SELECTIVE ATTENTION IN VISION: HEMIRETINAL FIELDS USED AS CHANNELS FOR DICHOPTIC VIEWING

Thesis presented in partial fulfillment
of the MA degree in Psychology

University of Manitoba

by
Leonard Dennis Chattaway

Autumn 1976

"SELECTIVE ATTENTION INVISION: HEMIRETINAL FIELDS USED AS CHANNELS FOR DICHOPTIC VIEWING"

by

LEONARD DENNIS CHATTAWAY

A dissertation submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

MASTER OF ARTS

© 1976

Permission has been granted to the LIBRARY OF THE UNIVERSITY OF MANITOBA to lend or sell copies of this dissertation, to the NATIONAL LIBRARY OF CANADA to microfilm this dissertation and to lend or sell copies of the film, and UNIVERSITY MICROFILMS to publish an abstract of this dissertation.

The author reserves other publication rights, and neither the dissertation nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

Table of Contents

Abstract

A survey of modes of stimulus presentation used in studies of selective attention, in both vision and audition, suggests that the two eyes or the two visual fields, considered as separate input channels, process information differently than the two ears. The difference between modalities can be attributed to the decussation of the optic nerve. Whereas each ear can be considered a separate channel to a respective contralateral hemisphere, each eye is better conceived as two channels, one to an ipsilateral and the other to a contralateral hemisphere. In order to study the conceptualization of the hemiretinal fields as channels, a new design for dichoptic viewing apparatus was developed, based on the horopter. The data were analyzed by Contralateral Contingent Event Analysis. When visual stimuli are considered in terms of input to hemiretinal fields, results similar to those for auditory stimuli are obtained. It was concluded that, not only are the hemiretinal fields analogous to the ears as separate input channels, but the apparatus was feasible for dichoptic viewing tasks in selective attention.

CHAPTER I

Introduction

Theories of Selective Attention

Selective attention is primarily concerned with the division of attention among "concurrent streams of mental activity" (Kahneman, 1973). Whether or not attention is divisible is pertinent to the argument of this thesis.

There are two anecdotal observations that bear on the problem of divisibility of attention. One of these observations is that often one does multiple overt activities simultaneously, such as talking while driving. One apparently divides his attention between the concurrent activities. A second observation is that, when two or more overt stimuli are presented simultaneously, often only one of them elicits a response and the others evoke no response. If all the stimuli elicit responses, the responses typically appear to be successive rather than simultaneous. To account for the apparent divisibility of attention, a number of theories have been developed.

Broadbent's filter theory. Broadbent's (1958) model of attention assumed a sequence of three elements: a short-term storage (S-system), a selective filter, and a limited capacity channel (P-system). Concurrent stimuli enter into the S-system in parallel, and are analyzed there for physical features. More elaborate perceptual analyses are carried out in the P-system. This system deals serially with those stimuli selected by the filter. The amount of time spent on each stimulus depends on the amount of information that the stimulus conveys. When the P-system has cleared, the filter allows a new stimulus to enter. Thus, when two stimuli are presented simultaneously, they can be handled successively, but only if the processing of the first is completed before the record of the other in the S-system has decayed.

Filter theory implies that attention cannot be divided between two stimuli at the same time because the P-system performs no parallel processing of discrete stimuli. The apparent division of attention in performance of concurrent activities is mediated by alternation between channels, defined as any physical characteristic for which the filter can be set, or between responses. According to Broadbent (1958), the rate of alternation is slow (the minimum dwell-time, the amount of time that the filter must remain on one channel before it can switch to another channel, of the filter being about 300- to 500-msec.). The processing of simultaneous complex messages fails when the processing of the first message that enters the P-system is so prolonged that the traces of the other message are lost from the S-system before they can be retrieved.

Although the model is based on the entire experimental evidence reviewed in Perception and Communication (Broadbent, 1958), the major experiment is the split-span experiment (Broadbent, 1954). Three simultaneous pairs of digits, or other signals, are presented to a subject, who is required to recall them. Typically, one of each pair is presented to one ear simultaneously with the presentation to the other ear of the other half of the pair. The criterion is the percentage of lists completely correct. Broadbent found that recall by ear (left-left-left-right-right-right [LLLRRR] or RRRLLL) was much superior (90% of the total number of lists correct as compared to less than 20%) to alternating recall by ear (LRLRLR or RLRLRL). He, thus, suggested that the only information that got through the filter was that from the accepted channel. Since subjects in the alternating recall task could give some correct lists, it is possible that some information from rejected messages does bypass the the filter.

