Attraction, Distraction, and Action: Multiple Perspectives on Attentional Capture
C. Folk and B. Gibson (Editors)

Final draft version (pre-copy-edited)

X Attentional Capture in the Spatial and Temporal Domains

Howard E. Egeth, Charles L. Folk, Andrew B Leber, Takehiko Nakama, and Sharma K. Hendel

In the study of what kinds of stimuli do or do not capture attention, one of the most productive ideas has been the notion of contingent capture–that is, that what captures attention depends on your current task goals. Most of this research has been conducted in the context of visual search studies either with or without precues indicating the likely location of the target. We begin with a discussion of an example of this paradigm that appears to challenge the notion of contingent capture. We then move on to an extension of the contingent capture idea to the temporal domain with the use of the RSVP paradigm, that is, the rapid serial visual presentation of stimuli. In this paradigm there is uncertainty about the time at which a target will occur. We first introduce some novel results bearing on the nature of the attentional blink, and then use a variant of this paradigm to study attentional capture. In our experiments, subjects will have the opportunity to resolve the uncertainty about time of arrival of the target by using a stimulus property–color–that just so happens to also define a spatially irrelevant stimulus. Under these conditions will subjects be able to ignore the irrelevant stimulus?

The Contingent Capture Hypothesis

Attentional allocation has often been described as having two sources of control: stimulus-driven and goal-directed selection. Stimulus-driven selection is characterized by allocation based entirely on the salience of features in the display; specific task demands or goals of the observer are irrelevant. Attentional allocation of this sort can be said to be involuntary and not under the observer's control (Remington, Johnston & Yantis, 1992; Yantis, 1993; Yantis & Jonides, 1990). Alternatively, attentional deployment may be a direct result of the objectives of the observer so that attention is allocated to visual elements that are task-relevant, familiar, or otherwise important to the observer (e.g., Francolini & Egeth, 1979; Posner, 1980; Yantis & Egeth, 1999).

Thus spatial shifts of attention have been thought to occur in two ways: either as an involuntary response to a salient stimulus (*attentional capture*) or as a voluntary process determined by the observer's goals. The fact that attention may be summoned involuntarily and automatically has been much studied with the goal of determining what kinds of stimuli can generate such shifts. To date, a variety of experiments reveal that abrupt stimulus onsets (such as new objects) will generate exogenous spatial shifts (e.g., Jonides & Yantis, 1988; Muller & Rabbitt, 1989; Remington, et al., 1992; Posner & Cohen, 1984; Yantis & Hillstrom, 1994; Yantis & Jonides, 1990).

A crucial question in the area of attentional allocation is whether salient features other than abrupt onsets can result in attentional capture. Jonides and Yantis (1988) examined whether color and intensity discontinuities, in addition to abrupt onsets, led to attentional capture and concluded that onsets were unique in this regard. Subsequently, some observations have cast doubt on the generality of that conclusion. For one thing, Yantis and Egeth (1999) found circumstances in which large or bright stimuli captured attention while moving or differently colored stimuli did not. More generally, Folk, Remington, and Johnston (1992) obtained evidence that suggested the deployment of attention depends critically on what the subject is set for. In a cuing paradigm they independently manipulated the nature of the target and the nature of a preceding cue. Subjects might have to locate the target on the basis of a color discrepancy (it was the only red element among three white elements), or they might have to locate it on the basis that it was the only element with a sudden onset (it was the only element in the display). Similarly, the cue might consist of one red element among three white elements, or it might be the only sudden onset element (again, it was alone in the display). The most important finding was that when cue and target were of the same type (i.e., both sudden onsets or both color singletons), subjects could not ignore the cue even when it was known to be 100% invalid.

A Challenge to the Contingent Capture Hypothesis

In a paper by Joseph and Optican (1996), the authors suggested that attention can be captured by an orientation discontinuity within a cue stimulus. What is particularly provocative about this finding is that there was no obvious relationship between the target and cue, and thus application of the contingent involuntary capture hypothesis seems strained. Joseph and Optican (1996) used a cuing design to assess the degree to which oriented bars captured attention. In Experiments 1 and 2, subjects were presented with a brief cue display consisting of a dense array of vertical (or horizontal) bars. One of the bars in the array was oriented differently (e.g., it was horizontal while the others were vertical) creating a "pop-out" stimulus. In order to assess the degree to which this stimulus captured attention, a subsequent target display was presented containing a similar array composed entirely of Ts except for a single target L; subjects were instructed to find the L and report its location. (Note that the Ts and Ls in these experiments were *not* randomly rotated.) Cue and target stimuli only appeared in one of four locations in the display -- the center element in each of the four quadrants of the array. The assumption was that if the cue led to a shift of attention to its location, it would result in improved accuracy for a target that subsequently appeared at same location (i.e., a *valid* cue). The cue, however, was not informative; when a cue was presented, it occurred equally often at the target location (valid) and the three other locations (invalid). On one-fifth of the trials all of the bars were of a uniform orientation; these were uncued trials. Figure 1 illustrates the stimuli and time course of a typical trial.

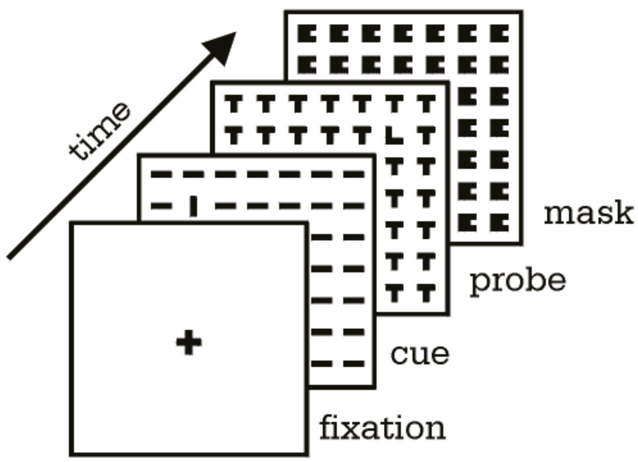

Figure 1. Overview of the sequence of trial events in experiments by Joseph and Optican (1996) and Hendel (1998/1999). In Joseph and Optican the fixation cross appeared for 500 ms and remained throughout the trial. Cue durations were varied over a range from 50 to 800 ms (across Exps. 1 and 2), and the probe duration was 50 ms. Event durations for Hendel's experiments are given in the text. In Exp. 1 of Joseph and Optican an array of adapter squares was shown prior to the cue display, but this had no effect on the pattern of results.

In these experiments, Joseph and Optican (1996) found that subjects' accuracy in locating the target was considerably greater when the target occurred at a location previously occupied by an orientation difference than when it appeared where there had been no orientation difference. Given this validity effect, they concluded that "orientation differences can cause shifts of visual attention, regardless of whether they convey any information relevant to the task at hand." p. 661. While it is certainly possible that the orientation cues used in their experiment resulted in shifts of attention, Joseph and Optican (1996) further suggested that these attentional shifts are involuntary, claiming that their experiments "…might be said to be probing bottom-up processing, which includes both the conventional notion of preattentive vision and the involuntary attentional shifts that can follow from it." p. 661.

A similar conclusion from a study using a rather different paradigm is due to Theeuwes (1992; see also Mounts 2000; Turatto & Galfano, 2001). Theeuwes (1992) proposed that early attentional allocation is mediated entirely by stimulus-driven characteristics of the display. In several experiments using a visual search design, he demonstrated that observers were slower to respond to a target when there was a salient singleton distractor in the display than when no singleton distractor was present. He argued that slower response in the presence of a distractor occurred because observers' attention was captured by the salient distractor (Theeuwes, 1991, 1992, 1994, 1996). In fact, he claimed that goal-directed selection was not possible during early preattentive visual search (Theeuwes, 1992).

