is expected if attention focuses completely on one task or the other in $\approx 15\%$ of the trials but is divided equally in the remaining $\approx 85\%$ of the trials **(Fig. 2.6)**. Another way of obtaining such a correlation is to divide attention on all trials, but to vary the exact division, favoring sometimes one task and sometimes another. The necessary degree of variation depends on the form of the PRF, specifically, on the parameter β. For $\beta = 1$, the division of attention would have to vary between approximately 30%/70% and 70%/30% (if both are equally likely), for $\beta = 2$ it would have to vary between approximately 39%/61% and 61%/39%, and for $\beta = 3$ the variation would have to be approximately 43%/57% to 57%/43%.

2.3.4 Earlier and later sessions

Pooled results for all task combinations are shown in **Fig. 2.9**. The distribution of double-task outcomes is rather broad, especially for dissimilar task combinations, and includes many blocks of trials well below the linear trade-off line, in which subjects failed to perform well on either task. This poor performance is puzzling, since subjects could have done better by simply ignoring one of the two tasks. The most plausible explanation of these data is that subjects failed to allocate attention to either task, perhaps because they were overwhelmed by the complexity of the double-task situation.

This interpretation is strengthened when the results of earlier and later sessions on any given task combination are considered separately (*i.e.*, sessions in the first and second half of all sessions on any particular task combination). Pooled results for earlier and later sessions are shown in **Fig. 2.10**. Blocks of trials in which subjects failed on both tasks occur almost exclusively in the earlier sessions. In later sessions, subjects are consistently better, apparently because they have learned to efficiently allocate attention to one or both tasks so that no attention is wasted. Considering only data from later sessions, the distribution of individual block results does not depart significantly from linear trade-off for either similar or dissimilar task combinations (χ^2-test, $p < .0001$). The average distance from the linear trade-off line is -0.26 ± 0.98 percentage points for similar and -0.23 ± 0.88 percentage points for dissimilar task combinations. Neither value is significantly different from zero (t-test, $p < .0001$).

The maximum likelihood analysis of pooled results was based only on later sessions, in order to reduce the proportion of data in which attention was less than fully allocated. The

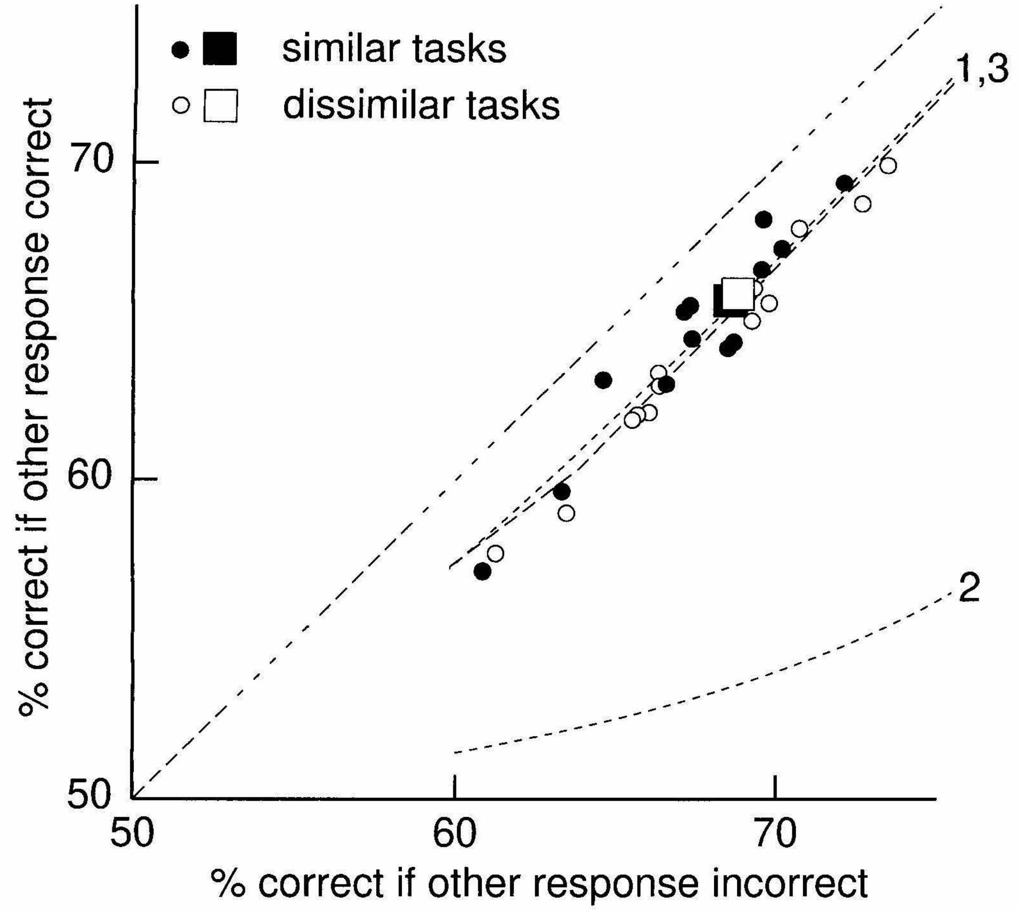


Fig. 2.8 Contingency analysis of pooled results. For a given task, we compute performance separately for trials in which subjects respond correctly and incorrectly on the other task, and plot the two values against each other. Open circles show results for similar task combinations (the open square represents the mean), and filled circles for dissimilar task combinations (the filled square represents the mean). On average, performance is about 3 percentage points lower when the other response is correct. For comparison, we show the expected contingencies if in 15% of the trials attention focuses on one task or the other instead of being divided (1), if this happens in all trials (2), and if attention is divided in all trials, but the exact division varies from trial to trial (sometimes 30% /70% and sometimes 70%/30%) (3).

results of this analysis are shown in **Fig 2.11**. The most likely parameter values were $\alpha_c =$ 1.0, $\alpha_p = 1.0$, $\beta_c = 1.2$, and $\beta_{=}1.0$ for similar and $\alpha_c = 0.98$, $\alpha_p = 1.0$, $\beta_c = 1.0$, and $\beta_{=}1.0$ for dissimilar task combinations. With confidence 0.95, $0.80 < \alpha_c \leq 1.00$, $0.85 < \alpha_p \leq 1.00$, $1.0 \leq< \beta_c < 4.0$, and $1.0 \leq \beta_p < 2.6$ for similar task combinations, and $0.89 < \alpha_c \leq 1.00$, $0.91 < \alpha_p \leq 1.00$, $1.0 \leq< \beta_c < 1.7$, and $1.0 \leq \beta_p < 1.5$ for dissimilar combinations. Thus, the outcome of similar and dissimilar task combinations is essentially the same: all investigated tasks require full or nearly full attention for optimal performance ($\alpha_c, \alpha_p \approx 1$) and when attention is divided the resulting trade-off in performance is linear or almost linear ($\beta_c \approx \beta_p$). This demonstrates that both similar and dissimilar task combinations compete for the same or almost the same pool of attentional resources. If there is any difference between similar and dissimilar task combinations, it is that for the latter the distribution of likely parameter values is consistently narrower than for the former (**Fig. 2.10**), which suggests that dissimilar tasks are if anything closer to the limiting case of linear trade-off than similar tasks.

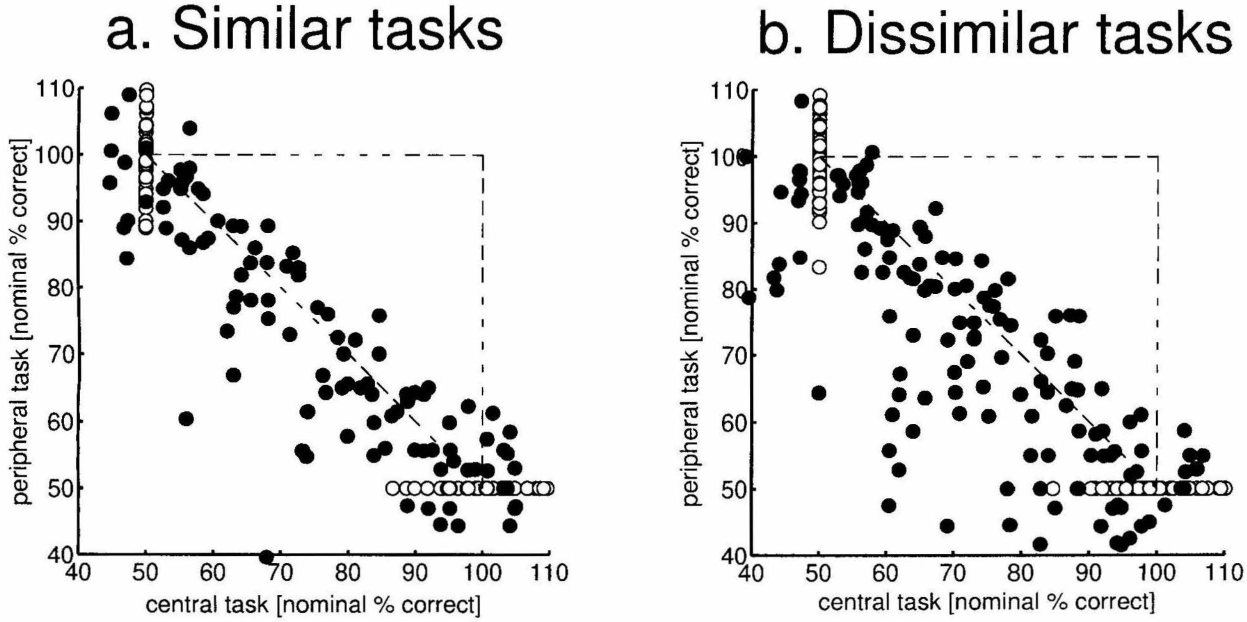


Fig. 2.9 Pooled results for similar and dissimilar task combinations, both subjects. See caption of **Fig. 2.3**. Note that numerous block results fall well below the linear trade-off line, especially for dissimilar task combinations.

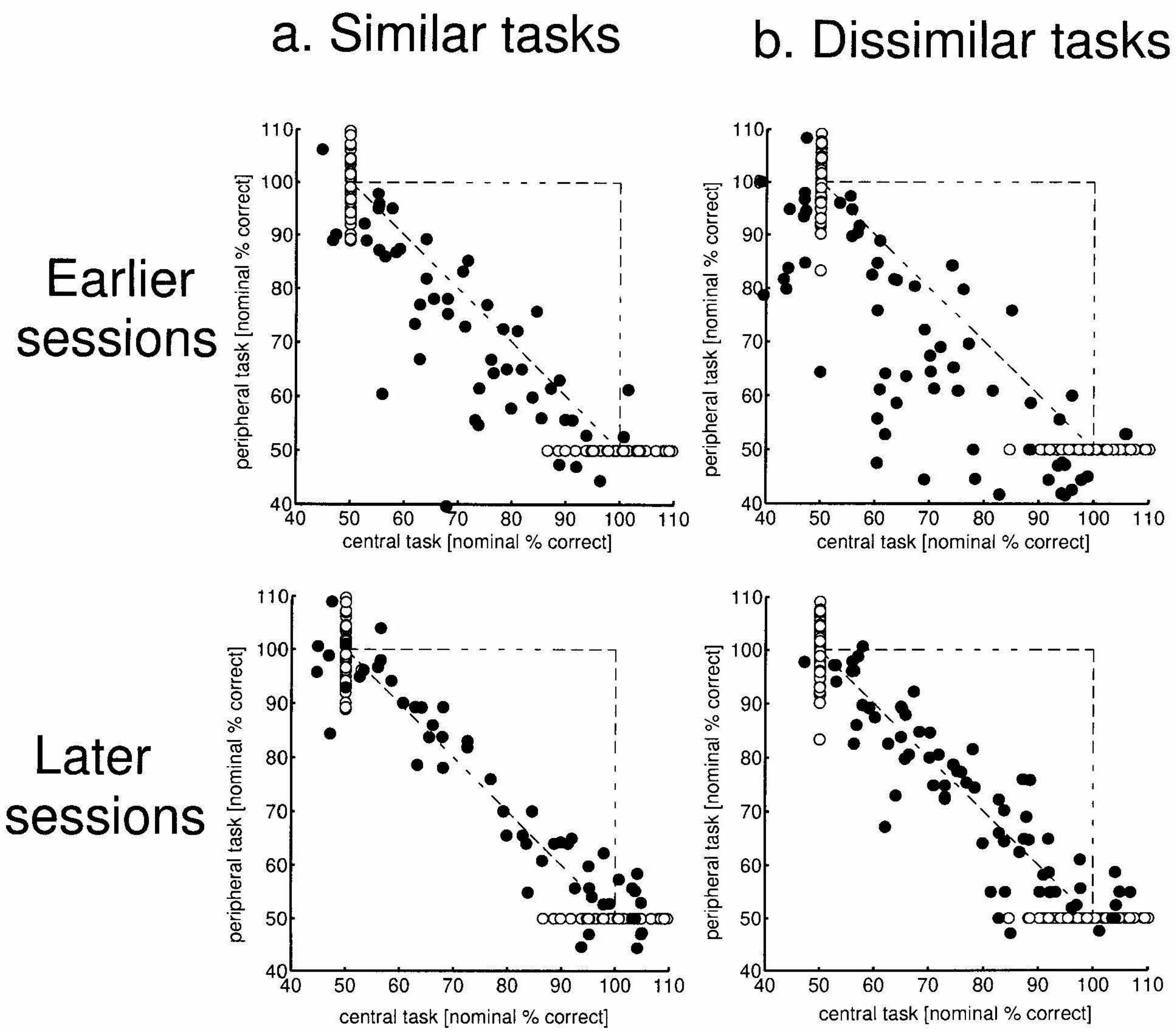


Fig. 2.10 Pooled results from earlier and later sessions with any given task combination, both subjects. See caption of **Fig. 2.2**. Most block results well below the linear trade-off line stem from earlier sessions. During later sessions, almost all block results cluster near the linear trade-off line, suggesting that subjects learned to allocate attention efficiently.

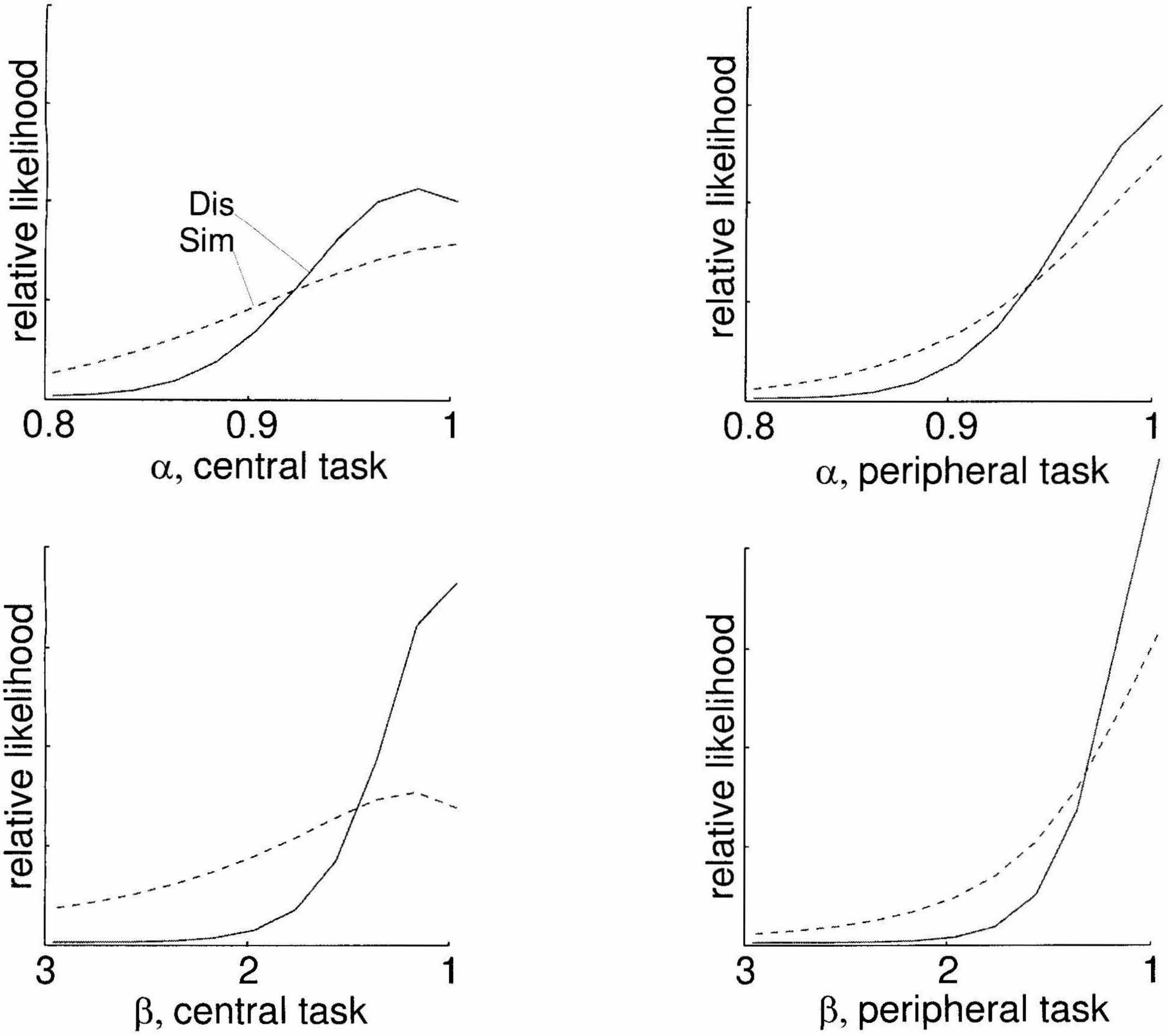


Fig. 2.11 Maximum likelihood analysis of pooled results from second (and subsequent) sessions. Relative likelihood is shown as a function of α_c **(a)**, α_p **(b)**, β_c **(c)** and β_p **(d)**. The other three parameters are set equal to unity in each case. Separate curves are shown for similar and dissimilar task combinations. Note that the distribution of likely values is consistently narrower for dissimilar task combinations.

2.4 Conclusion

2.4.1 Attentional capacity is divisible

Our results show that observers divide attention between tasks in almost every trial, rather than simply allocating attention sometimes to one task and sometimes to the other. This follows from the relatively small response correlation we observe: the average accuracy of one response is only about 3 percentage points better when the other response is incorrect rather than correct. The most likely cause of this small correlation is some degree of variability in the division of attention. Depending on the form of the PRFs, and especially the value of β, a relatively modest degree of variability would be sufficient to account for the observed response correlation (for $\beta = 1, 2, 3$, the division of attention would have to vary by approximately $\pm 20\%$, $\pm 11\%$, or $\pm 7\%$, respectively).

2.4.2 Attentional capacity is undifferentiated

The outcome of the present study could hardly have been clearer. Both similar and dissimilar task combinations resulted in an almost exactly linear performance trade-off. This was true for three similar tasks combinations which involved two discriminations of form (T *vs.* L), color (red-green *vs.* green-red), and motion (clockwise *vs.* counterclockwise), respectively, and for four dissimilar combinations which involved discriminations of color and form, color and motion, or motion and form. A more detailed analysis showed only minor differences between similar and dissimilar combinations. One of these differences concerned the first session with each task combination, during which observers sometimes performed poorly on both tasks. This happened more often with dissimilar than similar task combinations, suggesting that the initial difficulty of dividing attention may have been greater for dissimilar than for similar tasks. Another minor difference was the distribution of the most likely trade-off curves to account for the observed concurrent task performance. For dissimilar task combinations, this distribution was narrower, and thus the evidence for a strictly linear trade-off somewhat better, than for similar task combinations.

We conclude that similar and dissimilar tasks draw on exactly the same attentional capacity, in other words, that attentional capacity is entirely undifferentiated. Although the same conclusion has been reached by previous studies (Duncan, 1993; Duncan & Nimmo-Smith, 1996), we have considerably strengthened the evidence for it. By virtue of using tasks with high attentional cost, and by quantifying this cost, our paradigm was sufficiently sensitive to reveal intermediate outcomes (i.e., a partly differentiated capacity). The fact that actual outcome was extreme, rather than intermediate, must therefore count as highly significant. Our results do not support the suggestion (Duncan & Nimmo-Smith, 1996) that the discrimination of color enjoys a special status. We observe the same trade-off curves for task combinations with and without color discrimination. The reason why color discrimination yields different results for Duncan and colleagues is most likely that the particular discrimination they used carries little or no attentional cost (Braun & Julesz, 1998).

If a price had to be paid for the sensitivity of our paradigm, this price was the somewhat complex nature of our discrimination tasks. Two aspects of the design were responsible for the high attentional cost of our tasks. The first aspect was positional variability. Although this is not often emphasized, a long roster of divided attention studies show that performance decrements (that is, attentional costs) increase with positional variability (Braun & Sagi, 1991; Duncan, 1980, 1993; Duncan *et al.*, 1994; Shiffrin, Gardner & 1972). In the present case, positional variability for central tasks was achieved by presenting the five central targets in 35 distinct configurations (not counting rotational states), and for peripheral tasks by presenting the peripheral target at one of eight possible positions. The second aspect of the design which ensured high attentional cost was a requirement for joint discrimination of form (or color, or motion) and position. To report the form of a (randomly rotated) T or L, the observer has to discriminate both orientation and position of two component lines. Similarly, to distinguish red-green and green-red disks, the observer has to discriminate both the color and position of two half-disks, and to distinguish clockwise and counterclockwise motion, he has to discriminate both the direction of motion and position of two terminal points.

Thus, the discriminations used here are not "pure" discriminations of form, color, and motion, but always "mixed" discriminations of form, color, motion and *position.* It can be argued that this fact limits the generality of our conclusions, since it is possible that mixed discriminations engage different attentional capacities than pure discriminations. To address this question, it is necessary to compare similar and dissimilar combinations of pure discrimination tasks. Of course, the attentional cost of such discriminations will be lower, and the ability to characterize their PRFs correspondingly poorer. Nevertheless, such experiments are feasible.

Chapter 3

Attentional capacity for undemanding tasks

3.1 Background

In the last chapter, we established that for demanding tasks involving form, color and motion processing, the degree of interference does not change, no matter whether the task combination is similar or dissimilar. Thus, we conclude that attentional capacity is undifferentiated for tasks involving different stimulus dimensions. However, we use highly demanding tasks and find maximum linear trade-off under all conditions: when either task is performed at its best, the other task is performed at chance, any improvement in the performance of one task is at the expense of a corresponding reduction in the performance of the other task. Thus one question was raised often when we presented this study: is it possible that there exists some subtle difference between similar and dissimilar task combinations that is masked by the maximum ceiling trade-off? In order to find out whether this is the case or not, we carry out further experiments. Here we use less-demanding tasks (as a result, when either task was performed at its best, the other task was performed above chance but less than the single-task performance, and any improvement in the performance of one task was at the expense of a partial reduction in the performance of the other), and establish a complete trade-off curve for every task combination.

3.2 Methods

3.2.1 Equipment and procedure

Stimuli are generated by a Silicon Graphics Indigo^2 workstation and displayed on a high-resolution color monitor (1280 × 1024 pixels, 3 × 8 bit RGB, 60 Hz frame rate). Viewing is binocular, from a distance of about 120 cm, resulting in a display of approximately 12.5 × 16 deg of visual angle, with 1 deg corresponded to 80 pixels. Background luminance is 26.6 cd/m^2. Averaged over five subjects, luminance of display elements is $28.6cd/m^2$, $27.8cd/m^2$, and $91.5cd/m^2$, for red, green, and white elements, respectively, and ambient illumination level was ≈ 3 cd/m^2. Subjects fixated a cross at the center of the screen and initiated trials by pressing the "space" key. After viewing a sequence of stimulus and mask displays (see below and **Fig. 3.1**), they responded by pressing assigned keys on the computer keyboard. Incorrect responses were marked by auditory feedback.

Five subjects participated in the experiment. One was the author and the other four were volunteers paid for their participation, who were not informed about the purpose of the study. Subjects were well trained and performed 5-8 training sessions before 15 sessions of data collection (each session lasted about one hour and one session per day). Each session was interweaved with two or three task combinations, and included blocks of trials with all three instructions (left task only, right task only, both left and right task with different priorities, see below for details). For each task combination, at least two subjects were tested.

Unlike the demanding tasks discussed in the last chapter, performance of individual tasks continued to improve during the data collection despite the initial training. In order to keep performance of single tasks at a relatively stable level around 85%, we had to adjust the presentation time during the course of data collection (see Results). However, the interaction between the two tasks remained the same, as will be discussed in more detail later.