In 1953, Cherry introduced the technique of speech shadowing, the technique of asking a subject to verbally repeat a continuous message while hearing it. Cherry requested the subject to repeat a message through one ear while he played a different, distracting message into the other. He found that the subjects could repeat the accepted message, the message to be shadowed as designated by the experimenter. Although the subject was unable to recall the semantic content or the language of the rejected message (the distracting message), he could distinguish some general physical characteristics (e.g., whether or not the rejected message was words or music, was a prose passage or a list, or was spoken by a male or by a female). With a similar experimental task, Moray (1959) found that a subject's name could cause a shift in attention. Apparently, some information in rejected messages is able to bypass the filter. It appears, then, that selection is not an all-or-none process as Broadbent proposed.

Treisman's attentuation theory. In order to account for the fact that at least some of the rejected message does get past the filter, Treisman proposed a modification of Broadbent's (1958) theory. She suggested that the filtering was not all-or-none: the rejected message was merely attenuated, not eradicated. In her model, information enters the system through a number of parallel channels. At the level of a preprocessor, the messages are analyzed for crude physical properties, such as loudness, pitch, and spatial location. This information is available to conscious perception, and for reporting by the subject, regardless of what happens to the message beyond this point. The preprocessor can also act to attenuate the signal strength depending upon whether or not the signal has the requisite characteristic (e.g., a particular loudness, a particular pitch, or a particular spatial location). This means that the output from the preprocessor is differentially weighted.

The attenuated (rejected) messages and the one unattenuated (accepted) message pass deeper into the nervous system until they reach the main processor (the dictionary). At this level, according to Treisman, a message activates hypothetical "dictionary units" in memory. Each unit has a threshold that must be exceeded for perception to occur. The thresholds for highly significant stimuli (e.g., one's name) are permanently lowered. The threshold for a stimulus that the context makes probable is lowered temporarily. Even if such signals have been attenuated, they will trigger the appropriate dictionary units and perception will occur. Neutral attentuated signals are not able to trigger their appropriate dictionary units because their signal strength has been attenuated by the filter. However, unattenuated signals are able to do so and, hence, the message is perceived. In general, rejected signals do not activate the corresponding dictionary units, except when the threshold of one of these units is exceptionally low.

A postulate of the preceding two models is that there is a mechanism that receives the input in parallel and converts it to sequential form for processing. The parallel to sequential conversion suggests the image of a bottleneck, a stage at which information flow is restricted. It may be that the postulate of a structural bottleneck is not warranted.

Capacity theory. A capacity theory of attention provides an alternative to theories that explain man's limitations by assuming the existence of structural bottlenecks (i.e., the conversion from parallel to sequential processing). Instead of such bottlenecks, a capacity theory (as described by Kahneman, 1973) assumes that there is a general limit on man's capacity to perform work. It also assumes that this limited capacity can be allocated with considerable freedom among concurrent activities (Moray, 1967). To explain man's limited ability to carry out multiple tasks at the same time, a capacity theory assumes that the absolute amount of attention that can be deployed at any time is limited.

Both types of theory (structural and capacity) predict that concurrent activities are likely to be mutually interfering, but they ascribe the interference to different causes. In a structural model, interference occurs when the same mechanism (e.g., the filter) is required to carry out two incompatible operations simultaneously. In a capacity model, interference occurs when the demands of two concurrent activities exceed available capacity. Thus, a structural model implies that interference between tasks is specific (Kahneman, 1973), and depends on the degree to which the tasks call for the same mechanisms. In a capacity model, interference is nonspecific (Kahneman, 1973), and depends only on the demands of both tasks.

The capacity model would account for any physically possible behavior that is within the capacity limitations of the organism. Unfortunately, capacity limitation is the only real restriction of the model. Because this model can account for any physically or mentally possible behavior within the capacity limitations of the organism, including hypothetical behaviors that do not occur, the model is less useful for predicting behavior than the structural models, which are more restrictive in concept.

