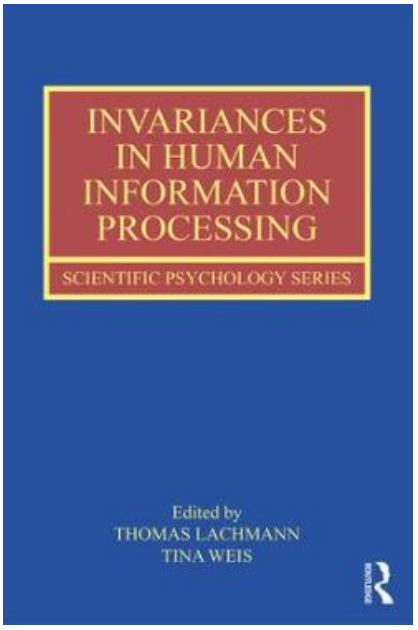
A brief overview of computational models of spatial, temporal, and feature visual attention

George Sperling

in



Free sample chapter.

Chapter 7, colored e-version

Sperling, G. (2018). A brief overview of computational models of spatial, temporal, and feature visual attention. In T. Lachmann & T. Weis (Eds.), *Invariances in Human Information Processing* (pp. 143-182). New York, NY: Routledge.

7 A Brief Overview of Computational Models of Spatial, Temporal, and Feature Visual Attention

George Sperling
UNIVERSITY OF CALIFORNIA, IRVINE, CA, USA

Abstract

Some representative experiments that span over fifty years of research on visual selective attention are reviewed here. They parametrically describe spatial attention to locations in space, temporal attention to intervals in time, and feature attention to particular visual features. The resulting data are accurately described by a class of lean computation models (i.e., few estimated parameters) in which attention is a represented as a gating (i.e., filtering) process. The models represent plausible brain functions utilizing components that represent basic neural transformations. Spatial, temporal, and feature attention are supported by parallel, independent brain processes that combine multiplicatively in the combinations so far tested.

7.1 Introduction

The recent celebration of the remarkable lifetime achievements of Prof. Geissler (Lachmann & Weis, 2018) was also a stimulus to review one's own achievements. This succinct overview of attention experiments, and the computational models that describe them, samples more than 50 years of working with a series of wonderful collaborators—too many to list all as co-authors. All examples are from published articles with citations and links to the source articles for readers who want more details.

Selective-attention tasks. The experiments and phenomena considered here, and for which we seek understanding and an explanation, are typically said to involve visual attention. What relates these phenomena is an instruction to subjects to selectively pay attention to some aspect of a visual stimulus—i.e., to selectively process only a fraction of the incoming visual information. However, the verbal instruction, while convenient, is not critical. Similar experiments can be conducted with nonhuman animals that are taught that a particular cue requires them to selectively respond to particular aspects of a display.

While we understand attention as an everyday concept, in the laboratory, when we study attention, we consider it in terms of experiments that involve

subjects. Here we consider only human subjects but, again, it is a very good exercise to see how the same experiment could be adapted to study attention in other species. The experiments considered here consist of trials on each of which a subject performs a specific task. A visual attention task consists of visual stimuli, other stimuli such as attention instructions (often called cues) that can be presented in any modality, responses that the subject makes, and a reward structure—i.e., feedback to the subject about the correctness or incorrectness of the response. Ideally, the feedback includes a specific value (called utility) that is associated with each particular stimulus-response combination (see, Sperling, 1984; Sperling & Dosher, 1986).

Some characteristics of the attention experiments considered here:

(i) For a particular condition—e.g., a particular exposure duration, all the visual stimuli are chosen randomly from the same urn of stimuli. What distinguishes the responses is the particular attention instruction on a trial. The experiments are designed so that any systematic differences in response can be attributed purely to different attention instructions and utility functions.
(ii) The instructions to the subject at the beginning of a block of trials are quickly irrelevant, what governs behavior is the reward structure—i.e., the utility function for each trial.
(iii) The explanation of performance is a model, for experiments after about 1980, a computational model.

A critical note on three related uses of the word attention: The primary use of the word attention refers to an instruction cue to the subject that indicates "selectively attend to X" and to the feedback utility function that defines the meaning of the cue. (A selective-attention utility function assigns high-value rewards to correct responses made to cued stimuli— e.g., to stimuli that occur in a cued region or stimuli that contain a cued feature—and assigns lower values to all other stimulus-response combinations.)

A second use of attention refers to aspects of a subject's response, as in "he succeeded in attending to X but failed to attend to Y". For example, in response to an instruction to shift attention from location X to location Y, we can talk about the space-time trajectory of "attention" just as we can talk about the space-time trajectory of the hand in response to an instruction to move the hand from X to Y. The trajectory of attention is the region in space-time within which stimuli are better processed in accordance with an attention instruction. An explanation for moving the hand is not that there is a hand-movement trajectory in the brain, rather that the brain interprets the instruction, computes representations of the starting hand position and of the intended final hand positions, computes a trajectory, and implements that trajectory with various muscle commands.

Explanations of the underlying brain processes in response to a demand to move attention are equally complex and are considered in sections 7.5–7.8. To foreshadow, in the computational models to be described herein, the demand for selective attention is implemented as the selective gating of information at various stages of processing.

There is a third use of the word attention when we talk, for example, about attention occurring at a particular stage or stages of processing. This is an adjectival, figurative use of attention. It is a shorthand way of saying that the brain, or whatever system is under consideration, responds differentially and selectively in response to situational demands; it does not imply any particular mechanism for accomplishing this. For example, in a spatially cued reaction-time task, Posner (1980) argued that the attention cue facilitated perceptual processing of the stimulus at the cued location; random-walk theorists (Smith & Ratcliff, 2009) argue that reaction-time facilitation typically occurs at a later decision stage. Brain imaging studies indicate that selective spatial attention can modulate responses of numerous regions in all parts of the cerebral cortex as well as in midbrain nuclei (Nobre & Kastner, 2014). It requires a very carefully designed experiment to focus the effect of selective attention primarily on just one particular level of processing.

Finally, using attention as the name of brain processes is best avoided. Explaining selective attention (behavior) with a selective-attention brain process sounds suspiciously like a tautology. Even in a computational model, in which what one calls processes doesn't affect their computations, using the same word for inputs and also for computations or their outputs makes it impossible to talk cogently about the model.

7.2 Spatial Attention: Iconic Memory and Partial Report

Figure 7.1 illustrates the procedure and describes results of an early attention experiment (Sperling, 1960). This was a spatial-attention experiment, but the emphasis of the study was (1) the demonstration that more letters were available to a partial report than were available to a whole report, and (2) that this implied a brief visual sensory memory that mostly decays in less than one-half second. Neisser, in his classic book *Cognitive Psychology* (1967), called it "iconic memory" and this term has survived.

In the partial-report experiment of Figure 7.1, the stimuli were hand lettered on cards, presented in a mirror tachistoscope, and the responses were written by the subject, unlike the computer-generated stimuli and typed responses that are common today. Therefore, immediate feedback of the correctness of a response was impractical. A tonal cue was the attention instruction to selectively attend to the indicated row. The bottleneck that required partial report and selective attention was the number of letters that could be retained in a short-term memory—now typically called working memory—for a sufficiently long time to enable writing them

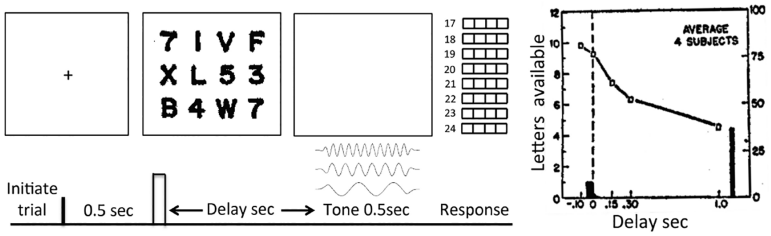


Figure 7.1 Selective attention to a spatial region: The partial report procedure. To initiate a trial, the subject fixates the fixation cross shown in the far left panel, and then presses a button. After a foreperiod of 500 ms (blank screen, not shown), the 12 item display flashes for 50 ms. Following a variable delay from the display offset (from −50 to +1,000 ms), a randomly selected low, medium, or high pitch tone sounds to indicate the report of only the bottom, middle, or top row. The subject writes a four-character response, guessing if necessary, in the answer sheet illustrated. The right-hand ordinate of the data graph illustrates the fraction f of correctly reported characters in the cued row as a function of the cue delay. The left-hand ordinate is the product 12f, the number of letters from the 12-letter stimulus available to the subject from which the designated row (a random sample of the available letters) was reported. The bar at far right indicates the fraction of 12 letters correctly reported in a separate block of trials in which there was no cue and the subject simply reported as may letters as possible of the whole stimulus (whole report). (In part after Sperling (1960), Figure 2, p. 3, and Figure 8, p. 11; APA as publisher, adapted by permission.)

down. Figure 7.2 is conceptual model that represents the memory processes involved in the iconic memory experiment.

The first component in Figure 7.2 is Visual Information Storage (VIS) in which the visual stimulus is stored iconically—i.e., essentially a retinotopic pixel image. The decay of information in VIS is measured by the partial report procedure (Figure 7.1), but this procedure cannot indicate the acquisition rate of VIS. Weichselgartner and Sperling (1987) measured the entire time course, rise and fall, of conscious visual persistence (versus information persistence) of a brief flash of a high-contrast grating. Subjects judged the grating's perceived contrast at the precise instant an auditory click occurred. The derived visual persistence for a subject with unusually long-duration persistence (700 ms from beginning-to-end) is illustrated as a function of time above the VIS component. Even this longest VIS is not long enough to enable the subject to locate and write the requested row of three or four letters. Reporting the letters from a brief flash, requires a rapid extraction from VIS before it decays, and the transfer of the attended letters to a Visual Short-Term Memory (VSTM), which, unlike VIS, has slow decay and is relatively immune to subsequent visual stimulation.

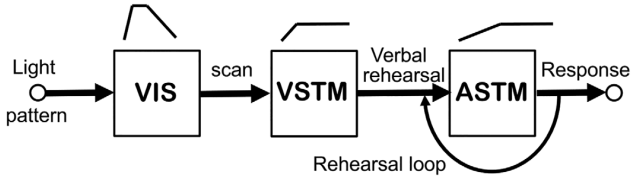


Figure 7.2 The memory processes involved in partial report (iconic memory) experiments. VIS is Visual Information Storage. The curve above VIS represents "visual persistence" (versus "information" persistence). VSTM is Visual Short-Term Memory. The representation of the contents of VSTM as a function of time, illustrated above VSTM, has a slower rise and lower height than VIS indicating both slower acquisition and a smaller capacity than VIS. Verbal rehearsal of the VSTM contents converts the visual contents to a representation in Auditory Short-Term Memory (ASTM). The initial acquisition slope of the contents-versus-time representation above the VSTM component indicates that verbal rehearsal occurs at a slower rate than the visual scan of VIS, but ASTM has a larger character capacity. (Adapted from Figure 6, p. 26, in Sperling (1963), Sage Publishing, by permission.)

Acoustic confusions in visual partial reports. The reason for including an auditory component in a model of a visual memory experiment is that subjects make acoustic confusions in their responses, for example between B and D, D and T (Sperling, 1960, p. 21). In fact, some subjects can be heard vocally rehearsing the to-be-reported items. Most subjects, however, rehearse subvocally. Acoustic confusions in experiments with visual displays were subsequently studied extensively by Conrad (1964) and by Sperling and Speelman (1970) who used visual stimuli that had either confusable or distinct acoustic representations. Sperling (1968) proposed a computational model of auditory memory that accounted for the effect of acoustic confusability in terms of the mnemonic inefficiency of repeating the same phoneme in the different confusable items (b,c,d,g,p,t, v,z). In a brilliant study, Scarborough (1972) demonstrated that his subjects' could store different to-be-remembered items independently in VSTM and in ASTM.