Another possibility is that subjects employ two different strategies depending on circumstances. Thus, subjects might rely on a strategy to look for the target-defining feature, as suggested by Folk et al. (1992) when the target is defined by a single simple feature such as "red" or "vertical". However, subjects might also use the strategy of trying to find a salient singleton, or odd-man-out, when there is no single feature that defines the target or when the target-defining feature does not easily segregate from the nontarget elements (e.g. L among Ts). The idea is that "singleton" itself can be the defining-feature of a target and observers can *choose* to search for a salient singleton (Bacon & Egeth, 1994; Pashler, 1988). Using a visual search experiment identical to that of Theeuwes (1991, 1992), Bacon and Egeth (1994) demonstrated that when observers searched for a target that was *always* a singleton, other singleton distractors impaired performance. Significantly, when observers searched for a target that was *not always* a singleton, a singleton distractor did *not* impair performance. In one experiment, the number of targets present on each trial varied from 1 to 3. When multiple targets were present they were not true singletons, which should have discouraged observers from setting their attention for singletons. In this condition, even when there was only a single target present on a given trial, the presence of a singleton distractor did not lead to slower responses. Bacon and Egeth (1994) took this as evidence that subjects could adopt a "singleton detection mode" strategy when appropriate (see also Pashler, 1988). In several other experiments, Bacon and Egeth demonstrated that observers relied heavily on singleton search when it was an efficient strategy. However, when singleton detection mode was inefficient, subjects used a feature search strategy.

New Evidence of Top-Down Control in Visual Search

It is arguable that the capture of attention found by Joseph and Optican (1996) was due to the voluntary adoption of singleton detection mode; after all, their target was the sole L in a field of canonically oriented Ts. However, it is not clear a priori whether this discrimination is sufficiently salient to invoke "singleton detection mode." One piece of evidence is that Egeth and Dagenbach (1991) found that two upright characters (T/L) could be searched in parallel, but the conditions of that study are very different from those of Joseph and Optican, especially in regard to the number of elements in the display (2 vs. 48). To explore whether an orientation singleton captures attention independently of the subject's search task, Hendel (1998/1999) conducted an extensive series of experiments, which will be summarized briefly here.

Hendel's first experiment was essentially a replication of Joseph and Optican's Exp. 2. She found that subjects were more accurate in locating a target when it occurred at the cued location. However, unlike Joseph and Optican, she also found evidence for a very large response bias effect, which rendered the cuing effect of dubious validity. When subjects were incorrect they tended to select the cued location much more frequently than would be expected on the basis of chance. In her second experiment she changed the paradigm to eliminate the opportunity for response bias. Now on each trial the L was presented either in its normal orientation or in its left-right mirror image orientation; subjects had to identify the orientation of the L rather than locate it. Here too she found a validity effect, as had been found by Joseph and Optican. In her next experiments she addressed the theoretical issue of whether the effect was due to bottom-up salience, as suggested by Joseph and Optican, or the adoption of a goal-directed process motivated by the subject's knowledge of target-relevant information. Such is the suggestion of Folk et al.'s (1992) contingent capture hypothesis.

A crucial test of the salience hypothesis is to determine whether the same (orientation) cue that gave evidence of capture when the target was an L in a field of Ts would also give evidence of capture when the target task is changed. Conversely, if the cue captures attention because of some feature similarity between the cue and the target, then a different target task, in which the target-defining feature is not the same as the cue-defining feature, should eliminate attentional shifts to the cue (Folk et al., 1992). In Hendel's Experiment 3a the target was defined in terms of color; it was the red element in a field of white elements. Specifically, the target display was a 7 x 7 matrix of Ls, half normal, and half mirror-image reversed. One of the Ls was red, and the remainder were white. The subjects task was to indicate the orientation of the red L. Two cue conditions were used. One was the orientation cue used earlier; all of the bars were horizontal (vertical), except for one which was vertical (horizontal). The other was a color cue the same color as the target; all of the bars were of the same orientation and one of them was red. As before, cue location and target location were independent, and one-fifth of the trials were uncued (i.e., no singleton was present in the cue display). The cue and target displays were each displayed for 100 ms with no interval between them. Thirteen subjects each received 320 color-cued trials and 320 orientation-cued trials, randomly mixed.

Recall that there were four possible locations where the cues and targets could appear. Consider, for example, the situation where the target appears in the upper left location. If the cue on that trial also appears in the upper left, it is a valid trial. There are three kinds of invalid trial, which are characterized with respect to the relative locations of the cue and target. If the cue appears in the upper right this is an invalid horizontal trial (i.e., the cue and target locations differ by only a horizontal translation). Similarly, if the cue appears in the lower left location this is an invalid vertical trial. If the cue appears in the lower right this is an invalid diagonal trial. If the cue display contained no singleton it is an uncued trial.

To determine whether the validity effect differed between the two cue types, percentage of correct responses were entered in a 2 x 5 repeated measures ANOVA. This analysis revealed a significant main effect of validity, $F(4, 48) = 11.64$, $p < .0001$, and a significant interaction between validity and cue type, $F(4, 48) = 6.72$, $p < .001$, however, there was no main effect of cue type itself, $F(1, 12) = 2.07$, $p > .17$.

As is evident from Figure 2a, the effect of validity is due predominantly to the color cue. In the color-cue condition, each cue validity was compared to the null cue. These comparisons confirm that the valid cue led to greater accuracy than the null cue, $F(1, 12) = 8.53$, $p < .05$, and the invalid cues led to lower accuracy than the null cue: horizontal, $F(1, 12) = 10.53$, $p < .01$, vertical, $F(1, 12) = 8.96$, $p < .05$, and diagonal, $F(1, 12) = 22.66$, $p < .001$.

Similar comparisons were made for validity in the orientation-cue condition. There were no reliable differences found between the null cue and any of the other validity conditions: null cue and valid cue, $F(1, 12) = 2.19$, $p > .16$, null cue and horizontal, $F(1, 12) = .157$, $p > .69$, null cue and vertical, $F(1, 12) = .001$, $p > .97$, null cue and diagonal, $F(1, 12) = .001$, $p > .97$.