Moray's switching theory. Once the input has entered the system in parallel, it encounters an unspecified structural bottleneck that reduces the processing to a sequential format. Moray, Fitter, Ostry, Favreau, and Nagy (1975) say that it does not matter whether the selective mechanism (the bottleneck) is a discontinuous (Broadbent, 1958) or a continuous (Treisman, 1960) filtering device; what matters is the proposition that sequential processing occurs. Moray et al. (1975) do not discuss whether or not the switching speed can change, but they do say that the amount of time spent on any one channel varies directly with the amount of time required to make a decision about the input. Thus, indirectly, the switching rate varies with decision time. In the situation in which Moray et al. (1975) were concerned, different pitch tone burst series, each series composed of both targets and nontargets, are presented to each of the ears. They describe the ensuing decision state:

> As soon as the first burst has been processed to the point at which the observer can decide whether to say "yes" or "no", processing of the first burst stops and processing of the second burst or its trace begins. Either when the

> second has been processed to a point where a decision can be made, or when time runs out and a decision, however, unsatisfactory to the observer, must be made the observer presses (or refrains from pressing, for a "no") the appropriate buttons. Since nontargets are more frequent than targets, they will be associated with a shorter decision time that the latter (Hyman, 1953): so short in fact that correctly to process a nontarget leaves enough time thoroughly to process the second burst, as though it alone had been processed. The longer decision time associated with the rarer target cuts down the time available for processing the second burst, and also results in processing a weakened trace for short bursts, and the detectability of the second target therefore falls (pp. 14-15).

There are two points in emphasize in the above statement. First, what Moray et al. (1975) refer to as the "second burst" is in in reality presented _simultaneously_ with the "first burst." It is called the second burst because it is the second input to be processed. This occurs merely because the switch is _not_ set for the particular channel on which that burst occurs. Because both bursts are presented simultaneously, processing of the second burst will always be on the trace of the second burst.

Second, the amount of time spent processing the second burst varies directly with the amount of time available after the completion of processing on the first burst. If the first burst does not require much processing time for a decision to be made, more time is available to process the second burst. As the amount of time required by the first burst to make a decision increases, the amount of time available for processing the second burst decreases. Thus, at least with the simple signals used by Moray et al. (1975), the first burst is always completely processed. The second burst may or may not be completely processed, depending on the amount of time taken by the first burst.

An analysis of this model shows that there is a structural bottleneck, which could be similar to either Broadbent's (1958) discontinuous or Treisman's (1960) continuous filter. Once past this bottleneck the processing is sequential. However, dwell-time on any one channel is construed as variable. The variability depends upon the amount of time that is taken to make a decision abount the stimulation on that channel. The variability of the switch dwell-time brings the switching theory more in line with the selective attention data. In a later section, differences in predictions between the four theories will be discussed.

The concept of channels. One of the assumptions of the above four models is the notion of a channel. As noted previously, a channel is defined as any physical characteristic for which the selecting device (or allocator in the case of the capacity models) can be set. Examples of channels include such manifest physical characteristics as spatial location, pitch, intensity, voice qualities, and timbre. In other words, the notion of channels assumes that the stimulus field is partitioned into segments or groups. The Gestalt laws of grouping would describe the operation of this stage; for example, successive sounds that originate in the same place would be more likely to be grouped as a unit than sounds from different places.

The concept of channels is very important to attention. It has been noted that effective search is possible only when all potential targets share a physical characteristic that permits these targets to be grouped together (Williams, 1966; von Wright, 1968, 1970). Kahneman (1973) states that "the properties that provide strong units also allow for the effective control of attention, because attention is most easily directed toward a natural unit" (p. 74) or Gestalten.

Experimenters in auditory selective attention have almost unanimously used different spatial locations as the stimulus input channels. Stimuli are typically presented dichotically in shadowing tasks or monitoring tasks, a task in which the subject is to respond to target stimuli within a continuous stream of nontarget stimuli no matter on which channel the target occurs. The stimulation presented to each ear is assumed to be processed separately. That is, there is no mixing of the auditory stimulation, presented to separate ears, along the ascending sensory pathways until the stimulation reaches the cortex.

With respect to the visual system, an interesting anatomical point is that each retina is functionally divided into two regions, which receive input from the left and right visual fields. The right visual field is mediated by both the right nasal hemiretina and the left temporal hemiretina. Similarly, the left visual field is mediated by both the left nasal and right temporal hemiretina. Correspondingly, ganglion cell axons from the right nasal and the left temporal hemiretinae project to the left cerebral hemisphere and ganglion cell axons from the left nasal and the right temporal hemiretinae project to the right cerebral hemisphere. There are interhemispheric connections through the commissure fibers crossing the corpus callosum, but there is no known interaction between the neural pathways of the two eyes at any level more pheripheral than the thalamic geniculates.