7.3 Post-Stimulus Noise Masking to Limit Iconic Persistence, Verification by Auditory Synchrony Judgments

To measure the rate of transfer of items from VIS to VSTM it seems logical to simply vary the exposure duration and determine how many letters can be extracted at each exposure duration. However, because of visual persistence (iconic memory) in the VIS component, the actual duration for which

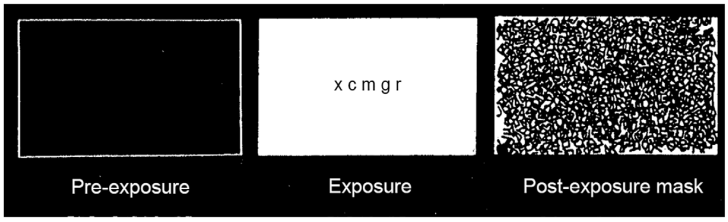


Figure 7.3 The post-exposure masking procedure as used to enable accurate measurement of the rate of information transfer from VIS (iconic memory) to VSTM (visual working memory). The post-exposure masking field overwrites the contents of VIS; thereby the time from the beginning of the stimulus exposure to the beginning of the post-exposure mask is the maximum time for which visual information is available for transfer to VSTM. Typically, three letters can be transferred within 50 ms. (After Figure 3, p. 24, Sperling (1963), Sage Publishing, by permission.)

information is visual available may vastly exceed the exposure duration. To resolve this problem, a post-exposure masking field is introduced (Figure 7.3). By having subjects on some trials adjust an auditory click to appear simultaneous with onset of visible letters and on other trials to appear simultaneous with termination of visible letters, Sperling (1967) showed that a post-stimulus masking procedure produced visual displays in which letters were visible for a duration (the interclick interval) that very closely approximated the duration from the onset of a visual stimulus to the onset of the post-stimulus masking field.

Figure 7.4 illustrates a detailed flow chart of the brain processes that were conceptually outlined in Figure 7.2 to account for performance in partial-report memory experiments.

Different partial report strategies? The early information processing experiments described earlier have several deficiencies that were corrected in subsequent investigations. Because of the primitive nature of the apparatus, no immediate feedback after a trial was possible. Therefore this is a Type 2 (versus a Type 1)[1] experiment (Sperling, 1992; Sperling, Dosher, & Landy, 1990). For the same reason, the different tonal cue delays were run in different blocks. This raises the possibility that a subject could adopt a different, optimal strategy in each block of trials. A mixed list design in which any delay can occur randomly on any trial is necessary to ensure that the subject's strategy mix is the same, independent of the cue delay. Gegenfurtner and Sperling (1993) investigated this possibility by running a partial report procedure similar to Figure 7.1 in which there were only two cue delays: short (zero ms from stimulus onset) and long (400, 800, or 1,000 ms, depending on the subject). In different blocks of

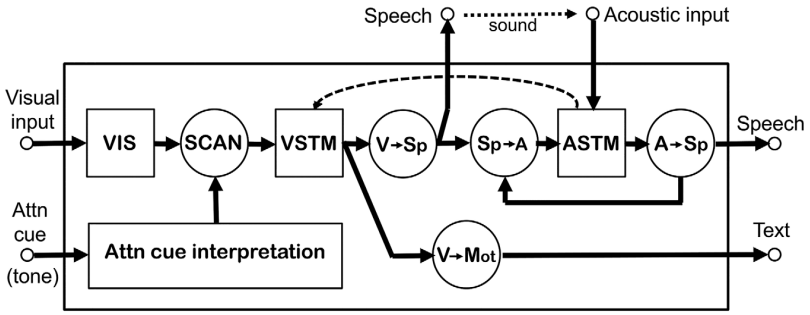


Figure 7.4 Outline of a possible computational model of the brain processes involved in the partial report procedure. The visual input is a 3×4 letter-and-number stimulus matrix, that is briefly stored in VIS. The attention cue is high, medium, or low-frequency auditory tone that is interpreted to direct the scan process to transfer the corresponding row to VSTM. The content of VSTM could, in principle be converted directly into a response, however, writing the response produces visual interference with the contents of VSTM. Therefore, the contents of VSTM are converted into auditory images either by vocal rehearsal (V→Sp, visual representations converted to motor movements of speech, the listened to) or by subvocal rehearsal (Sp→A, the motor movements of silent speech are converted to auditory images to be stored in ASTM, which has a greater and more durable capacity for letters than VSTM. The contents of of ASTM are either spoken (A→Sp, auditory images converted to motor movements of speech), or rehearsed (re-entered into ASTM via Sp→A), or converted into written responses. The conversion of auditory images to written letters probably involves reactivation of the letter images in VSTM (the dotted connecting line) and then using the Visual-to-Motor conversion component (V-Mot) to write letters (Text). (Adapted from Figure 4, p. 290, Sperling (1967), Elsevier, by permission.)

trials, the probability of the short delay was 0, 0.1, 0.5, 0.9, 1.0. Performance was the same in all blocks indicating that, indeed, subjects employed the same strategy in all cue delays: Fill memory with letters from an idiosyncratic set of default locations until a cue to report is heard and interpreted, then fill memory (VSTM) with letters from the cued locations.

A third deficiency of the original iconic memory experiments was that the estimated duration of VIS was confounded with the times to interpret the cue and to shift attention to the cued row. The shift-time and other characteristics of the attention shift were quantitatively measured by Shih and Sperling (2002), described next. A fourth deficiency is that the data were not rich enough to enable the estimation of parameters in a minimal model of performance (e.g., Figure 7.3) in the partial report procedure. The following experiments provide the data for the first computational attention model.

7.4 Attention Reaction Time and the Window of Attention

This and next two sections describe the procedures for precisely measuring shifts of attention and the brain mechanisms that implement these attention shifts. We begin with an analogy, the grabbing response. Figure 7.5a illustrates an indirect procedure for measuring the reaction time of a motor response without a clock capable of measuring short durations. A conveyor belt is stacked with billiard balls so that each second ten balls pass an opening that gives the subject access to the belt. Each ball has a number painted on it. The numbers on the balls are random but the experimenter knows the precise time at which each numbered ball is available at the aperture. The subject is instructed that when a critical stimulus—e.g., the letter C, appears on the screen he is monitoring—he is to reach in as quickly as possible to grab the earliest ball he can, and then to tell the number to the experimenter. Although a numerical report may not occur until several seconds after the grabbing response, it nevertheless indicates the particular time interval during which the subject grabbed the billiard ball—i.e., the reaction time of the grabbing response to within 1/10 second.

Attention reaction time, ART. An analogous grabbing response can be used to measure an attention reaction time. A subject views two adjacent streams: On the left, a stream of letters each one on top of the previous one and, on the right, a stream of numbers, each one on top of the previous one. A fixation point is in the middle (Figure 7.5b). The subject's task is to attend to the stream on the left until a target letter, C, appears and then to switch attention to the stream on the right and attempt to report the adjacent numeral, typing it on a keyboard. In any case, the subject is to report the earliest numeral that he/she was able to report after the concurrent numeral. This is analogous to the grabbing response in Figure 7.5a although in this case it is an attention grab of a numeral.

The rate of letter presentation was chosen to require the subject to pay full attention to the letter stream in order to detect a target chosen, in different blocks of trials, from among C, U, outline square. Typically, when numerals in the numeral stream occur at 9 per second, the subject reports the third or fourth letter after the simultaneous letter. The distribution of positions of the reported letter defines the distribution of ART (switching attention from monitoring the letter stream to grabbing a letter from the numeral stream) illustrated as ART in Figure 7.5c. In other blocks of trials, the subject merely presses a key as soon as he/she detects the target. This produces a conventional motor reaction time distribution, illustrated as MRT in Figure 7.5c. Evidently, it is possible to obtain as complete data about an "unobservable" implicit attention reaction time as about a MRT, and for this letter "C" detection task, ART and MRT are remarkably similar.

To specify a model, more data are better. Instead of requesting subjects to report just the single numeral that was simultaneous with the target

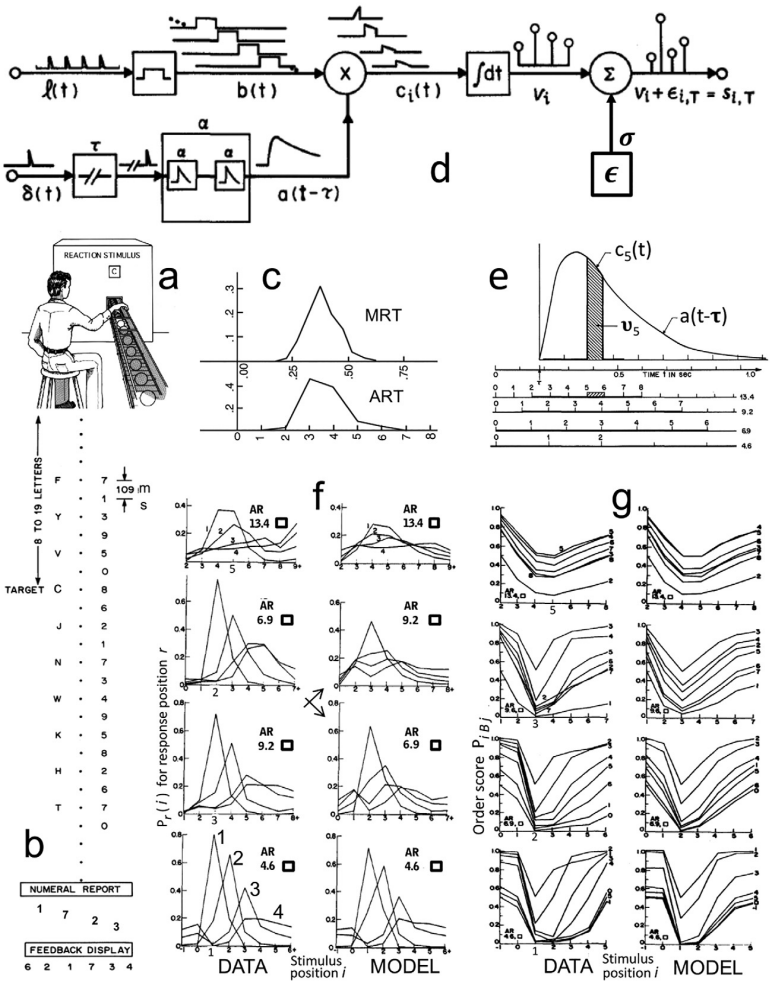


Figure 7.5 Measuring a temporal window of attention in the Attention Reaction Time (ART) procedure: Theory, data, model. (a) The grabbing response. Upon detecting the target "C" in a letter stream beneath the heading "REACTION STIMULUS", the subject grabs the first ball from the conveyor belt. The code on the ball indicates it's position in the stream and thereby the reaction time of the subject's grabbing response. (b) A sequence of stimuli in the attention switching-time analog to the grabbing response. The subject fixates between two streams, letters on the left, numerals on the right. His task is to detect the target letter (and to report the earliest possible concurrent and subsequent four numerals from the numeral stream. Complete feedback is given (see bottom). (c) Reports from a preliminary set of trials in which the the target is

the letter "C" and the subject reports only one numeral. In ART, the abscissa indicates the position of the reported numeral in the sequence and thereby it's time of occurrence (0 is the onset time of the numeral concurrent with target). The ordinate indicates the frequency of such reports. In MRT (motor reaction time) the abscissa indicates the time in seconds of a finger press, the ordinate indicates the frequency of that reaction time. (d) Flowchart of the computational attention model. In the letter input path, $l(t)$ represents the luminance of the input stream of numerals as a function of time; $b(t)$ represents the output of a visual sensory store that extends the visual persistence of each input numeral until it is overwritten by the subsequent numeral. In the numeral input path, $\delta(t)$ represents the occurrence of the target at time t = 0, τ, represents the time taken to interpret the target and begin the process of opening an attention window, the attention window function $a(t)$ is arbitrarily chosen as a second-order Gamma function produced by two consecutive exponential decay filters each with time-constant α. Attention gating occurs at the $\times$ component where the processed input $b(t)$ is multiplied (gated) by the attention function $a(t)$ to produce the attention-gated output $c(t) = a(t)b(t)$. (e) The attention function $a(t - \tau)$. The dark rectangle within $a(t - \tau)$ illustrates the product $c(t)$ for the fifth item i_5 in the processed 13.4 items/sec input stream $b(t)$ which is shown immediately below the abscissa, just above the three slower streams. (d, continued). An integrator computes the total area—the value v_i of $\int c_i(t)$ for each item i; a normally distributed noise sample $\varepsilon_{i,T}$ with mean zero and standard deviation σ is added to every item v_i on every trial T. Items are then output in order of v_i, the item with the largest value v_i is first. (f) The response data for subject AR. The target for these data was the "outline square" (a square surrounding whatever letter occurred at time 0), it is shown to the right of the numeral presentation-rate in each panel. Ordinate: $p_r(i)$, the proportion of reports of items at stimulus position i shown as a function of stimulus position i (abscissa) for each of the four response positions r (curve parameter). The larger-font numeral under the abscissa indicates the strongest stimulus position (max v_i). The presentation rate (number of numerals per second) of items in the to-be-reported stream is indicated in each panel. The right column of four panels represents the output of the three-parameter model in (d). (g) Outcomes of the *iBj* paired comparisons. When a stimulus item from position i is reported in an earlier response position than an item from position j, it wins a paired comparison competition. The results of all possible data competitions are shown. The abscissa is i, the ordinate is iBj, the fraction of time i was reported before j, the curve parameter is j. The top curve is the strongest stimulus position in that condition, followed in strength-order by the other positions. Insofar as the curves are laminar (don't cross) a strength model describes the data. ((a) is adapted from Figure 2.34, p. 2-57 in Sperling and Dosher (1986), Wiley, by permission; (c) is adapted from Figure 17.1, p. 351 in Sperling and Reeves, (1980); (b) is adapted from Figure 1, p. 182, (d) and (e) are adapted from Figure 15 p. 195, (f) and (g) are adapted from Figures 16 and 17, p. 197, all in Reeves and Sperling (1986), APA as publisher, by permission.)

letter, in subsequent sessions, subjects were requested to report the four most contemporaneous numerals in their correct order. Subsequently, the subject's four-numeral response was shown to on the display screen as it was produced, followed by a display of the most relevant six-numeral stimulus sequence, thereby giving the subject complete feedback (Figure 7.5b). In a block of trials, numerals were presented at just one of the four rates tested (shown in panel inserts Figure 7.5f).