3.2.2 Stimuli and tasks

In order to make the tasks less attentional demanding, several changes were made based on central-periphery task paradigm described in chapter 2: the asymmetrial configuration

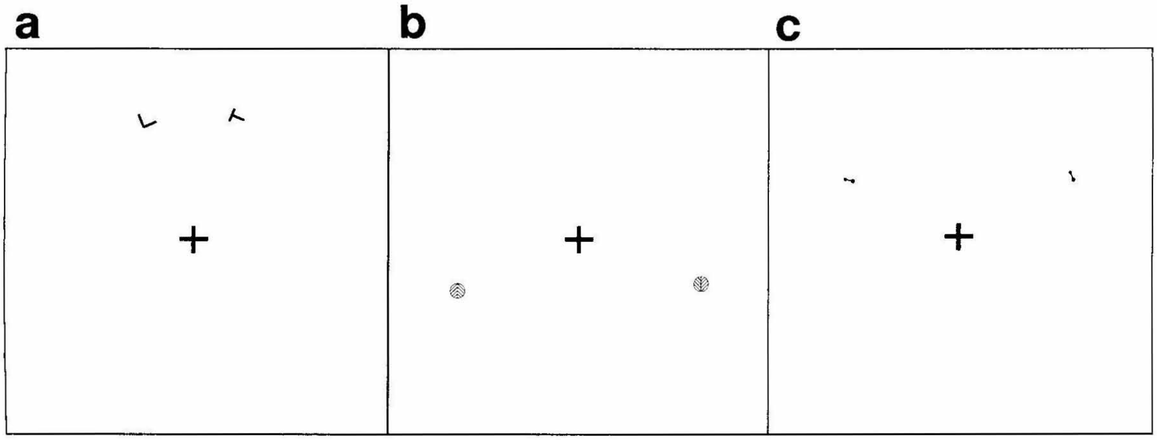


Fig. 3.1 Examples of stimulus displays (schematic, drawn to scale). Left and right form **(a)**, color **(b)** and motion **(c)** discrimination. Form elements are T- or L-shaped and randomly rotated. Color elements are vertically bisected discs with red and green halves (indicated by hatching). Motion elements are dumbbell forms rotating either clock- or anti-clockwise around their center of gravity. In each case, left and right stimulus appears at equal elevation and equal eccentricity of 4 degrees from the fixation point. With respect to each task, the subject reports which one of the two alternatives appears. Each peripheral element is masked at appropriate (and usually similar) times to obtain performance levels between 80% and 90% correct.

was changed into a symmetrical one, the number of elements that need to be processed was reduced, and positional uncertainty was reduced as well. All displays comprised a left and a right component, which posed two independent visual tasks. Depending on instructions, subjects performed either the "left task" or the "right task" or both. The left and right components were both single elements at one of four possible locations. For right stimulus, polar coordinates are (r, θ) with $r = 4.0°$ and $\theta = -67.5°, -22.5°, 22.5°, 67.5°$. The left element always appeared at the same elevation and same eccentricity in the contralateral side as the right stimulus.

The basic tasks were the same as described in the last chapter. For form discrimination, we used randomly rotated T- and L-shaped elements of white color. For color discrimination, we used vertically bisected discs with red and green halves **(Fig. 3.1b)**. The colors were isoluminant (as determined by flicker photometry) and the color order was either green left and red right (GR), or red left and green right (RG). For motion discrimination, we used white dumbell rotating around its center **(Fig. 3.1c)**. The angular velocity of the dumbbell was chosen randomly in the range from $1.5rps$ (revolutions per second) to $3.5rps$, and the rotation was either clockwise (CW) or counterclockwise (CCW). Instead of asymmetrical central and peripheral tasks described in the last chapter, here the two peripheral tasks were symmetrical, both involving reporting which of the two alternatives had been present.

For form and color discrimination, the visual persistence of stimulus elements was curtailed by masking, and *stimulus onset asynchrony* (SOA) was defined as the time between stimulus and mask onset. For motion discrimination, no masking was used, as there was no discernible visual persistence of dumbbell rotation. Here, SOA was defined as the time between stimulus onset and offset. For each task and subject, the SOA was adjusted individually to obtain a performance level of $\approx 85\%$ correct when the task in question was carried out alone. SOAs for different subjects and different tasks ranges from $\approx 60ms$ to $\approx 100ms$.

Compared to tasks discussed in the last chapter, all tasks in the present study are relatively less attention demanding for several possible reasons. First, the subject only needs to process one single element for the peripheral task instead of five elements for the previous central task. The removing of the crowding effect (the fact that discrimination of one object is degraded when it is flanked by similar objects) make the task much less demanding. Also, because the stimulus configuration is symmetrical, it is more natural to divide attention than in previous central-peripheral paradigm. For the peripheral task, reducing positional uncertainty (from eight possible locations to four possible locations) also contributes to reduce attentional cost.

3.2.3 Concurrent-task procedure

Experiments were conducted and analyzed as described in the last chapter. To obtain the full range of outcomes, we used three variations of the basic double-task instruction: (i) "perform both tasks but give priority to the left task", (ii) "perform both tasks but give priority to the right task", (iii) "perform both tasks and give equal priority to both tasks". Note that the display remained the same for all instructions and always contained both left and right components.

3.2.4 Differentiated and undifferentiated capacity

To formalize the distinction between differentiated and undifferentiated capacity and quantify the results, we treat attentional capacity as a divisible resource. We postulate a function which describes visual performance as the fraction of attention increases from zero to unity ("performance-resource function", PRF). Although we cannot observe PRFs directly, we can infer them from the "attention-operating characteristic" (AOC or, more informally, as the trade-off curve). We use the same method to fit the PRF as described in the last chapter. Details of the method can be found in section 2.2.5 and 2.2.6.

If attentional capacity is differentiated, the PRF of a given task will depend on the task with which it is paired. Dissimilar task pairs will result in more relaxed trade-off curves and more rapidly saturating PRFs than similar task pairs. If attentional capacity is undifferentiated, however, a given task will have the same PRF, no matter with which other task it is paired. In this case, similar and dissimilar task pairs will result in the same PRFs.

3.3 Results

3.3.1 Similar task combinations

The first series of experiments concerned concurrent performance of similar task pairs. Three pairs of tasks were investigated: discrimination of form, color, and motion at left and right hemisphere. Attentional cost of the tasks was assessed by these experiments.

Results are shown in **Fig. 3.2**, and were comparable for different subjects and all task combinations. Generally, double-task performance was equal or inferior to single-task performance. When double-task performance of one task approached the single-task level, double-task performance of the other task was somewhere between chance level and single-task performance. And any improvement in the performance of one task was at the expense of a partial reduction in the performance of the other. A contingency analysis was performed after pooling blocks for each subject and task combination to determine whether subjects had maintained a stable division of attention. No significant ($p < .05$) response correlation was found.

A maximum likelihood analysis of these results reveals the attentional requirements of the tasks in question (see Methods in chapter 2). As the AOC curve is roughly compatible with a straight line above the diagonal,we expect the most likely values of α_l and α_r to be smaller than unity, and the most likely values of β to be close to 1. (The subscript 'l' and 'r' denotes parameters of the left and right task, respectively.) The most likely values of α_l and α_r are between 0.75 and 0.79 for all three task combinations. The most likely values of β_l and β_r are between 1.14 and 1.20 for all task combinations. It is not surprising that roughly α_l=α_r and β_l=β_r as we are dealing with similar task combinations and the two tasks are exactly the same. However, α and β for the three different tasks involving form, color and motion are also similar, reflecting that these three tasks share similar RPF.

3.3.2 Dissimilar task combinations

A second series of experiments established concurrent performance of dissimilar tasks. Three task combinations were investigated: form vs. color discrimination, form vs. motion discrim-

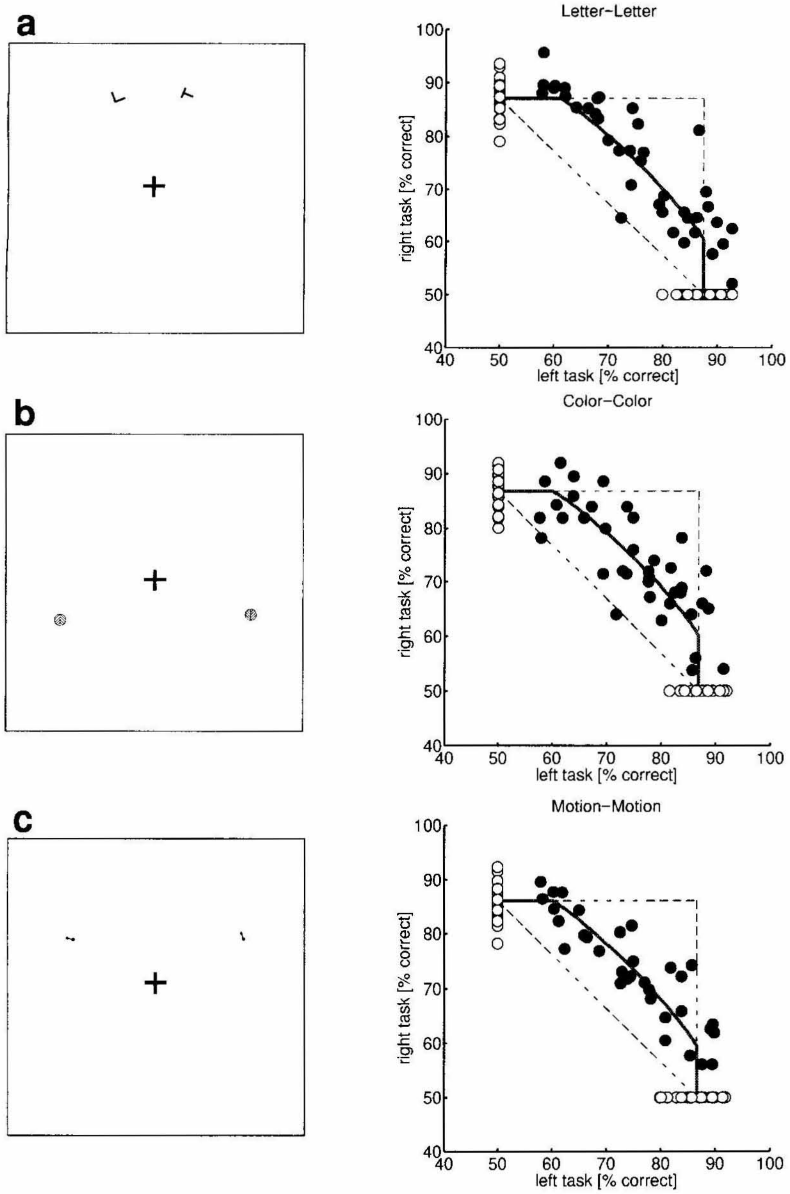


Fig. 3.2 Displays and results for combinations of similar tasks. Stimulus and mask displays are **not** drawn to scale. Red and green color is indicated by hatching. Results are presented in the form of an attention-operating characteristic (AOC), with left task performance on the horizontal and right task performance on the vertical axis. Open circles indicate single-task and filled circles double-task performance. Each circle represents one block of 50 trials. Dashes lines indicate the theoretical extremes of the AOC. The solid line represents the most likely trade off function, as determined by maximum likelyhood analysis. **(a)** Left and right form discrimination. **(b)** Left and right color discrimination. **(c)** Left and right motion discrimination. In all three experiments, double-task performance falls inside the theorectical extreme of the AOC, indicating intermediate trade-off.

ination and color vs. motion discrimination for tasks in left and right hemisphere. Subjects were counterbalanced for left and right tasks.

Results are shown in **Fig. 3.3**, and were comparable for different subjects and all task combinations. Double-task performance was characterized by an intermediate trade-off between tasks: when either task was performed at its best, the other task was performed above chance but less than the single-task performance, and any improvement in the performance of one task was at the expense of a partial reduction in the performance of the other. A contingency analysis was performed after pooling blocks for each subject and task combination to determine whether subjects had maintained a stable division of attention. No significant ($p < .05$) response correlation was found.

Again, we use maximum likelihood analysis to reveal the attentional requirements of the tasks in question. The AOC curves are roughly compatible with a straight line above the diagonal, and they are silimar to AOC curves for similar task combinations. The most likely values for αs amd βs comfirm this. The most likely values of α_l and α_r are between 0.74 and 0.79 for all three task combinations. The most likely values of β_l and β_r are between 1.14 and 1.21 for all task combinations. The αs and βs obtained here are not significantly different from values obtained in similar task combination. The shows that the tasks in this set require similar amount of attention, whether they are paired with a similar task or not.

3.3.3 Earlier and later sessions

We observed singnificant learning at the single task level. For three of the five subjects, we prolonged the training until the performance stablized in order to avoid further complication. For the other two subjects, we paid a closer look at the learning process. During the 15 sessions of data collection, in order to keep single task performance around 80% - 90%, we had to reduce the presentation time (SOA) in later sessions. The average SOA reduction is about 24 ms. Maximum likelihood analysis is performed to determine the PRF. There is no significant change for *alpha* and *beta* values in early and later sessions. Thus, single task performance improves over time and double task performance improves accordingly, but there is no significant change in attentional interaction, unlike what we observed for demanding tasks in the last chapter, where subjects learned to allocate attention more efficiently in later

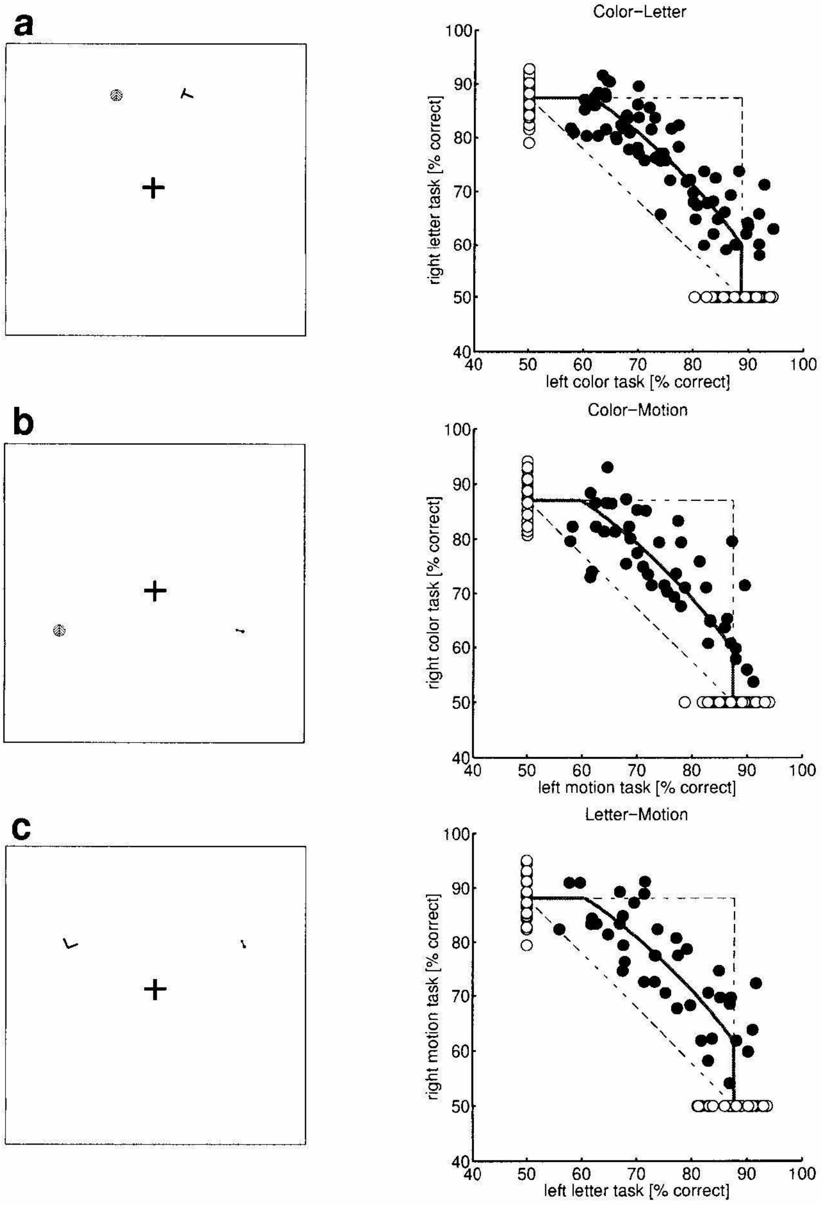


Fig. 3.3 Displays and results for combinations of dissimilar tasks. See caption of **Fig. 3.2**. **(a)** Left color and right form discrimination. **(b)** Left color and right motion discrimination. **(c)** Left letter and right motion discrimination. All three task combinations have intermediate trade-off.

sessions. It is probably because this set of tasks is less-demanding, and with symmetrical display, it is more natural for observers to divide attention.

3.4 Discussion and conclusion

3.4.1 General remarks

Attention is widely considered to be first and foremost a means of controlling the flow of information between different levels of processing. Stimuli selected by attention gain access to higher levels of processing which include recognition, awareness, memory, and the generation of voluntary responses, while stimuli rejected by attention are denied such access. This contrasts with lower levels of processing, which represent all stimuli indiscriminately, whether selected by attention or not. Lower processing levels are presumed to have spatially parallel architecture and to be free of capacity limitations ("early vision"). A strong implication of this view is that effects of attention are more evident at higher than at lower levels of processing (*e.g.*, Broadbent, 1971; Kahneman, 1973; Kahneman & Treisman, 1984; Nakayama & Joseph, 1998; Neisser, 1967; Pashler, 1997).

A somewhat different possibility, which has recently been gaining ground, is that attention affects all levels of processing. In this view, attention has the direct consequence of enhancing or attenuating responses at lower levels of processing, and the indirect consequence of facilitating access to higher levels of processing. This view is easier to reconcile with the neurophysiological finding that attention modulates neural responses in throughout visual cortex (reviewed by Desimone & Duncan, 1995; Maunsell, 1995). Attention has been observed to modulate visual responses in inferotemporal cortex (Chelazzi, Miller, Duncan, & Desimone, 1993; Miller, Li, & Desimone, 1993; Ungerleider, 1995), in extrastriate cortex (Moran, Desimone, 1985; Motter, 1994; Treue & Maunsell, 1996) and to some extent even in striate cortex (Motter, 1993). Another fact which points to a modulation of lower processing levels is that in many cases attention enhances stimulus discriminability d', suggesting an altered sensory representation (*e.g.*, Bonnel & Miller, 1994; Braun, 1994; Downing, 1988; Nakayama & Mackeben, 1989; Lee, Koch, & Braun, 1997ab).

An important implication of this modified view is that certain stimuli may reach higher levels of processing even without being selected by attention. For the necessary and sufficient condition for access to higher levels is *not* selection by attention but simply a suitably strong response at lower levels. Of course, most stimuli elicit a suitable response only as a result of being selected by attention. However, when a visual scene contains stimuli that are "salient" (Koch & Ullman, 1985; Robinson & Petersen, 1992), observers are subjectively aware of, and able to respond to, such stimuli even when attention is engaged by a demanding task elsewhere in the scene (Braun, 1994; Braun & Sagi, 1990; Braun & Julesz, 1998). Thus it would appear that "salient" stimuli elicit a suitably strong response to reach higher levels of processing even without being selected by attention.

Its "limited capacity" is the characteristic of attention. Although it is well established that only a limited number of stimuli can benefit from attention at any one time, it is far from evident why this should be so (Pashler, 1997, p. 226ff). A reason often cited is that the limitation does not reflect so much attention but the higher levels of processing to which attention provides access (e.g., Kinsbourne, 1981). In this view, attention can (and occasionally does) select a large number of visual stimuli, but the results are counterproductive: visual performance suffers because higher levels of processing are equipped to handle only a few stimuli at a time. An alternative possibility is that the limitation is inherent in the process of selection, which may be so constituted as to allow only a certain number of stimuli to be selected at any one time. Perhaps selection involves the formation visual objects and is limited by Gestalt rules (e.g., Kanwisher & Driver, 1992). Although the present study does not directly address the reason for the limited capacity of attention, its results nevertheless have some bearing on the matter.

3.4.2 Choice of paradigm

Over the years, various paradigms have been employed to measure interference between concurrent visual tasks. In many cases, an important consideration was the ability to distinguish attentional limitations from the well-known limitations of short-term memory (*e.g.*, Potter, 1976; Sperling, 1960). For example, one can compare discrimination of two stimuli that are presented either simultaneously or successively. Since memory limitations are expected with both modes of presentation but attentional limitations only with simultaneous presentation, the comparison should distinguish between limitations of memory and attention (Duncan, 1980; Duncan *et al.*, 1994; Fisher, 1984; Kleiss & Lane, 1986; Shiffrin & Gardner, 1972; Ward *et al.*, 1996).

Another way of identifying *attentional* interference between concurrent discriminations are experiments involving either the same or different visual objects (Duncan, 1984; Treisman, 1969; Vecera & Farah, 1994). In these experiments, observers either discriminate two attributes of one object (ignoring the other object), or one attribute of each object (ignoring the other attributes). Interference is found only in the latter case, the discrimination of one attribute of each object, leaving little doubt that interference originates at the visual level (*i.e.*, the level of attention). Incidentally, these results also highlight the intimate connection between attentional selection and visual objects (Kanwisher & Driver, 1992).

Besides memory and attention, yet another potential source of interference between concurrent tasks is *response selection* (Johnston, McCann, & Remington, 1995; Pashler, 1991, 1994). It is readily distinguishable from attentional interference, since it is independent of the exact timing of stimulus presentation, occurs even when two tasks concern different modalities, and is highly sensitive to the characteristics of the response (*e.g.*, order, timing, complexity). Another difference may be that the resources of response selection cannot be shared in a graded fashion between two concurrent tasks (Pashler, 1997, pp. 311ff). Indeed, interference at the response level may involve subcortical rather than cortical structures (Ivry, Franz, Kingstone, & Johnston, 1998; Pashler, Luck, O'Brien, Mangun, & Gazzaniga, 1995). These factors have led Johnston and colleagues to argue for two distinct types of attention — "input attention" and "central attention" — which produce interference at different levels (Johnston *et al.*, 1995).

The present experiments were designed to measure task interference at the level of visual attention. The two components formed two distinct visual objects, which were presented simultaneously and masked, while responses were unspeeded. Responses involved one or two binary choices, well below the capacity of visual short-term memory (*e.g.*, Luck & Vogel, 1997). Control experiments show that interference disappears when central and peripheral targets appear successively rather than simultaneously, and that interference is independent of response order (Braun & Julesz, 1998; Julesz & Braun, 1992). This makes clear that any interference obtained with this paradigm does not reflect limitations at the level of memory or response selection.

The most surprising finding with this paradigm is that widely disparate outcomes are obtained with seemingly similar task combinations. As discussed further below, some task combinations produce little or no interference while other, similar task combinations produce severe interference (Braun, 1994; Braun & Julesz, 1998). Thus, the linear trade-off curves reported in chapter 2 are not necessarily typical. For example, there is little or no interference between either central form discrimination (*i.e.*, the very task used in chapter) and peripheral localization of a uniquely oriented Gabor element ("pop-out"), or between central form discrimination and *two* peripheral hue discriminations. In both cases, the peripheral task is carried out comparably well with and without the central task. In other instances with the same outcome, the peripheral tasks involved up to four binary choices, which demonstrates among other things that the task combinations studied here fall well within the capacity of short-term memory.

3.4.3 Attentional capacity is divisible

Our results show that observers divide attention between tasks in almost every trial, rather than simply allocating attention sometimes to one task and sometimes to the other. This follows from the relatively small response correlation we observe: the average accuracy of one response is only about 4 percentage points better when the other response is incorrect rather than correct. The most likely cause of this small correlation is some degree of variability in the division of attention.

Of course, a major unsolved question is *how* the division of attention is accomplished. Two possibilities are evident. One is that attention concurrently "selects" the stimuli relevant to both tasks, but does so only "partially". The how and why of such a "partial selection" is far from obvious. The other possibility is that attention shifts from task to another at some point during the trial. In this case only one set of stimuli is selected at any one moment, but each set is selected only for about half the time. The problem with this possibility is that attention would have to shift essentially without delay between the two sets of stimuli.

Current evidence does not rule out either of these possibilities (Braun, 1998; but see Duncan *et al.*, 1994; Ward *et al.*, 1996). When subjects are as highly practised as they were in the present study, a "rapid shift" certainly appears no less likely than a "partial selection".

3.4.4 Attentional capacity is undifferentiated

The outcome of the present study could hardly have been clearer. Both similar and dissimilar task combinations resulted in an almost same amount of intermediate performance trade-off. This was true for three similar task combinations which involved two discriminations of form (T *vs.* L), color (red-green *vs.* green-red), and motion (clockwise *vs.* counterclockwise), respectively, and for three dissimilar combinations which involved discriminations of color and form, color and motion, motion and form.

We conclude that similar and dissimilar tasks draw on exactly the same attentional capacity; in other words, that attentional capacity is entirely undifferentiated. We have considerably strengthened the evidence for it. By virtue of using tasks with intermediate attentional cost, and by quantifying this cost, our paradigm was sufficiently sensitive to reveal intermediate outcomes (i.e., a partly differentiated capacity). The fact that we obtained similar amount of intermediate trade-off for similar and dissimilar task combinations strenthened our conclusion of the last chapter: attentional capacity is undifferentiated.

If attentional capacity is undifferentiated, this fact may provide some clues as to the nature of this capacity. As mentioned, it has never been evident what limits attentional capacity in the first place. Many writers assume that the limitation has nothing to do with attention proper, that is, the process of selecting some stimuli and rejecting others, but that it is a limitation at the level of classification or recognition, which can only deal with a limited amount of information (*e.g.*, Mozer, 1991). However, classification and recognition presumably involve different cortical sites depending on the stimulus being recognized or classified, and it is difficult to see why all of these sites should be subject to the same capacity limitation. In other words, if the limitation arises at higher levels of processing, one would expect attentional capacity to be at least partially differentiated. It is easier to understand the results of the present study if capacity limitations arise from the step of selection itself. If attention can select only a limited number of stimuli at any one time, it follows that attentional capacity will be as undifferentiated as we have observed. An important consequence of this is that cortical sites representing different stimulus dimensions will be tightly coupled by attention, so that any stimulus selected at one site will also be selected at the others (Duncan, Humphreys & Ward, 1997). We conclude that the present results are more consistent with capacity limitations of selection than of subsequent levels of processing.