In this experiment in which the target was defined by a color discrepancy, the orientation cue did not lead to a reliable shift of attention, calling into question the generality of the effect reported by Joseph and Optican. However, the color cue did lead to a reliable shift of attention. This pattern is consistent with the contingent capture hypothesis and suggests that the shifts of attention found in this experiment

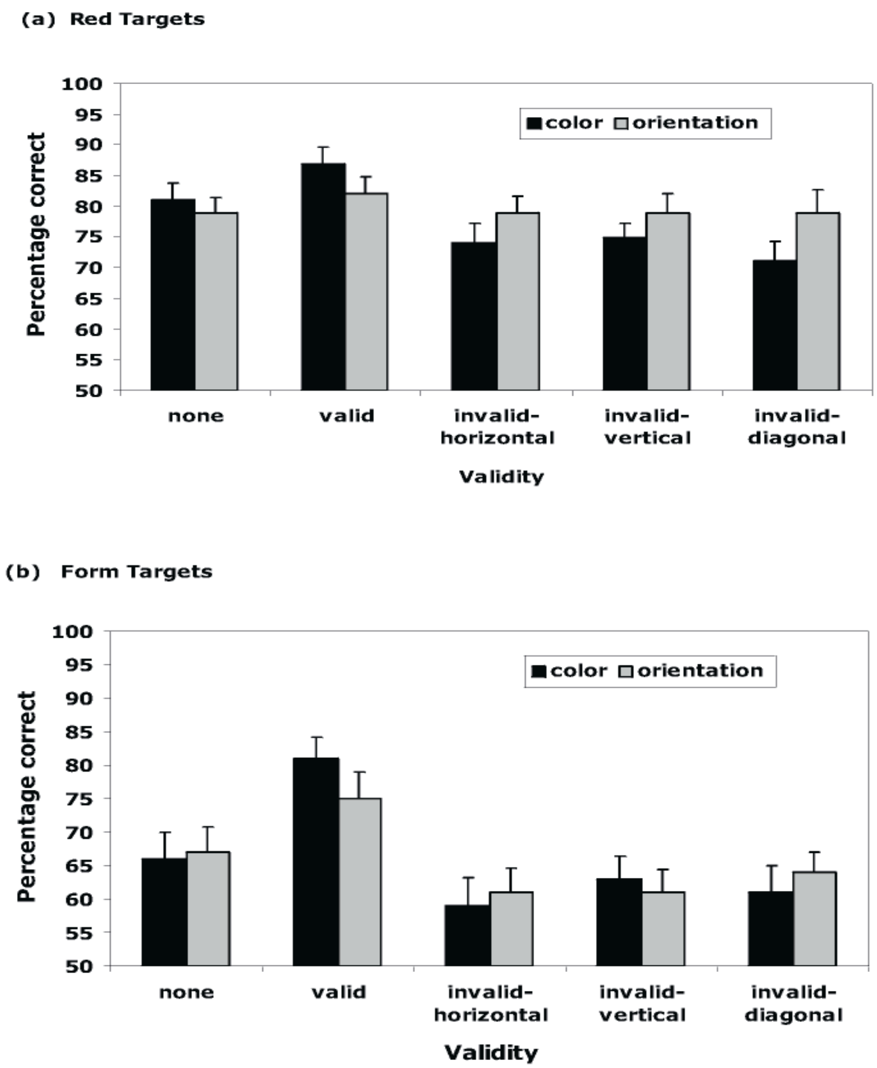

Figure 2. Mean percentage of correct determinations of target orientation for the various levels of cue type (color and orientation) and cue-target validity in

Experiment 3 of Hendel (1998/1999). See text for a full description of the validity conditions. Experiment 3a (red target among white nontargets) is represented in the top panel and 3b (L or backwards L target among nontarget Ts) in the lower panel.

may be due to the similarity between cue and target. This finding prompted re-examination of why the original task (identifying an L among Ts) showed a validity effect when an orientation cue was used. One possibility is that subjects shifted attention to the orientation cue because they were searching for singletons. If this were the case, then we might expect that both color cues and orientation cues would capture attention in the L among Ts task.

In Hendel's Experiment 3b the target was a forwards or backwards L among upright Ts. Cues were the same as in Exp. 3a, consisting of either an orientation or a color singleton (as before, one-fifth of the trials were uncued). In all other respects the experiment was identical to the previous one. Thirteen subjects participated.

The results are shown in Figure 2b. Analysis of variance revealed a significant main effect of validity, $F(4, 48) = 19.18$, $p < .0001$, but neither cue type, $F(1, 48) = .06$, $p > .80$, nor the interaction between cue type and validity were significant, $F(4, 48) = 1.88$, $p > .12$. In the absence of both a main effect of cue type and an interaction between cue type and validity, contrasts were analyzed collapsed across the two cue types. The increase in accuracy on valid as compared to null trials was significant, $F(1, 12) = 26.11$, $p < .001$, as was the decrease for horizontal trials relative to the null cue, $F(1, 12) = 7.05$, $p < .05$. However, the differences between the vertical and null cues and the diagonal and null cues did not reach significance, $F(1, 12) = 3.06$, $p > .10$, and $F(1, 12) = 1.63$, $p > .22$, respectively.

In Hendel's Experiment 3b, both the red cue and the orientation cue led to shifts of attention. This result is expected on the contingent capture model assuming subjects are in singleton detection mode. It is also expected on the salience model, if we can assume that both the color and the orientation cue displays contain a salient discontinuity. However, the results of Hendel's Exp. 3a suggest that salience alone cannot explain the observed pattern of validity effects. In that study, significant validity effects were obtained only when the target-defining feature and the cue-defining feature were the same (red), suggesting that observers maintained a set for red stimuli.

Taken as a whole, Hendel's Experiments 3a and 3b are more consistent with the contingent capture hypothesis than the salience hypothesis for the following reasons. As we have argued, the salience hypothesis predicts that subjects will shift their attention to the most salient element in the cue display. In each of these experiments, on any given cue display, there is only one salient difference -- the cue. Consequently, it was expected that a shift of attention to the most salient element would always be to the cue, leading to a cue-validity effect for both cues in

Experiments 3a and 3b. As no validity effect was observed with the orientation cue in Experiment 3a, the predicted outcome for the salience hypothesis was not obtained. Nonetheless, a different interpretation of these data may be entertained. Theeuwes (1994) argued that the typical brief presentation of stimuli in the cueing paradigm may lead subjects to integrate the cue and the target displays. Thus attention would be likely to go to the most salient element in the *entire trial sequence*. If this were the case, then the salience hypothesis could be interpreted to predict that attention would be likely to shift to the target, rather than the cue, if the target were more salient than the cue. In Experiment 3a, it is possible, then, that the target -- a red L among white Ls -- was more salient than the oriented bar cue presented on half of the trials. The red bar cue, on the other hand, may have been *even more* salient than the red target, leading to the validity effects observed for the red cue. Note that this interpretation requires that we assume that the red bar cue was the most salient element in the trial sequence, followed by the red target L. The orientation cue is assumed to be less salient than the red cue and the red target. This interpretation would be ruled out if one could find a target task in which the orientation cue leads to attentional capture but the color cue does not. This was the rationale for Hendel's Experiment 4.

In Experiment 4 the cue types used previously (a color singleton and an orientation singleton) were used in conjunction with a new target task in which 24 subjects were set for a stimulus of a particular orientation. Subjects had to detect the presence of a Q-like character among Os. The "stem" of the Q was a horizontal bar on either the left or the right of the target. Subjects were asked to report whether the stem was on the left or right side of the character. The rationale was that this task would lead subjects to set their attention for horizontal bars.

The experiment was similar to the preceding ones except that the orientation cue was now always a matrix of vertical bars with a single horizontal bar; as before the color cue consisted of a single red vertical bar among white vertical bars.

The results are shown in Figure 3. The crucial result is that contrast analysis showed that for color cues there were no reliable differences between the null cue and any of the other cues (all p values >.15) However, the orientation cue led to a reliable improvement in the valid cue condition compared to the null cue condition (p <.01), and to reliable decrements for the three invalid cue conditions relative to the null condition (all p values <.01).

This study confirmed that attention was most likely to be drawn to cues that matched the defining features of the target. Hendel went on to discuss an alternative interpretation of such results proposed by Theeuwes (1994). According to Theeuwes, disengagement of attention from a cue that contains a target-defining feature may take longer than disengagement of attention from a cue that does not contain a target-defining feature. Thus cue-validity effects are observed in the latter case because recovery (i.e., disengagement) occurs sufficiently quickly so that attention is available to shift to the target location. Hendel provides a detailed rebuttal of this argument, but its details are beyond the scope of the present paper.