Figure 7.5f shows the complete results of the four-item report procedure, the distribution of times of occurrence of the numerals reported in the first, second, third, and fourth response position for subject AR with the target "outline square." According to the model (Figure 7.5d) detection time τ for the outline square was 222 msec, whereas τ for the "C" target data in Figure 7.5c) was 265 msec. The first numeral reported has the same distribution as when the subject reports just one (not shown). At the slowest numeral rate, 4.6 items/sec, the subject reports the numeral occurring immediately after the simultaneous-to-target numeral on about 80% of the trials and the next-occurring numeral on the remaining trials. In the second response position, the subject reports the +2 numeral on about 70% of the trials. Subsequent response positions have a greater spread of reports. In blocks with faster numeral streams, the initially reported numeral occurs later and later in the stream and the distribution of reports becomes more variable, ultimately becoming quite chaotic at numeral rates of 13.4/sec. The right-hand set of graphs in Figure 7.5f represents the output of the three-parameter model (Figure 7.5d) that was presented with the same display sequences as the subject.

No order information within an attention glimpse: The method of paired comparisons. The subject was asked to report items in order of their occurrence in the stimulus. Does the subject have any knowledge of the true order in which reported items actually occurred in addition to simply their strength in memory? Using the method of paired comparisons, this section shows that the answer to this question is "no".

Strength in the model is determined simply by the probability of an item appearing somewhere in the response, it is unconcerned with it's order in the response. Consider two response items taken from two stimulus positions of equal strength, one from an early position (i.e., before the position of maximum strength) and the other from a later position. Is there a tendency for the earlier-occurring item to be written in the response before the later occurring item? A powerful, independent analysis of order information in the data was undertaken with the method of paired comparisons, a method that is widely used, for example, to compare the relative strengths of chess or tennis players. Let i and j be the stimulus positions of two numerals that appear in a subject's four-item report. If the numeral representing position i appears before the numeral representing position j in the subject's report, i wins that particular competition and we say iBj (i before j) on that trial. There are six competitions in each four-item

response. All the possible competitions in all the responses in a condition yield a set of *iBj* outcomes, the fraction of times the item in position *i* was reported before the item in position *j* in that particular condition (i.e., numeral rate and particular target). As item strength is measured by the probability of an item appearing in the response, and order strength is measured between items when both appear in the response, these two measures could in principle yield two quite different strength orders.

The *iBj* data are shown for one target (outline square), one subject, and all numeral rates in Figure 7.5g. The *iBj* model predictions are derived from the same Monte Carlo trials that were used to generate the serial position predictions. Within measurement error, the *iBj* curves in Figure 7.5g are laminar (don't cross) which means that a order-strength model can describe the data. The derived order-strength is not significantly different from item-strength, a single value of strength predicts both sets of data.

Extreme efficiency of the computational attention model. The model shown in Figure 7.5d and 7.5e has only three parameters (detection-interpretation time of the target τ = 222 ms, width of the attention window $\alpha\sqrt{2}$ = 188 ms, and σ, the added noise that describes how much strength values must differ in order to produce a reliable difference in order of report. There are 112 data points in the serial position graphs plus 84 *iBj* data points yielding 196 data points for target outline square. The data points for the other targets (C, U) are modeled by just one additional parameter τ, the detection-interpretation time for each target. These two additional two parameters form a five-parameter model that accounts 83% of the variance in 3 targets × 196 data per target = 583 data points. The two other subjects in this extensive experiment require two more data points corresponding to detection times that vary with the presentation rate of the to-be-reported stream. Predictions of their data were equally good. While none of these predictions is perfect, the predictions demonstrate that a very simple window-of-attention model efficiently captures the essence of this huge data set.

Summary: Subjects were required to detect a target at location A and then to switch attention to location B and report the concurrent or earliest following events at B. The model of an attention glimpse that accounts for their performance is equivalent to (1) the subject takes a mental photo of a rapid stream of letters with a shutter (like the Compur Rapid camera shutter) that opens and closes to admit light in proportion to the value of the attention window $a(t)$; subsequently (2) the subject reports the letters in order of their intensity in the recorded image. The shutter begins to open access to VSTM about 1/4 sec after the cue and then closes gradually for a net exposure duration of about 1/3 sec. The subject has no knowledge of of the order of events in the input sequence other than their strength in the remembered image.

Although the window-of-attention seems like a simple concept, according to Sperling and Weichselgartner (1995), the attention window actually

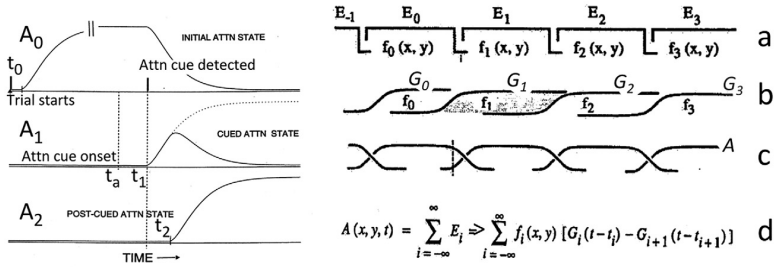


Figure 7.6 Constructing a single attention window and a sequence of attention windows. A_0 is the initial attention state (alertness for the cue to switch attention) which is initiated at time t_0, the start of the trial. The attention cue occurs at time t_a; it is detected at time t_1 at which time the attention window begins to open to initiate the attention state A_1 and thereby to terminate A_0. Very shortly afterward, attention state A_2 begins, closing the attention window thereby terminating A_1. The preparation of the attention-captured items for response output begins in A_2. (a) A sequence of intended attention states i, $(-\infty,i,+\infty)$ that have intended time-course $E_i(t)$ and spatial distribution $f_i(x,y)$. (b) The achieved (versus intended) time course of attention states is represented by a cumulative probability distribution G_i. (c) Each new G_{i+1} terminates the previous G_i thereby producing the sequence of attention states shown in (c). (d) A simple expression represents the sequence of visual attention states from birth to death: $f_i(x,y)$ represents the spatial distribution of visual attention in attention state i, $(G_i - G_{i+1})$ represents its time course. (Reproduced from Figures 2 and 3, pp. 505 and 507, in Sperling and Weichselgartner (1995), APA as publisher, by permission.)

is generated by three consecutive processes (Figure 7.6). The first attention process is searching for the attention-shift target in the cue stream. Detecting the target initiates the second attention process which is opening a window to admit information to VSTM. The third process is closing the attention window and preparing a motor response. The right side of Figure 7.6 illustrates how a series of intended attentions states (Figure 7.6a) is implemented (Figure 7.6b, c) and can be expressed in a simple equation (Figure 7.6d).

7.5 Movements of the Spotlight of Spatial Attention: Continuous or Discrete?

The previous section described the "movement" of subject's attention (from location A where a target was detected in a letter stream to location B where items were grabbed from a numeral stream. This was modeled as a gate

opening to enable items to enter memory, but the process by which attention moved from A to B was not explicitly considered. The movement of attention could have been modeled as an attention spotlight in which stimulus information is available only from locations illuminated by the spotlight. The spotlight is initially pointed at A until a target is detected then continuously moved to B without turning off during movement, but turning off 1/3 sec after reaching B to prevent memory overload. Figure 7.7a illustrates this continuous spotlight model. In the Reeves and Sperling (1986) attention window experiment (Figure 7.5), attention was implicitly modeled as a pair of nonmoving spotlights in which spotlight-1 was pointed at A to illuminate the search array. Once a target was detected, spotlight-1 turned off and spotlight-2 aimed at location B turned on with the time course of the gating-attention function to enable items to enter memory (Figure 7.7b). This experiment does not discriminate between the two models.

In a subsequent study that used the difference between MRTs and ARTs to measure the duration of attention shifts as a function of the distance traversed, Sperling and Weichselgartner (1995), found that the time course of attention shifts was independent of the distance traversed. Independence of distance suggests discrete versus continuous movement of attention. However, Shulman, Remington, and Mclean (1979) had data that indicated attention moved like a continuous spotlight. So, Sperling and Weichselgartner (1995) undertook a reanalysis of the conflicting data.

Figure 7.7a, c shows the formal representation of the continuous spotlight model; (Figure 7.7b, d) shows the discrete (sometimes called quantal) spotlight model. The lower part Figure 7.7 deals with the Posner-style experiment by three of his associates, Shulman et al. (1979). Subjects are required to respond as quickly as possible to the flash of a spot of light that can occur in one of four possible locations (Figure 7.7e). Prior to a light flash (with a foreperiod, t), an arrow is shown that indicates the location of the flash with the probabilities shown in Figure 7.7e. Trials with the left-pointing arrow are interleaved with symmetrical trials with a right-pointing arrow (not shown). The combined data, the facilitation (speed-up) of the RTs due to the four cues are shown in Figure 7.7g as if all arrows pointed left. For very brief foreperiods, there is no benefit of the arrow cue; all flash locations produce approximately the same RT. That is the control condition. As the foreperiod t increases, there is a benefit $\triangle RT$—i.e., a reduced RT relative to the control condition. Figure 7.7g shows the $\triangle RT$ (reduction in mean RT is graphed upward) as a function time for the four locations. Although the leftmost point occurs with probability 0.7, the greatest $\triangle RT$ occurs at the intermediate-left point. The authors interpreted this as indicating that attention moves in the direction of the arrow but never reaches the far location. Interestingly, there is an RT benefit even when the arrow points left and the far right location flashes—i.e., where an RT "cost" is expected.

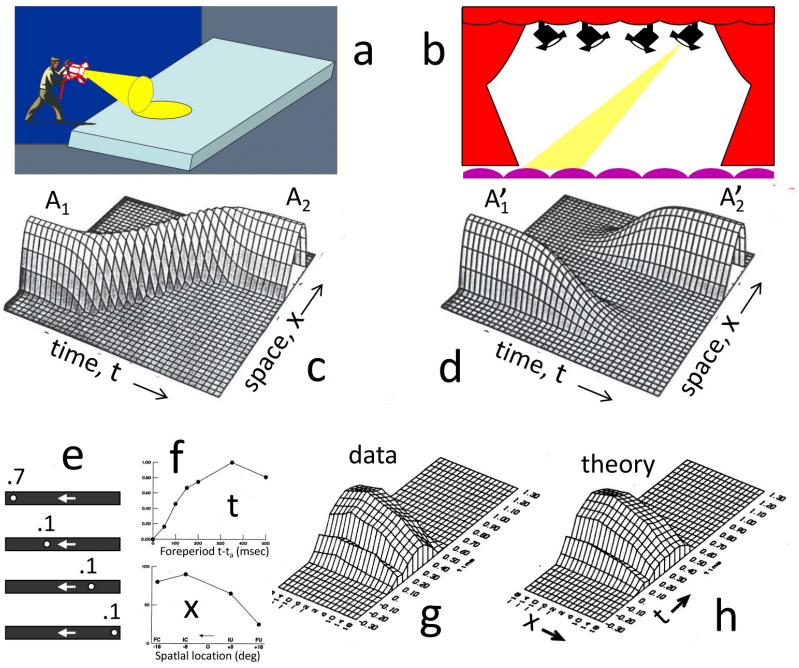


Figure 7.7 Continuous versus discrete shifts of spatial attention. (a) Continuous spotlight controlled by a worker. (b) Pre-aimed spotlights, typically controlled by a computer program. (c) A space-time plot showing light intensity in one spatial dimension, x, as a function of time, t, as the spotlight is moved by a human controller from location A_1 to A_2. (d) A discrete (quantal) spotlight. A space-time plot showing light intensity in x, t when a spotlight pointed at A_1 is turned off simultaneously with the turn-on of a spotlight pointed at location A_2. Because spotlights have thermal inertial, although the power turnoff is instantaneous, the light at A_1 fades out slowly; similarly, the light at A_2 turns on gradually in a fraction of a second, not instantaneously. (e) Schematic illustration of four conditions in the go/no-go reaction-time experiment of Shulman, Remington, and Mclean (1979). In each condition, a central arrow (attention cue) was presented t ms before the flash of a spot of light (the white dot) that occurred at one of four location with the indicated probability (i.e., 0.7, 0.1, 0.1, 0.1). The subject's task was to press a button as quickly as possible upon detecting the light. (f) The temporal, **t**, and spatial **x** attention functions of Sperling and Weichselgartner (1995) to account for the RT data. The vertical axes are $\triangle RT$, the speedup (facilitation) of RT in the attention-cued condition (relative to a control condition) due to temporal t and spatial x facilitation. The horizontal axes are t, the time from the arrow attention-cue to the stimulus flash, and space x (the four possible flash locations). (g) The observed $\triangle RT$ as function of x, t. (h) The $\triangle RT$ predicted as

the product of the spatial and temporal RT attention functions shown in (f). (Panels (a), (b), (e) Copyright G. Sperling, by permission; other panels from Figures 1, 4, 5, pp. 504, 510, 511 in Sperling and Weichselgartner (1995), APA as publisher, by permission; colored version of this chapter: http://www.sowi.uni-kl.de/lachmann/sperling.pdf).