None of these considerations address the more fundamental issue as to whether capacity limitations have a functional role or simply reflect biophysical limitations. Our guess is that capacity limitations serve a useful function in connection with the Gestalt principles governing the formation of "visual objects". We surmise that selection involves an interaction between attentional and grouping processes, and that the result of this interaction is necessarily a single "attended" visual object. However, this remains mere speculation.

Finally, it is worth mentioning that different sensory modalities seem to draw on different attentional capacities. Concurrent words or syllables are more readily discriminated when they are presented in different modalities (*e.g.*, visually and aurally) rather than in the same modality (Duncan, Martens & Ward, 1997; Treisman & Davies, 1973). Similarly, an attentionally demanding visual tracking task can be carried out concurrently with an auditory, but not a visual, word recognition task (Wen, Koch & Braun, 1995). Although instances of comparable inter- and intra-modal interference have been reported as well (Lindsay, Taylor & Forbes, 1968; Massaro & Warner, 1977), these seem to be more in the nature of exceptions to the rule (Pashler, 1997, pp. 157ff).

3.4.5 Attentional cost of discrimination

The paradigm of the present study has also been used in two previous studies to measure the attentional cost of discrimination tasks (Braun, 1994; Braun & Julesz, 1998). Part of the results are presented in **Fig. 3.4** and **Fig. 3.5**. When one compares the results, one is struck by large differences in the attention cost of different tasks. One way to conduct this comparison is to use the parameter α as a measure of attentional cost, that is, the fraction of attention at which peripheral performance reaches ceiling. After re-analyzing the results of previous studies accordingly, one obtains $\alpha \approx 0.3$ for the discrimination of triangular and circular forms, $\alpha \approx 0.2$ for the discrimination of subtle hues, and $\alpha \approx 0.25$ for the simultaneous discrimination of color and orientation **(Fig. 3.6)**. Of course, α is between 0.7 and 0.8 for all discriminations used in the present study.

Why this large difference in attentional cost? The tasks of the present study are certainly not more "demanding" than those of previous studies, inasmuch as they are performed comparably well at comparably long SOAs. The three tasks used here are performed between 82% and 90% correct at SOAs between 60ms and 100ms, while the three tasks studied previously are performed between 85% and 91% correct at SOAs between 72ms and 126ms. One possibility is that the present tasks are more "complex" than those used previously, in that they involve the discrimination of two target components (i.e., two line elements, two half-disks, or two terminal points). Another possibility is that the increased attentional cost is due to the need to jointly discriminate attributes and their positions (i.e., the orientation and position of two lines, the color and position of two half-disks, or the motion and position of two terminal points). This latter possibility appears particularly attractive since recent single-unit recordings in area V4 of the awake macaque suggest that attention helps encode the relative position of stimulus attributes (Connor, Preddie, Gallant, & Van Essen, 1997; Salinas & Abbott, 1997). This is clearly an issue which deserves further study.

3.5 Conclusion

We investigated three types of tasks with intermediate attentional costs and quantified the extent to which they interfere when observers attempt to perform them concurrently. We find that similar and dissimilar task combinations result in exactly the same degree of interference. This shows that all tasks investigated here engage precisely the same attentional capacity, strongly suggesting that there *is* only one such capacity. In other words, attentional capacity is undifferentiated. This confirms similar finding presented in the last chapter, which is regarding tasks of high attentional cost. Thus, the notion that attention is a unitary process is further strengthened in this study.

Translated into neural terms, an undifferentiated attentional capacity means that visual cortical areas processing different stimulus dimensions must be tightly integrated, since at-

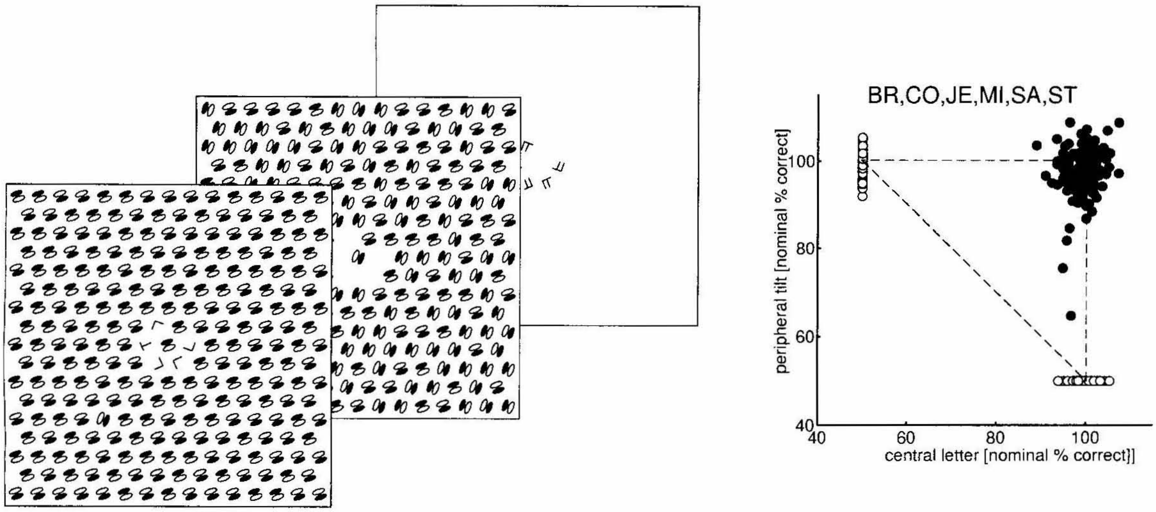


Fig. 3.4 Central form discrimination and peripheral pop-out localization (after Braun & Julesz, 1998). The central task is identical to the one used in the present study. The peripheral task involves localizing a uniquely oriented Gabor element in a dense array of uniformly oriented Gabor elements. Observers report whether the uniquely oriented element appeared in the upper of lower half of the display. Results are shown in the format of the present study. It is evident that there is little or no interference, and that both tasks together are performed as well as each task alone. Note that the peripheral localization is comparably demanding than the peripheral tasks used in the present study (i.e., SOAs and single-task performance are comparable).

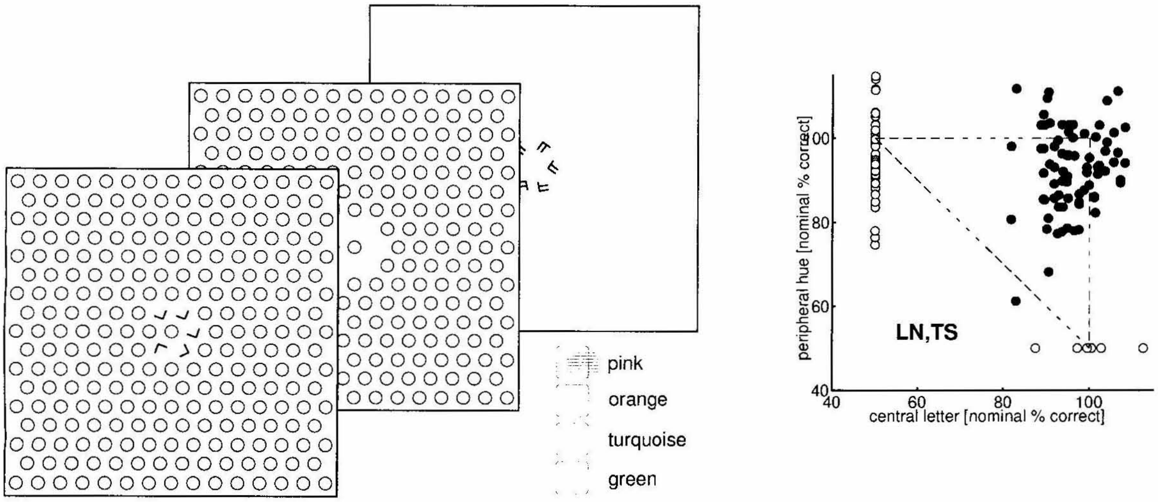


Fig. 3.5 Central form discrimination and peripheral hue discrimination (after Braun & Julesz, 1998). The central task is again identical to the one used in the present study. The peripheral task involves discriminating the hues of two chromatic targets in a dense array of isoluminant nonchromatic elements. Observers report whether the target in the upper half of the display was "pink" or "orange" and whether that in the lower display half was "turqoise" or "green". Results are shown in the format of the present study. In this case as well there is little or no inference. Note that the concurrent task situation requires three responses, and that the results show the average of both hue responses.

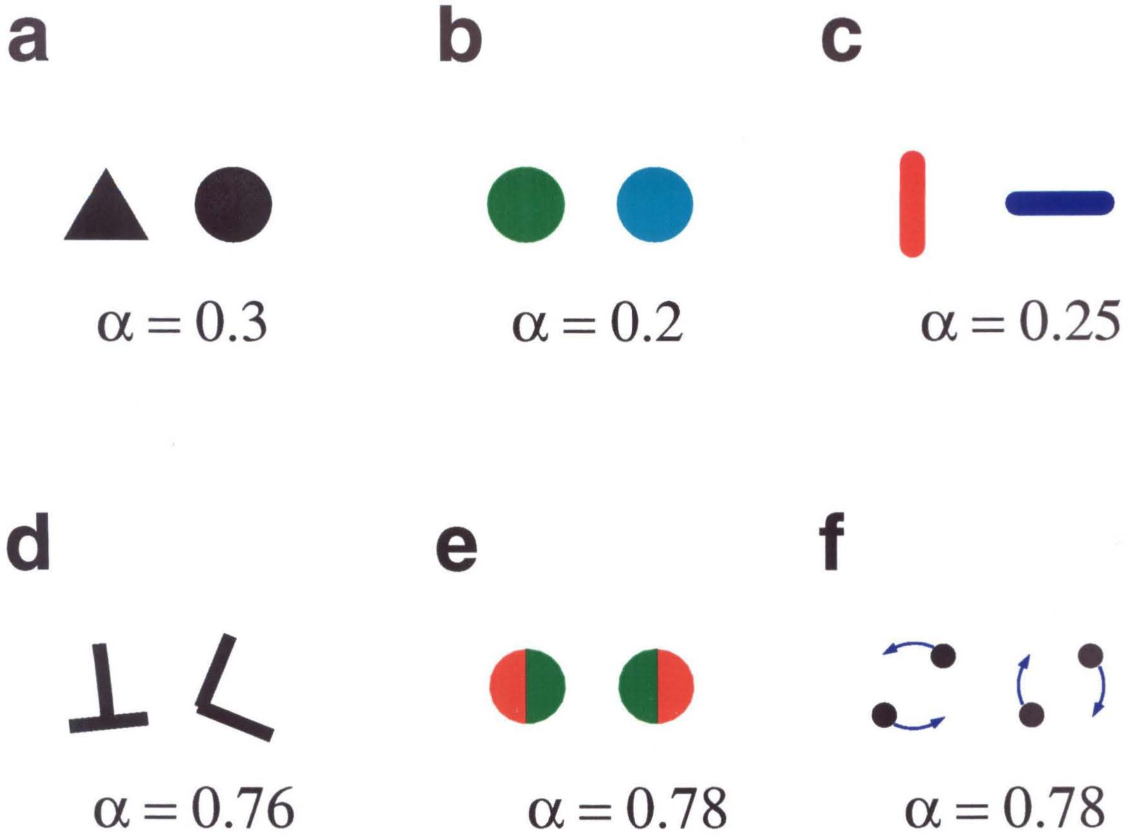


Fig. 3.6 Attentional cost of discrimination tasks. Comparison of six discrimination tasks whose attentional cost was measured with the same paradigm (*i.e.*, by engaging attention with a letter discrimination at display center). Attentional cost is expressed in terms of the parameter α (see **Fig. 2.2**) **(a)** Discrimination of form (triangular/circular)(Braun, 1994). **(b)** Discrimination of hue (orange/pink and turqoise/green)(see**Fig. 3.5**). **(c)** Simultaneous discrimination of orientation (horizontal/vertical) and color (red/green/blue/yellow)(Braun, Julesz, 1998). **(d)** Discrimination of letter form (rotated T or L). **(e)** Discrimination of coloring of bisected disks (red-green/green-red). **(f)** Discrimination of sense of rotation (clockwise/counter-clockwise).

tention appears to select the same region of visual space in all of these areas (Duncan *et al.*, 1997a). It further means that capacity limitations are likely to be an integral aspect of selection and thus truly a limitation of attention. This contrasts with the widely held notion that capacity limitations reflect shortcomings of recognition and other higher levels of processing (Kinsbourne, 1981; Pashler, 1997).

Chapter 4

Spatial vision thresholds in the near absence of attention

4.1 Spatial vision thresholds: Detection vs. discrimination

4.1.1 Background

Although it has long been recognized that visual processing is strongly influenced by attention (Helmholtz, 1850; James, 1890), the precise nature of this influence remains unclear. Most would allow that attention does more than simply select from among the visual information that is made available by early stages of visual processing and, indeed, it would seem that attention actively shapes the early visual processing of attended information to suit current behavioral requirements. In the terminology of signal-detection-theory (SDT; Green and Swets, 1966; Macmillan and Creelman, 1991), attention can alter the sensitivity or d' of a visual discrimination rather than merely its criterion or β [1]. For example, when the amount of attention paid to a particular stimulus is manipulated with visual "cueing", the d' for discriminating, say, a simple shape tends to be significantly larger at cued than at uncued locations (Bashinski and Bacharach, 1980; Shaw, 1984; Mueller and Findlay, 1987; Downing,

---

[1] Note that this β is different from the β used in the PDF function in chapter 2 and chapter 3

1988; Nakayama and Mackeben, 1989). Even stronger evidence for attentional effects on d' comes from experiments in which attention is divided in various proportions between two concurrent visual tasks (Sperling and Melchner, 1978; Duncan, 1984; Braun and Sagi, 1990, 1991; Bonnel, Possami, and Schmitt, 1987; Bonnel and Miller, 1994; Braun, 1994; Braun and Julesz, 1996).

Neurophysiological studies confirm that attention affects almost all levels of visual processing. In the visual cortical areas of the so-called "object" pathway (areas V2, V3, V4, and inferotemporal areas; Desimone and Ungerleider, 1989; Felleman and VanEssen, 1991), up to half of the neurons respond more strongly to a stimulus with respect to which the animal carries out a visual task than to a stimulus that is viewed passively (presumably because the task-relevant stimulus is attended; Moran and Desimone, 1985; Spitzer, Desimone, Moran, 1988; Haenny, Maunsell, and Schiller, 1988; Chelazzi, Miller, Duncan, Desimone, 1993; Motter, 1994; Maunsell, 1995). In some cases even neurons in area V1 exhibit attentional effects of this kind (Motter, 1993; Press, Knierim, and Van Essen, 1994). Functional imaging studies leave little doubt that similar attentional effects operate in humans (Corbetta, Miezin, Dobmeyer, Shulman, Petersen, 1991; Maunsell, 1995).

Many psychophysical studies of attention have relied on the closely related paradigms of visual search and visual texture processing (Treisman and Gelade, 1980; Julesz, 1981, 1991; Treisman, 1991, 1992; Watt, 1992; Duncan and Humphreys, 1989, 1992; Wolfe, 1994). A somewhat different approach takes advantage of the fact that reportable visual experience does not cease in the near absence of attention. For example, sensitivity (d') for a luminance increment is rather similar at attended and unattended locations (Bashinski and Bacharach, 1980; Shaw, 1984; Mueller and Findlay, 1987; Bonnel, Stein, and Bertucci, 1991). The same is true for sensitivity d' for a stimulus with a unique feature (*i.e.*, shape, color, motion, *etc.*) which is embedded in a sufficiently dense and uniform array of stimuli lacking this feature (Nakayama and Mackeben, 1989; Braun and Sagi, 1990; 1991; Braun, 1993, 1994). The visual response to stimuli that are "unique" in this sense is thought to be particularly strong because of pervasive lateral inhibition between similar features at nearby visual locations, which attenuates responses to all other stimuli in such a display (Sagi and Julesz, 1985, 1987; Koch and Ullman, 1985; Malik and Perona, 1990; Rubenstein and Sagi, 1990). Thus, lateral inhibition explains why stimuli distinguished by a unique feature are visually salient and "pop out" from the display. As this lateral inhibition operates pre-attentively and "in parallel", stimuli rendered salient by a unique feature can guide eye movements and shifts of attention (Julesz, 1981; 1991). Accordingly, it should come as no surprise that such stimuli are readily reported even when attention focuses elsewhere in the display (Braun and Sagi, 1990, 1991; Braun, 1993, 1994).

If visual information can be reported in the near absence of attention, it is of evident interest to compare psychophysical performance under this condition with performance when attention is fully available (Braun, 1994; Braun and Julesz, 1996). This comparison must necessarily throw light on the ways in which visual experience is altered and augmented by attention.

This chapter investigates early visual processing in the near absence of attention. We chose to measure contrast thresholds as well as orientation thresholds for sinusoidal gratings (blurred with gaussian profile at the edge), partly because these thresholds are well characterized under normal conditions when attention is fully available (Rovamo and Virsu, 1979; Virsu and Rovamo, 1979; Orban, Vandenbussche and Vogels, 1984), and partly because these thresholds reflect the spatial frequency and orientation tuning of the visual filters that characterize the first stage of visual processing (reviewed in Spillman and Werner, 1990; Regan, 1991). In addition, we studied unidirectional and bidirectional offset thresholds for Vernier targets, because these thresholds may also relate to the tuning properties of visual filters (Westheimer and McKee, 1977; Wilson, 1986, 1991; Fahle, 1991; Waugh, Levi, and Carney, 1993; Harris and Fahle, 1995). By re-measuring these thresholds in the near absence of attention, we hoped to learn whether either the visual filters themselves or the interactions between visual filters are affected by attention.

To measure thresholds in the near absence of attention, we ask subjects to carry out two concurrent visual tasks, one of them designed to be highly demanding of attention ("primary task", see details in chapter 2). As a result, optimal performance on this task is reached only when attention is almost fully focused on it and thus almost completely withdrawn from the other task ("secondary task"). Performance on the primary task is monitored to ensure that subjects maintain this highly unequal division of attention. Thus the concurrent task paradigm ensures that substantially less attention is available for the secondary task than would be available without the primary task. Of course, it does not necessarily ensure that attention is entirely withdrawn from the secondary task. For this reason we speak of the "near" absence, rather than the absence, of attention. Further details on the concurrent task paradigm can be found elsewhere (Braun, 1994; Braun and Julesz, 1996; chapter 2 of this thesis).

Using this approach, we have shown that the near absence of attention exacerbates visual search asymmetries (Braun, 1994). A qualitatively identical pattern of results was encountered by Schiller and Lee (1991) following a lesion in area V4. Thus is appears that the absence of attention produces behavioral deficits that, at least in some respects, are comparable to those produced by a lesion in area V4. As mentioned, we have also shown that near absence of attention does not interfere with the detection of stimuli rendered salient by a unique feature, and that this is true even in the threshold region ($d' \approx 0.3$) (Braun, 1994; Braun and Julesz, 1996). In general, the residual visual experience in the near absence of attention seems to be considerably richer than hitherto appreciated, and permits even the discrimination of simple features of salient stimuli (Braun and Julesz, 1996).

4.1.2 Methods

Stimuli were generated by a Silicon Graphics computer system and displayed on a high resolution color monitor (1000 × 1280 pixels). Lightness and color of each pixel were determined by 3 × 8 bit RGB values. The frame rate was 72 Hz. Viewing was binocular, from a distance of about 120 cm, resulting in a display of approximately 12.5 × 16 deg of visual angle, with 1 deg corresponded to 80 pixels. Average screen luminance was 26.6 cd/m^2. For the contrast range used in the contrast sensitivity experiment, luminance was a linear function of pixel grey level (accuracy 2%). The room luminance was about 5 cd/m^2.

Three subjects participated in the experiment. Each subject was trained and tested for more than 30 hours. They were Caltech students and received $10 per hour for participating in the experiment. Not all subjects participated in all experiments, but every condition was investigated with at least two subjects. All subjects had normal or corrected to normal vision.

We used an adaptive staircase method to measure thresholds, specifically, the up-down transformed-response (UDTR) method suggested by Levitt (1970). Changes of the stimulus were made to depend on the outcome of two preceding trials. The intensity of the stimulus (that is, luminance contrast, orientation difference, or vernier offset, depending on the experiment) was increased with each incorrect response and decreased after two successive correct response (1-up/2-down, or 2-step). The upward and downward steps were of the same size. Levitt calculated the target probability converging to 0.707. This value is derived from the probabilities which are expected on the basis of a binomial distribution of correct and incorrect responses. Analysis of our experimental data showed that the performance at threshold is around 70% correct.

As in previous concurrent task studies (Braun, 1994; Braun and Julesz, 1996), we avoid using completely unpractised subjects since their results tend to vary greatly between individuals. For example, one subject may succeed immediately at performing two tasks concurrently while another subject may do so only after one or two days of practice. However, after two or three practice sessions (that is, two to three hours spead over days) subjects generally converge to a uniform result. A pragmatic reason for using practised subjects is our reliance on within-subject comparisons which makes it necessary to conduct tens of thousands of trials with each subject.

Although subjects were practised, performance on all tasks generally continues to improve somewhat. To ensure that all critical comparisons were based on comparable states of practice, we measured any given threshold both with and without the concurrent task during each session (one hour session per day). Thus the reported effects do not in any way reflect differences in practice level. Data were collected in blocks of 80 trials and every threshold measurement reported below was based on at least six blocks of trials. Each session consisted of alternating blocks with and without the concurrent task.

Displays always contained both central targets for the attention-demanding task (see below) and a peripheral target for the threshold measurement (sinusoidal grating or Vernier target). As a result, the only difference between situations with and without the concurrent task lay in the instructions provided to the subject and in the number of responses collected after each trial **(Fig. 4.1)**. The central targets were the same in all experiments, while the peripheral target was different in each experiment.

Attention-demanding Task

Form identification tasks in general, and letter identification tasks in particular, are thought to present a significant demand for the attentive resources of an subject (Bergen and Julesz, 1983; Kroese and Julesz, 1989; Duncan, Ward, and Shapira, 1994). Here we used an identification/search task involving 5 T- or L-shaped targets with randomized positions and orientations. Maximal luminance contrast was used, and after approximately 200ms the target elements were replaced by the elements of a perceptual mask. Details of the task were described in chapter 2, where we showed it engages all or most of visual attention no matter if the secondary task involves form, motion or color perception.

The five letter targets could appear at seven possible locations: the exact center of the display and six locations at 0.9 deg eccentricity, spaced evenly around the center. On any given trial, five T- or L-shaped elements were distributed randomly among the seven possible locations, as well as rotated randomly and independently, resulting in a large number of possible configurations. There were either five Ls, five Ts, four Ts and one L, or four Ls and one T. Subjects were instructed to report whether all elements were the same (five Ts, five Ls) or whether one was different from the other four (four Ts and one L, four Ls and one T). The masking pattern for the letter task consisted of F-shaped elements. Five such elements appeared at the same locations as the five target elements of the stimulus pattern, but in different states of rotation.

The choice of this task is not critical, as shown in results in chapter 2 and chapter 3. Since attentional capacity is undifferentiated, other attention-engaging tasks could have been substituted without changing the outcome. The present task was chosen simply because its ability to engage attention is better documented than that of other tasks (Braun, 1994; Braun and Julesz, 1996). Specifically, the attention-operating characteristic (Sperling and Melchner, 1978) of two letter discrimination tasks shows that either task engages attention *completely* (within measurement precision) (Braun and Bauer, 1993; Braun and Julesz, 1996).