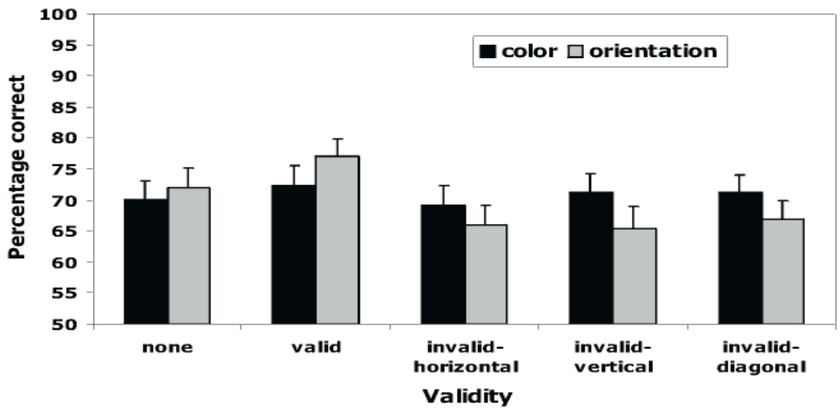

Figure 3. Mean percentage of correct determinations of target orientation for the various levels of cue type and cue-target validity in Experiment 4 (Q-like target among nontarget Os) of Hendel (1998/1999). This figure represents combined data for the 13 subjects who participated in her Experiment 4a and the 11 who served in her Experiment 4b.

The series of studies by Hendel provide converging evidence for the existence of some top-down control in visual search. Her data show, using an accuracy measure, that the relationship between target-defining features and distractor features is crucial to eliciting involuntary shifts of attention. In addition, these studies have succeeded in uncovering a specific task in which salient cues do reliably capture attention (Exp. 3a). This indicates that when subjects are in singleton detection mode, the salience of the cue is critical.

The RSVP Paradigm and the Attentional Blink

Before turning to our research on the applicability of the contingent capture hypothesis to stimuli appearing in rapidly presented streams, we discuss some research that speaks to the nature of the capacity demands of identifying targets in such streams. Part of the appeal of the rapid serial visual presentation (RSVP) paradigm is that it seems subjectively to be very attentionally demanding. A typical RSVP task is to present a stream of letters to a subject that is all black except for one that is green. The subject has to name the green letter. At a rate of 10 letters per second the task feels difficult. Overall accuracy rate will be high, but distinctly less than perfect, say 85-90% correct. Moreover, if subjects have to identify a second element in the stream following identification of the green letter, they may perform very poorly for a period of up to a half second or more, a refractory period that has been dubbed the *attentional blink* (e.g., Broadbent & Broadbent, 1987; Raymond, Shapiro, & Arnell, 1992).

An interesting example of an attentional blink appears in a paper by Joseph, Chun, and Nakayama (1997). They presented an RSVP stream of black letters at fixation. There were two targets. The first was a green letter embedded in the stream; this was followed after a variable interval by the second target, which was a ring of Gabor patches surrounding (at 5.3 deg eccentricity) one of the later letters in the stream. Subjects had to name the colored letter and indicate if the patches were all oriented in the same direction or if one was misoriented by 90 deg. In a control condition subjects could ignore the RSVP stream and just indicate if the ring of Gabor patches did or did not contain an orientation oddball. Performance in the control condition was in excess of 90% correct and was independent of the lag between the green letter and the Gabor patches. In the experimental condition performance was poor (about 60% correct detections) when the ring of Gabor patches was simultaneous with the green target letter (i.e., lag=0), and improved monotonically to nearly 90% correct when the lag between the target letter and the Gabors was 700 ms.

This monotonic improvement over time is one of the standard forms of the attentional blink, with recovery taking over half a second (cf. Visser, Bischof, & DiLollo, 1999). What makes this study so interesting is the implications it may have for the study of visual attention. Specifically, Joseph et al. (1997) pointed out that this result is something of an embarrassment for the notion of preattention. Detection of an orientation singleton is usually thought to be preattentive (e.g., Sagi & Julesz, 1985; Treisman & Gormican, 1988). Why, then, should there be such a huge dual-task decrement? Indeed, Braun and Sagi (1990; 1991) have conducted conceptually similar experiments and found seemingly contrary results. In their studies, identification of a single element at fixation was used as the central primary task. This target element was followed by a masking stimulus, thus this situation is equivalent to the case in which the number of letters in an RSVP stream is two. No decrement was found in the detection of an orientation singleton in a field of elements surrounding fixation. The difference between the two studies may result in part from the length of the RSVP stream. There are several possible reasons why the attention demand of RSVP may be greater when the RSVP stream contains more letters. To cite just one, it may become more difficult to identify just which letter is the colored target when there are more opportunities for falsely conjoining color and identity. There are, of course, numerous differences between the Joseph et al. (1998) and the Braun and Sagi (1990, 1991) studies, and so a direct test is required to evaluate our suspicion that stream length may be important. In particular, Braun (1998) has observed that naive subjects were used in the Joseph et. al. experiments, whereas highly trained psychophysical observers were used by Braun and Sagi. Braun (1998) has shown that the RSVP version of the experiment also yields no attentional decrement when sufficiently experienced observers are tested. The question remains, however: When an attentional blink *is* found, what determines its magnitude?

The Role of Stream Length in the Attentional Blink

Egeth and Nakama (1999) explored the role of stream length in naive subjects. The central RSVP string consisted of a stream of letters presented at fixation on a gray screen. Each letter was presented for 67 ms; there was an 84 ms interval between successive letters. Thus, letters were presented at the rate of approximately 6.6 per sec. One major independent variable was the length of the RSVP string. In one condition that approximates the typical attentional blink paradigm, the white target letter appeared in the middle of a 19-letter sequence; the other 18 letters were all green. In another condition, the white target letter appeared as the first of two letters, the second being green; i.e., this was a sequence of length two. The two stream lengths were not randomly mixed, but were tested in separate conditions. The second major independent variable was the temporal lag (SOA) between the white target letter and the display of line segments; these lags were 0, 200, and 500 ms and varied randomly from trial to trial. The stimulus for the singleton detection task consisted of a ring of eight bars centered at fixation. The outside diameter of this array was 12 deg of visual angle, and each bar was 1.2 deg in length. On half of the trials all of the bars were oriented in the same direction (either +20 or -20 deg from vertical); on the other half of the trials one of the bars was oriented +20 while the others were at -20 deg, or vice versa. Each array of oriented bars was displayed for 150 ms and was followed by a masking field.Note that the line segments never preceded the white letter. The 0 ms lag is typical of the Braun and Sagi experiments (e.g., 1990); the remaining two lags permit us to sketch out the time course of any attentional blink that may be generated in our experimental conditions. Finally, half of the subjects were assigned to the dual task version of the experiment, while the other half were instructed to ignore the letter stream and just perform the singleton detection task. In the dual task the importance of identifying the target within the RSVP stream was emphasized.