7.6 Spatial and Temporal Attention Combine Independently

Sperling and Weichselgartner (1995) propose a formal model for Shulman et al.'s RT facilitation: The observed RT facilitation $\triangle RT$ is the product of a spatial attention function and a temporal attention function (Figure 7.7f). The moving spotlight model requires that there be a diagonal portion of $\triangle RT(x,t)$ in the 3-D representation of RT facilitation—e.g., as in Figure 7.7c. The quantal spotlight theory as implemented by the product of a spatial-attention function and a temporal attention function requires that all ridges be parallel to one of the axes, as indeed they are. When the RT facilitation data are plotted in 3-D, it is obvious that attention facilitation by the arrow cue has a spatial and a temporal component that act independently. The lack of a diagonal ridge in the 3-D RT facilitation plot together with the ridges perfectly parallel to the axes is unambiguous confirmation of a discrete attention-movement model for data that, without a formal model, were originally incorrectly interpreted as supporting a continuous attention-movement model.

A curious aspect of the data is the decline of temporal facilitation at the longest interval. This probably occurs because there are six equally likely foreperiods and a 0.15 proportion of catch trials with no light flash. When the target light flash has not occurred by the time the sixth longest foreperiod has passed, the probability of a catch trial has increased from 0.15 at the start of the trial to > 0.51 during the sixth foreperiod; this hazard naturally makes observers cautious, and consequently slower at the sixth foreperiod.

Neural implementation of the quantal attention model (Figure 7.8). In the course of practice for an attention experiment, and in the course of real-life experience, a subject learns to respond to a particular cue with a particular state of attention, here instantiated as a set of weights that selectively control the transmission of information between two levels of neural processing, analogous to the attention gate in Figure 7.5d. In a spatial attention experiment, each to-be-attended location corresponds to locations illuminated by a spotlight or combination of spotlights illustrated in Figure 7.7b. During practice trials, the subject learns the appropriate distribution of spatial processing resources for each attention cue. Learning consists of creating a cue-dependent set of weights—i.e., a template of weights α_i in Figure 7.8, that controls the transmission of neural signals between two levels of processing.

We know from anatomy that all sensory pathways from the midbrain and higher, are two-way—there are as many neurons that transmit in the reverse as in the forward direction. From fMRI observations, we learn

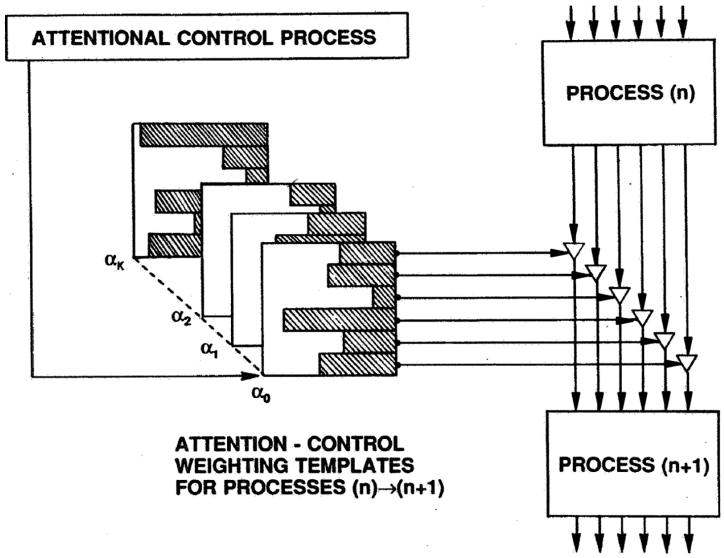


Figure 7.8 Learned templates α_i for the control of attention as represented in a neural network architecture. The attention weight of each item in a template (indicated by the gray area) determines the amount of control signal passed to the gates (triangles) between neural process(n) and process($n + 1$). An Attention Control Process implements the changeover from one learned template α_i to another α_j with a time course as defined in Figure 7.7f. Whereas in a cleverly designed experiment, control may be exerted primarily at a single level of processing, in most situations, many levels simultaneously come under attention control. (Reproduced from Figure 14, p. 527 in Sperling and Weichselgartner (1995), APA as publisher, by permission.)

that these neural feedback pathways are indeed used, that attention-driven signals from the frontal lobes exert effects all the way to the midbrain. So, even in experiments designed to focus on one particular level of processing, neurons at many levels are likely to be involved. The representations in Figure 7.8, and in the various attention models considered here, are conceptual simplifications of more complex brain processes.

7.7 Two Consecutive Attention Events

The examples up to this point have been concerned with a single attention glimpse. What happens when the subject reports from two successive attention events? In the case considered here, the subject is instructed to attend to a single event but is unable to gather the required information from that event and so is implicitly required in a report from two consecutive attention

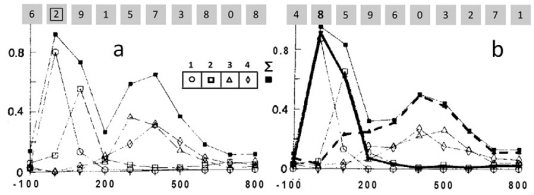


Figure 7.9 Successive attention glimpses. Top: A rapid stream of items (10/sec) from which a subject attempts to report the indicated target item ("2" surrounded by an outline square in (a), "8" highlighted relative to other items in (b)) plus the next three items. (a,b) The abscissa represents the time t of occurrence of items. The target item occurs at $t = 0$. The ordinate indicates the probability of report of items occurring at time t. The symbols indicating the response positions of the four reported items within the response are indicated in the insert; the solid curve connecting filled squares is the cumulative sum of all reported items. Individual data are shown for two subjects. The reported items fall into two groups. The first group of typically two items is centered around the target item, the second group of items has the same distribution as those reported following a cued shift of attention (e.g., Figure 7.5c and Figure 7.5f(9.2)). (b) One subject's self-report of the group (1 or 2) from which an item was retrieved is shown by the heavy lines. (Adapted from Figure 2, p. 779 in Weichselgartner and Sperling (1987), AAAS, by permission.)

events. Figure 7.9 shows the procedure and data from Weichselgartner and Sperling (1987) who's subjects viewed a rapid stream of digits and were required to report a target numeral and the next three numerals from the stream—i.e., to report four items. Targets were either a numeral surrounded by an outline square or a numeral that was brighter than the others. The histogram of reported digits' times of occurrence is clearly bimodal. The first mode consists of the target and the following item, the second mode is virtually the same as in the Reeves and Sperling (1986) attention reaction time experiment, consisting of items clustered around 300–400 ms after the outline square target, and 400–500 ms following the highlighted numeral target. As shown in Figure 7.9b, this subject was able to perfectly identify to which attention glimpse (one or two) the reported item occurred. These data inspired the attention blink experiments (e.g., Raymond, Shapiro, & Arnell, 1992) that investigate the circumstances under which subjects can and cannot report items that follow closely in a rapid stream (RSVP).

7.8 Tracing the Dynamic Trajectory of an Attention Shift

Donald Glaser who shared an office with my undergraduate adviser (Cy Levinthal) was awarded a Nobel prize for inventing the bubble chamber, a significant improvement over the Wilson Cloud Chamber for tracing

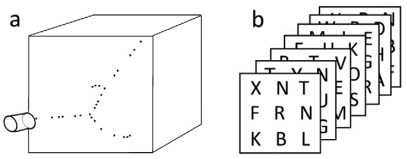


Figure 7.10 (a) Hypothetical trajectory of an unknown subatomic particle that entered the opening of a Glazer bubble chamber (an x,y,z cube) which was embedded in perpendicular magnetic and electric fields. Dots indicate microscopic bubbles produced where the entering particle interacted with water molecules; two splits into component particles are visible. (b) Space-time (x,y,t) cube of consecutively displayed letter arrays. The few letters a subject can report represent the space-time trajectory of attention. ((a) Modified from Figure 1, p. 261 Shih and Sperling (2002), APA as publisher, by permission; (b) Copyright G. Sperling, by permission.)

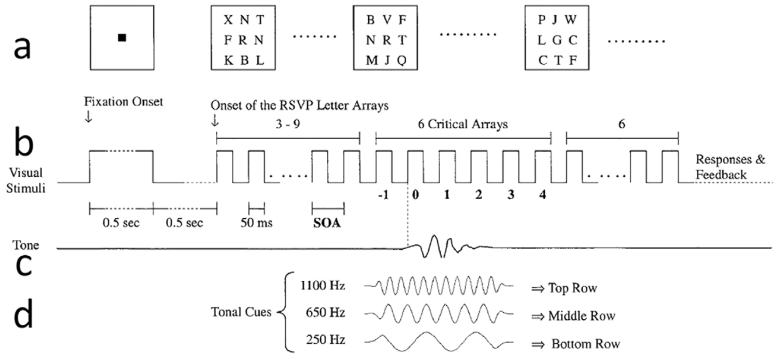


Figure 7.11 Filling an *x,y,t* cube filled with letters to measure the *x,y,t* trajectory of attention. (a) A sequence of 15–21 3 × 3 stimuli. (b) The time course of stimulus exposures. (c) Time course of the tonal cue (schematic). (d) The set of possible tonal cues (duration 200 ms). Procedure: In response to a tonal cue to shift attention, the subject reports the earliest possible three letters from the cued row (top, middle, or bottom). The three reported letters (from among the ≥135 letters presented) are analogous to the bubbles in the bubble chamber; they represent the space-time trajectory of attention (After Figure 3, p. 265 in Shih and Sperling (2002), APA as publisher, by permission.)

the trajectory of subatomic particles. Many years later, he became interested in visual motion perception, we became friendly, he attended AIC conferences, and there I demonstrated to him that tracing the trajectory of attention involved a very similar process to tracing particles in the bubble chamber. To trace the trajectory of attention, one fills a space-time

cube with letters (Figure 7.10). Figure 7.11 illustrates Shih and Sperling's (2002) rapid sequence of 15 to 21 3×3 letter arrays. A subject is instructed to maintain fixation at the middle of the array. At a random time during the sequence, a high, middle, or low frequency tone is presented, and the subject attempts to report the letters in the indicated row from the stimulus that was concurrent with the tone. After the response, the subject is shown a comprehensive feedback display including the number of points earned on that trial—the closer in time the reported letters are to the target array, the more points earned. The letters that the subject actually reports reveal the trajectory of attention.

Figure 7.12 shows the average responses for 7,376 trials of one subject in the attention trajectory experiment. Attention shifts from the fixation point to the top, middle, and bottom row with almost the same time constant, and it shifts simultaneously to all locations. Further analysis showed that letters reported from the different locations within a row are virtually independent—e.g., reporting a letter from the +250 array in the far left position does not make it more likely that the reported letter in the center position on that trial will also be from the +250 array versus the +450 array.