Before the two modalities can be appropriately compared, there is a question of whether or not there is an approriate visual analogue to auditory input channels. When the two modalities are compared, it becomes apparent that some methods of stimulus presentation in vision may be more analogous to dichotic stimulus presentation in audition than other methods of presentation. The first possible appropriate analogue is the conceptualization of each eye as an input channel. A second appropriate analogue is the consideration of the right and left visual fields as the input channels. The final conceptualization of visual input channels is the hemiretinal fields. Each hemiretinal field would be considered a separate channel. It is submitted that the most appropriate conceptualization of visual input channels is that of the hemiretinal fields.

Experimental Survey of Visual Input Channels

In the last section, it was suggested that the concept of the hemiretinal fields as input channels was the most appropriate analogue to the concept of the ears as separate input channels. This suggestion would be much stronger if it were supported by the findings in the literature.

Before exploring the literature, however, the definition of "visual field" should be clarified. Sanders (1963) has made a relevant distinction among three orthogonal types or subdivisions of the visual field: the headfield, that part of the visual field that can be sampled by moving the head; the eyefield, that part of the visual field that can be sampled by moving the eyes, but with the head held stationary; and the stationary field, that region that can be sampled while holding both head and eyes stationary.

Sander's distinctions provide a useful division for work on visual attention. Moving the head may totally prevent visual input from some parts of the environment, but it changes merely the direction of auditory stimuli relevant to the ears. Moving the eyes will affect the retinal location of the stimulus. The auditory receptors cannot be moved independently of head movement. The stationary field seems to be more relevant to visual selective attention than either the headfield or the eyefield, especially since dichotic stimuli are typically presented to the separate ears via stereophonic earphones, which ensure that moving the head will not affect the perception of the direction of the stimulation. All of the following studies have controlled for head and eye movement (i.e., have used the stationary field).

Sampson and his colleagues (Sampson, 1964; Sampson & Horrocks, 1967; Sampson & Spong, 1961a, 1961b) conducted a series of studies using a slide projector to back-project stimuli onto a screen. An opaque division prevented the left eye from seeing the right side of the field and the right eye from seeing the left side of the field. In some of the experiments (Sampson & Spong, 1961a, 1961b) fixation points were provided. By fusing them, a standard direction of gaze was maintained. They used the split-span method of presentation, in which signals were presented simultaneously to each eye. The subjects were required to recall the signals after the entire series had been presented. Although, in the original split-span experiments in audition, Broadbent (1954) had found that listeners tended to recall all the stimuli from one ear followed by all those from the other ear, Sampson and Spong (1961a, 1961b) did not find this. They found that viewers tended to recall the presented digits as simultaneous pairs, not separated by eye. This result does not seem surprising, however, since, it two digits appear side by side in the visual field, reading bias alone would give a tendency for them to be read as single two-digit numbers. To read them otherwise might be expected to require practice.

In order to overcome the positive effects of reading bias, Sampson (1964) presented digits to one eye and colored patches to the other. Although there were great individual differences, at least some of the subjects tended to report all digits together and then all colors together rather than pairing digits and colors. In addition, digits were recalled with fewer errors than colors, and had a shorter response time.

One possible interpretation of these results is that response factors were involved rather than attentional ones. Most people have probably had greater practice at naming numbers than color patches in everyday life, and, thus, perhaps show a reduced latency for numbers (Moray, 1970). Sampson did not control for this effect.

In 1967, Sampson and Horrocks explored the importance of the different regions of the visual field in more detail. The method of presentation was the same as in the earlier experiments (Sampson, 1964; Sampson & Spong, 1961a, 1961b), but the arrangement of the stimuli in the first of their series of experiments was as shown in Figure 1. They found that the most frequent pairing of responses was upper-lower; that is, the upper of a pair was given before the lower. The only exception was the last condition in which the left side (lower) was given before the right side (upper). The subjects spontaneously reported in pairs. Again, as in the previous studies, these results appear to show the effects of reading bias.

....
Insert Figure 1 about here
....