The results are shown in Figure 4. They indicate two important points. For length=19 (the right panel), there was a significant attentional blink; attending to a 19-item stream severely hampered detection of a misoriented bar in the

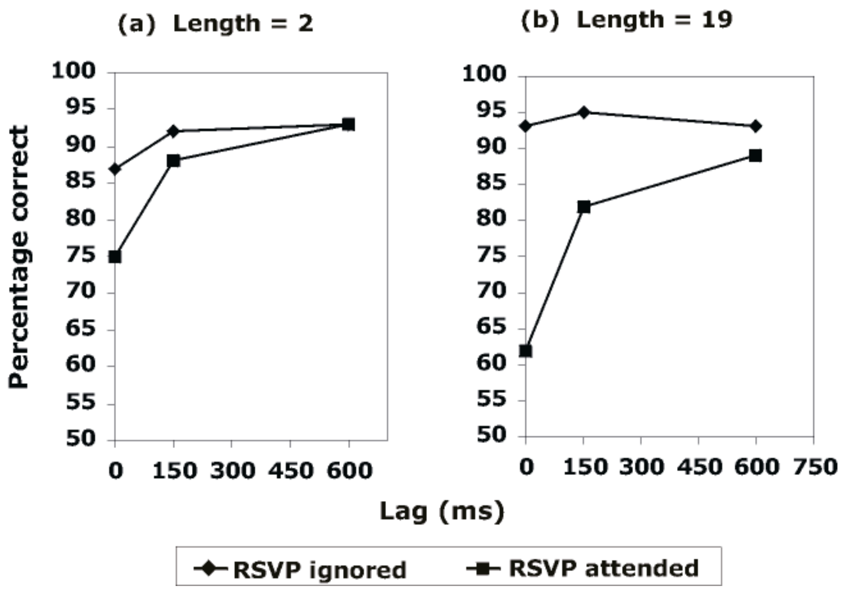

Figure 4. Mean percentage correct detections of an orientation singleton as a function of the delay (SOA) between presentation of white target letter in foveal RSVP stream and presentation of display of oriented bars. Filled circles show performance for the RSVP-ignored control condition, open circles show performance for the RSVP-attended dual task. Upper panel represents data for the condition in which a white target letter is followed by a single green nontarget letter. Lower panel represents data for the condition in which a white target letter appears in the middle of a 19-letter stream (the other 18 letters being green).

periphery. For length=2 (left panel) there was relatively little difference between the RSVP-attended and RSVP-ignored condition. (In another experiment in which stream length was 1 rather than 2, there was literally *no* attentional blink. That is, there was no difference between the RSVP-attended and RSVP-ignored conditions.) Interestingly, these results tend to support important features of both the Braun and

Sagi (1990, 1991) studies and the Joseph, Chun, and Nakayama (1997) study; attending to a 2-letter RSVP "stream" has a small effect on singleton detection in the periphery, whereas attending to a 19-item stream has a large effect.

Second, note that for the 2-item stream there is a decline in performance for SOA=0, even when the RSVP task is unattended. (This was also the case for the 1-item stream.) This struck us as a surprising finding. It is known that alphanumeric characters may be processed to the point of identification even when subjects do not intend to do so. This is the basis of the Stroop effect among other phenomena. Teichner and Krebs (1974) argued that even simple RT tasks may be subject to "compulsive encoding" effects when alphanumeric stimuli are used. With this in mind we simplified the preceding experiment by testing only RSVP-ignored trials, and using a 1-element "stream" that consisted of a small filled white square presented at fixation. The results were striking: at SOA=0 performance was 85% correct; at SOA = 150 and 600 ms performance was 95% and 94%, respectively. The effect is not due to the subject being alerted to the upcoming line discrimination task by the onset of the white square. When the white square was replaced by a brief tone, accuracy was independent of SOA at about 90% correct.

The decrement at SOA=0 most likely does not reflect simple masking, as the separation between the central target element and the peripheral bars was greater than 5 deg (e.g., Breitmeyer, 1984). The effect is reminiscent of the interfering effect of strong transients (or "mudsplashes") found in the literature on change blindness (e.g., O'Regan, Rensink, & Clark, 1999). However, when strong transients have interfered with the perception of other items they have typically been presented in advance of those other items; that was not the case here, as the SOA was zero. It may be, then, that the effect is an instance of cognitive masking, like the filtering cost discussed by Treisman, Kahneman, and Burkell (1983).

With respect to the focus of the present chapter, the point of the first set of experiments is that the attentional blink paradigm, with its rapidly streaming alphanumeric characters is attentionally demanding. A 2-element central stream (i.e., a target character followed by a mask) as used by Braun and Sagi among many others, may not be sufficiently demanding to interfere with detection of a peripheral orientation singleton, but a lengthy central stream apparently is. However, we do not know just what it is about the attentional blink paradigm that results in the performance deficit on the second target. Presumably, it is the attentional demands of having to identify the first target that makes it difficult to detect or identify the second target (see, e.g., models proposed by Chun & Potter, 1995 and Jolicoeur, 1998).

An Attentional Blink at Negative SOA?

At this point we reflected on the fact that Joseph, Chun and Nakayama (1997), following standard practice in the field, had explored only positive lags, with the presumably attention demanding letter identification task coming first. What would happen if the supposedly preattentive task came first? Using the same paradigm described earlier, Nakama and Egeth (1999) tested both positive and negative lags. When letter identification came first the results were much like those of Joseph, Chun, and Nakayama although in our study recovery wasn't complete even by 600 ms. When the orientation discrimination came first we also found a large decrement in the ability to detect an orientation singleton. This suggests that the dual-task decrement was not caused solely by identification of the green letter; much of the deficit may be caused by the need to monitor the stream for the target letter. This was confirmed in our next study in which we omitted the green letter on some trials. On these trials performance stayed low throughout the trial (see Figure 5). Again, the importance of identifying the target in the RSVP stream was emphasized.

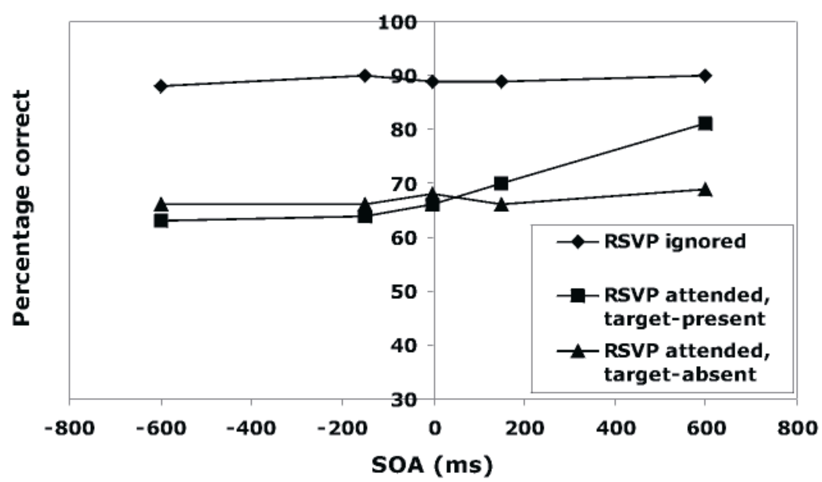

Figure 5. Mean percentage correct detections of an orientation singleton as a function of the delay (SOA) between presentation of white target letter in foveal RSVP stream and presentation of display of oriented bars. Negative SOA means bar display preceded target letter. Open circles show data for standard dual task condition. Filled circles show data for condition in which the RSVP stream was all green and thus did not contain a target letter. Triangles show data for a control condition in which subjects are instructed to ignore the RSVP stream.