The data in Figure 7.12 provide a remarkable confirmation of the quantal spotlight model (Figure 7.7b). Where a spotlight is pointed does not influence how long it takes to switch from the current spotlight location to the next spotlight location. In terms of the attention model in Figure 7.6, the subject is initially in the attention state of alertness for the tonal signal and remains in that state until the tonal signal is detected and interpreted. There are three possible subsequent states, open the attention gate to admit letters into short-term memory from the top, middle or bottom stimulus row. According to Figure 7.8 (the neural template instantiation of the spotlight model), there is no a priori reason for any of these gate openings to differ significantly in their temporal characteristics, even though the middle row does not require a vertical shift of attention. The quantal trajectory of attention is strikingly different from the sampled continuous trajectories of particles in the Glaser Bubble Chamber.

Dynamics of attention. Summary. The ART experiment illustrated that ARTs can have a very similar distribution of latencies as MRTs. The gating model for temporal attention showed that the window of attention began to open about 1/4 sec after a cue to shift attention, and that the attention window remained open for about 1/3 sec. The analysis of spatially cued RTs showed that a shift of attention corresponded to turning a spotlight off at the initial location and turning a differently directed spotlight on at the second location. In a further analogy with spotlights, the distance between the points at which spotlights are directed is completely irrelevant to the dynamics of the switch from one spotlight to another. The analogous property—independence of distance traversed—was observed for attention by comparing MRTs and ARTs. Independence of distance traversed also observed in the letter-chamber experiment in which tones of different pitch

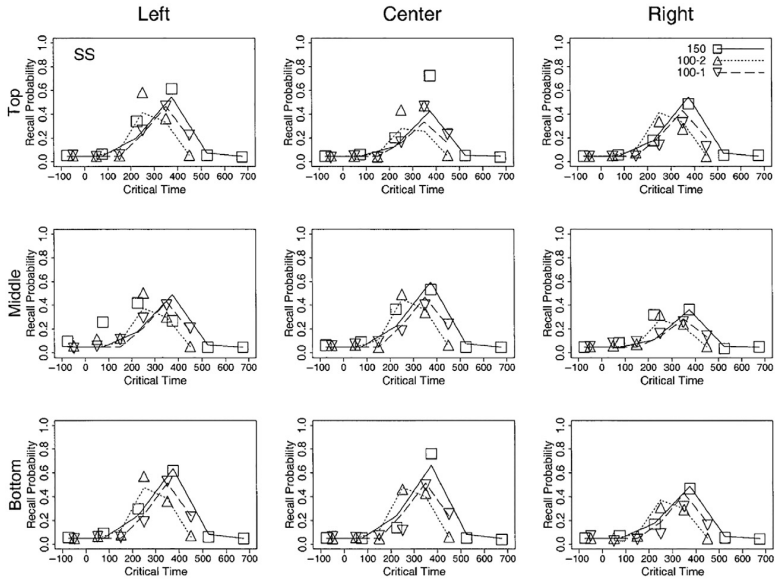


Figure 7.12 Full 3-D-representation of three trajectories of visual attention in response to tonal cues to—as quickly as possible—shift attention to the top, middle, or bottom row. Each row indicates responses to just one of the three tonal cues. Each column indicates the response letters the subject wrote in just one of the three response positions: left, center, and right. The curves represent that proportion of response letters reported from a display that occurred at the indicated "critical time". (The critical time of a display is the midpoint of the interval during which it is visually available—i.e., the midpoint of the time between the onset of the display and the onset of the following display.) Insert at top right identifies different symbols for the different rates of stimulus presentation in terms of the SOA (stimulus onset asynchrony, the time from the onset of one display to the onset of the next); 100-1 and 100-2 refer to the 100 ms SOA stimuli before and after extensive practice. Curves are best-fits from a model similar to that in Figure 7.5d. The overlap of the data within a panel and the similarity of the response curves between panels indicate that attention shifts to the top, middle, and bottom rows with the same temporal dynamics, and that it shifts simultaneously to each stimulus location (left, center, right)—i.e., there is no left-to-right attention scan. (Reproduction of Figure 12, p. 266 in Shih and Sperling (2002), APA as publisher, by permission.)

cued attention shifts to different locations. Whether attention remained at the initial location (middle row, zero distance traversed) or shifted to a different location (top or bottom row) made no difference in the dynamics of the shift. And, although items within a row are typically read from left-to-right and reported from left-to-right, attention shifted concurrently to all items in the to-be-attended row. The experiment on two attention glimpses was consistent with the aforementioned and further showed that two consecutive attention episodes follow after about 1/3 sec.

Saccadic eye movements typically present the visual system with three new images per sec. All the earlier experiments were conducted without intentional eye movements, but the dynamic properties of attention that were observed seem perfectly in tune with the dynamics of saccadic eye movements. It seems quite likely that visual spatial attention evolved to normally coordinate with eye movements, and perhaps also vice versa. This concludes the analysis of the dynamics of attention, the subsequent sections deal with the filtering properties of selective attention.

7.9 The Attention Operating Characteristic

In the experiments considered so far, subjects attended to one area or one instant in time and attempted to ignore everything else. When a subject attempts to perform two tasks concurrently, how does performance change as the attention is selectively allocated to one task or the other or equally to both? Historically, the two ways of investigating performance in multiple tasks are illustrated by two search tasks, searching for a target in area A, searching for a target in area B. In a *compound* task, the search is for a target that can occur in either A *or* B; in a *concurrent* task, the search is for two targets, one in A *and* one in B (Sperling and Dosher, 1986). The compound task has greater uncertainty (noise) than either of its component tasks, therefore, even an ideal observer has a deficit. Therefore, to determine that a human observer's deficit also has an attentional component requires an ideal observer model. This considerable complication is bypassed in concurrent tasks. Examples: Driving a car while listening to the radio, driving while texting on a smartphone.

The prototypical example of visual spatial attention that is considered here uses a search task and stimuli generally similar to those of the previous experiments herein. Subjects search the rapid stream of stimuli (Figure 7.13) with attention directed primarily to the inside or the outside or equally to both streams to find the two target items in an embedded critical array, one inside target, one outside target. In control conditions the subject search for only one target. All control conditions were designed to be approximately equally difficult so that difference between concurrent conditions could be attributed to the compatibility of the two tasks.

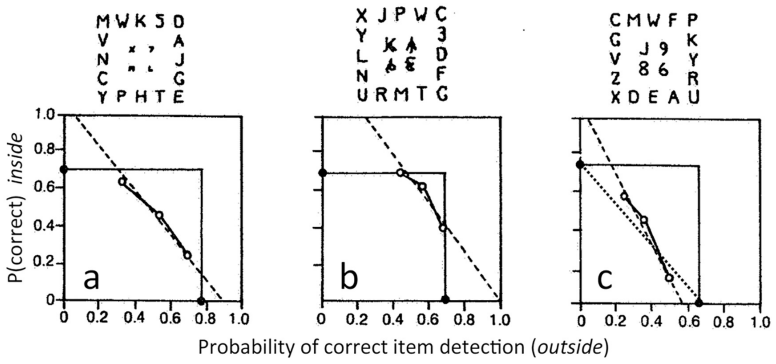


Figure 7.13 Attention operating characteristics, AOCs, determined for three different search arrays. Subjects view a rapid stream of 20–25 arrays that are composed entirely of random letters except, in the middle of the stream, there is a target array (shown) that has one numeral in the inside group and another numeral in the outside group. (a) The inside letters are much smaller than the outside letters. (b) Inside letters are the same size as outside letters but masked by a noise character. (c) The role of letter and numbers is reversed only in the inside where the target is a letter among numerals. The task is to detect both targets. For each set of stimuli, each of three attention conditions was run in a homogenous block of trials: Most attention outside, equal attention to both, most attention inside. In control blocks, the subject reported only the inside (or only the outside) target throughout the entire block. Data are shown for one subject: the percent of correct detections in attention conditions (open circles) and correct detections in control conditions (filled points on axes). The intersection of the lines from the control conditions represents the independence point, where tasks are performed perfectly without interference. The solid lines are the experimentally determined AOCs. The dashed lines are straight-line fits. The closer the AOC lies to the independence point, the more compatible are the two concurrent search tasks. The dotted line connecting the two control conditions in (c) represents the performance expected when the subject performs only one task on a trial and flips a coin that has probability p, $0 \leq p \leq 1$ to determine whether to do task 1, otherwise do task 2—i.e., total incompatibility. (After Figure 2.15, p. 2–28 in Sperling and Dosher (1986), Wiley, by permission; originally from Figure 1, p. 216 in Sperling and Melchner (1978).)

Figure 7.13 shows the data of one practiced subject. In Figure 7.13b (noise-masked, same-size letter in inside), performance is close to the independence point (the corner of the embedded rectangle where there is no loss in dual- versus single-task performance). When attending primarily to inside or outside, in this condition, the subject does as well on the primary task as in the control condition, and has impaired but still good performance

on the secondary task. By analogy to driving and listening to the radio, giving primary attention to driving would leave driving unimpaired even when the radio is playing, listening would be somewhat impaired.

Figure 7.13a shows that searching for small targets on the inside and large targets on the outside are considerably less compatible tasks than searching for the same-size targets in both locations (Figure 7.13b). Even when attention is directed mostly to one location, there is a performance deficit at the attended location relative to the control condition.

Figure 7.13c shows data from the most extreme case, total incompatibility when searching for a numeral among letters in the outside location and a letter among numerals in the inside. The data fall nearly on the line representing an arbitrary choice on each trial of which single task to perform in the dual task experiment. The deviation above the line in attending to the inside task (detect letter among numbers) suggests that the subject can perform slightly better on the overlearned detect-a-number-among-letters task than expected from the fraction of resources devoted to that task. This situation represents the extreme case in which the two tasks are so incompatible that while performing one, the other is impossible. To perform in a continuing dual task situation, one would have to alternate attention between the tasks. For some persons, driving and texting might represent such an extreme situation. Rapid alternation between the tasks would enable some degree of joint average performance, but at any given moment, only one of the two tasks can be performed. Task alternation would involve the dynamics of attention switching that were considered in the previous sections.

Three additional facts. (1) When the two targets in the incompatible dual task (Figure 7.13c) do not occur in the same display, performance recovers to control condition levels when the SOA (time between the onsets of the two target-containing displays) is 240–480 ms. This task-switching time (from detecting target of type 1 to detecting target of type 2) perfectly matches the attention reaction time ART data in Figure 7.5c, which measures the switch from searching for and detecting a target in one stream and then "grabbing" an item from an adjacent stream. (2) When only a single target frame is presented and followed by a noise frame (at the same SOA as in the rapid streams), the data are statistically identical (unpublished). (3) The AOC is isomorphic to the ROC (receiver operating characteristic of Signal Detection Theory (Sperling, 1984; Sperling & Dosher, 1986). In conventional ROC graphs the "conventional origin" (worst performance) is at the lower right whereas, in AOC graphs, the origin is, as usual, at the lower left.

7.10 Resources and Utility in Dual-Task Performance

The AOC Figure 7.14 illustrates the possible outcomes when two tasks compete for the same limited resource. Competition for limited resources

is universal, and so are the algorithms that describe it. A simple example that applies universally is illustrated competition in Figure 7.14 which illustrates a limited time resource in the case of two overlapping classes, one from noon to 3:00 p.m., the other from 2:00 to 4:00 p.m. The shared resource, needed for both classes, is the overlap time from 2:00 to 3:00 p.m. How a student allocates this resource to each class depends on the utility function: Getting the highest average grade requires a different allocation than passing both courses, i.e., getting the same, maximum grade in each course. This example illustrates one important reason to specify an explicit utility function for subjects, rather than to simply telling them, for example, to allocate 90% of your attention to task 1 and 10% to task 2.

Strategy microprocesses: Alternating versus sharing. In the classroom Attending Operating Characteristic example, the shared resource, time, was divided between the overlapping classes on every day. Suppose that both classes met for 15 days during the term. To get the maximum average grade, the student has to attend all of Class-2 so the strategy is the same whether the classes occur just once or many times. However, to pass both courses with a grade of 80%, the critical resource, the

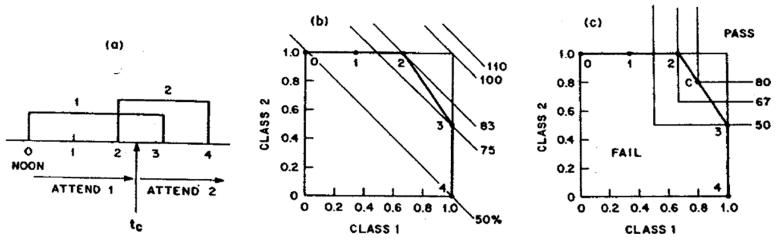


Figure 7.14 Resource allocation in dual attention tasks: The classroom attendance example. (a) Class-1 is offered from noon to 3:00 p.m., Class-2 is offered from 2:00 to 4:00 p.m., a student needs to attend both. The strategy is to attend Class-1 until time t_c, then to quickly switch to Class-2. (b) The AOC, the fraction of the total class time the student spends in each class as a function of t_c. Assume the grade achieved in each class is directly proportional to the fraction of time attended. The diagonal lines are equi-utility contours that represent the average grade (assuming the two grades are weighted equally) associated with that point on the graph. Switching at 2:00 p.m., yields the highest achievable average grade, 83%. (c) Assume that grading is pass-fail. A utility functions divides the graph into pass and fail regions that depend on the passing grade. The highest passing grade that still enables the student to pass both courses is 80% which is achieved by switching classes at precisely 2:24 p.m.—i.e., by attending 80% of each class. (After Figure 2.2, p. 2–8 in Sperling and Dosher (1986), Wiley, by permission.)