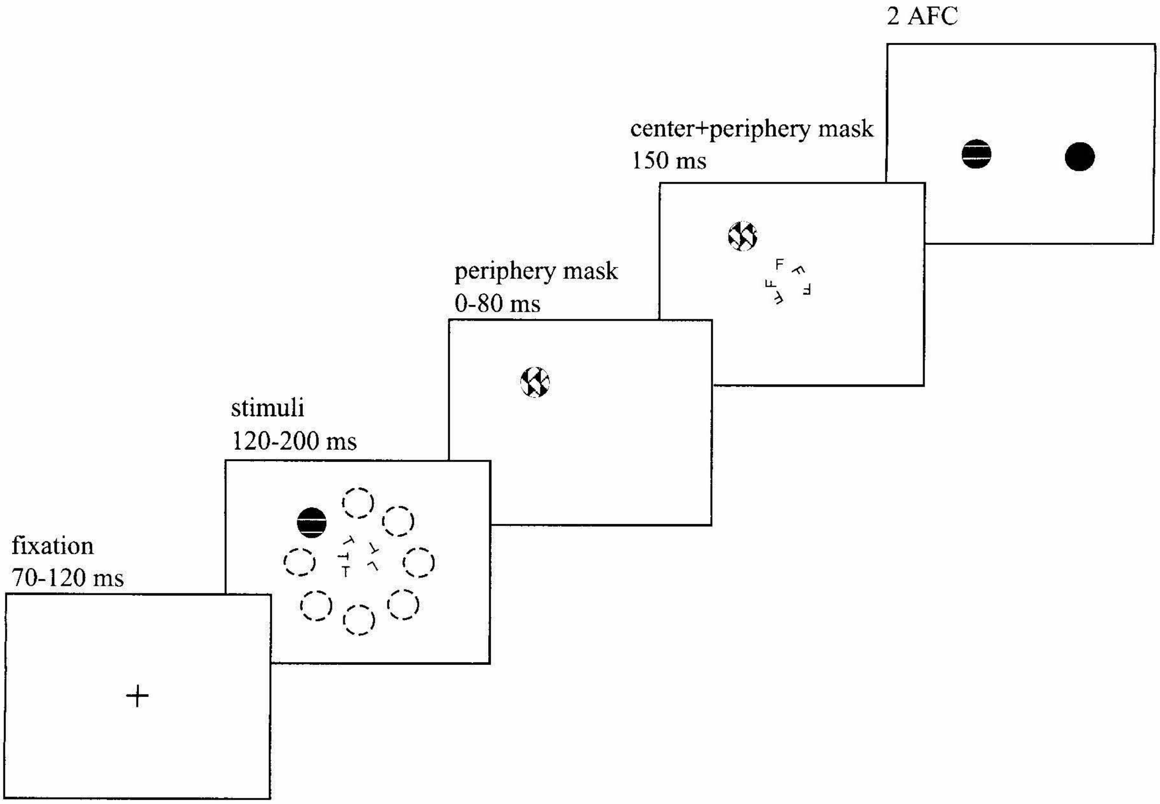


Fig. 4.1 Illustration of a trial sequence. The sequence begins with a fixation display, and is continued by a stimulus display, two mask displays, and a display which prompts the observer's response. The reason for having two mask displays is that this permits the central and peripheral parts of the stimulus display to be masked independently. The central part of the stimulus display consists of five letter-shaped elements, and the peripheral part consists of either a test grating (shown schematically) or a Vernier target (not shown). The peripheral test grating or Vernier target appears at one of eight possible locations (shown as dashed outlines) in the stimulus display.

Threshold measurements

- **Contrast thresholds:** We measured contrast thresholds for sine wave gratings as a function of spatial frequency. Designating the mean luminance by L and the spatial frequency by f, a vertically oriented sine wave grating of contrast C is described by the intensity distribution $I(x)$:

  $$I(x) = L * [1.0 + C * sin(f * x)]\,. \tag{4.1}$$

  To measure the contrast sensitivity function, the value of C was adjusted from trial to trial according to the staircase method disucssed above.

  The sine wave grating measured 1.5° × 1.5°, and at the margins of this area the luminance transition was blurred with a space constant of 0.1°. The grating appeared with equal probability at one of 8 random locations at 4 degree of eccentricity (presentation time 200 ms, viewing distance 120 cm). Its orientation was either vertical or horizontal. Following presentation of the mask (a plaid formed by two superimposed sine wave gratings of identical spatial frequency and orthogonal orientation), two gratings of the same size, one horizontal and one vertical, appeared at the bottom left of the display (away from all 8 positions at which the grating could appear during the trial) and the subject chose one of the two by clicking the mouse on it. We measure threshold with a staircase method (see above) and this procedure was repeated for 5 spacial frequencies between 1 cycle per degree (cpd) and 11.4 cpd.

- **Orientation thresholds:** To measure orientation thresholds, we presented a sine wave grating of 4 cpd and size 1.5° × 1.5° at one of 8 locations at 4° of eccentricity (presentation time 200 ms). The grating was either exactly vertical or slightly tilted to the left or right of vertical. The amount of tilt varied with the status of the staircase. Following presentation of the stimulus and mask, two gratings appeared at the bottom left of the display, one exactly vertical and one tilted, but otherwise identical to the grating in the stimulus. The difference in tilt reflected the status of the staircase. Subjects reported which of the two gratings had appeared in the stimulus by clicking the mouse on it. We measured thresholds for three levels of luminance contrast, 25%, 50%, and 100%, 100%, all three well above the threshold contrast measured in the previous experiment.

- **Unidirectional Vernier thresholds:** To measure unidirectional Vernier thresholds, we presented a pair of lines (each 80 pixels or $\approx 1^\circ$ in length and 1 pixel in width) forming a Vernier target at one of 8 possible locations at 4 deg eccentricity. Presentation time was 120 ms. The lines were tilted 20° from vertical, either to the left or right, in order to reduce aliasing due to finite pixel size. The lines were either precisely aligned, or exhibited a Vernier offset of an amount which varied with the status of the staircase (schematic stimuli shown in **Fig. 4.4**). After the stimulus and mask (lines parallel to the Vernier target and spaced by 1° covering the entire display except the center), two pairs of lines appeared at the bottom left, one aligned and one offset, but otherwise identical to the pair in the stimulus. The difference in Vernier offset reflected the status of the staircase. Subjects reported which of the two pairs had appeared in the stimulus by clicking a mouse on it.

- **Bidirectional Vernier thresholds:** To measure bidirectional Vernier thresholds, we presented Vernier targets which exhibited either a left or a right offset (schematic stimuli shown in **Fig. 4.4**). In all other respects, they were identical to those described above. As a result, it was no longer sufficient to simply report the presence or absence of a Vernier offset and observers were required to report the direction of the offset.

Experimental Procedure

Subjects were instructed to fixate a cross at the center of the display before initiating each trial. The trial sequence began with a blank interval of a duration chosen randomly in the range of 70-120 ms, continued with the stimulus presentation (120-200 ms, depending on the experiment), and concluded with the mask presentation (150 ms; see **Fig. 4.1**). Central and peripheral targets were masked separately, so that different presentation times could be obtained for different parts of the display. The random duration of the blank interval at the beginning of the trial sequence prevented planned saccades (which could have defeated the masking). Although eye movements were not monitored, we are confident that the relatively short presentation time and the random location of the peripheral stimulus prevented a second fixation. Both central and peripheral masks were designed to be as effective as possible, so that relatively large differences in performance are obtained from relatively small changes in the stimulus-onset-asynchrony (SOA = interval between stimulus and mask onset). As visible persistence near contrast threshold is likely to be short (Coltheart, 1980), the necessity for masking is unclear. Accordingly, some experiments were conducted both with and without a peripheral mask. The central mask was always used, however.

In separate blocks of trials, subjects were asked to report on both central and peripheral targets and to ignore the central targets and report only the peripheral target. In the first case two responses were collected after each trial, in the second case only one response was required. In both cases, every mistaken response elicited immediate auditory feedback. When subjects reported on both central and peripheral targets, the two tasks were ranked, with the central task being designated primary and the second target secondary. Subjects were told that they might encounter a trade-off between central and peripheral task performance and that, in this case, they should absolutely favor the central task. A baseline performance level was established for the central task by running one or two blocks of trials in every session which the peripheral task was ignored. When performance in the concurrent task situation fell significantly below this baseline, the block (80 trials) was rejected and the peripheral task performance was not counted towards the determination of the associated thresholds. This ensured that thresholds reflected a situation in which attention was nearly absent from the peripheral targets.

4.1.3 Results

For the three tested subjects, average performance for the primary task alone was 87% after some practice (not shown) while chance level is 50%. All subjects reported the primary task to be a highly demanding task and that it required considerable effort and concentration. In the concurrent task situation, subjects were encouraged to maintain a comparable level of performance for the primary task. Blocks of trials with primary task performance less than 80% were rejected and the peripheral task performance in these blocks was not considered in the computation of thresholds (see Methods).

Contrast Thresholds

Peripheral contrast thresholds were measured without the central task, for sine wave gratings with spatial frequencies of 1.0, 2.0, 4.0, 8.0, and 11.4cpd. We assumed that, in this situation, attention was fully devoted to the peripheral sine wave grating and its discrimination. Peak sensitivity was observed at 2-4 cpd and declined towards higher spatial frequencies, consistent with previous studies (Rovamo and Virsu, 1979; Virsu and Rovamo, 1979).

With the same subjects (and in the same sessions, see Methods), we measured peripheral contrast sensitivity while the primary task was being carried out concurrently. We assumed that, in this situation, attention was almost fully devoted to the primary task and, thus, almost completely absent from the sine wave grating and its discrimination. Contrast sensitivity obtained with and without the primary task was comparable for all subjects at all spatial frequencies. Although thresholds were consistently somewhat higher with the concurrent task, the difference does not reach significance for any subject or spatial frequency (t-test, $p > .05$). This was true both when the peripheral sine wave grating was left unmasked (three subjects, **Fig. 4.2a**) and when it was masked (two subjects, **Fig. 4.2b**). There was no significant effect of masking. A 2 (with or without primary task) $\times$ 3 (subjects) $\times$ 5 (spatial frequencies) analysis of variance was carried out for the results obtained with masking. Combining data from all spatial frequencies for each subject, performance of the primary task had a significant effect in two of three subjects ($F(1,4) = 15.20, 25.56; p < 0.01; F(1,4) = 2.25, p = 0.21$). Combining data for all subjects, the effect of the primary task was not significant ($F(1,20) = 0.53; p = 0.47$), probably due to the large performance differences between subjects.

Orientation Thresholds

After finding that near absence of attention has little or no effect on contrast sensitivity, we asked how absence of attention would affect another important aspect of spatial vision — orientation discrimination. To investigate this question, we used sine wave gratings of

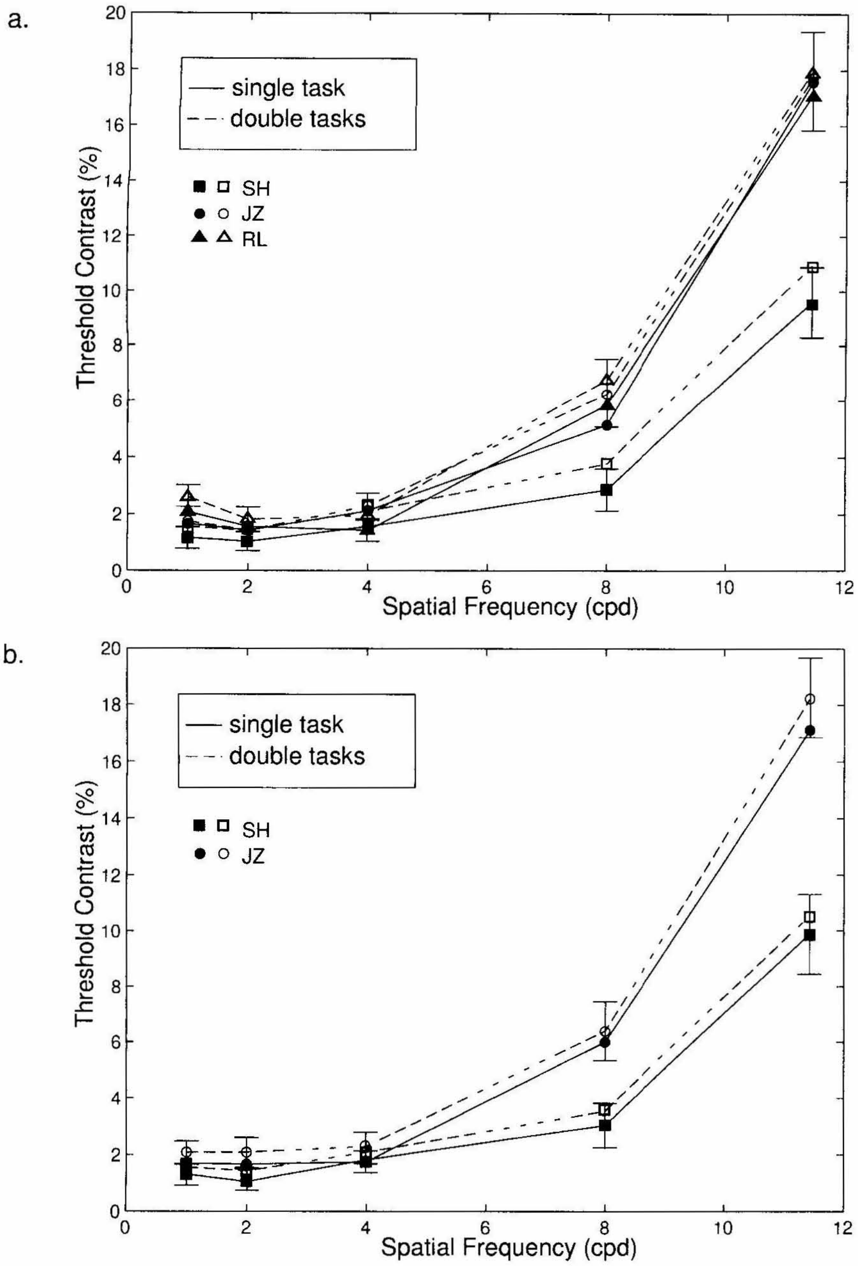


Fig. 4.2 Contrast thresholds in the full presence and near absence of attention. Threshold contrast is plotted as a function of spatial frequency of the test grating. Thresholds obtained when attention is fully available ("single task") are plotted with solid symbols and solid lines. Thresholds obtained when attention is nearly absent ("double task") are plotted with open symbols and dashed lines. (a), Contrast thresholds without peripheral masking (three observers); (b), Contrast thresholds with peripheral masking (two observers). Error bars represent the average standard error at each spatial frequency and were computed separately for single and double tasks.

4 cpd, as gratings of this spatial frequency exhibited the highest contrast sensitivity in the previous experiment. The threshold for orientation discrimination was measured in terms of the difference in orientation between vertically and off-vertically oriented gratings. This threshold was established for gratings with a luminance contrast of 25%, 50%, and 100%. These contrast thresholds are approximately 14, 28, and 56 times the threshold contrast determined in the previous experiment.

Thresholds were determined both with and without the primary task, that is, both in the near absence and the full presence of attention. **Fig. 4.3a** shows the results of three subjects on displays in which the peripheral grating was not masked. **Fig. 4.3b** shows results of two subjects when the peripheral grating was masked by a plaid formed by superimposing two gratings of different orientation. The overall effect of a peripheral mask is not significant. Otherwise, the results with and without the peripheral mask are very similar, in that thresholds for orientation discrimination are elevated 2.9 to 5.0-fold in the near absence of attention. Specifically, thresholds are between 1 and 2° in the presence of attention and there is no discernible dependence on contrast level, as has been shown by previous studies (Orban, Vandenbussche, and Vogels, 1984). In the absence of attention, thresholds increase to between 4 and 6°. This difference is significant for each subject and contrast level (t-test, $p < 0.01$).

Uni- and Bidirectional Vernier Thresholds

With three subjects, we determined unidirectional Vernier thresholds (thresholds for discriminating the presence or absence on a Vernier offset) in both the presence and near absence of attention. The Vernier target was always masked, and results are shown in **Fig. 4.4**. In the presence of attention, thresholds were in the range of $2.5'$ to $3.5'$. This would seem to be roughly consistent with threshold values of $\approx 1'$ reported for considerably brighter targets ($960 cd/m^2$) that are presented without positional uncertainty (Levi, Klein, and Aitsebaomo, 1985). In the near absence of attention, thresholds were elevated slightly, by a factor ranging between 1.10 and 1.20. Although this threshold elevation was not significant for any individual subject (t-test, $p < 0.01$), a 2 (with and without primary task) × 3 (subjects) analysis of variance showed that it did reach significance when data from all subjects were combined

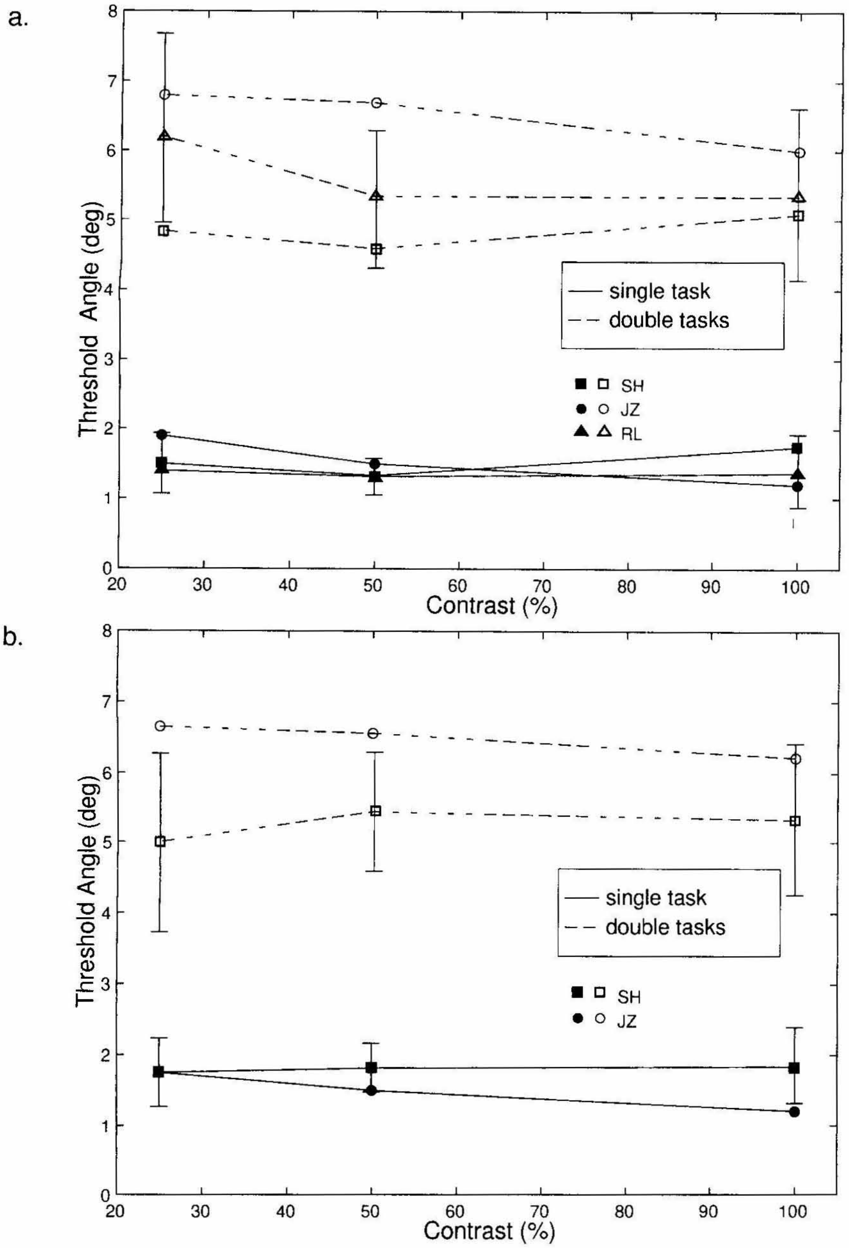


Fig. 4.3 Orientation thresholds in the full presence and near absence of attention. Threshold orientation is plotted as a function of luminance contrast of the test grating. Thresholds obtained when attention is fully available ("single task") are plotted with solid symbols and solid lines. Thresholds obtained when attention is nearly absent ("double task") are plotted with open symbols and dashed lines. (a), Orientation thresholds without peripheral masking (three observers); (b), Orientation thresholds with a peripheral mask (two observers). Error bars represent the average standard error at each luminance contrast and were computed separately for single and double tasks.

$(F(1,2) = 149.50; p < 0.01)$.

With the same three subjects, we determined bidirectional Vernier thresholds (thresholds for discriminating left or right Vernier offsets) in both the presence and near absence of attention. Again the Vernier target was masked. As shown in **Fig. 4.4**, thresholds in the presence of attention were in the range of $3.0'$ to $4.0'$ and thus somewhat larger than the unidirectional thresholds. This difference is consistent with reports that extrafoveal vision poorly discriminates spatial phase (Rentschler and Treutwein, 1985). In the near absence of attention, thresholds were elevated by a factor ranging from 1.80 to 1.90 across observers. This difference was significant for each observer (t-test, $p < 0.01$). This shows that near absence of attention has a markedly different effect on bi-directional than on uni-directional Vernier thresholds.

4.1.4 Discussion

We measured a number of spatial vision thresholds in both the full presence and near absence of attention. We find that thresholds can be established with traditional staircase methods in both situations. This is consistent with previous finding that residual visual experience in the near absence of attention is considerably richer than hitherto appreciated and can be readily studied with appropriate psychophysical paradigms (Braun and Sagi, 1990, 1991; Braun, 1994; Braun and Julesz, 1998).

The most interesting aspect of the present results is that the investigated thresholds differ substantially in the degree to which they depend on attention. For example, near absence of attention has at most a small effect on contrast thresholds for sine wave gratings and on unidirectional Vernier thresholds, but a rather large effect on orientation thresholds for sine wave gratings and bidirectional Vernier thresholds.

Note that the observed small effect of attention on contrast and unidirectional Vernier thresholds may actually be an overestimate of the true effect. When two tasks are performed concurrently, as was the case in the present experiments, one may expect some interference at post-perceptual levels of processing (i.e., response encoding and execution). Such post-perceptual interference would compound any perceptual (i.e., attentional) inter-

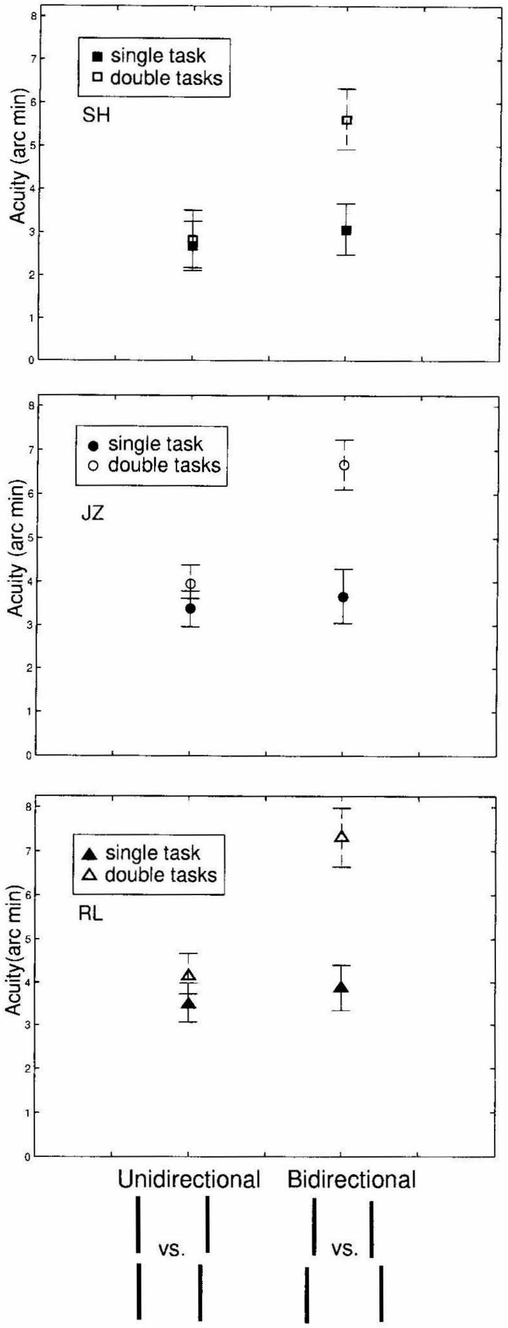


Fig. 4.4 Uni- and bidirectional vernier thresholds in the full presence and near absence of attention. Threshold vernier offset is plotted for unidirectional Vernier discrimination (observers report "offset" or "no offset") and for bidirectional Vernier discrimination (observers report "left offset" or "right offset"). Thresholds with attention fully available ("single task") are plotted with solid symbols and thresholds with attention nearly absent ("double task") are plotted with open symbols. (a),(b),(c), Data for three observers.

ference (Allport, 1980; Duncan, 1980; Pashler, 1991). Accordingly, the small elevation of orientation and bidirectional Vernier thresholds by the concurrent task may well have been due to post-perceptual rather than perceptual interference.

What could account for the differential dependence on attention of the investigated spatial vision thresholds? The conventional view is that "discrimination" judgments depend on attention to a greater extent than "detection" judgments (e.g., Bashinski and Bacharach, 1980; Shaw, 1984; Mueller and Findlay, 1987; Downing, 1988; Bonnel and Miller, 1994; Bonnel, Stein, and Bertucci, 1994). The intuition behind this distinction is that the perceptual distinction between larger or smaller sensory signals ("detection") poses a much simpler problem for the visual system than the distinction between sensory signals that are equally large but differ in qualitative ways ("discriminination").

Indeed, our results on sine wave gratings can be understood in terms of this distinction. Contrast sensitivity, which exhibits little or no dependence on attention, almost certainly represents a "detection" threshold. Even though observers reported grating orientation (vertical or horizontal) the most demanding aspect of the task was probably not the discrimination of its orientation but the detection of the grating at its varying peripheral location. Thus, performance was determined primarily by the ability to distinguish between larger and smaller sensory signals (grating location and empty locations, respectively). On the other hand, the visual differentiation of grating orientation (vertical or tilted), which exhibits a pronounced dependence on attention, is almost certainly a "discrimination": assuming that gratings of all orientations elicit a response of comparable size, this differentiation concerns responses that differ qualitatively rather than quantitatively.

A similar account can be given for our results on Vernier targets, provided we make certain assumptions about the mechanisms that underlie Vernier acuity judgments (Wilson, 1991; Harris and Fahle, 1995). It is thought that certain visual filters exhibit spatial response properties that can be loosely described as "end-stopped", and that these filters detect stimulus configurations such as the abrupt line termination in a Vernier target (Rosenthaler, Kubler, Heitger, Von der Heydt, 1992; Heitger, Rosenthaler, Von der Heydt, Peterhans, Kubler, 1992). End-stopped filters would signal the presence, but not the direction, of a Vernier offset and could therefore mediate unidirectional but not bidirectional Vernier discrimination. Because end-stopped filters would operate in parallel across the field of view, Vernier discrimination is especially likely to be based on such filters when there is spatial uncertainty about the position and orientation of the target (as in our case). In other situations, for example when the position and orientation of the Vernier target is known, *other mechanisms* seem to come into play (Wilson, 1991; Waugh, Levi, and Carney, 1993; Harris and Fahle, 1995; see below).