The main point we draw from this is that, regardless of whether it is the identification of the target letter itself or simply preparation for the identification that is the basis of the effect, the RSVP task is indeed attentionally demanding. As a secondary point, we speculate that the need to monitor a rapid stream of stimuli at the fovea may both center attention at that location and narrow it to the approximate size of the letters that appear there (e.g., LaBerge, 1983), at least when the RSVP task is given high priority. We turn next to an exploration of whether peripheral stimuli may capture attention when attention is narrowly focused by a foveal RSVP task.

Attentional Capture in the RSVP paradigm

We earlier described the contingent involuntary orienting hypothesis that assumes attentional control settings are a function of the behavioral goals of the observer, such as searching for a red letter. Another form of top-down control over attentional capture involves the degree to which attention is spatially focused prior to the presentation of a salient irrelevant stimulus. Yantis and Jonides (1990) found that when the location of a target letter in a visual search task was uncertain, the presence of an irrelevant abrupt onset letter produced evidence of attentional capture. However, when subjects were given a 100% valid precue regarding the subsequent location of a target, capture effects were eliminated. Similar results were reported by Theeuwes (1991). Such findings support the widely held belief that when spatial attention is in a highly focused state, salient stimuli (such as abrupt onsets) are no longer capable of capturing spatial attention.

A potentially important aspect of these studies, however, is that the use of a 100% valid spatial precue not only eliminated uncertainty about the target location, it also eliminated any uncertainty about which object in the display was the target. That is, on any given trial, only one object ever occurred at the cued location. One could imagine a situation in which the target location is known, but *multiple* objects appear at that location, producing uncertainty with regard to which object is the target. This is precisely the situation in an RSVP paradigm in which all of the letters appear at fixation. Although all of the letters in the stream are spatially attended, there is still uncertainty about which is the target. This uncertainty is resolved through an additional act of non-spatial selection based on properties such as color or shape (e.g., report the white letter in the stream). A important question then, is whether the elimination of attentional capture by events outside the focus of attention still holds when there remains uncertainty about which object within the focus of attention is the target. In other words, is the spatial focusing of attention *sufficient* to override attentional capture?

It is likely that the non-spatial selection of a target involves the establishment of attentional control settings for the defining property of the target. For example, determining the identity of a red letter in a sequence of white letters would presumably require an attentional control setting for the color red. An interesting issue concerns the extent to which this attentional control setting for non-spatial selection would influence the allocation of attention in space. Specifically, what effect would an irrelevant distractor have if it appeared outside the focus of spatial attention, but matched the attentional control setting (e.g., "red") for non-spatial selection from a temporal sequence? On the one hand, if the focusing of spatial attention is sufficient to eliminate capture, then peripheral events should not interfere with the identification of targets at the focused location. On the other hand, if attentional control settings for non-spatial selection influence the allocation of spatial attention, then, consistent with the contingent involuntary orienting hypothesis, we might expect an irrelevant stimulus that matches the attentional control setting to capture attention even if the stimulus occurs outside the focus of attention.

To address this issue, we used a variant of the rapid serial visual presentation (RSVP) paradigm. In our task, subjects were required to monitor a centrally presented stream of letters for a target letter of a particular color, and to report the identity of that letter. It was assumed that this task would require attention to be tightly focused at fixation. However, instead of an additional target in the stream, a task-irrelevant, peripheral distractor was presented at different temporal positions relative to the target. As the most interesting data derive from the conditions in which the peripheral distractor preceded the central target, one can think of that distractor as playing the role of T1 in the standard attentional blink paradigm. However, it is important to keep in mind that there was no task associated with that stimulus; it was an irrelevant distractor. The critical manipulation was whether this distractor shared the color that defined the central target or not. We reasoned that if a peripheral distractor matching the color of the central target captures attention, then even under the focused attentional state required by the central stream, a decrement in the identification of the centrally presented target should obtain.

In our initial study with peripheral distractors in an RSVP stream (Folk, Leber, & Egeth, submitted) subjects were shown a stream of 15 letters centered at fixation coming at a rate of about 12 characters per second (42 ms on, 42 ms blank ISI interval). The letters were all gray except for one that was colored (see Figure 6). The colored target character appeared equally often in positions 8-12 of the sequence. For 17 of the subjects the one colored letter was always red; for the remaining 16 subjects it was green. The task was to name the colored letter. On trials containing a distractor, one of the letters in the series was surrounded by four #'s whose inner edges were 5.2 deg above, below, left, and right of the center of the letter. Depending on the distractor condition, the #'s were either all gray, or three were gray and one was red or green.

Subjects received four different distractor conditions that occurred randomly and equally frequently within blocks. In the no-distractor condition each letter in the RSVP stream appeared alone at the center of its frame. In the other three conditions, on one frame four #'s appeared along with the central letter. In the four-gray distractor condition the #'s were all gray. In the same-color distractor condition, one frame contained three gray #'s plus one that was the same color as the target letter. In the different-color distractor condition one frame

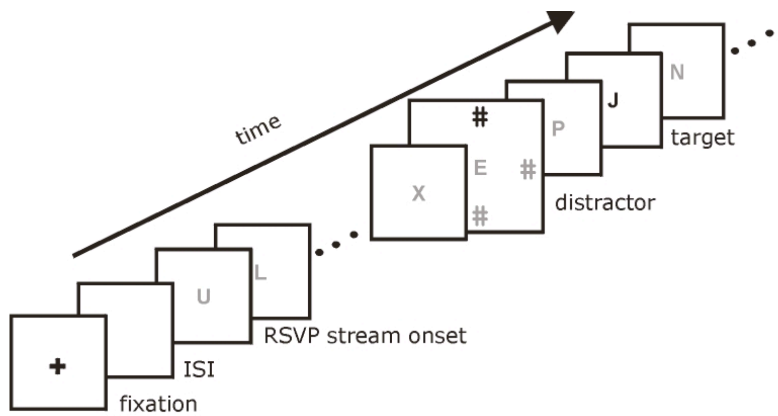

Figure 6. Representation of stimuli and sequence of events on a trial with a distractor-target lag of 2. The characters printed in black were actually red or green (see text for details). In the RSVP stream blank frames were inserted between successive letters. The durations of the letter and blank displays for each experiment are given in the text.

contained three gray #'s and one # that was different in color from the target. (If the subject was searching for a red target, the distractor would be green, and vice versa.)

Distractors could appear at any of four temporal lags with respect to the target. The target could appear two frames after the distractor (lag 2), one frame after the distractor, (lag 1), simultaneous with the distractor (lag 0), or one frame before the distractor (lag -1). For the purpose of analysis, each trial in the no-distractor condition was assigned a lag value, but the distractor was omitted from the sequence.

Subjects were fully informed about the nature of the trials and were urged to ignore the distractor, if possible. There were 24 practice trials and 320 experimental trials.