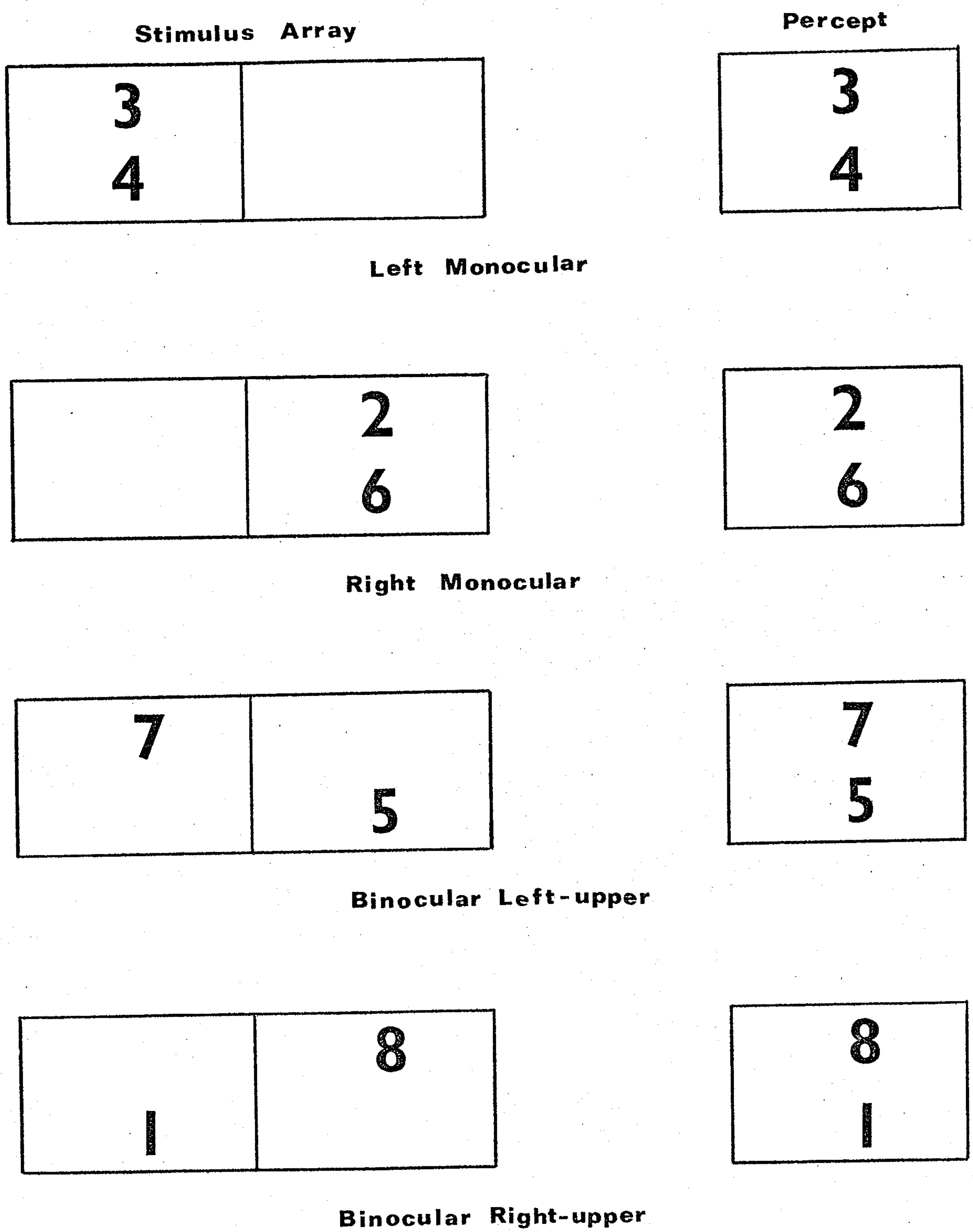
In a second experiment, reported in the same paper, they investigated the effect of partially overlapping the stimuli, either binocularly or monocularly (Figure 2). Recall was more accurate when viewing monocularly overlapped stimuli than when the overlapping was caused by binocular fusion. The stimuli tended to be recalled pair by pair, the left hand digit first. This tendency was more marked in the monocular recall condition.

....
Insert Figure 2 about here
....

A final experiment investigated the possible role of binocular rivalry. Slides, similar to the binocular overlapping condition of the previous experiment (Figure 2), were presented and viewers were asked to call out after each slide what they had seen. The stimuli presented to the left eye were recalled with a level of accuracy of 74.2% and those to the right eye with 85%. These stimuli apparently were viewed using no designated fixation points because there were no differences in accuracy of recall between the eyes in those conditions in which the subjects were required to fixate binocularly. Implicit in the Sampson studies (Sampson, 1964; Sampson & Horrocks, 1967; Sampson & Spong, 1961a, 1961b) was that there was an appropriate analogy between perception in the auditory system and perception in the visual system. They considered each eye as an input channel. In other words, the eyes were conceived as functioning simultaneously as independent information sources.

Two difficulties with apparatus design, however, tend to

Figure 1. Stimulus display from Sampson and Horrocks (1967).


Stimulus Array
Percept
3
4
3
4
Left Monocular
2
6
2
6
Right Monocular
7
5
7
5
Binocular Left-upper
8
1
8
1
Binocular Right-upper