If it is true that unidirectional Vernier thresholds reflect the differential response of end-stopped filters to targets with and without an offset, then this differentiation would represent a "detection". This would account for the fact that reduced attention leaves undirectional thresholds almost unchanged and that visual search for a target with offset is independent of the number of targets in the display ("parallel search", Fahle, 1991; see also Wolfe, Yee, and Friedmann-Hill, 1992). Bidirectional Vernier thresholds, however, would have to be based on the responses of additional mechanisms, and would thus represent a "discrimination". This would account for the fact that reduced attention raises bidirectional thresholds, and that visual search for an offset in a particular direction requires more time when there are more targets in the display ("serial search"; Fahle, 1991).

Although these considerations show that the observed attentional demands are consistent with a detection/discrimination account, they also expose the essential weakness of this account: the detection/discrimination distinction depends on which mechanisms are presumed to underlie visual performance and, since these are generally unknown, is of limited predictive value. For example, it has also been proposed that Vernier offset judgments are based on visual filters sensitive to a range of orientations, especially orientations at 15° to either side of the axis of the Vernier target (Wilson, 1986, 1991; Waugh, Levi, and Carney, 1993). If this were the case, then left offset, no offset, and right offset elicit visual responses of comparable strength and any differentiation between these alternatives would be a "discrimination". Indeed, there is evidence that a mechanism of this type is sometimes used: in some situations, left and right offset are approximately twice as discriminable than presence or absence of offset (Harris and Fahle, 1995). This illustrates how ignorance of the mechanisms that underlie a visual judgment blurs the detection/discrimination distinction even in exhaustively researched instances such as Vernier offset judgements.

The deeper question would seem to be in what way attention alters the distribution of responses across visual filters. Does attention selectively enhance or attenuate the responses of individual filters? Or does attention simply strengthen or weaken certain interactions between filters, for example, the inhibitory interactions between filters at the same visual location suggested by Heeger and others (Heeger, 1993; Carandini and Heeger, 1993), or the competitive interactions between filters at distant visual locations postulated by Koch, Desimone and others (Koch and Ullman, 1985; Desimone and Duncan, 1995; Niebur and Koch, 1996)? Or perhaps attention simply attenuates responses outside an attended area that is defined in anatomical terms, for example, the area covered by the receptive fields of a certain number of hypercolumns (Hubel and Wiesel, 1977; Blasdel, 1992)? We will return to this question in section 4.2.4.

Invasive studies of non-human primates would seem to have contributed relatively little to the resolution of these issues. In visual cortical areas V2 and V4, it has been reported that attention sharpens orientation tuning of neurons (Spitzer, Desimone, Moran, 1987), that orientation tuning remains unchanged but response levels increase (McAdams and Maunsell, 1996), and that response levels remain roughly the same for attended stimuli but decline for unattended stimuli (Moran and Desimone, 1985; Reynolds, Chelazzi, Luck, Desimone, 1994). In visual cortical area V1, where receptive field properties correspond most closely to psychophysically defined visual spatial filters, attentional effects are rather difficult to observe but are consistent with a suppression of responses to unattended stimuli (Motter, 1993; Press, Knierim, and Van Essen, 1994).

Thus it would appear that appropriately designed psychophysical paradigms coupled with quantitative modeling remain the most promising approach to understanding attention and its effect on early levels of visual processing. We will describe such an approach now.

4.2 Spatial vision thresholds: a closer look

4.2.1 Background

In the previous section, we established that for some tasks, attention has little or no effect on performance (e.g., contrast detection), while for other tasks, attention dramatically improves performance (e.g., orientation discrimination). While a simple distinction of detection vs. discrimination task offers a straightforward explanation, it usually depends on which mechanisms are presumed to perform the task. Since the mechanisms are generally unknown, and many tasks may fall somewhere on the continuum between the two extremes instead of being classified as detection or discrimination task, such dichotomy offers little predictive value. Is there some general characteristics of visual attention mechanisms that underlie basic spatial vision which can explain the differential attentional effects we observed? In order to answer this question, we carried out further experiments as well as some modeling work (which forms a significant part of the forthcoming thesis by my colleage Laurent Itti and is briefly discussed in the Discussion).

Here we report how attention alters thresholds for discriminating contrast, orientation, and spatial frequency of simple patterns. Based on results from the last section, we expect markedly different effects on different threholds. In addition, we describe how attention changes thresholds for detecting target pattern in the presence of another, superimposed mask pattern of different orientation or spatial frequency. By using the color bit-stealing technique introduced by Tyler (Tyler, 1997), we are able to reveal some small but significant effects.

4.2.2 Methods

Stimuli were generated on a Silicon Graphics Indigo with a 1280×1024 pixels color monitor. Viewing was binocular at $\approx$ 120 cm distance, such that $1^\circ \approx 80$ pixels. Room luminance was $3cd/m^2$, average screen luminance was $30cd/m^2$.

We use a new technique called "color bit-stealing" to achieve better grey scale resolution. The human eye can resolve grey scales with an accuracy of about 0.2 percent, which is much finer than the grey-scale precision of the 8-bit monitor we use for our experiments (about 1.5 percent when phosphor nonlinearities are taken into account). In order to overcome the equipment limitation, we use the color bit-stealing technique introduced by Tyler (Tyler, 1997). The technique "steals" bits from the color variation to increase the precision of the *luminance* variation in each pixel. The 8-bit display has only 256 levels of luminance at a constant color ratio. By jittering the three color guns independently within 1 bit, we achieve a much finer grey scale (about 0.3 percent) at the cost of an undetectable color variation caused by small change in color ratios. This technique enables us to measure some subtle effects, such as the well known detection facilitation at the sub-threshold level.

We measure attentional modulation of spatial vision thresholds using double-task paradigm (Sperling and Melchner, 1978; Lee *et al.*, 1997). The stimuli consist of central and periphery components which represent two independent visual task. The trial sequence is shown in **Fig. 4.5**. Central targets appeared at $0 - 0.8°$ eccentricity and measured $0.4°$ across. Peripheral targets appeared at $4°$ eccentricity, in a circular aperture of $1.5°$. For orientation and spatial frequency discrimination, the peripheral targets were sinusoidal gratings (**Fig. 4.7** and **4.8**). For masking experiments and contrast incremental threshold measurement, the targets were vertical stripes whose luminance profile was given by the 6th derivative of a Gaussian and mask patterns were generated by superimposing 100 Gabor filters, positioned randomly within the circular aperture (stimuli shown in . **Fig. 4.6**, **4.9** and **4.10**). When the spatial frequency was not varied, it was *4cpd*. The mask contrast was 0.5 when it was constant (**Fig. 4.9** and **4.10**). Thresholds were established with an adaptive staircase method (80 trials per block), *i.e.*, by adjusting target contrast, orientation, or spatial frequency in each trial according to the success or failure of previous trials (Watson and Pelli, 1983). In the double-task situation, observers were required to match or exceed a certain level of central performance (the level achieved when the central task is carried out alone). Approximately 15% of double-task blocks were discarded because of poor central performance. In both single- and double-task situations, observers fixated the display center, ensuring identical visual stimulation. The brief presentation effectively precluded shifts of fixation towards the peripheral target (Fischer *et al.*, 1993).

An important concern in concurrent task experiments is the level of processing at which

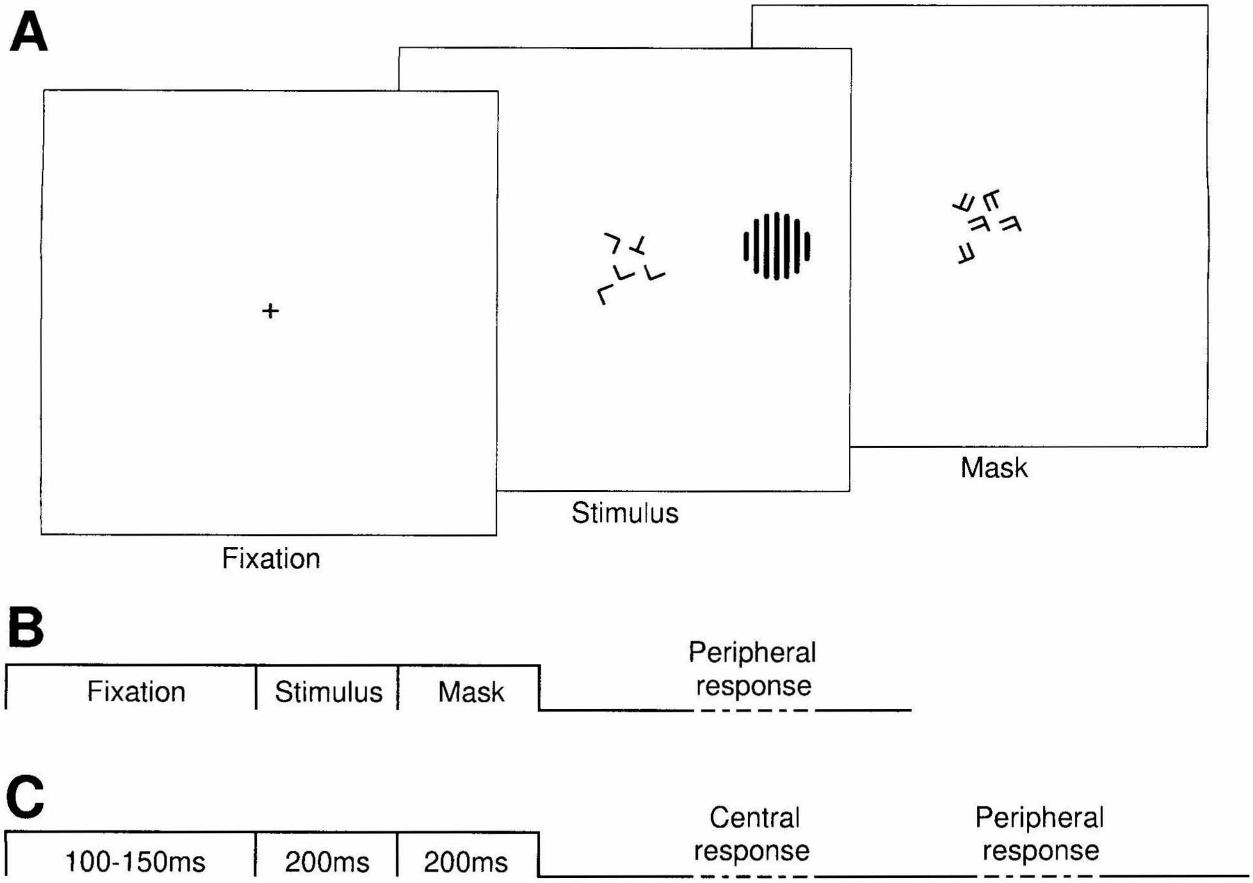


Fig. 4.5 Measurement of visual thresholds with either full or poor attention. **(A)** Sequence of fixation, stimulus, and mask displays (schematic). Observers fixate the center of all displays. The stimulus comprises a central and a peripheral component, which appear at varying locations of constant eccentricity. The central component consists of 5 Ts and/or Ls ("central targets") and observers report "same" (*i.e.*, 5 Ts or 5 Ls) or "different" (*i.e.*, 4 Ts + 1 L or 4 Ls + 1 T). The peripheral component consists of the luminance-modulated patterns shown in Fig. 2A-E ("peripheral target"). For example, the peripheral component might be a grating pattern of vertical or tilted orientation, in which case observers would report "vertical" or "tilted." The mask display limits visual persistence of central targets. **(B)** Single-task (peripheral target 'fully attended'): observers fixate the center but respond only to the peripheral task (see Fig. 2). **(C)** Double-task (peripheral target 'poorly attended'): observers fixate the center and respond first to the central task and second to the peripheral task.

the two tasks interfere. In general, interference can reflect limitations of attention, memory, and/or response generation (Duncan, 1984; Pashler, 1994; Braun, 1998). In the present experiments, observers have ample time to respond to each task in turn, so that response generation can be excluded as a limiting step. Limitations of memory are also unlikely, since interference between the kind of tasks used here disappears when central and peripheral targets are presented successively (e.g., with an onset asynchrony of $200ms$ or more) (Duncan *et al.*, 1994; Braun and Julesz, 1998; Joseph *et al.*, 1997; Braun, 1998). Further evidence that the limiting factor is attention is that the degree of interference is independent of the nature of the central task (as long as this task is sufficiently demanding to engage attention). For example, comparable interference results from central tasks based on form, color, and motion discrimination (Lee *et al.*, 1999), as well as from an "attentional blink" (Braun, 1998).

4.2.3 Results

Exp. 1: Incremental contrast threshold

Observers report the presence or absence of a vertical target stripe from a circular mask pattern. The profile of the central target strip is defined by 6th derivative of gaussian, which has a narrow distribution in frequency domain (Wilson, 1980). The major spatial frequency of the strip is 4cpd. The circular mask pattern is superposition of one hundred randomly positioned gabor patches with 4 cpd spatial frequency and vertical orientation. The stochastic nature of the mask makes sure that there is no specific feature resulted from the superposition of the target and mask pattern. The contrast of the mask pattern is varied between 0% and 50%.

Two alternative stimuli and data for two subjects are shown in **Fig. 4.6**. When peripheral targets are fully attended, contrast detection thresholds (zero mask contrast) are about 20% lower and contrast discrimination thresholds (mask contrast between 1% and 50%) about 40-50% lower than when peripheral targets are poorly attended. In addition, the decrease of the discrimination threshold as mask contrast increase from zero to subthreshold level (the well known "dipper") is only evident when targets are fully attended.

Note that positional uncertainty is known to reduce the dipper (Palmer, 1995; Solomon *et al.*, 1997; Foley *et al.*, 1998). Since the target location varies from trial to trial, our data may underestimate the depth of the dipper.

Exp. 2: Orientation discrimination

A sinusoidal wave grating of 4 cpd and diameter 1.5° is presented randomly at one of 8 possible locations at 4° of eccentricity. The presentation time is around 200 ms. The grating is either vertical or slightly tilted to the right. Following the stimulus, two alternative gratings appear and subject reports which one has appeared in the stimulus by mouse clicking. Staircase procedure is used to find the minimum tilt angle the subjects can discriminate. Thresholds are measured at five levels of luminance contrast, 5%, 10%, 20%, 40% and 80%.

Fig. 4.7 shows the two alternative stimuli and results of two subjects. Thresholds for orientation discrimination are about 70% lower when the periphery stimulus is fully attended. Thresholds at 5% and 10% contrast are greater because the stimulus is faint and hard to perceive, the error bars are much bigger as staircase does not converge so well because of the noisy response. Consistent with the results from the last section, thresholds remain essentially constant for constrast values above 20%.

Exp. 3: Spatial frequency discrimination

A sinusoidal wave grating of diameter 1.5° is presented randomly at one of 8 possible locations at 4° of eccentricity. The presentation time is around 200 ms. The grating is vertical and is either of 4 cpd or slightly higher spatial frequency. Following the stimulus, two alternative gratings appear and the subject reports which one has appeared in the stimulus by clicking the mouse. Staircase procedure is used to find the minimum spatial frequency difference the subjects can discriminate. Thresholds are measured at five levels of luminance contrast, 5%, 10%, 20%, 40% and 80%.

Fig. 4.8 shows the two alternative stimulus and results of two subjects. Thresholds for spatial frequency discrimination are about 60% lower when periphery stimulus is fully attended. Again, thresholds at 5% and 10% contrast are greater and have big error bars

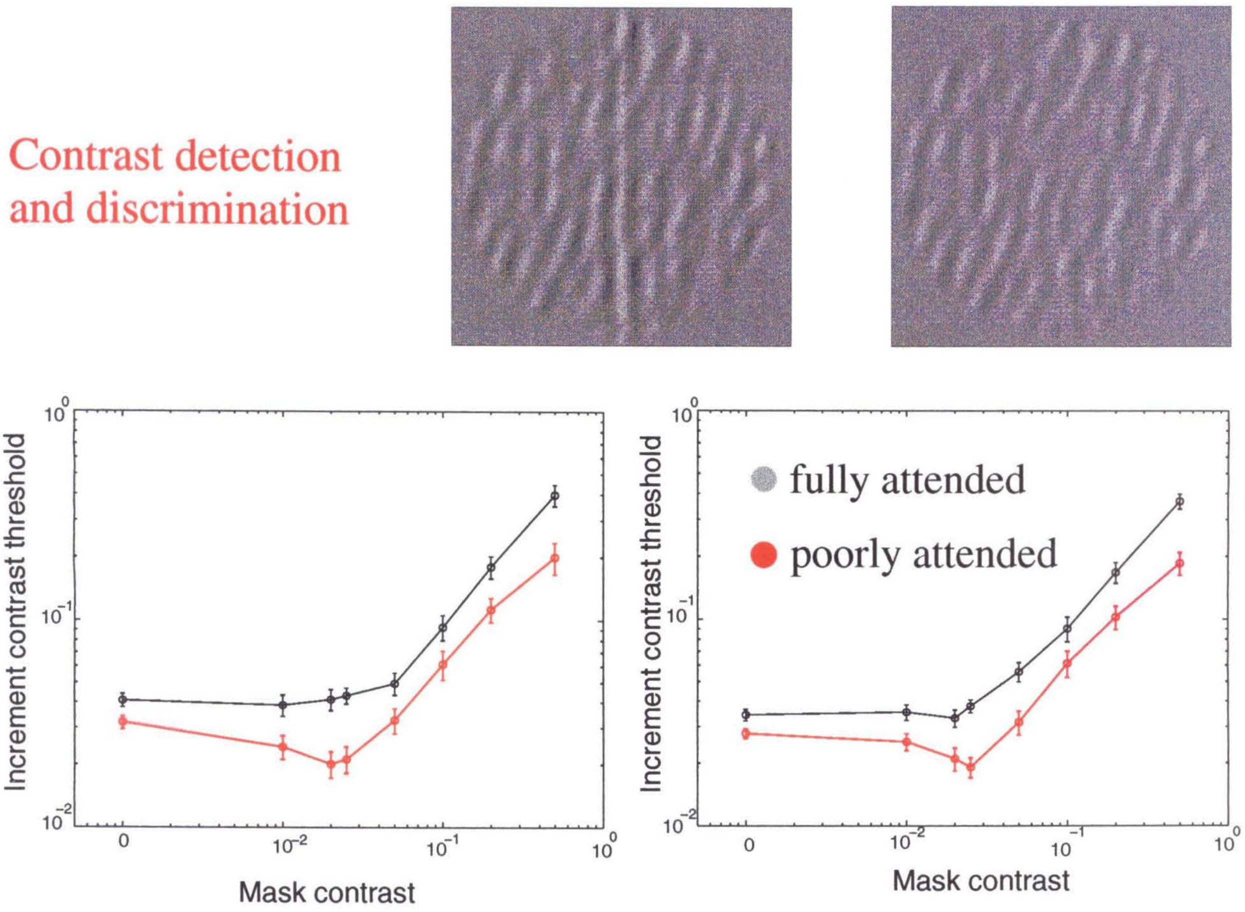


Fig. 4.6 Exp. 1: contrast detection and discrimination thresholds. Red and black data points represent fully attended (single-task) and poorly attended (double-task) thresholds. Observers report the presence or absence of a vertical target stripe from a circular masking pattern (contrast range 0.0-0.5).

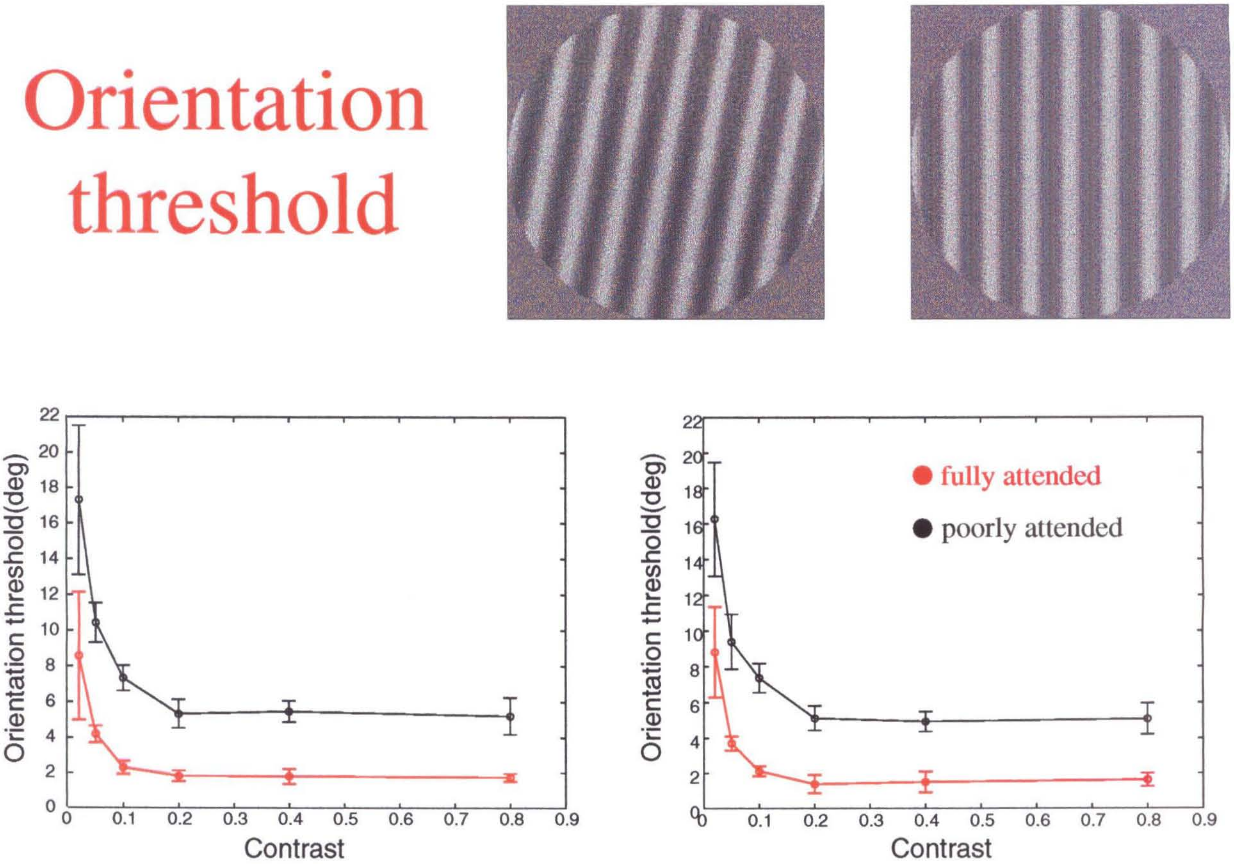


Fig. 4.7 Exp. 2: orientation discrimination. Observers report whether a circular target grating is vertical or tilted to the right. Contrast ranges from 2% to 80%. As in **Fig. 4.6**, red dots represent thresholds for fully-attended stimuli while black dots correspond to poorly-attended stimuli.

because the stimulus is faint and hard to perceive. Thresholds remain essentially constant for constrast values above 20%.

Exp. 4: Orientation masking

The target and mask pattern are the same as described in Exp. 1 except that the contrast of the mask pattern is fixed at 50% and its major orientation is varied between 0° and 90° from the vertical orientation. The subject's task is to discriminate whether the target is present or not.

Fig. 4.9 shows the two alternative stimuli and data for two subjects. When the major orientation of the mask is vertical (target orientation), the threshold elevation is biggest and attention lowers the maximum threshold by about 50% (consistent with **Fig. 4.6**, mask contrast 50%). As target and mask orientation become progressively different, fully and poorly attended thresholds decrease towards the similar baseline level. The baseline is comparable to thresholds without mask (**Fig. 4.6**, mask contrast 0.0), indicating minimal interaction between targets and masks of very different orientation.

Exp. 5: Spatial frequency masking

The target and mask pattern are the same as described in Exp. 1 except that the contrast of the mask pattern is fixed at 50% and its major spatial frequency is varied between 2 cpd and 8 cpd. The major orientation component of the mask is 15 degrees tilted to the right. The subject's task is to discriminate whether the target is present or not.

Fig. 4.10 shows the two alternative stimuli and data for two subjects. When target and mask pattern have similar major spatial frequency (4cpd), the threshold elevation is biggest and attention lowers the maximum threshold by about 50% (consistent with **Fig. 4.6**, mask contrast 50%). As target and mask spatial frequency become progressively different, fully and poorly attended thresholds decrease towards the similar baseline level. The baseline is comparable to thresholds without mask (**Fig. 4.6**, mask contrast 0.0), indicating minimal interaction between targets and masks of very different spatial frequency.