Table 7.1 Contingency table for the joint distribution of responses in tasks 1 and 2 of a dual task experiment. *P1* and *P*2 are the probabilities of correct performance; the probabilities in the cells represent independent performances in Task 1 and 2; *K* is an indicator of the deviation from independence. A negative *K* indicates a negative correlation between performances in the two tasks, typically produced by alternating the access to a shared resource from trial-to-trial between task 1 and task 2 ("switching"). (After Table 1, p. 317 in Sperling and Melchner (1978), AAAS, by permission.)

		TASK 1	
		Wrong	Right
TASK 2	Wrong	$(1-P_1)(1-P_2)$ $+K$	$P_1\,(1-P_2)$ $-K$
	Right	$(1-P_1)(1-P_2)$ $-K$	P_1P_2 $+K$

$1-P_2$ (row 1) P_2 (row 2)

$1-P_1$ (col 1) P_1 (col 2)

overlapping hour had to be shared 36/60 for Class-2 and 24/60 for Class-1. Whereas this division was accomplished by switching daily at 2:24 p.m., it could also be accomplished by attending all of Class-2 (switch at 2 p.m.) on three-fifths of the days (e.g., Mon, Wed, Fri) and all of Class-1 (switch at 3 p.m.) on the remaining days (Tue, Thu). Conventionally, this alternating strategy is called switching.

Contingency analysis. Switching and sharing strategies in the classroom example can be distinguished by looking at performance on test questions that ask about the material presented during the overlap hour, the shared resource. Suppose tests in each class always ask at least one question about material that was presented at random times in the overlap hour. Except by chance, the alternating student will never answer both questions correctly because he was never in both classrooms on the same day. However, the sharing student will have a three-fifths chance of answering the question from Class-2 and a two-fifths chance of answering the questions from Class-1.

Attending two classes on the same day is analogous to a a subject making two responses in a dual-task experiment. It makes possible some very useful analyses. Consider, for example, a 2×2 contingency table (Table 7.1) an alternating student would be expected to exhibit a perfect negative correlation, whereas a sharing student would exhibit complete independence. When this contingency analysis was applied to the dual task divided attention data in Figure 7.13, the major mechanism for dividing attention was switching. In the classroom example, alternation occurred between two states: "Attend all of Class-1" and "Attend all of Class-2". For the data in Figure 7.13, Sperling

and Melchner (1978) reject the hypothesis that the alternation occurs between just two states, but the data are not rich enough to yield a more detailed analysis. One methodological conclusion, demonstrated repeatedly earlier, is that obtaining multiple responses on each trial makes possible analyses that would otherwise be impractical or impossible.

7.11 Using Ambiguous Motion Displays to Measure the Amplification of Attention

Figure 7.13 illustrates two ambiguous motion displays. The motion in these stimuli is invisible to the first-order motion system, which computes motion on a space-time map of local Weber contrast, a quantity that varies from -1 to $+1$ that is often mistakenly called luminance. That is, all the stimuli in these experiments have been calibrated to be equiluminant for (and therefore invisible to) the first-order (Fourier) motion system (e.g., Anstis & Cavanagh, 1983; Lu & Sperling, 2001a). However, motion in these displays is easily perceived by the third-order motion system which computes motion on a space-time map of salience, i.e., a salience field in which areas that are perceived as figure versus areas that are perceived as ground are assigned different values (Lu & Sperling, 1995, 2001b).

In the stimulus of Figure 7.15a, the saturated green stripes in frames 1, 3, 5 differ more from the yellow background (which consists of an equal mixture of red and green) than the unsaturated red stripes, and therefore the green stripes have higher salience. In frames 2 and 4, the areas of high-contrast texture have greater salience. Therefore, when these frames follow on top of each other in rapid succession, motion from right-to-left is perceived between these consecutive high-salience areas, as indicated in the figure. In the stimulus of Figure 7.15c, the color grating's stripes are of equal saturation and therefore of approximately salience. Typically, the perceived motion direction, if any, of these stimuli is ambiguous. However, Lu and Sperling (1995) showed that attending to a feature causes that feature to be more salient in motion paradigms. When the subject is instructed to attend to green and to follow the motion of the green stripes in the stimuli of Figure 7.15c, motion is perceived from right to left, just as in Figure 7.15a. Attention to the color green has increased the salience of green. Note that attending only to green stripes without automatic attention to the texture would yield ambiguous motion; it is the combination of the high salience green with the high salience texture that produces consistent right-to-left motion in these stimuli. Similarly, in the stimulus of Figure 7.15c, attending to the red stripes produces perceived motion in the opposite direction from attending to green.

Blaser, Sperling, and Lu (1999) varied the relative saturation of the red and the green stripes in stimuli like those of Figure 7.15 over the full range. In some stimuli, red was more saturated (red advantage) and others the green

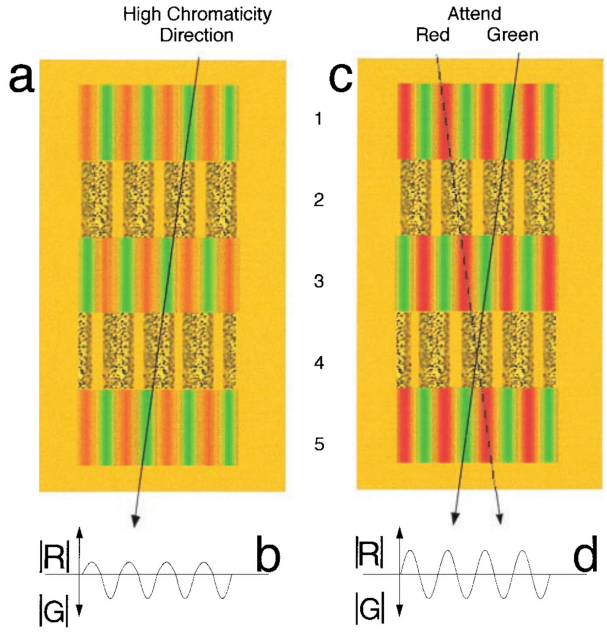


Figure 7.15 Manipulating color salience by varying color saturation. (a,c) Two ambiguous equiluminous motion displays, one in which the color saturation of the red component of the red-green color grating is low, another with a equally saturated red and green components. A stimulus consists of a single grating (color or texture); the numbers represent the sequence in which the stimuli are displayed one on top of the other. The lines represent the direction (in space and time) in which the stripes of highest salience in each successive grating are correlated to produce apparent motion. With neutral attention (no instruction to attend to a particular color), the stimulus sequence (a) produces apparent motion to the left whereas (b) is ambiguous. Attending to red or to green in (a) or (b) increases the salience of the attended color by about 25% and thereby increases the tendency to perceive apparent motion in the indicated direction. (b,d) A graphical representation of the magnitude of saturation (difference of color from the neutral yellow background) of the stripes in the chromatic grating. (After Figure 1, p. 11682 in Blaser et al. (1999), Copyright (1999) National Academy of Sciences, USA, by permission; colored version of this chapter: http://www.sowi.uni-kl.de/lachmann/sperling.pdf.)

was the more saturated (green advantage). Subjects judged the direction of apparent motion under three attention conditions: First a neutral condition (no attention instruction given), then sessions were conducted with attention to a particular color, and, finally, sessions were conducted with attention to the opposite color. In each attention condition, the subject viewed the stimuli from four distances, beginning with the nearest, then doubling the distance three times to produce visual stimuli of 0.5, 1, 2, and 4 cycles/deg. The data in Figure 7.16 show that the perceived direction of the ambiguous motion stimulus depends on the attention instruction. The exception is that two of the three subjects could not perceive consistent stimulus motion in the smallest stimuli—i.e., at the furthest viewing distance.

All the data in Figure 7.16 are very well described by a simple motion model (Blaser et al., 1999, Figure 3, p. 11684). The input to the motion computation is salience as a function of space (x) and time (t). Salience is equal to one for the local areas of greatest texture contrast and for color areas of maximum saturation, and is proportional to texture contrast and to color saturation in other local areas on a scale (0,1), where zero is the salience of the yellow background. The input is multiplied by two factors: (1) a spatial filter that determines how the input is reduced as viewing distance increases and (2) an attentional gain factor that is 1 for neutral attention and achieves the values shown in Figure 7.16 for red and green attention—i.e., red gain is a factor of 1.29, green attention gain is 1.46 for the expert observer shown in Figure 7.16. Finally, noise is added before the motion computation. A subject's model parameters are the same for the four viewing distances, three attention instructions, and five values of red-green stimulus red advantages, and the model accounts for 99% of the variance of the data.

Salience versus appearance. Insofar as apparent motion is concerned, dull red stripes of a stimulus under red attention behave like more intensely saturated red stripes under neutral attention. This is due to an attention-produced change in salience, not to a change in appearance. For stimuli that are well above threshold, attention makes stimulus judgments more accurate. Attending to a pink stimulus makes it more prominent but it does not make it appear more red, on the contrary, it makes the judgment of its pinkness more accurate (Prinzmetal, Amiri, Allen, & Edwards, 1998).

Trained salience persists. Tseng, Gobell, and Sperling (2004) trained some subjects to search briefly flashed displays for target characters on red squares among distractor squares of other colors, and other subjects to search for targets on green squares. Training to asymptotic performance took four to seven hours. Subsequently, when viewing ambiguous motion displays similar to those in Figure 7.15, the red- and green-trained subjects perceived motion in opposite directions. In some subjects the selective increase in salience of the trained color persisted for more than a month and had to be reversed by search-training the opposite color. This further

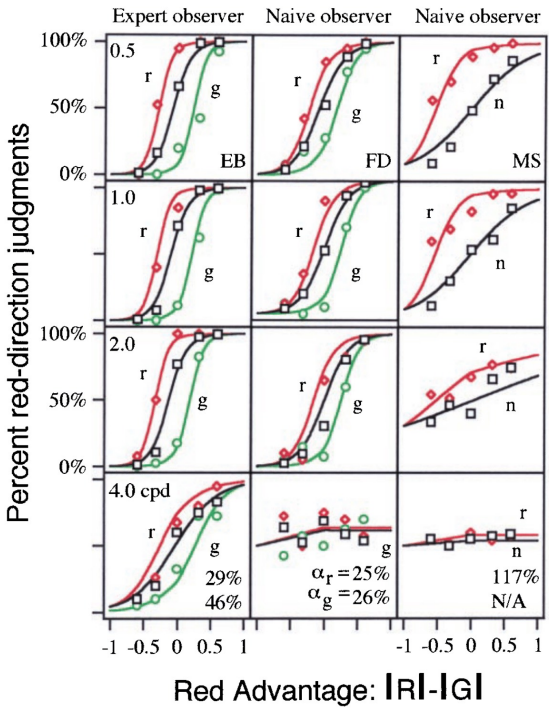


Figure 7.16 How attending to red (diamonds), to green (circles), or being attentionally neutral determines the perceived direction of ambiguous red-green motion displays. Each panel shows the data from one observer in three attention conditions. Each row represents gratings of particular retinal size (spatial frequency) produced by varying the viewing distance. All stimuli are equiluminant. The abscissa is the "Red Advantage", the color saturation of red minus the saturation of green. The ordinate is the percent of motion judgments that were in the red direction. The colors of the points indicate the attention condition (diamonds = red, circles = green), black represents neutral attention. The curves are derived from a simple model that subjects the input stimulus to a high spatial frequency cutoff filter, introduces a small level of random noise, and then computes motion direction based on the surviving salience-strength of the stimulus areas. The numbers in the bottom panels indicate the increase in color salience produced by the attention instruction. All the curves for a subject are produced by the model for that subject—i.e., all panels use precisely the same parameters (a spatial filter, attentional gain, and internal noise). (After Figure 2, p. 11683 in Blaser et al. (1999), Copyright (1999) National Academy of Sciences, U.S.A, by permission; colored version of this chapter: http://www.sowi.uni-kl.de/lachmann/sperling.pdf).

demonstrates that the ambiguous motion paradigm is an extremely sensitive measure of salience and thereby also of attention.