The results are shown in Figure 7. The four gray distractors condition yielded performance that was independent of lag and essentially identical to the no-distractor condition. In contrast, both the same- and different-color distractor conditions yielded substantial and equal interference that increased as lag varied from -1 to 2. Analysis of variance yielded significant main effects of distractor condition, $F(3,93) = 23.53$, $p < .0001$, and distractor-target lag, $F(3, 93) = 34.56$,

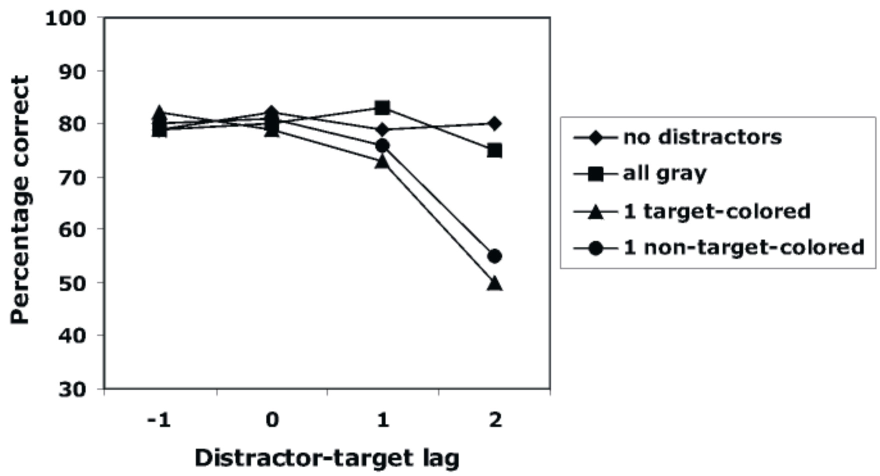

Figure 7. Mean percentage of correct target identification as a function of distractor condition and distractor-target lag. For some subjects targets were red, for others green. The four distractor conditions were: (1) no distractors; (2) four gray number signs (#s); (3) one target-colored number sign and three gray number signs; (4) one number sign colored differently from the target (e.g., green if the target was red) and three gray number signs. In this study all nontarget stream letters were gray. Lag was dummy coded for the no-distractor condition; i.e., each of these trials was assigned one of the lag values arbitrarily.

$p < .0001$, and a significant distractor condition by distractor-target lag interaction, $F(9, 279) = 14.25$, $p < .0001$.

The fact that interference was no greater for the four gray distractor condition than for the the no-distractor condition indicates that the disruption in the two conditions with colored distractors was not due to the mere presence of abrupt onsets in the periphery. It is interesting to compare this aspect of the data with the results of Egeth and Nakama (1999) described earlier. There, it appeared that sudden onset at the fovea disrupts peripheral processing. There are many differences between the studies, but the discrepancy in outcomes is provocative.

Feature Search *vs.* Singleton Detection

The results make it clear that top-down attentional control settings influence the allocation of attention. Although the spatial location of the target was known with certainty, selection of the target required an attentional set for color. When the distractor matched this attentional set it disrupted performance even though its presence and location were irrelevant to the task at hand. In this connection, the fact that the interference was the same in the same- and different-distractor conditions is intriguing. Keep in mind that a given subject only saw one of these colors as the target. The fact that both colors were equally effective as distractors suggests that subjects may have been operating in what Bacon and Egeth (1994; see also Pashler, 1988) have referred to as singleton detection mode. That is, despite the fact that subjects were supposedly looking for just a specific color, it appears that they may have been doing something more like looking for any non-gray item to name. The non-gray distractor was irrelevant to the task both by dint of its shape and its location, but nevertheless subjects could not effortlessly ignore it. An alternative possibility is that the disruption in performance observed in this task does not have anything to do with attentional control settings, but instead reflects capture of attention by any color discrepancy in the periphery (i.e., by a spatial singleton). One way to distinguish these accounts is to set subjects to look for a particular color letter in the stream, but to use several different colors in the stream. With a heterogeneous stream singleton detection mode should not permit target acquisition. In this circumstance, the same-color distractor should show a performance decrement, but the different-color distractor should not.

In our next experiment, we again tested at lags of -1, 0, 1, and 2, but now the RSVP stream consisted of variously colored items. For subjects searching for red targets, the other colors in the stream were green, blue, purple, and gray. For subjects searching for a green target the other colors in the stream were red, blue, purple, and gray. The task was somewhat more difficult, and so the duration of letters was increased from 42 to 56 ms for a total frame duration of 98 ms. Note that the distractor conditions were the same as in the preceding experiment.

The results are shown in Figure 8. Analysis of variance showed that the main effects of distractor condition and distractor-target lag were significant, $F(3,90) = 51.80$, $p < .0001$, and $F(3, 90) = 37.10$, $p < .0001$, respectively, as was the distractor by lag interaction, $F(9, 279) = 27.33$, $p < .0001$. At lag 1 a Tukey test showed that only the same-color distractor differed from the no-distractor condition, while at lag 2 all three distractor conditions differed significantly from the no-distractor condition. However, in contrast to the preceding experiment, the mean for the same-color distractor was significantly lower than the mean for the different-color distractor.

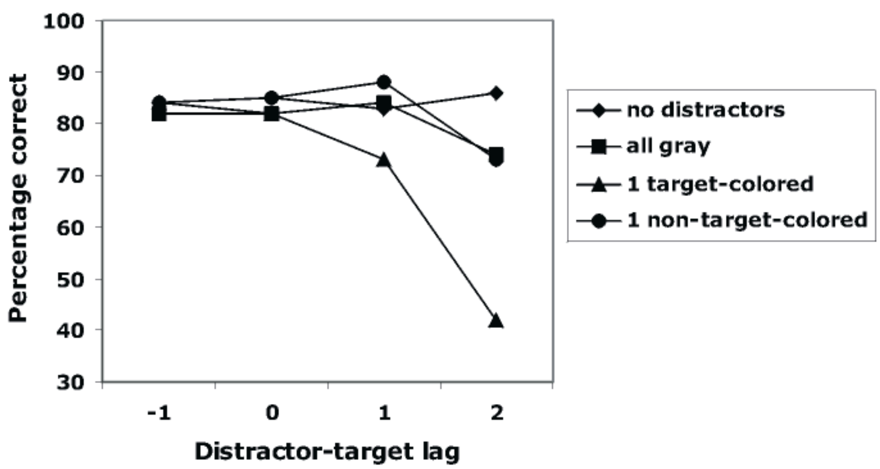

Figure 8. Mean percentage of correct target identification as a function of distractor condition and distractor-target lag. For description of distractor conditions see text or Figure 7 caption. In this study, letters of various colors were presented in the RSVP stream.

These results establish two important points. First, it would appear that the subjects in the preceding experiment had adopted an attentional control setting for any singleton, and not just the specific color of the target for which they were instructed to search. Second the results also make it clear that absent the adoption of singleton detection mode, attentional capture is not simply produced by any color discontinuity in the periphery. What these experiments do not tell us is *why* subjects adopt the strategy of looking for any singleton, especially since adoption of such a strategy leaves one susceptible to interference from salient distractors. This puzzle awaits the results of further research.

Spatial Capture *vs.* Filtering Cost

An implicit assumption underlying our experiments on capture by peripheral distractors in RSVP experiments is that the deficits in detection of the central target are due to the spatial capture of attention by the involuntary orienting of spatial attention to an irrelevant spatial location containing a color singleton. However, it is possible that the deficit is not spatial at all, and might even reflect a process as general as the "filtering cost" of Treisman, Kahneman, and Burkell (1983; see Folk & Remington, 1998, for evidence of both spatial and nonspatial forms of attentional capture). To address this issue the design of the preceding experiment, which used a heterogeneous RSVP stream, was changed in several ways. First, the four number-sign (#) distractors were replaced by four boxes. (As before, there were four gray; three gray and one red; or three gray and one green; there was also a no-box condition.) Second, in the frame immediately following the boxes, four gray letters appeared, each one at the center of the space previously occupied by a box. On each trial one of the four peripheral letters (the prime) was the same as the target letter for that trial. The position of the prime was varied systematically across trials. In the same-color and different-color distractor conditions the position of the prime varied such that it appeared at the location of the colored peripheral singleton box on an unpredictable 1/4 of the trials and at the location of one of the three gray non-singleton boxes on 3/4 of the trials. (Note singleton vs. non-singleton status here refers to the color of the boxes.) If spatial attention is drawn to the location of a same-color peripheral singleton, then the likelihood that the gray letter that follows immediately at that location will be identified should be increased, and, if that character is the prime, then we might expect that identification of the central target would be more likely (via perceptual priming, for example).