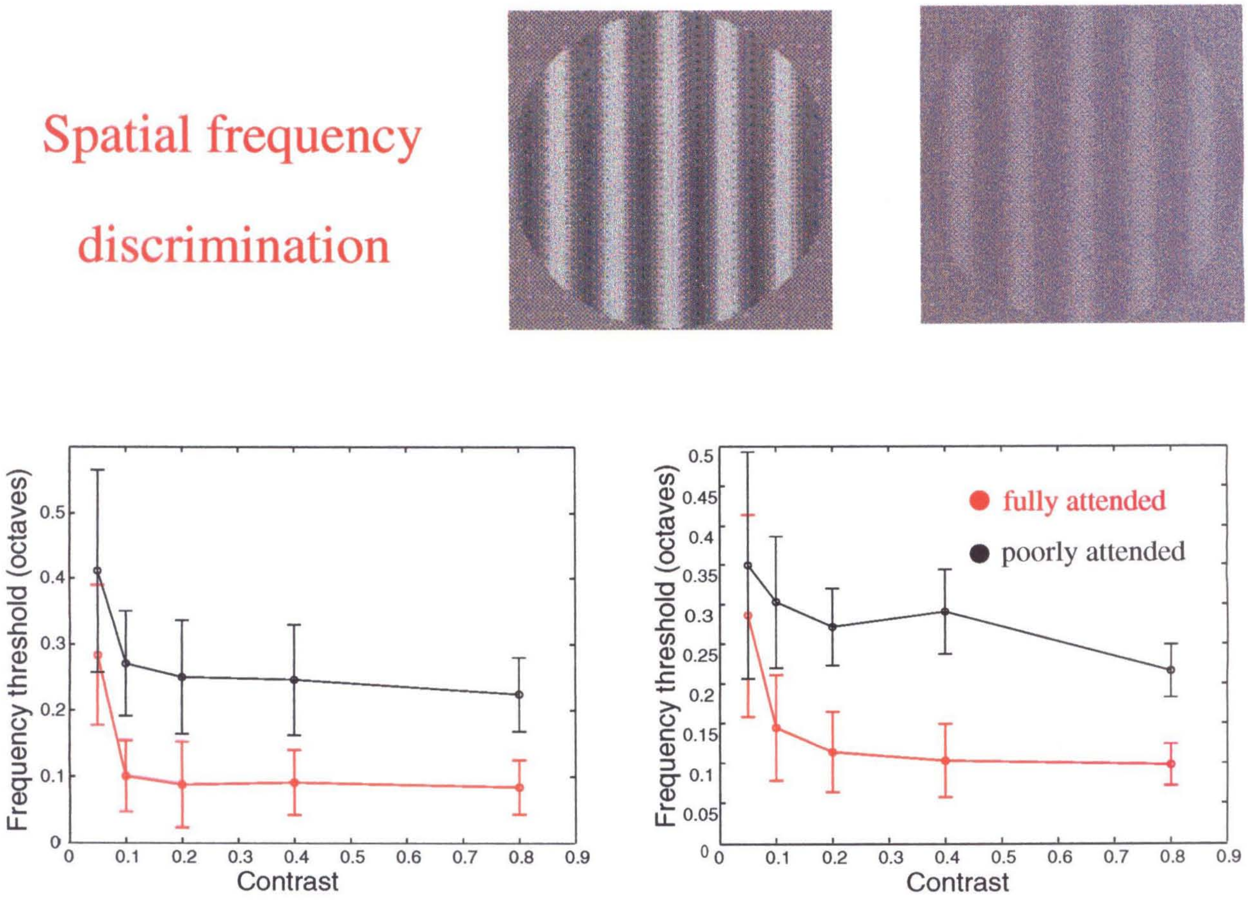


Fig. 4.8 Exp. 3: spatial frequency discrimination. Obersers report whether a circular target grating exhibits higher or lower spatial frequency. As in **Fig. 4.7**, contrast ranges from 2% to 80%. Red and black sympols represent single- and double-task thresholds.

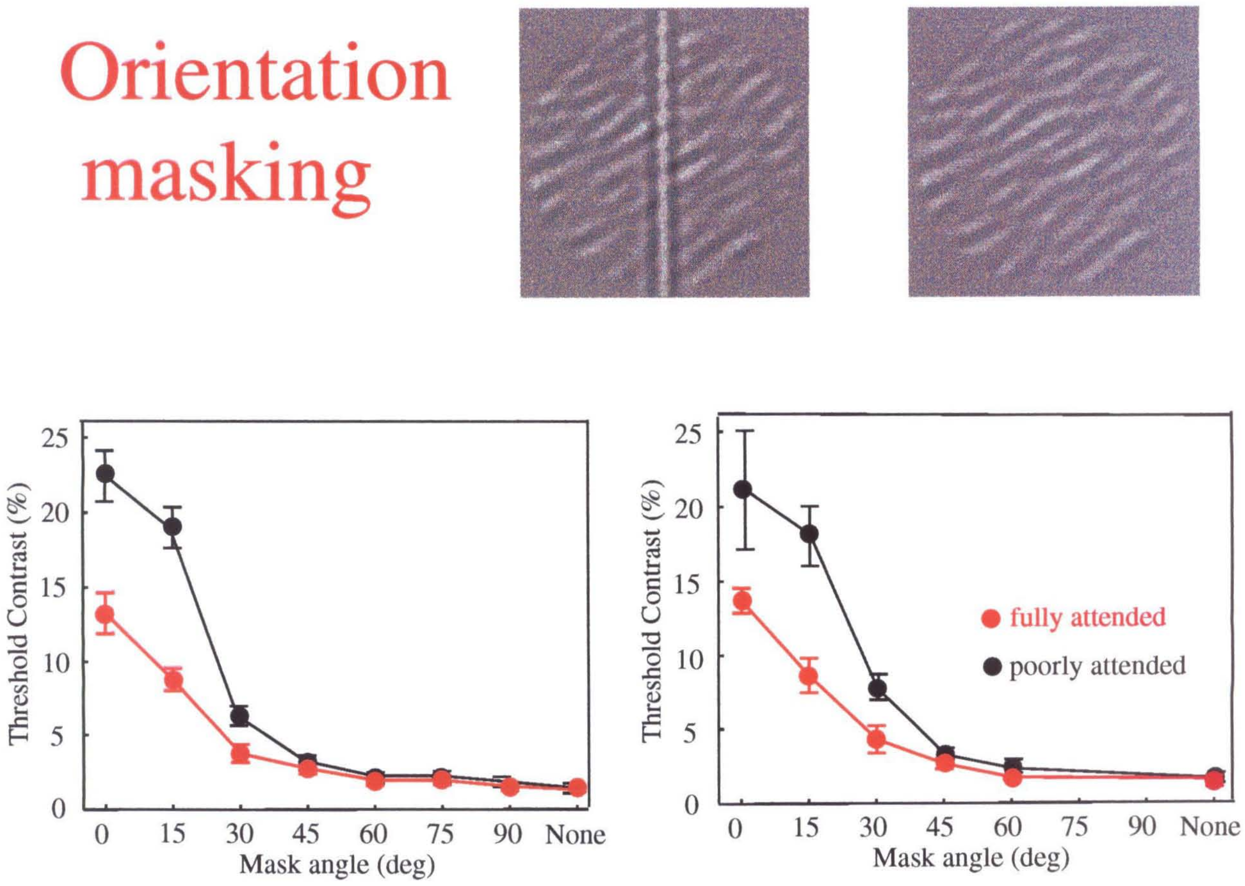


Fig. 4.9 Exp. 4: orientation masking. Observers report the presence or absence of a vertical target strip from circular masking pattern (which contrast is 50%) of different orientations. The orientation difference between target and mask ranges from 0 degree to 90 degrees. The data points labeled with "none" are detection thresholds without any masking. Red and black dots correspond to single and double task conditions respectively.

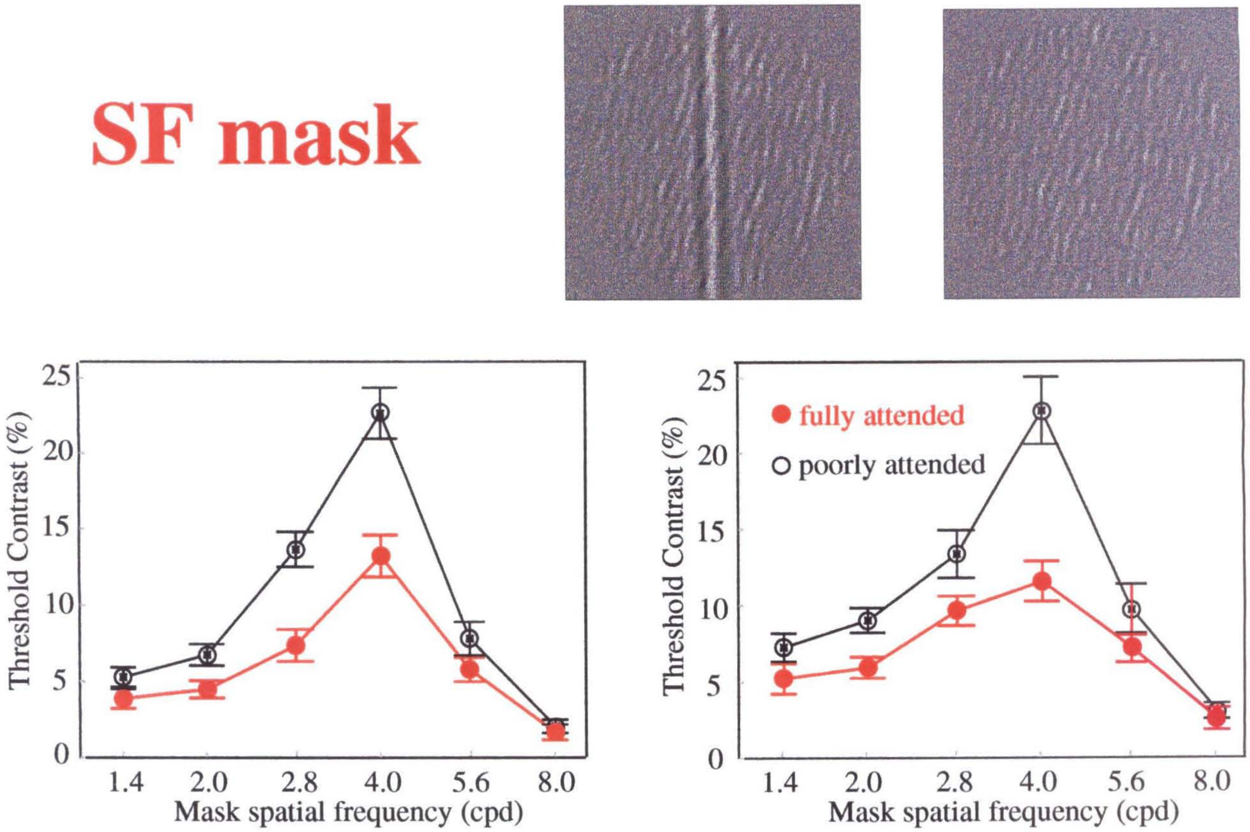


Fig. 4.10 Exp. 5: spatial frequency masking. Observers report the presence or absence of a vertical target strip from circular masking pattern (which contrast is 50%) of different spatial frequencies. The major orientation component of the mask is 15 degrees tilted to the right. The spatial frequency difference ranges from -1 to +1 octave. Red and black dots correspond to fully and poorly attended stimuli respectively.

4.2.4 Discussion

We measure thresholds for discriminating the contrast, orientation, and spatial frequency of simple patterns that are either fully or poorly attended. For fully attended stimuli, we observe improvement of about 20% in contrast detection thresholds, 40-50% in contrast discrimination thresholds (and appearance of the dipper), 60-70% in orientation and spatial frequency discrimination thresholds, and up to 50% in orientation and spatial frequency masking thresholds. Instead of the simple dichotomy we observed in the study presented in last section–small or no attentional effect for detection tasks, large attentional effects on discrimination tasks, we here see a continuum of attentional modulation.

Previous studies of attentional changes in visual thresholds are broadly consistent with our results, even though our effects are larger. This includes reports that attention reduces contrast thresholds by 17% (Lu and Dosher, 1998), orientation acuity by 15% (for an individual target without distractors) (Morgan *et al.*, 1998), and size acuity by 20% (Yeshurun and Carrasco, 1999). However, the fact that these studies manipulated attention with a spatial cue rather than with a concurrent task complicates quantitative comparison. We believe that a concurrent tasks detains attention more consistently than spatial cuing; certainly concurrent tasks induce substantially larger changes in thresholds.

These observations have been paralleled by electrophysiological studies of attention. In the awake macaque, neuronal responses to attended stimuli can be 20% to 100% higher than responses to otherwise identical unattended stimuli. This has been demonstrated in visual cortical areas V1, V2, and V4 (Spitzer *et al.*, 1988; Motter, 1993; Maunsell, 1995; Luck *et al.*, 1997) when the animal discriminates stimulus orientation, and in areas MT and MST when the animal discriminates the speed of stimulus motion (Treue and Maunsell, 1996). Even spontaneous firing rates are 40% larger when attention is directed at a neuron's receptive field (Luck *et al.*, 1997). Whether neuronal responses to attended stimuli are merely enhanced (Treue and Maunsell, 1996) or whether they are also more sharply tuned for certain stimulus dimensions (Spitzer *et al.*, 1988) remains controversial. Very recently, fMRI studies have shown similar enhancement (as measured with BOLD contrast) in area V1 of humans, specifically at the retinotopic location where subjects had been instructed to focus their attention to (Gandhi, *et al.*, 1998; Somers, *et al.*, 1998).

All of these observations directly address the issue of the "top-down" computational effect of attentional focusing onto early visual processing stages. This issue should be distinguished from that of the "bottom-up" control of visual attention (Koch and Ullman, 1985), which studies which visual features are likely to attract the attention focusing mechanism (e.g., pop-out phenomena and studies of visual search). Top-down attentional modulation happens after attention has been focused to a location of the visual field, and most probably involve massive feedback circuits which atatomically project from higher cortical areas back to early visual processing areas.

With a closer look at our experimental data, we can draw some qualitative conclusions about how attention appears to alter the neuronal representation of contrast, orientation and spatial frequency. For example, attention evidently does not act primarily by reducing background noise as the data show that the smallest attentional effects are obtained for stimuli with the lowest contrast (**Fig. 4.6**). A mere reduction of background noise would produce largest attentional effects for stimuli with lowest contrast level, the exact opposite of experimental data. Another qualitative conclusion concerns the contrast gain of neural response. The data in **Fig. 4.6, 4.9 and 4.10** show that attention reduces by about 50% the threshold elevation caused by a superimposed mask. The easiest way to account for the reduced threshold elevation is to postulate higher contrast gain, since this would increase the incremental response obtained when the mask is added to the target. A strong qualitative conclusion can also be drawn about the orientation and spatial frequency tuning of neural responses. The data in **Fig. 4.7 and 4.8** show a substantial vertical shift in thresholds with attention. Since increased contrast gain can produce only a horizontal shift, it follows that attention sharpens the turning for orientation and spatial frequency. Finally, the data in **Fig. 4.9 and 4.10** suggest that the range of orientation and spatial frequency over which different tuned channels interact is fairly constant with and without attention. This is the implication of the fact that threshold elevation with full and poor attention retain the same relative size, no matter how similar or different the orientation and spatial frequency of the target and mask. If attention would produce a substantial change in the range of interactions, we would not expect this simple proportionality of the results.

With all these intuitions suggested by the data, the next question we ask is: is there a unifying mechanism that might underlie all of our obvervations? On the basis of the experimental data presented in this chapter, Laurent Itti, Jochen Braun and Christof Koch tried to quantify the modulatory effect of attention using a model of early visual processing. Although attention could modulate virtually any visual processing stage (e.g., the decision stage, which compares internal responses from different stimuli), our basic hypothesis here – supported by electrophysiology and fMRI study (Spitzer *et al.*, 1988; Motter, 1993; Maunsell, 1995; Treue and Maunsell, 1996; Luck *et al.*, 1997; Gandhi *et al.*, 1998; Somers *et al.*, 1998) – is that this modulation might happen very early in the visual processing hierarchy. Given this basic hypothesis, we investigated how attention should affect early visual processing in order to quantitatively reproduce the psychophysical results. The model is based on non-linearly interacting visual filters and statistically efficient decision (Itti *et al.*, 1997 NIPS; Itti *et al.*, 1998 NIPS; Lee and Itti *et al.*, 1999). It is similar to several other models (Wilson and Humanski, 1993; Foley, 1994; Zenger and Sagi, 1996; Carandini, *et al.*, 1997) and comprises three stages (**Fig. 4.11**): *(i)* a population of overlapping linear filters responsive to different orientations and spatial frequencies at one visual location, *(ii)* non-linear interactions amongst this population to carry out the normalization (recurrent excitation and inhibition), and *(iii)* an "ideal observer" decision which discriminates between stimulus alternatives on the basis of the maximum *a posteriori* likelihood and is limited only by noise.

When we fit this model (10 free parameters) separately to poorly attended thresholds and fully-attended thresholds, we obtain good fit between predicted and phychophysical data. In order to isolate influences of different parameters, we fit the model simultaneously to single- and double-task data, while allowing only δ and γ (two exponents which represent the power law of recurrent excitation and inhibition) to change depending on attention. This assumes attention only changes the second stage of the model: interaction between filters. We achieve acceptable fits with physiologically feasible parameters value with this restriction. On the contrary, if we allow all other parameters except δ and γ to change with and without attention, we cannot get any reasonable fit even with physiologically unrealistic values. Averaged data are shown in **Fig. 4.12** together with the fits. A thorough and detailed discussion about this can be found in Laurent Itti's thesis.

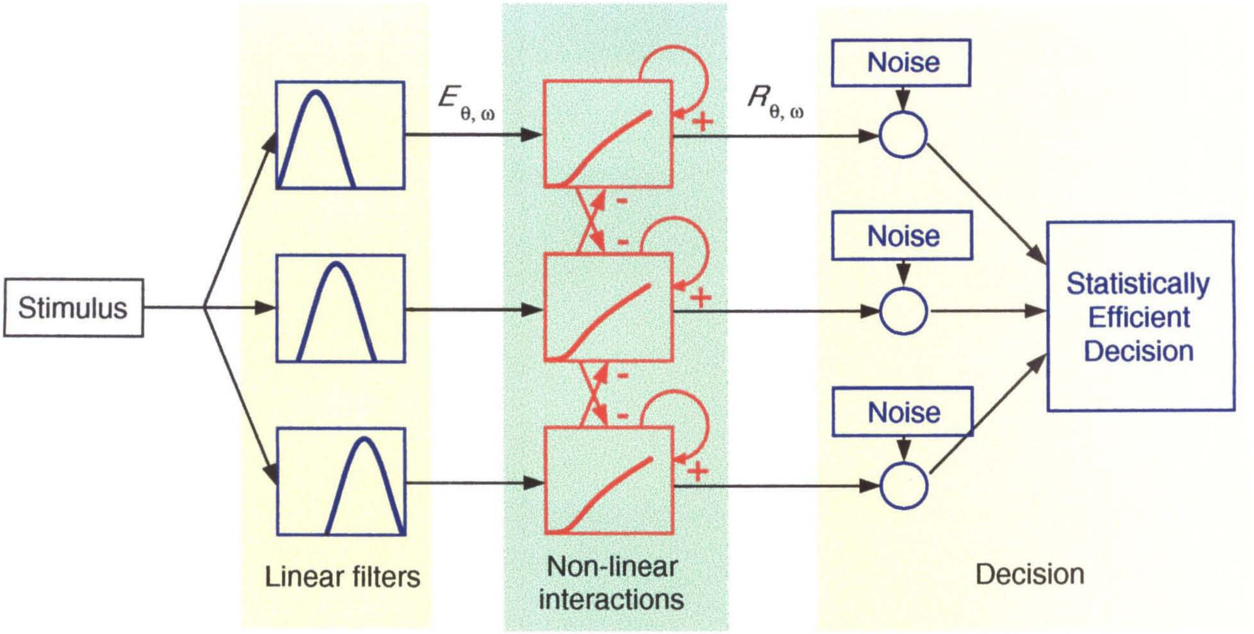
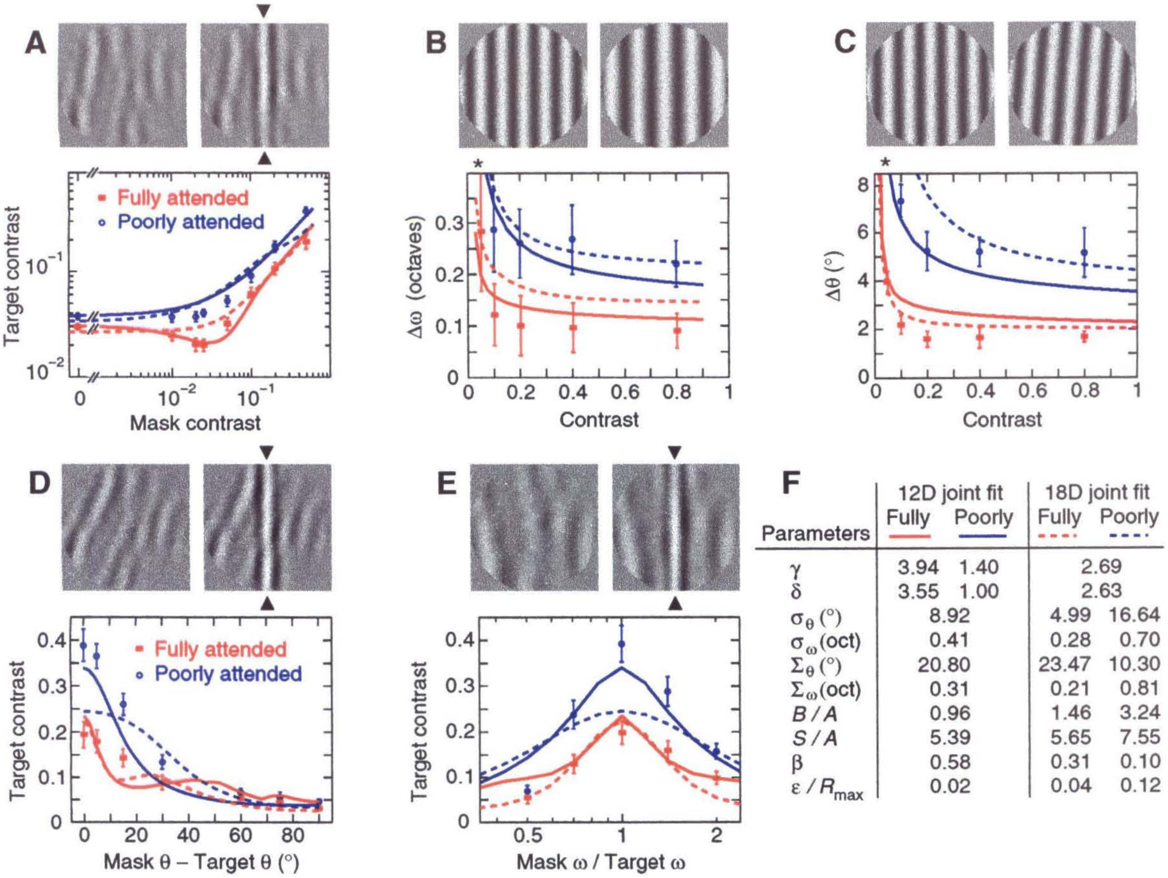


Fig. 4.11 Three-stage model of visual filters and their interactions (schematic). Each stimulus location is analyzed by linear filters sensitive to different orientations and spatial frequencies. Filter responses are subjected to excitatory and inhibitory interactions in the form of amplification and divisive normalization. The decision stage assumes that non-linear responses exhibit a variance similar to that of cortical neurons and chooses between stimulus alternatives on the basis of maximal likelihood. See details in Itti *et al.*, 1998 and Itti's thesis. The results suggest that attention strengthens non-linear interactions between filters (red), but does not affect other parts of the model (blue).



Parameters	12D joint fit Fully	12D joint fit Poorly	18D joint fit Fully	18D joint fit Poorly
γ	3.94	1.40	2.69	
δ	3.55	1.00	2.63	
σ_θ (°)	8.92		4.99	16.64
σ_ω(oct)	0.41		0.28	0.70
Σ_θ (°)	20.80		23.47	10.30
Σ_ω(oct)	0.31		0.21	0.81
B/A	0.96		1.46	3.24
S/A	5.39		5.65	7.55
β	0.58		0.31	0.10
ε/R_{max}	0.02		0.04	0.12

Fig. 4.12 Predicted thresholds when attention changes some model parameters but not others. Experimental data are averaged from individual data present in **Fig. 4.6** to **Fig. 4.10**. The solid curves represent a simultaneous fit to both single- and double-task data, in which only the exponents γ and δ take different values depending on attention (12 free parameters). Observed and predicted thresholds agree reasonably well, and parameters values are physiologically plausible (two leftmost columns in F). The dashed curves represent the optimal joint fit when all parameters *except* the exponents γ, δ take different values depending on attention (18 free parameters). Neither the dipper (A) nor the maximal extent of contrast masking (D,E) are predicted, and parameter values are unrealistic (two rightmost columns in F). See details in Laurent Itti's thesis.

In short, the attentional modulation we observed can be explained in a quantitative manner by changing the exponents of the power law, but not by changing other parameters of the model. Response distribution change caused by this power law change is shown in **Fig. 4.13**. Filters that respond relatively well to a given stimulus are enhanced by attention, while filters that respond relatively poorly are supressed. By strengthening a winner-take-all competition among overlapping visual filters, attention restricts responses to the filters best tuned to the stimulus.

Finally, we do not want to claim that attentional effect is only restricted to local interactions at early visual processing. More than likely, attention acts at all levels of visual cortex and acts on both local and long-range interactions. Nevertheless, our results show that the activation of winner-take-all competition among local filters can explain many basic attentional effects on spatial vision. The details of the modeling work will be discussed in Laurent Itti's forthcoming thesis.