7.12 The Resolution of Spatial Attention

The resolution-quality of lenses is characterized by how accurately they can image gratings of different spatial frequencies. Although human visual acuity typically is clinically evaluated by the ability of subjects to identify letters of different retinal sizes on a Snellen chart, visual acuity is better studied with the same gratings used to study lenses. The advantage of studying a linear system, like a lens imaging system, with gratings of different spatial frequencies is that, once the responses to gratings of the relevant range of spatial frequencies is known, the response to any other stimulus can be calculated. According to Fourier's theorem, any periodic stimulus can be represented as a sum of orthogonal sine waves. In a so-called linear system, the response to any arbitrary stimulus is simply the sum of the system's responses to each of the sinewaves of which the stimulus is composed. Gobell, Tseng, and Sperling (2004), applied this logic to derive the spread function of spatial attention. Figure 7.17, shows the method and results.

A spatial-attention instruction-cue to a subject is a two-color display (Figure 7.17a) in which one color denotes the areas to be attended, the other the areas to be ignored. Subsequently, the subject is briefly shown a 12×12 array of disks in which a target (a larger disk) occurs in a to-be-attended area. To insure that the subject doesn't simply attend uniformly to the whole field, ten false targets (disks identical to the target) occur in the to-be-ignored areas. The subject has to scrunch attention into the to-be-attended areas and to suppress inputs strongly for the to-be-ignored areas. The attention cues are gratings of four different spatial frequencies. There are two gratings of the lowest spatial frequency (one cycle per stimulus) representing two different phases, one with the attended area in the center, the other with the attended area in the flanks. On a trial, a random 1 of the 20 possible attention cues is presented and target location is random. All attention cues have equal to-be-attended and to-be-ignored areas.

In Figure 7.17d (Results), the percent of target detections is shown separately for vertical and horizontal attention cues for one typical subject. Performance declines as spatial frequency increases, i.e., as the to-be-attended area is divided into more and more smaller and smaller subareas. The spatial attention model (Figure 7.18) is essentially the same as the temporal attention model except that the stimuli and the attention cues are spatial instead of temporal. It gives a good account of the data (Figure 7.17d) which are the percent correct target detections in each of the 72 target locations for each of the 20 different attention cues (1,440 data points). Only highly aggregated data are illustrated here.

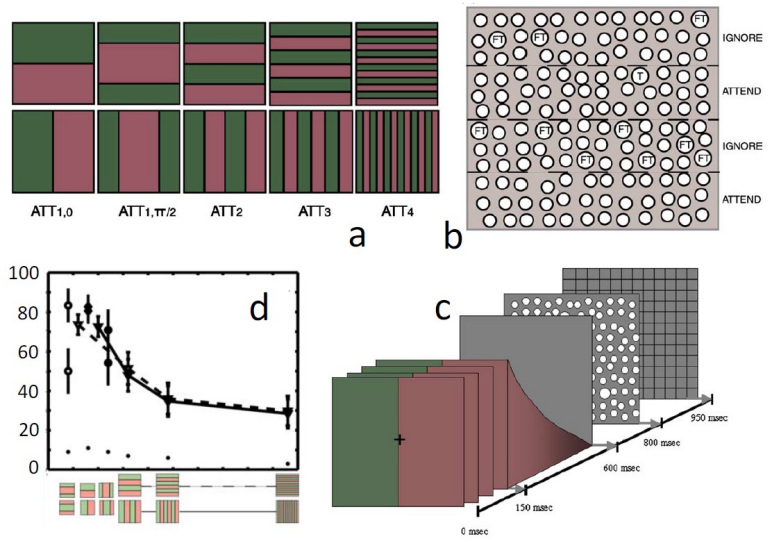


Figure 7.17 The procedure for measuring the spatial resolution of attention by determining its spatial frequency response and thereby the spatial attention system's impulse response. (a) Ten of the 20 attention cues. The other ten cues simply have the colors reversed. For half the subjects, the target appears only in a red area (light gray), the other subjects attend to a target in green areas (dark gray). (b) A sample stimulus shown schematically. The search target *T* is a slightly larger disc that appears in an area colored in the attended color. Ten *FT* (false targets) are disc identical to the target but located in unattended areas. The false targets force subjects to shape their spatial attention to the to-be-attended areas. The dashed lines and labels are for illustration only and are not present in the displayed stimulus. (c) Schematic illustration of a single trial. The attention cue display fades gradually because when it is turned off suddenly, the subject sees an afterimage in the opposite colors of the cue. The attention cue is followed by a 200 ms blank interval, the target-containing stimulus, a response grid, and, not shown, a feedback display. (d) Data for one typical subject: The percent of correct target detections as a function of the spatial frequency of the attention cue. Error bars represent 95% confidence intervals. The continuous curves are from an attention model, the solid curve is for vertical attention gratings, the dotted line is for horizontal attention gratings. Dots at bottom of graph indicate chance performance. (After Figure 1, p. 1277 and Figure 2, p. 1280 in Gobell et al. (2004), Elsevier, by permission; colored version of this chapter: http://www.sowi.uni-kl.de/lachmann/sperling.pdf.)

The model applies to any requested spatial distribution of attention. Once the parameters of the model of Figure 7.18 have been determined—e.g., for the stimuli illustrated, the model can process any arbitrary requested distribution of attention for disks of the kind illustrated here, and with modifications could presumably be applied to a much wider variety of search tasks. In preliminary experiments (Hsu, Scofield, & Sperling, 2006; Sperling, Scofield, & Hsu, 2008) quite complex random patterns of requested attention distribution were tested, two of which are shown in Figure 7.19. Although the model makes precise parameter-free predictions for the 72 possible target locations in these stimuli, the predictions capture only about half the data variance for the simpler stimulus (Figure 7.19b), and are worse for the more complex stimulus (Figure 7.19c). The conclusion is that the simple spatial-resolution-of-attention model captures some important elements of spatial attention, but spatial attention is more complex than the model.

Attention is a feature. Minor variations of the paradigm provide profound insights into the nature of attention. In particular, when the colored cue stimulus remains visible throughout the stimulus exposure, the task is reduced to finding the target on the attended color, not merely on a remembered area previously designated by the color. In terms of performance, making the color available throughout the exposure provides a benefit but it is confined primarily to the higher spatial frequency displays (Gobell et al., 2004, Exp. 4, pp. 1285–1287). The spatial resolution of an active perceptual system is greater than the spatial resolution of remembered areas-to-be-attended. However, comparing the tasks also suggests that, at a neural level, selective attention is a feature, just like color. The task of the nervous system is to give more weight to an item that is a large disk AND some other property; that other property can be "on a red area" or "on an attended area". The subsequent decision process is concerned only with the salience weight of an item, not how it was computed. Other, parallel processes, of course, do take note of the color of the background, of an item's location, and many other aspects of the display and the situation.

7.13 Measuring Feature Attention Filters With the Centroid Paradigm

So far, the focus has been on temporal and spatial attention: Here, the focus is on attention to features, describing the mechanisms of feature-based attention (FBA) and measuring them quickly and efficiently. The particular instance of feature-based attention is attention to a color. The measurement method is the centroid paradigm (Drew, Chubb, & Sperling, 2010), which was extensively elaborated specifically for measuring FBA by Sun, Chubb, Wright, and Sperling (2016a). It is well established that FBA operates broadly across space, heightening sensitivity to the attended feature even

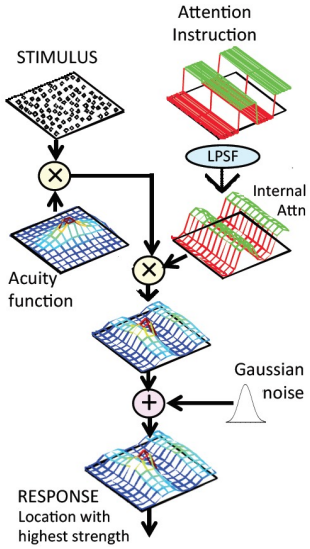


Figure 7.18 A model of the spatial resolution of visual attention applied to the results of Gobell et al. (2004). As with all attention models, there is a path for the stimulus to an attention gate (×) where the selective-attention process initiated by the attention cue determines what information passes and what is attenuated—i.e., an attention filter. The 144-disk stimulus flashes for 200 ms to enter the stimulus path where it persists briefly. Each stimulus distractor disk has a weight 1, the target and the ten false-target disks (slightly larger than the distractors) have weights of $1 + w_t$, all stimulus weights are multiplied by the spatial acuity function which has weight 1 in the center and less elsewhere. The Attention Instruction cue assigns a weight of one to the attended areas and zero to the others. The Low Pass Spatial Filter (LPSF) attenuates the amplitudes of high spatial frequencies in the Attention Instruction cue. The effect of low pass frequency filtering on the requested Attention Instruction cue is illustrated schematically as Internal Attention, which represents the achieved weights of attended and unattended areas. The net effect is to reduce the difference w_t between targets and distractors, particularly near attend-ignore boundaries. The resulting Internal x,y map of filter weights is stored in memory and applied continuously to the Attention gate where it multiplies the input weights. Independent normally distributed random numbers are added the 144 disks, and the Decision process chooses the location of the disk with the highest net value (target or non-target) as the model output. Model performance is shown in Figure 7.17d. (Modification of Figure 9, p. 1289 in Gobell et al. (2004), Elsevier, by permission; colored version of this chapter: http://www.sowi.uni-kl.de/lachmann/sperling.pdf.)

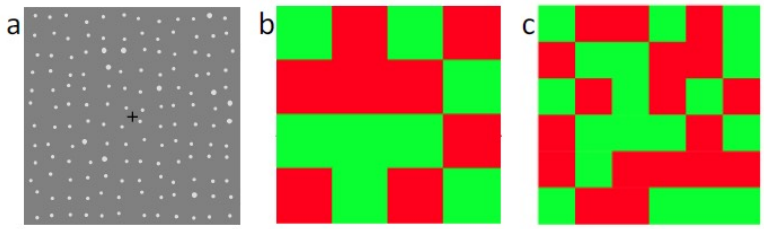


Figure 7.19 Two examples of complex attention attention cues. (a) A 12 × 12 jittered array of disks with 133 distractor disks and, slightly larger, 1 target disk and 10 false-target disks. (b) An attention cue indicating that the target will occur in one color (e.g., red (dark gray) in a red-observing session) and the false targets in the other color. (c) An even more complex attention cue. (Copyright G. Sperling, by permission; colored version of this chapter: http://www.sowi.uni-kl.de/lachmann/sperling.pdf.)

at locations that are irrelevant to the task at hand (for references, see Sun et al., 2016a). This global property of FBA was exploited by Sun, Chubb, Wright, and Sperling (2016b) in a centroid paradigm to derive human attention filters for a single color among a wide range of distractor colors. Just as the physical description of a color filter describes the relative transmission of the filter for each wavelength of light, a color-attention filter describes the relative effectiveness with which each color in the retinal input ultimately influences performance. Here we consider a simpler instance of feature attention, namely the extent to which attention to dark dots (targets) among light dots (distractors) enables subjects to ignore the light dots, and vice versa.

In the centroid paradigm, subjects are briefly shown a cloud of items, in this instance, dots, and asked to move a mouse-controlled pointer to the centroid of the dots. The centroid is defined as the Euclidean center of mass. If the dots were coins on a weightless sheet, the centroid is the point where the sheet would be perfectly balanced if placed on a fulcrum. Judged centroids are an instance of a statistical summary judgment, a property of the display that is perceptually available even though the individual components—in this case the dots—are not. Figure 7.18 shows the procedure. When subjects are given immediate feedback, they learn the task in just a few trials, and reach asymptotic performance in several hundred trials. As a trial takes just 3 to 4 sec, learning takes less than 15 minutes.