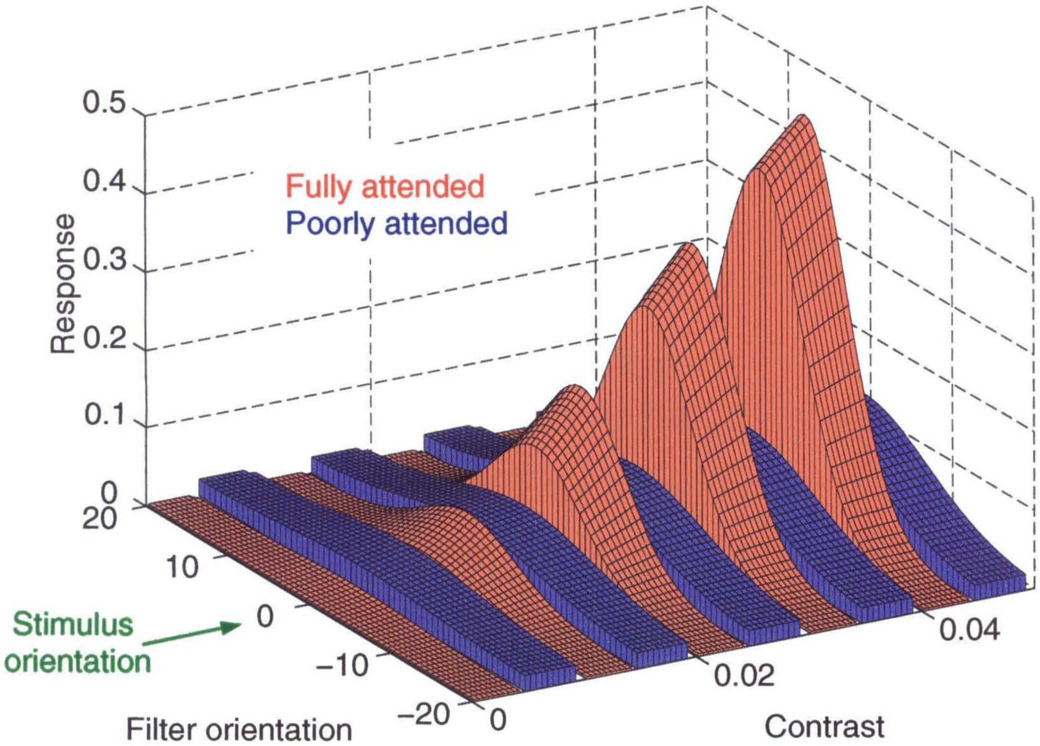


Fig. 4.13 Attentional change in the response distribution. Predictions based on 12-dimensional joint fit in **Fig. 4.12**. Responses $R_{\theta,\omega}$ of filters tuned to orientations between $-20°$ to $+20°$ to a grating stimulus of orientation 0° and contrast between 0 to 0.05 (threshold regime). Responses to fully and poorly attended stimuli are represented by the red and blue surfaces, respectively (shown interleaved for clarity). By strengthening a winner-take-all competition amongst visual filters, attention restricts responses to the filters tuned best to the stimulus at hand.

Chapter 5

Preliminary experiment about attentional effect on motion perception

[1]

5.1 Background

The influence of attention on motion perception is poorly understood. It has been reported that attention influences motion aftereffects (Chauhuri, 1992; Lankheet and Verstraten, 1995) and image segmentation by color enhances discrimination of motion in visual noise (Croner and Albright, 1997). Previous studies also have demonstrated attentional modulation of MT neurons: some motion-sensitive neurons increase their firing rates by 50% during a visual search task (Buracas and Albright, 1995), and by 100% during a speed and direction discrimination task (Treue and Maunsell, 1996; Thiele and Hofmann, 1996). However, it is less clear how focal attention affects motion perception *per se*.

In order to investigate the interaction of attention and motion perception, we use a visual stimulus and experimental paradigm used widely in recent studies of motion processing

---

[1] This preliminary work is done with collaboration of Dr. Giedrius Buracas at Salk Institute. Buracas is now continuing this work.

(Willams and Sekuler, 1984; Newsome and Pare, 1988; Downing and Movshon, 1989; Newsome *et al.*, 1989; Britten *et al.*, 1992, 1993). The stimulus consists of a random dynamic array of dots, a variable fraction of which move coherently from frame to frame and constitute a motion signal, while the rest of the dots move randomly and constitute motion noise. Previous studies of observers' ability to discriminate signal direction revealed a consistent relationship between motion signal strength (coherence) and performance (Downing and Movshon, 1989; Britten *et al.*, 1992, 1993), enabling the measurement of discrimination thresholds. To investigate attentional effect, we put the motion stimulus at periphery and introduce a RSVP (rapid serial visual presentation) task at the center to engage attention.

5.2 Methods

5.2.1 Visual stimuli

We used a RSVP (rapid serial visual presentation) task to engage attention. At the center of the display, successive letters were presented for a fixed duration without any blank interval between letters. The speed was 10-15 letters per second, adjusted for each subject so that the single task performance for the RSVP task was above 70%. There were 1 or 2 numbers randomly embedded in the letter stream during the 1 second presentation. Due to the well known transient blindness for the second target after the detection of first target, the distance of two successive numbers was at least 4 letters apart. After the presentation the subject had to report what number or numbers they had seen.

For the peripheral task, we used a dynamic dot stimulus in which dot positions were manipulated so that a motion signal of variable strength was embedded in dynamic noise (Newsome and Pare, 1988; Britten *et al.*, 1992, 1993). The motion signal was generated by randomly selecting a percentage of the dots ("signal dot") from each frame to be replotted at a location shifted 0.15 deg in a single direction after a delay of 50 msec (3 frames). Each dot was extinguished during the intervening frames, as in previous studies (Britten *et al.*, 1992), yielding apparent motion of 3 deg/sec. The remaining dots ("noise" dots) were replotted at random positions after the same temporal delay, yielding dynamic noise. The strength of the motion signal was controlled by changing the proportion of "signal" dots. The random dot pattern was 2.5 degrees in diameter, and appeared at 1 of 8 equally possible loctions at 3 degree eccentricity for one second.

5.2.2 Psychophysical tasks

We used three variants (RSVP task only, motion task only, both tasks) to measure the attentional modulation of motion perception. 3 subjects were tested.

RSVP task

A previous study(Chaudhuri, 1992) showed that RSVP task is attention demanding. The difficulty level of the task was adjusted by varying the duration of character presentation

(67-100 msec/character) so that the subject's performance was between 70-90% correct. This provided a baseline for estimating how the concurrent motion task affected performance of the RSVP task.

Motion discrimination task

Observers performed a 2-AFC task discriminating direction of motion. Thresholds were determined by a staircase procedure, whereby the threshold variable was increased by a fixed step after an error trial and decreased by the same fixed step after two consecutive correct trials. The thresholds were determined with and without concurrent RSVP task.

Dual task: Motion discrimination and RSVP

Attention was manipulated by requiring subjects to perform the motion discrimination task simultaneously with the demanding RSVP task. Subjects were asked to give first priority to the RSVP task and try to perform it as well as they were asked to perform RSVP task alone. This reassured that attention was nearly absent for the peripheral motion task.

5.3 Preliminary results

5.3.1 Exp. 1: discriminate up/down motion

In Exp. 1, the motion direction of “signal” dots was either upward or downward, coherence level was varied according staircase methods to measure detection thresholds. The threshold elevation due to withdrawal of attention is between 30% and 40%. Data for two subjects are shown in **Fig. 5.1**.

5.3.2 Exp. 2: discriminate motion direction

In Exp. 2, motion direction was tilted towards either right or left (deviation angle from the straight-up was the same) and the task was to report whether it was left or right tilted. Signal coherence was fixed during the block and direction discrimination thresholds were determined by a staircase procedure, where the deviation of motion direction from straight-up was increased by 2 degrees after an error trial and decreased by 2 degrees after two consecutive correct trials. Allocation of attention away from motion stimulus significantly elevated direction discrimination thresholds in most cases. The attention withdrawal effect was stronger for stimuli with lower signal coherence. For 30% coherence signal, thresholds decreased by an average of 6.1 degrees with attention; and for 15% coherence signal, they decreased by 10.6 degrees. Data for three subjects are shown in **Fig. 5.2**.

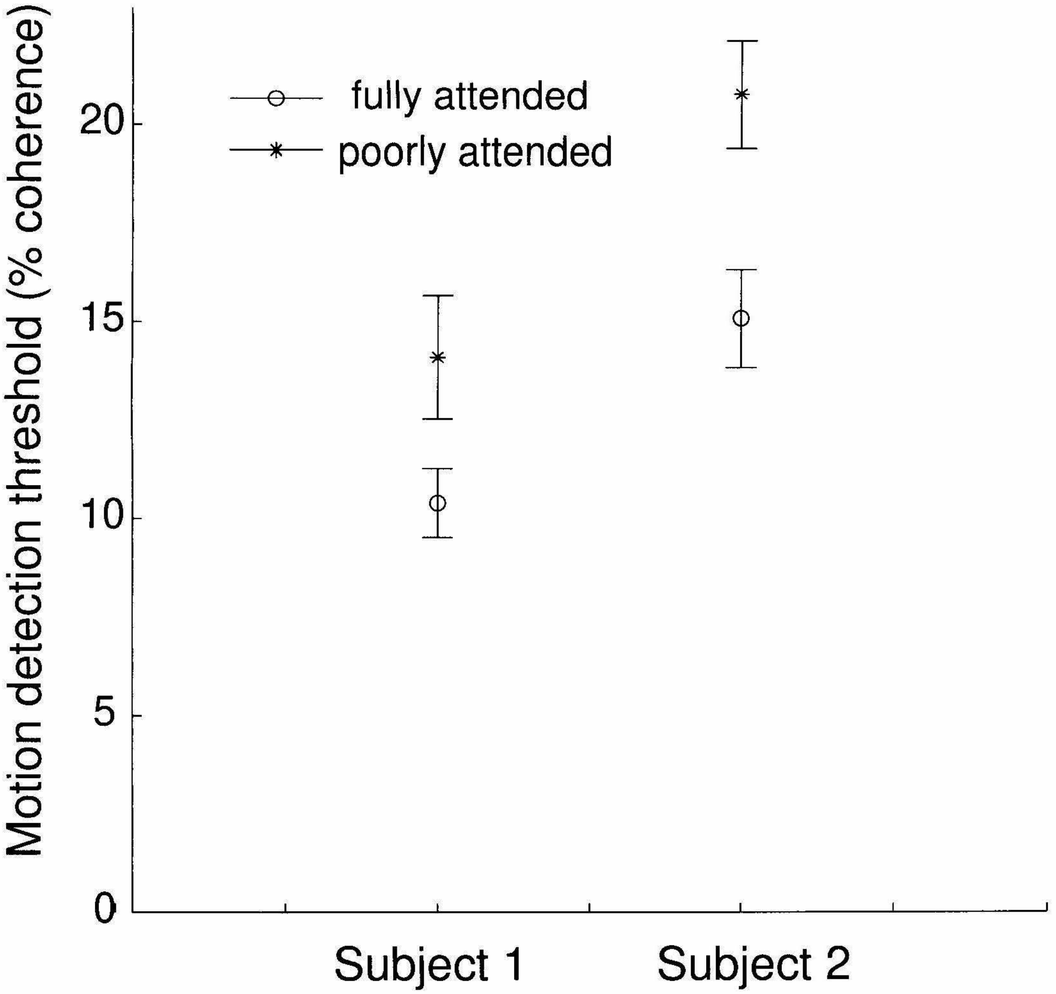


Fig. 5.1 Motion detection thresholds for two subjects. Open circles represent single-task thresholds. Star symbols represent double-task thresholds.

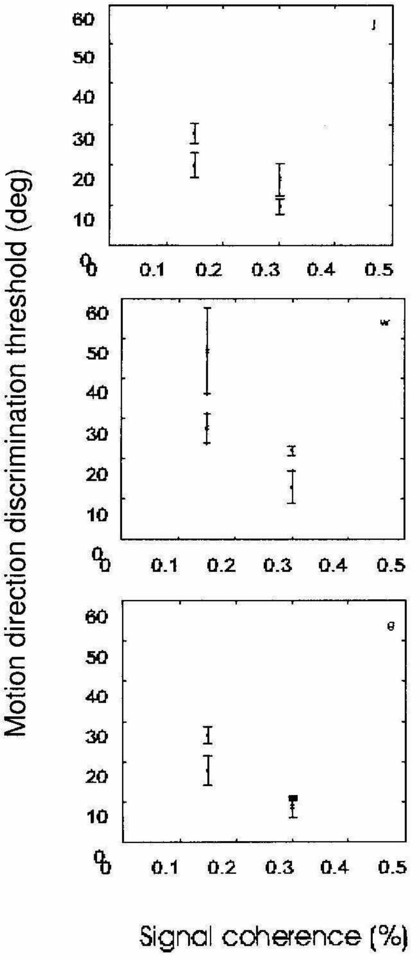


Fig. 5.2 Motion direction discrimination thresholds for three subjects. Coherence level was fixed at 15% and 30%.

5.4 Discussion

Our preliminary data shows that withdrawal of attention increases both detection and discrimination thresholds for motion perception. We would like to make a comparison between this data and data presented in chapter 4, which showed how attention changes detection and discrimination thresholds for spatial vision. In order to do that, we need to do more experiments and collect more data. Dr. Giedrius Buracas at Salk Institute is now continuing this work.

Another possible line of work is to investigate how attention changes motion thresholds for homochromatic and heterochromatic stimuli. Croner and Albright introduced a simple but critical change to random dot motion stimuli: making the signal and noise dots different colors. This manipulation profoundly enhanced human subject's ability to discriminate signal direction, decreasing thresholds by, on average, a factor of six (Croner and Albright, 1997). Further physiological study in macaques (Croner and Albright, 1999) suggested that heterochromatic motion signals are processed in a fundamentally different way than homochromatic motion signals. They proposed that color segmentation of the motion signal allows attention to be directed to signal dots, thus freeing the decision from concern with noise dots. By manipulating attention with double task paradigm, we can test their hypothesis and investigate possible attentional mechanisms behind homochromatic and heterochromatic motion perception.

Chapter 6

Summary

6.1 What facts have we learned?

Visual attention alters perception in several quantitative and qualitative ways. To characterize the different perceptual consequences of attention, we compare "attentive vision" with "ambient vision" that remains with poor attention. Ambient vision is surprisingly robust and supports many discrimination task performances far above chance. The results suggest that the richness of visual experience does not derive exclusively from a narrow focus of visual attention, but also reflects a simultaneous awareness of the entire visual space, based on the topographical activity maps of visual cortex. Many psychophysical studies show that focusing attention narrowly on one location reduces, but does not totally eliminate visual performance with respect to other objects in a visual scene. Thus it appears that observers enjoy a significant "residual" awareness of poorly attended stimuli, especially when these are salient and "pop out" from the scene. We propose the term "ambient vision" to describe this visual performance with respect to poorly attended but salient stimuli.

In order to study ambient vision, one has to induce observers to focus attention on one part of the display and thus at least partially withdraw attention from other parts of the display. Then the question arises that if one task is equally effective in engaging attention for all different kinds of tasks or the effectiveness depends on whether the two concurrent tasks involve same stimulus dimension or not. This is the question we try to answer in chapter 2 and chapter 3.

Both monkey physiology and human fuctional imaging studies show silimar functional specializations to stimulus dimensions such as form, color and motion in visual cortex. As the processing appears to concentrate in separate subsystems, it is quite natural that some people argue that attention is differentiated, two tasks compete more vigoriously when they involve the same stimulus dimension and have to compete for resource in the same subsystem (*e.g.*, Sperling & Dosher, 1986; Kinsbourne, 1981; Pashler, 1997). On the other hand, Duncan and colleagues stand on the contrary point of view: the different subsystems are so integrated that "as a winning object emerges in one [sub]system, it tends also to become dominant in others" (Duncan, Humphreys & Ward, 1997). In order to decide the matter, we select discrimination tasks involving form, color or motion and carry out experiments to make a quantitative comparison of similar and dissimilar task combinations.

In chapter 2, we design tasks with high attentional cost and use them to investigate whether the interference of two concurrent tasks depends on similar or dissimilar task combinations. By quantifying the attentional cost, our paradigm is sufficiently sensitive to reveal intermediate outcomes (*i.e.*, a partly differentiated capacity). However, the actual outcome shows that there is no significant difference between similar and dissimilar task combinations. All the trade-off curves we observe are strictly linear along the diagonal line. We thus conclude that similar and dissimilar tasks draw on exactly the same attentional capacity, in other words, that attentional capacity is entirely undifferentiated. In chapter 3, we continue a similar set of experiments with tasks of less attentional cost and observe intermediate performance trade off in both similar and dissimilar task combinations. By carefully quantifying the attentional cost, again, we find no significant difference between similar and dissimilar task combinations. This verify our finding in chapter 2: attention is indeed an undifferentiated resource. This is not specific for the high attentional cost tasks we use in chapter 2; instead, it is a more general rule.

It is worthy pointing out that when we talk about attentional cost of a certain task, it should not be confused with the difficulty level of the task. Contrary to some popular belief, the attentional cost of a task is not necessarily positively correlated with its difficulty level. Specifically in experiments reported in chapter 2 and 3, we adjust the presentation time (SOA) that the average single-task performance is around 85%. In this sense, all the tasks are equally difficult. Some of them are highly attentional demanding while some only have an intermediate attentional cost.

Having established that attention is a unitary process, we proceed to use concurrent task paradigm to investigate how attention affects many fundamental properties of spatial vision, such as contrast, orientation and spatial frequency. We are interested in spatial vision because it directly reflects properties and interactions of "filters" which corresponds to neurons in early visual processing. Spatial visual thresholds have been extensively studied for several decades (*e.g.*, Nachmias & Sansbury, 1974; Wilson, 1980; Legge & Foley, 1980) and lots of effort has been put into modeling work trying to account for these thresholds quantitatively. Our goal is to compare behavioral threshols for fully and poorly attended stimuli. Combined with modeling work, this comparison should reveal in detail how attention modulates the early stages of processing which underlie spatial vision.

We conduct five separate experiments to compare thresholds under single- and double-task conditions. When peripheral targets are fully attended, contrast detection thresholds (zero mask contrast) are about 20% lower and contrast discrimination thresholds (mask contrast greater than zero) about 40-50% lower than when peripheral targets are poorly attended. In addition, the well known "dipper" (the decrease of the discrimination threshold as mask contrast increases from zero to subthreshold level) dissappears when targets are poorly attended. The effects of attention on orientation and spatial frequency discrimination are even larger. Orientation thresholds are about 70% lower and spatial frequency thresholds are about 60% lower when peripheral targets are fully attended, compared to thresholds when they are poorly attended. Both types of thresholds remain essentially constant for contrast values above 20%. Interactions between superimposed stimuli of different orientation or spatial frequency are also modulated by attention. When target and mask have similar orientation or spatial frequency, attention lowers the maximal thresholds by about 50%. As target and mask become progressively more different, fully and poorly attended thresholds decrease towards the same baseline level. The baseline is comparable to thresholds without mask, indicating minimal interactions between targets and masks of very different orientation or spatial frequency.

The visual thresholds measured here are thought to reflect the activity of a population of "noisy filters" selective for stimuli of different orientation and spatial frequencies (Wilson, 1980; Legge & Foley, 1980; Watt & Morgan, 1985). In collaboration with Laurent

Itti and Jochen Braun, we develop a model to account for our observation quantitatively. The modeling work is mainly done by my colleague Laurent Itti and will be presented in detail in his forthcoming thesis. The critical finding is that a simple model of response normalization implemented by divisive inhibition accounts for all the attentional changes to contrast, orientation, and spatial frequency thresholds we have observed. In the model, attention activates a winner-take-all competition amongst overlapping visual filter, which accounts quantitatively for all attentional modulation we observe. The model predicts that the effects of attention on visual cortical neurons include increased contrast gain as well as sharper tuning to orientation and spatial frequency.

6.2 What do these facts mean?

Recent evidence leaves little doubt that visual attention modulates the activity of all levels of visual cortex, including primary visual cortex (area V1). In the visual cortex of human observers, functional imaging reveals a dramatic attentional modulation of hemodynamic activity, which is comparable in size to the effect of visual stimulation (Brefczynski & De Yoe, 1999). Attentional modulations are evident also at the level of single-unit responses, and have been reported in visual cortical areas V1, V2, V4, and MT/MST (Moran & Desimone, 1995; Motter, 1993; Treue & Maunsell, 1996; Luck, Chelazzi, Hillyard & Desimone, 1997; Roelfsema, Lamme & Spekreijse, 1998). Even at the single-unit level, the effect of attention can be quite large (*e.g.*, responses in area V1 may double, given an appropriate stimulus and task)(Ito & Gilbert, 1999).

The fact that attention acts at many cortical levels is consistent with the multiplicity of attentional modulations we observe psychophysically. Visual discrimination of elementary stimulus attributes such as contrast, orientation, spatial frequency, or color are performed readily even in unattended or poorly attended parts of the display. Sometimes the perception is degraded (*e.g.*, threshold elevation), but it is not eliminated. On the other hand, more complex discrimination, especially those involving spatial relationships (*i.e.*, red-green disk vs. green-red disk), are performed well only with full attention. Contrary to a widely held belief, the extent to which a visual discrimination depends on attention is unrelated to task "difficulty". In other words, discrimination with comparable psychometric functions (*e.g.*, performance as a function of stimulus presentation time) may exhibit wildly different requirements for attention.

These results suggest that visual experience derives from two sources, one which depends critically on attention ("attentive vision") and another which does not ("ambient vision"). Ambient vision is limited in scope, and only provides information about the salient stimuli of a display and, furthermore, only about elementary attributes of such stimuli. Nevertheless, a wide range of demanding visual discriminations are performed readily on the basis of ambient vision. Thus, ambient vision cannot be considered a "subliminal" or "implicit" faculty. In general, visual thresholds measured with ambient vision are somewhat higher (i.e., worse) than those measured with attentive vision. Thus, attention can be said to improve ambient vision in a quantitative manner. Therefore, the distinguishing characteristic of ambient vision is that it is not entirely uninfluenced by attention. Ambient vision differs from the related notion of "parallel, preattentive processing" in being a direct source of visual experience and in being modulated by attention.

Attentive vision adds fundamentally new and qualitatively different aspects to visual experience. One example is spatial relationships. Even the most elementary discriminations involving relative position (e.g., green-red vs. red-green, T vs. L) are performed at chance when attention is focused elsewhere in the display. Indeed, optimal performance of such discriminations appears to require full attention. Although from a computational point of view it may seem surprising that such elementary discriminations are not performed "in parallel", the concurrent-task results presented here bear out earlier work on visual search, according to which attention is required to associate stimulus attributes and stimulus location.

If attention has both quantitative and qualitative consequences for visual perception, it may well be that these disparate effects reflect the influence of attention at different levels of visual cortex. For example, the quantitative effects of attention on visual thresholds for contrast, orientation, and spatial frequency can be understood almost entirely in terms of a inhibitory interaction among neurons with overlapping receptive fields, presumably at the level of visual cortical area V1 and/or V2. As a result of this interaction, the response of each neuron is "normalized" relative to the total response of the population. The changes in thresholds caused by attention are consistent with the possibility that attention intensifies this interaction, and thereby activates a winner-take- all competition among overlapping visual filters. Thus, attention appears to merely modulate existing cortical circuits, at least at this relatively early level of processing. On the other hand, the qualitative aspects of attention may involve different levels of processing. Recent single-unit results suggest that neurons in visual cortical area V4 encode the spatial relationship between a visual stimulus and the focus of attention (Connor, et al., 1996). Results of this kind can be modeled in terms of attentional "gain fields" which modulate receptive fields and further differentiate the neural population by providing selectivity to spatial relationships (Salinas, Abbott, 1997). In this case, attention would not only act at a different neural level but would endow neural responses with qualitatively new information.

References

[1] D A Allport. Parallel encoding within and between elementary stimulus dimensions. *Perception & Psychophysics*, 10:104–108, 1971.

[2] D A Allport. Attention and performance. In G Claxton, editor, *Cognitive Psychology: New Directions*. Routledge and Kegan Paul, London, 1980.

[3] H S Bashinski and V R Bacharach. Enhancement of perceptual sensitivity as the result of selectively attending to spatial locations. *Percep Psychophysics*, 28:241–248, 1980.

[4] M B Ben-Av, D Sagi, and J Braun. Visual attention and perceptual grouping. *Perception & Psychophysics*, 1952:277–294, 1992.

[5] R. Ben-Yishai, R. L. Bar-Or, and H. Sompolinsky. Theory of orientation tuning in visual cortex. *Proc Natl Acad Sci U S A*, 92:3844–8, 1995.

[6] J Bergen and B Julesz. Parallel versus serial processing in rapid pattern discrimination. *Nature*, 303:696–698, 1983.

[7] A-M Bonnel and J Miller. Attentional effects on concurrent psychophysical discriminations: investigations of a sample size model. *Perception & Psychophysics*, 55:125–247, 1994.

[8] A M Bonnel, C A Possamai, and M Schmitt. Early modulations of visual input: A study of attentional strategies. *Quart J Exper Psychol*, 39A:757–776, 1987.

[9] A-M Bonnel and W Prinzmetal. Dividing attention between the color and shape of objects. *Perception & Psychophysics*, 60:113–124, 1998.

[10] A-M Bonnel, J-F Stein, and P Bertucci. Does attention modulate the perception of luminance changes? *Quart J Exp Psychol*, 44A:601–626, 1992.

[11] S F Bowne. Contrast discrimination cannot explain spatial frequency, orientation or temporal frequency discrimination. *Vision Res.*, 30:449–461, 1990.

[12] J Braun. Focal attention and shape-from-shading. *Perception*, 19:A112, 1990.

[13] J Braun. Shape-from-shading is independent of visual attention and may be a 'texton'. *Spatial Vis*, 7:311–322, 1993.

[14] J Braun. Visual search among items of different salience: removal of visual attention mimics a lesion in extrastriate area V4. *J Neurosci*, 14:554–567, 1994.

[15] J Braun. Divided attention: narrowing the gap between brain and behavior. In R Parasuraman, editor, *The Attentive Brain*, pages 328–351. MIT Press, Cambridge, MA, 1998.

[16] J Braun. Vision and attention: the role of training. *Nature*, 393:424–425, 1998.

[17] J Braun and B Julesz. Withdrawing attention at little or no cost: Detection and discrimination tasks. *Perception & Psychophysics*, 60:1–23, 1998.

[18] J Braun and D Sagi. Vision outside the focus of attention. *Perception & Psychophysics*, 48:45–58, 1990.

[19] J Braun and D Sagi. Texture-based tasks are little affected by a second task which requires peripheral or central attentive fixation. *Perception*, 20:483–500, 1991.

[20] J A Brefczynski and E A DeYoe. A physiological correlate of the spotlight of visual attention. *Soc. Neurosci. Abstr.*, 24:493.7, 1998.

[21] D E Broadbent. *Decision and Stress*. Academic Press, London, 1971.

[22] G Buracas, D K Lee, C Koch, T Albright, and L Croner. Attentional modulation of motion perception. *Invest Ophth Vis Sci*, 39:in press, 1998.

[23] M Carandini and D J Heeger. Summation and division by neurons in primate visual cortex. *Science*, 264:1333–1336, 1994.

[24] M. Carandini, D. J. Heeger, and J. A. Movshon. Linearity and normalization in simple cells of the macaque primary visual cortex. *J Neurosci*, 17:8621–44, 1997.

[25] M Cheal and D R Lyon. Attention in visual search: multiple search classes. *Quart. J. Exp. Psychol. A*, 47:49–69, 1994.

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

[27] M Coltheart. Iconic memory and visible persistence. *Perc Pscyhophys*, 27:183–228, 1980.

[28] C E Connor, D C Preddie, J L Gallant, and D C Van Essen. Spatial attention effects in macaque area V4. *J Neurosci*, 17:3201–3214, 1997.

[29] M Corbetta, F M Miezin, S Dobmeyer, G L Shulman, and S E Petersen. Selective attention modulates neural processing of shape, color, and velocity in humans. *Science*, 248:1556–1559, 1990.

[30] M Corbetta, G L Shulman, F M Miezin, and S E Petersen. Superior parietal cortex activation during spatial attention shifts and visual feature conjunction. *Science*, 270:802–805, 1995.

[31] R Desimone. Visual attention mediated by biased competition in extrastriate visual cortex. *Phil. Trans. R. Soc. Lond. B*, 353:1245–1255, 1998.

[32] R Desimone and J Duncan. Neural mechanisms of selective visual attention. *Annu Rev Neurosci*, 18:193–222, 1995.

[33] R Desimone and L G Ungerleider. Neural mechanisms of visual processing in monkeys. In F Boller and J Grafman, editors, *Handbook of Neuropsychology, Vol. II*, pages 267–299. Elsevier, Amsterdam, 1989.

[34] R L DeValois and K K DeValois. *Spatial Vis.* Oxford University Press, New York, NY, 1988.

[35] C. J. Downing. Expectancy and visual-spatial attention: effects on perceptual quality. *J Exp Psychol Hum Percept Perform*, 14:188–202, 1988.

[36] J Duncan. The locus of interference in the perception of simultaneous stimuli. *Psychol Rev*, 87:272–300, 1980.

[37] J. Duncan. Selective attention and the organization of visual information. *J Exp Psychol Gen*, 113:501–17, 1984.

[38] J Duncan. Similarity between concurrent visual discriminations: dimensions and objects. *Percept Psychophysics*, 54:425–430, 1993.

[39] J Duncan and G Humphreys. Beyond the search surface: visual search and attentional engagement. *J Exp Psychol: Hum Perc Perf*, 18:578–588, 1992.

[40] J Duncan, G Humphreys, and R Ward. Competitive brain activity in visual attention. *Curr Op Neurobiol*, 7:255–261, 1997.

[41] J Duncan and G W Humphreys. Visual search and stimulus similarity. *Psychological Review*, 96:433–458, 1989.

[42] J Duncan, S Martens, and R Ward. Restricted attentional capacity within but not between sensory modalities. *Nature*, 387:808–810, 1997.

[43] J Duncan and I Nimmo-Smith. Objects and attributes in divided attention: Surface and boundary systems. *Perception & Psychophysics*, 58:1076–1084, 1996.

[44] J Duncan, R Ward, and K Shapiro. Direct measurement of attentional dwell time in human vision. *Nature*, 369:313–315, 1994.

[45] J T Enns and R A Rensink. Influence of scene-based properties on visual search. *Science*, 247:721–723, 1990.

[46] D C Van Essen and J L Gallant. Neural mechanisms of form and motion processing in the primate visual system. *Neuron*, 13:1–10, 1994.

[47] M. Fahle. Spatio–temporal interpolation. *Concepts in Neuroscience*, 1:253–273, 1990.

[48] M. Fahle. Parallel perception of vernier offsets, curvature, and chevrons in humans. *Vision Research*, 31:2149–2184, 1991.

[49] D J Felleman and D C Van Essen. Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex*, 1:1–47, 1991.

[50] B Fischer, H Weber, M Biscaldi, F Aiple, P Otto, and V Stuhr. Separate populations of visually guided saccades in humans: reaction times and amplitudes. *Exp. Brain. Res.*, 92:528–541, 1993.

[51] D L Fisher. Central capacity limits in consistent mapping, visual search tasks: Capacity scanning in visual search. *Cognitive Psychology*, 16:449–484, 1984.

[52] J. M. Foley. Human luminance pattern-vision mechanisms: masking experiments require a new model. *J Opt Soc Am A*, 11:1710–9, 1994.

[53] J M Foley and W Schwarz. Spatial attention: effect of position uncertainty and number of distractor patterns on the threshold-versus-contrast function for contrast discrimination. *J. Opt. Soc. Am. A*, 15:1036–1046, 1998.

[54] S P Gandhi, D J Heeger, and G M Boynton. Spatial attention in human primary visual cortex. *Inv. Ophth. Vis. Sci. (Suppl)*, 39:5194, 1998.

[55] W S Geisler and D G Albrecht. Visual cortex neurons in monkeys and cats: detection, discrimination, and identification. *Vis. Neurosci.*, 14:897–919, 1997.

[56] D M Green and J A Swets. *Signal detectability and psychophysics*. Wiley, New York, NY, 1966.

[57] P E Haenny, J H R Maunsell, and P H Schiller. State dependent activity in monkey visual cortex: 2. Retinal and extraretianl factors in V4. *Exp Brain Res*, 69:245–259, 1988.

[58] J.P. Harris and M. Fahle. The detection and discrimination of spatial offsets. *Vision Research*, 35(1):51–58, 1994.

[59] D J Heeger. Modeling simple-cell direction selectivity with normalized, half-squared linear operators. *J Neurophysiol*, 70:1885–1898, 1993.

[60] H Helmholtz. Vol. 3. In J P C Southall, editor, *Handbuch der Physiologischen Optik*. Dover, New York, NY, 1850/1962.

[61] L Isenberg, M J Nissen, and L C Marchak. Attentional processing and the independence of color and orientation. *J Exp Psychol Human*, 16:869–878, 1990.

[62] L Itti, J Braun, D K Lee, and C Koch. A model of early visual processing. In M C Mozer, M I Jordan, and T Petsche, editors, *Advances in Neural Information Processing Systems, Vol 9*, pages 173–179. MIT Press, Cambridge, MA, 1997.

[63] L Itti, J Braun, D K Lee, and C Koch. Attentional modulation of human pattern psychophysics reproduced by a quantitative model. In *Advances in Neural Information Processing Systems, Vol 10*. MIT Press, Cambridge, MA, 1998.

[64] R B Ivry, E A Franz, A Kingstone, and J C Johnston. The psychological refractory period effect following callosotomy: uncoupling of lateralized response codes. *Journal Experimental Psychology: Human Perception Performance*, 24:463–480, 1998.

[65] W James. *The Principles of Psychology*. Harvard University Press, Cambridge, MA, 1890/1981.

[66] J C Johnston, R S McCann, and R W Remington. Chronometric evidence for two types of attention. *Psychological Science*, 6:365–369, 1995.

[67] J C Johnston, E Ruthruff, and M Monheit. Dependence by any other name smells just as sweet: Reply to van der Velde and van der heijden (1997). *J Exp Psychol Human*, 23:1813–1818, 1997.

[68] J S Joseph, M M Chun, and K Nakayama. Attentional requirements in a 'preattentive' feature search. *Nature*, 387:805–807, 1997.

[69] B Julesz. Textons, the elements of texture perception and their interactions. *Nature*, 290:91–97, 1981.

[70] B Julesz. Towards an axiomatic theory of preattentive vision. In G M Edelman, W E Gall, and W M Cowan, editors, *Dynamic aspects of neocortical function*, pages 585–612. Neurosciences Research Foundation, Washington, DC, 1984.

[71] B Julesz. Early vision and focal attention. *Rev Mod Phys*, 63:735–772, 1991.

[72] B Julesz and J Braun. Early vision: dichotomous or continuous. *33rd Psychonomic Soc*, page 50, 1992.

[73] D Kahneman. *Attention and effort.* Prentice Hall, Englewood Cliffs, NJ, 1973.

[74] D Kahneman and A Treisman. Changing views of attention and automaticity. In D Parasuraman, R Davies, and J Beatty, editors, *Varieties of attention*, pages 29–61. Academic, New York, NY, 1984.

[75] N Kanwisher and J Driver. Objects, attributes, and visual attention: which, what and where. *Current Dir Psychol Sci*, 1:26–31, 1992.

[76] M Kinsbourne. Single channel theory. In D Holding, editor, *Human Skills*, pages 65–89. Wiley, Chichester, England, 1981.

[77] J A Kleiss and D M Lane. Locus and persistence of capacity limitations in visual information processing. *J Exp Psychol: Hum Perc Perf*, 12:200–210, 1986.

[78] C Koch and S Ullman. Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiol*, 4:219–227, 1985.

[79] B Kroese and B Julesz. The control and speed of shifts of attention. *Vision Res*, 29:1607–1619, 1989.

[80] D K Lee, L Itti andC Koch, and J Braun. Attention activates winner-take-all competition among visual filters.

[81] D K Lee, C Koch, and J Braun. Attentional modulation of contrast masking. *Invest Ophth Vis Sci*, 38:5457, 1997.

[82] D K Lee, C Koch, and J Braun. Spatial vision thresholds in the near absence of attention. *Vision Res*, 37:2409–18, 1997.

[83] D K Lee, C Koch, and J Braun. Attentional modulation of contrast masking (continued). *Invest Ophth Vis Sci*, 39:630, 1998.

[84] D K Lee, C Koch, and J Braun. Visual attention is undifferentiated also for less demanding tasks. *Invest Ophth Vis Sci*, 40:in press, 1999.

[85] D K Lee, C Koch, and J Braun. Visual attention is undifferentiated: concurrent discrimination of form, color, and motion. *Perception & Psychophysics*, 61:in press, 1999.

[86] G. E. Legge and J. M. Foley. Contrast masking in human vision. *J Opt Soc Am*, 70:1458–71, 1980.

[87] H. Levitt. Transformed up-down methods in psychoacoustics. *J Opt Soc Am A*, 33:467–476, 1970.

[88] P H Lindsay, M M Taylor, and S S Forbes. Attention and multidimensional discrimination. *Perception & Psychophysics*, 4:113–117, 1968.

[89] Z L Lu and B A Dosher. External noise distinguishes attention mechanisms. *Vision Res.*, 38:1183–1198, 1998.

[90] S. J. Luck, L. Chelazzi, S. A. Hillyard, and R. Desimone. Neural mechanisms of spatial selective attention in areas v1, v2, and v4 of macaque visual cortex. *J Neurophysiol*, 77:24–42, 1997.

[91] S J Luck and E K Vogel. The capacity of visual working memory for features and conjunctions. *Nature*, 390:279–281, 1997.

[92] N A Macmillan and C D Creelman. *Detection Theory: A User's Guide.* Cambridge University Press, Cambridge, 1991.

[93] J Malik and P Perona. Preattentive texture discrimination with early vision mechanisms. *J Opt Soc Am A*, 7:923–932, 1990.

[94] D W Massaro and D S Warner. Dividing attention between auditory and visual perception. *Perception & Psychophysics*, 21:569–571, 1977.

[95] J H R Maunsell. The brains visual world – representation of visual targets in cerebral cortex. *Science*, 270:764–769, 1995.

[96] C J McAdams and J H R Maunsell. Attention enhances neuronal responses without altering orientation selectivity in macaque area V4. In *Soc. Neurosci. Abstr.*, volume 22, page 475.2, 1996.

[97] E K Miller, L Li, and R Desimone. Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *J Neurosci*, 13:1460–78, 1993.

[98] M Monheit and J C Johnston. Spatial attention to arrays of multidimensional objects. *J Exp Psychol Human*, 20:691–708, 1994.

[99] J. Moran and R. Desimone. Selective attention gates visual processing in the extrastriate cortex. *Science*, 229:782–784, 1985.

[100] M J Morgan, R M Ward, and E Castet. Visual search for a tilted target: tests of spatial uncertainty models. *Quart. J. Exp. Psychol. A*, 51:347–370, 1998.

[101] B C Motter. Focal attention produces spatially selective processing in visual cortical are as V1, V2, and V4 in the presence of competing stimuli. *J Neurophysiol*, 70:909–919, 1993.

[102] B C Motter. Neural correlates of attentive selection for color or luminance in extrastriate area V4. *J Neurosci*, 14:2178–2189, 1994.

[103] M C Mozer. *The Perception of Multiple Objects: A Connectionist Approach.* MIT Press, Cambridge, MA, 1991.

[104] H J Müller and J Findlay. Sensitivity and criterion effects in the spatial cuing of visual attention. *Perception & Psychophys*, 42:383–399, 1987.

[105] J. Nachmias and R. V. Sansbury. Letter: Grating contrast: discrimination may be better than detection. *Vision Res*, 14:1039–42, 1974.

[106] K Nakayama and J Joseph. In R Parasuraman, editor, *The Attentive Brain.* MIT Press, Cambridge, MA, 1998.

[107] K Nakayama and M Mackeben. Sustained and transient components of focal visual attention. *Vision Research*, 29:1631–1647, 1989.

[108] D Navon and D Gopher. On the economy of the human processing system. *Psychological Rev*, 86:214–255, 1979.

[109] U Neisser. *Cognitive Psychology.* Appleton Century Crofts, New York, NY, 1967.

[110] E. Niebur and C. Koch. A model for the neuronal implementation of selective visual attention based on temporal correlation among neurons. *J Comput Neurosci*, 1:141–58, 1994.

[111] M J Nissen. Accessing features and objects: Is location special? In M I Posner and O S Marin, editors, *Attention & Performance XI*, pages 205–219. Erlbaum, Hilldale, NJ, 1985.

[112] D A Norman and D G Bobrow. On data-limited and resource-limited processes. *Cognit Psychology*, 7:44–64, 1975.

[113] G. A. Orban. *Neuronal Operations in the Visual Cortex.* Springer, Berlin, 1984.

[114] J Palmer. Attention in visual search: distinguishing four causes of set-size effects. *Curr. Dir. Psychol. Sci.*, 4:118–123, 1995.

[115] H Pashler. Shifting visual attention and selecting motor responses: distinct attentional mechanisms. *J Exp Psychol: Hum Perc Perf*, 17:1023–1040, 1991.

[116] H Pashler. Dual-task interference in simple tasks – Data and theory. *Psychology B*, 116:220–244, 1994.

[117] H Pashler. *The psychology of attention.* MIT Press, Cambridge, MA, 1997.

[118] H Pashler, S Luck, S O'Brien, R Mangun, and M Gazzaniga. Sequentail operation of disconnected cerebral hemispheres in 'split-brain' patients. *Neuroreport*, 5:2381–2384, 1995.

[119] M Potter. Short-term conceptual memory for pictures. *J Exp Psychol: Hum Learn Mem*, 2:509–521, 1976.

[120] A. Pouget, K. Zhang, S. Deneve, and P. E. Latham. Statistically efficient estimation using population coding. *Neural Comput*, 10:373–401, 1998.

[121] W A Press, J J Knierim, and D C Van Essen. Neuronal correlates of attention to texture patterns in macaque striate cortex . *Soc Neurosci Abstr*, 20:838, 1994.

[122] W A Press and D C Van Essen. *Soc Neurosci Abstr*, 20, 1997.

[123] W H Press, S A Teukolsky, W T Vetterling, and B P Flannery. *Numerical Recipes in C, 2nd ed.* Cambridge University Press, Cambridge, 1992.

[124] J Reynolds, L Chelazzi, S Luck, and R Desimone. Sensory interactions and effects of selective spatial attention in macaque area V2. *Soc Neurosci Abstr*, 20:1054, 1994.

[125] J Reynolds, T Pasternak, and R Desimone. Attention increases contrast sensitivity of cells in macaque area V4. *Inv. Ophth. Vis. Sci. (Suppl)*, 38:3206, 1997.

[126] D L Robinson and S E Petersen. The pulvinar and visual salience. *Trends Neurosci*, 15:127–132, 1992.

[127] P R Roelfsema, V A Lamme, and H Spekreijse. Object-based attention in the primary visual cortex of the macaque monkey. *Nature*, 395:376–381, 1998.

[128] J. Rovamo and V. Virsu. An estimation and application of the human cortical magnification factor. *Exp Brain Res*, 37:495–510, 1979.

[129] B S Rubenstein and D Sagi. Spatial variability as a limiting factor in texture-discrimination tasks: implications for performance asymmetries. *J Opt Soc Am A*, 7:1632–1642, 1990.

[130] D Sagi and B Julesz. Where and what in vision. *Science*, 228:1217–1219, 1985.

[131] D Sagi and B Julesz. Short-range limitation on detection of feature differences. *Spatial Vision*, 1:39–49, 1987.

[132] E Salinas and L F Abbott. Invariant visual responses from attentional gain fields. *J Neurophysiol*, 77:3267–3272, 1997.

[133] H. S. Seung and H. Sompolinsky. Simple models for reading neuronal population codes. *Proc Natl Acad Sci U S A*, 90:10749–53, 1993.

[134] M L Shaw. Division of attention among spatial locations: a fundamental difference between detection of letters and detection of luminance increments. *Attention and performance X*, Bouma, H and Bouwhuis, D G:109–121, 1984.

[135] R M Shiffrin and G T Gardner. Visual processing capacity and attentional control. *J Exp Psychol*, 93:78–82, 1972.

[136] J A Solomon, N Lavie, and M J Morgan. Contrast discrimination function: spatial cuing effects. *J. Opt. Soc. Am. A*, 14:2443–2448, 1997.

[137] D. C. Somers, S. B. Nelson, and M. Sur. An emergent model of orientation selectivity in cat visual cortical simple cells. *J Neurosci*, 15:5448–65, 1995.

[138] G Sperling. The information available in visual presentations. *Psychological Monographs*, 74(11, Whole No 498), 1960.

[139] G Sperling and B Dosher. Strategy and optimization in human information processing. In K Boff, L Kaufman, and J Thomas, editors, *Handbook of Perception and Performance*, pages 1–65. Wiley, New York, NY, 1986.

[140] G Sperling and M J Melchner. The attention operating characteristic: some examples from visual search. *Science*, 202:315–318, 1978.

[141] H. Spitzer, R. Desimone, and J. Moran. Increased attention enhances both behavioral and neuronal performance. *Science*, 240:338–40, 1988.

[142] A Treisman. Strategies and models of selective attention. *Psychol Rev*, 76:282–299, 1969.

[143] A Treisman. Search, similarity, and integration of features between and within dimensions. *J Exp Psychol [Hum Perc Perf]*, 17:652–676, 1991.

[144] A Treisman. The perception of features and objects. In A Baddeley and L Weiskrantz, editors, *Attention: Selection, Awareness, and Control*, pages 1–35. Clarendon Press, Oxford, 1993.

[145] A Treisman and A Davies. Dividing attention to ear and eye. In S Kornblum, editor, *Attention and performance iv.*, pages 101–117. Academic Press, New York, NY, 1973.

[146] A Treisman and G Gelade. A feature integration theory of attention. *Cognit Psychology*, 12:97–136, 1980.

[147] A Treisman and S Gormican. Feature analysis in early vision: evidence from search asymmetries. *Psychol Rev*, 95:15–48, 1988.

[148] A Treisman and J Souther. Search asymmetry: a diagnostic for preattentive processing of separable features. *J Exp Psychol [Gen]*, 114:285–310, 1985.

[149] S. Treue and J. H. Maunsell. Attentional modulation of visual motion processing in cortical areas mt and mst. *Nature*, 382:539–541, 1996.

[150] J K Tsotsos, S Culhane, W Wai, Y Lai, N Davis, and F Nuflo. Modeling visual attention via selective tuning. *Artificial Intelligence*, 78:507–547, 1995.

[151] C. W. Tyler. Colour bit-stealing to enhance the luminance resolution of digital displays on a single pixel basis. *Spat Vis*, 10:369–77, 1997.

[152] L G Ungerleider. Functional brain imaging studies of cortical mechanisms for memory. *Science*, 270:769–775, 1995.

[153] F van der Velde and A H C van der Heijden. On the statistical independence of color and shape in object identification. *J Exp Psychol Human*, 23:1798–1812, 1997.

[154] D C VanEssen and H A Drury. Structural and functional analyses of human cerebral cortex. *J Neurosci*, 17:7079–7102, 1997.

[155] S Vecera and M J Farah. Does visual attention select objects or locations? *J Exp Psychol: General*, 123:146–160, 1994.

[156] V. Virsu and J. Rovamo. Visual resolution, contrast sensitivity, and the cortical megnification factor. *Experimental Brain Research*, 37:475–494, 1979.

[157] R Ward, J Duncan, and K Shapiro. The slow time-course of visual attention. *Cognitive Psychol*, 30:79–109, 1996.

[158] R Ward, J Duncan, and K Shapiro. Effects of similarity, difficulty, and nontarget presentation on the time course of visual attention. *Perception & Psychophysics*, 59:593–600, 1997.

[159] A. B. Watson and D. G. Pelli. Quest: a bayesian adaptive psychometric method. *Percept Psychophys*, 33:113–20, 1983.

[160] R Watt. Seeing texture. *Cognit Neurosci*, 1:137–139, 1991.

[161] R J Watt and M J Morgan. A theory of the primitive spatial code in human vision. *Vision Res.*, 25:1661–1674, 1985.

[162] J Wen, C Koch, and J Braun. Visual tracking of multiple moving-objects requires modality-specific attention. *Inv Ophth Vis Sci (Suppl)*, 36:900, 1995.

[163] J Wen, C Koch, and J Braun. Orientation, but not contrast, discrimination deteriorates without focal attention. *Inv Ophth Vis Sci (Suppl)*, 37:2440, 1996.

[164] G. Westheimer and S.P. McKee. Spatial configurations for visual hyperacuity. *Vision Research*, 17:941–947, 1977.

[165] H R Wilson. A transducer function for threshold and suprathreshold human vision. *Biol. Cybern.*, 38:171–178, 1980.

[166] H R Wilson. Responses of spatial mechanisms can explain hyperacuity. *Vision Res*, 26:453–69, 1986.

[167] H R Wilson. Model of peripheral and amblyopic hyperacuity. *Vision Research*, 31(6):967–982, 1991.

[168] H R Wilson and R Humanski. Spatial frequency adaptation and contrast gain control. *Vision Res.*, 33:1133–1149, 1993.

[169] H R Wilson, D Levi, L Maffei, J Rovamo, and R DeValois. The perception of form: Retina to striate cortex. In L Spillmann and J S Werner, editors, *Visual Perception: The Neurophysiological Foundations*, pages 231–272. Academic Press, San Diego, CA, 1990.

[170] H R Wilson and F Wilkinson. Evolving concepts of spatial channels in vision: from independence to nonlinear interactions. *Perception*, 26:939–960, 1997.

[171] J. Wolfe, A. Yee, and S.R. Friedman-Hill. Curvature is a basic feature for visual search tasks. *Perception*, 21:465–480, 1992.

[172] J M Wolfe. Guided search 2.0: a revised model of visual search. *Psychon Bull Rev*, 1994. in press.

[173] Y Yeshurun and M Carrasco. Spatial attention improves performance in spatial resolution task. *Vision Res.*, 39:293–305, 1999.

[174] B. Zenger and D. Sagi. Isolating excitatory and inhibitory nonlinear spatial interactions involved in contrast detection. *Vision Res*, 36:2497–513, 1996.