Figure 7.20 shows the procedure and results of an experiment (Drew et al., 2010) in which subjects judged the centroids of 8 and of 16-dot clouds under three attention conditions: Give equal weight to all dots, give equal weight to all dark dots and zero weight to light dots, give equal weight to all light dots and zero weight to dark dots. Subjects

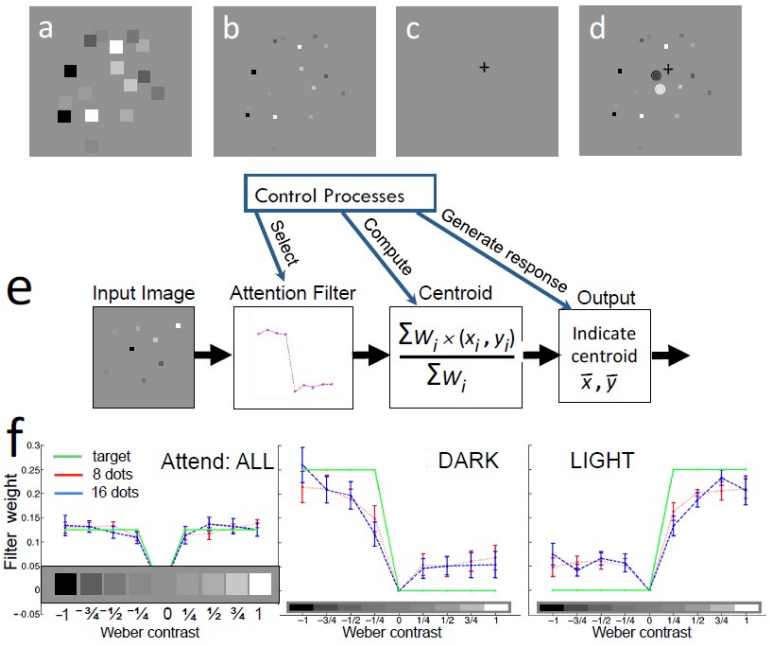


Figure 7.20 Using the centroid paradigm to measure attention filters for features: dark versus light dots. (b) A sample 16-dot stimulus that has 2 dots of each of 8 Weber contrasts. (a) A representation of the stimulus in (b) with the dots replaced by large squares to better indicate the contrasts of the dots. (c) After a brief flash of the stimulus (b), a cross-hairs cursor display appears. The task of the subject is to use a mouse to move the cursor to the remembered centroid of the stimulus. (d) Feedback display. The illustration shows the subject's judged centroid, the correct centroid, and the stimulus. The two centroids shown in (d) represent the centroid of only the dark dots and the centroid of only the light dots. However, only the one task-relevant centroid is displayed to the subject—i.e., the centroid of all dots, only the dark dots, only the light dots. (e) Model for computing an attention filter. The stimulus is represented perfectly as an input to the model—i.e., the exact location and contrast of each dot is given. An attention filter assigns a weight to each color of item in the input stimulus. Illustrated here is an estimate of an attention filter that attempts to give equal weight to all dark dots and zero weight to all light dots. The model then computes a perfect centroid using the dot weights assigned by the filter. The output of the model is then compared to the subject's response for that stimulus. Simple linear regression gives the optimum attention filter for predicting the subject's response. (f) The average attention filter for 11 subjects in three attention conditions: Give equal weight to all dots, give equal weight to all dark dots and zero weight to light dots, give equal weight to all light dots

and zero weight to dark dots. Shown are the ideal attention filters and those achieved for 8-dot and 16-dot displays. Error bars are 95% confidence intervals for the population means. (Panels (a–d) Copyright G. Sperling, by permission, panels e, f after Figures 5 and 7, pp. 5 and 7 in Drew et al. (2010), ARVO, by permission; colored version of this chapter: http://www.sowi.uni-kl.de/lachmann/sperling.pdf.)

performed the centroid task quite well and achieved reasonable attention filters, weighting attended (target) dots about three to four times more than distractor dots. More detailed measures of performance are described in Sun et al. (2016a)—e.g., efficiency (the fraction of stimulus dots an ideal detector would need in order to match a subject's performance), filter fidelity, selectivity (the ratio of the weights of attended to unattended dots), data driveness (weight of a Bayesian prior versus stimulus information).

The beauty of the centroid paradigm is (1) that the attention filter is a complete quantitative description of selective attention to a feature in a particular context. The filter assigns an attention (salience) weight not only to the target feature but also to each of the unattended features in the display. (2) Deriving an attention filter with the centroid paradigm is an order of magnitude quicker than traditional attention paradigms that yield less detailed measures of feature attention.

7.14 Summary and Conclusions

Data and computational models for spatial, temporal, and feature attention were reviewed. The mechanism of selectively attention is the selective transmission of only a subset of the information between stages at different levels of information processing. Although attention normally simultaneously affects many stages of neural information processing, in the experiments considered here, the influence of selective attention was largely focused on one particular stage of processing. In the spatial attention experiments, information was selected from a particular region, not necessarily a connected region, for subsequent entry into memory (Shih & Sperling, 2002; Sperling, 1960) or for detection (Gobell et al., 2004), or for speeded reaction time (Shulman et al., 1979). Temporal attention to a brief interval in time enabled selected items to enter memory Reeves and Sperling (1986) and to speed RT (Sperling & Weichselgartner, 1995). Feature attention enabled (1) selected-feature-containing areas to influence the direction of apparent motion (Blaser et al., 1999) and (2) computation of the centroid of only the selected items (Drew et al., 2010). When both temporal and spatial attention were involved, the selection was determined by the product of independent spatial and feature filters (Shih & Sperling, 2002; Sperling & Weichselgartner, 1995). When a subject attempts to perform two concurrent tasks, an AOC describes

the performance as a function of the amount of attention allocated to each task (Reeves & Sperling, 1986; Sperling & Dosher, 1986). The brain processes by which attention is accomplished are quite complex. Nevertheless, the processes of attention listed above were accurately and efficiently described by the computational models herein.

Information

A colored full version of this chapter is available as free download: http://www.sowi.uni-kl.de/lachmann/sperling.pdf. All articles of which Sperling is an author are available as free downloads from http://www.cogsci.uci.edu/~whipl/staff/sperling/

Note

A Type 1 trial has a right answer that can be rewarded and a wrong answer: e.g., which of two identically colored light patches, the one on the right or the one on the left, is brighter, or which has greater area. Once a subject has mastered the Type 1 procedure—i.e., learned the meaning of brighter or of area in this context, a Type 2 procedure without systematic reward is possible: Which of two lights of different colors seems brighter? Type 2 is also used to study illusions and preference: Which of the two odd shapes seems to have more area? Which is more beautiful? (Sperling, 1992; Sperling, Dosher, & Landy, 1990).

References

Anstis, S., & Cavanagh, P. (1983). A minimum motion technique for judging equiluminance. In J. D. Mollon & E. T. Sharpe (Eds.), *Colour vision* (pp. 155–166). New York, NY: Academic Press.

Blaser, E., Sperling, G., & Lu, Z.-L. (1999). Measuring the amplification of attention. *Proceedings of the National Academy of Sciences, 96*(20), 11681–11686. doi: 10.1073/pnas.96.20.11681

Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology, 55*(1), 75–84. doi: 10.1111/j.2044-8295.1964.tb00899.x

Drew, S. A., Chubb, C. F., & Sperling, G. (2010). Precise attention filters for weber contrast derived from centroid estimations. *Journal of Vision, 10*(10), 1–16. doi: 10.1167/10.10.20

Gegenfurtner, K. R., & Sperling, G. (1993). Information transfer in iconic memory experiments. *Journal of Experimental Psychology: Human Perception and Performance, 19*(4), 845–866. doi: 10.1037/0096-1523.19.4.845

Gobell, J. L., Tseng, C.-h., & Sperling, G. (2004). The spatial distribution of visual attention. *Vision Research, 44*(12), 1273–1296. doi: 10.1016/j.visres.2004.01.012

Hsu, A., Scofield, I., & Sperling, G. (2006). A computational model for the distribution of spatial attention. *Journal of Vision, 6*(6), 507. doi: 10.1167/6.6.507

Lachmann, T., & Weis, T. (Eds.). (2018). *Invariances in Human Information Processing*. New York, NY: Routledge.

Lu, Z.-L., & Sperling, G. (1995). Attention-generated apparent motion. *Nature, 377*(6546), 237–239.

Lu, Z.-L., & Sperling, G. (2001a). Sensitive calibration and measurement procedures based on the amplification principle in motion perception. *Vision Research, 41*(18), 2355–2374. doi: 10.1016/S0042-6989(01)00106-7

Lu, Z.-L., & Sperling, G. (2001b). Three-systems theory of human visual motion perception: Review and update. *Journal of the Optical Society of America A, 18*(9), 2331. doi: 10.1364/JOSAA.18.002331

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

Nobre, K., & Kastner, S. (2014). *The oxford handbook of attention*. Oxford and New York: Oxford University Press.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3–25. doi: 10.1080/00335558008248231

Prinzmetal, W., Amiri, H., Allen, K., & Edwards, T. (1998). Phenomenology of attention: I. Color, location, orientation, and spatial frequency. *Journal of Experimental Psychology: Human Perception and Performance, 24*(1), 261–282. doi: 10.1037/0096-1523.24.1.261

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance, 18*(3), 849–860. doi: 10.1037/0096-1523.18.3.849

Reeves, A., & Sperling, G. (1986). Attention gating in short-term visual memory. *Psychological Review, 93*(2), 180–206. doi: 10.1037/0033-295X.93.2.180

Shih, S.-I., & Sperling, G. (2002). Measuring and modeling the trajectory of visual spatial attention. *Psychological Review, 109*(2), 260–305. doi: 10.1037/0033-295X.109.2.260

Shulman, G. L., Remington, R. W., & Mclean, J. P. (1979). Moving attention through visual space. *Journal of Experimental Psychology. Human Perception and Performance, 5*(3), 522–526. doi: 10.1037/0096-1523.5.3.522

Smith, P. L., & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychological Review, 116*(2), 283–317. doi: 10.1037/a0015156

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied, 74*(11), 1–29. doi: 10.1037/h0093759

Sperling, G. (1963). A model for visual memory tasks. *Human Factors, 5*(1), 19–31. doi: 10.1177/001872086300500103

Sperling, G. (1967). Successive approximations to a model for short term memory. *Acta Psychologica, 27*, 285–292. doi: 10.1016/0001-6918(67)90070-4

Sperling, G. (1968). Phonemic model of short-term auditory memory. *Proceedings, 76th Annual Convention of the American Psychological Association, 3* (pp. 63–64) Washington, DC, USA: American Psychological Association.

Sperling, G. (1984). A unified theory of attention and signal detection. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 103–181). New York: Academic Press.

Sperling, G. (1992). *Type 1 and type 2 experiments*. Retrieved from http://www.cogsci.uci.edu/~whipl/Type_1_and_Type_2_Expts.pdf

Sperling, G., & Dosher, B. A. (1986). Strategy and optimization in human information processing. In K. Boff, L. Kaufmann, & J. Thomas (Eds.), *Handbook of perception and human performance* (pp. 2-1–2-65). New York, NY: Wiley.

Sperling, G., Dosher, B. A., & Landy, M. S. (1990). How to study the kinetic depth effect experimentally. *Journal of Experimental Psychology: Human Perception and Performance, 16*(2), 445–450. doi: 10.1037/0096-1523.16.2.445

Sperling, G., & Melchner, M. (1978). The attention operating characteristic: Examples from visual search. *Science, 202*(4365), 315–318. doi: 10.1126/science.694536

Sperling, G., & Reeves, A. (1980). Measuring the reaction time of a shift of visual attention. In R. Nickerson (Ed.), *Attention and performance VIII* (pp. 347–360). Hillsdale, NJ: Erlbaum.

Sperling, G., Scofield, I., & Hsu, A. (2008). Computational model of the spatial resolution of visual attention. *Journal of Vision, 8*(6), 396. doi: 10.1167/8.6.396

Sperling, G., & Speelman, R. G. (1970). Acoustic similarity and auditory short-term memory: Experiments and a model. In D. A. Norman (Ed.), *Models of human memory* (pp. 149–202). New York, NY: Academic Press.

Sperling, G., & Weichselgartner, E. (1995). Episodic theory of the dynamics of spatial attention. *Psychological Review, 102*(3), 503–532. doi: 10.1037/0033-295X.102.3.503

Sun, P., Chubb, C., Wright, C. E., & Sperling, G. (2016a). The centroid paradigm: Quantifying feature-based attention in terms of attention filters. *Attention, Perception & Psychophysics, 78*(2), 474–515. doi: 10.3758/s13414-015-0978-2

Sun, P., Chubb, C., Wright, C. E., & Sperling, G. (2016b). Human attention filters for single colors. *Proceedings of the National Academy of Sciences of the United States of America, 113*(43), E6712-E6720. doi: 10.1073/pnas.1614062113

Tseng, C.-h., Gobell, J. L., & Sperling, G. (2004). Long-lasting sensitization to a given colour after visual search. *Nature, 428*(6983), 657–660. doi: 10.1038/nature02443

Weichselgartner, E., & Sperling, G. (1987). Dynamics of automatic and controlled visual attention. *Science, 238*(4828), 778–780. doi: 10.1126/science.3672124