Fifteen subjects searched for a red target, and fifteen searched for a green target. Only two lags were studied, 0 and 2. Note that lag refers here (as before) to the separation between the distractor boxes (not the gray peripheral letters) and the target.

Mean percentage of correct target identifications as a function of distractor condition, distractor-target lag, and prime status are presented in Figure 9. Note first that there is a substantial decrement in performance at lag 2 when the distractor box is the same color as the target letter. This decrement is consistent with the effect of distractors observed in the immediately preceding experiments. Analysis of variance showed distractor condition and target-distractor lag both produced significant main effects, $F(3, 84) = 39.65$, $p < .0001$, and $F(1, 28) = 6.93$, $p < .01$, respectively. These two variables also entered into a significant interaction, $F(3, 84) = 16.96$, $p < .0001$. As is evident in the figure, this interaction is driven by a deficit in performance that is specific to lag 2 in the same-color distractor condition.

Consider next the influence of the peripheral prime on central target identification. Prime status (i.e., whether the prime letter was at the location of a singleton or a nonsingleton distractor square) produced a significant main effect, $F(1, 28) = 4.51$, $p < .05$, and interacted significantly with distractor condition. Simple effects analyses of prime status at each distractor condition yielded a

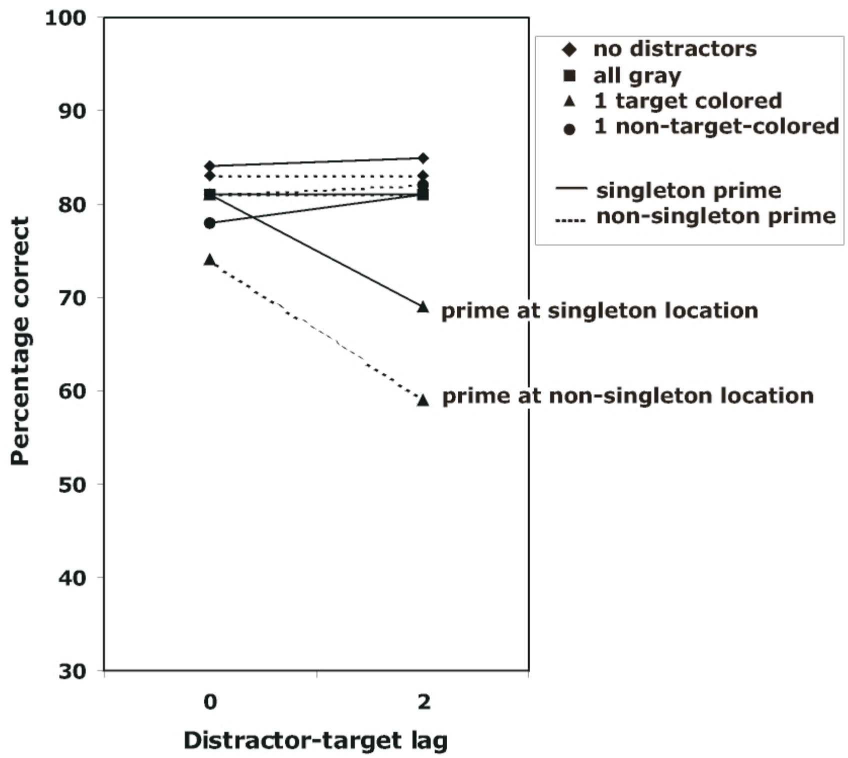

Figure 9. Mean percentage of correct target identification as a function of distractor condition and distractor-target lag and prime status. In this study peripheral boxes appeared in one frame, followed by a frame containing peripheral gray letters, one of which was identical to the target. The distractor conditions refer to the relation between the presence and color of the boxes and the color of the target. The novel finding here is that performance was significantly better when the prime letter was at a singleton as opposed to a non-singleton location in the condition in which there was a target-colored singleton distractor in the stream (denoted by triangles in the figure).

significant effect in the same-color condition only, $F(1, 28) = 7.68$, $p < .01$. In this condition, gray primes at the singleton location produced a significant enhancement in central target identification relative to trials on which primes appeared at a non-singleton location.

In an effort to obtain converging evidence for the spatial capture of attention by distractor squares that were the same color as the target, we examined error trials. On error trials, subjects report a letter other than the target letter in the central stream. If attention is shifted to the spatial location of the singleton, this should increase the likelihood that on error trials subjects will report the letter at the singleton location. The percentage of error trials on which subjects incorrectly reported the singleton letter instead of the central target letter is shown in Figure 10.

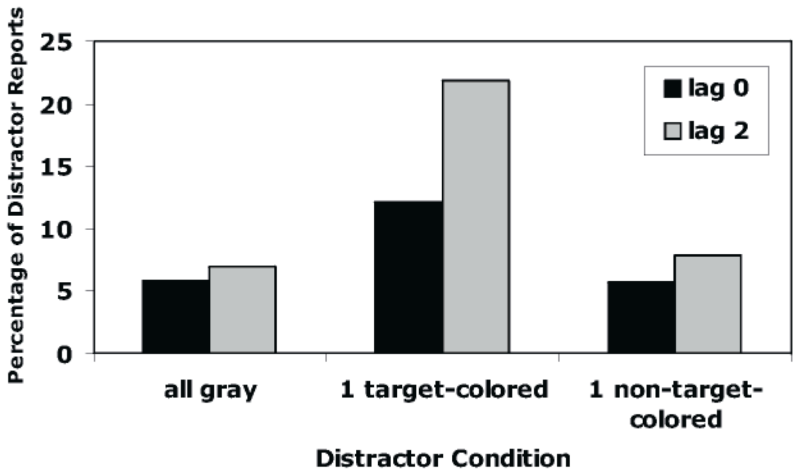

Figure 10. On error trials, mean percentage of erroneous reports of the letter appearing in the same location as a singleton distractor square as a function of distractor condition and distractor-target lag. In the four gray square condition, where there was no singleton, on each trial one randomly selected location was treated as if it had contained a color singleton. The other two distractor conditions represent cases where the singleton square was the same color as the target or a different color from the target.

(Dummy coding was used in this analysis; In the four gray square condition, where there was no distractor, on each trial one randomly selected location was treated as if it had contained a color singleton.) Note that there are substantially more errors of this type in the same-color condition than in either of the other two distractor conditions, and that this is particularly evident at lag 2. An ANOVA yielded significant main effects of both distractor condition and lag, $F(2, 56) = 7.45$, $p < .01$, and $F(1, 28) = 4.57$, $p < .05$, respectively. Although the interaction between distractor condition and target-distractor lag just failed to reach significance, focused comparisons yielded a significant effect of lag in the same color condition only, $F(1, 28) = 11.73$, $p < .01$.

General Discussion

The general thrust of our conclusions should already be clear. The contingent involuntary orienting hypothesis of Folk et al. (1992) is able to account for a wide variety of data. It was originally developed to account for data in cuing experiments in which reaction time was the chief independent variable. Hendel's (1998/1999) work has shown that similar results can be obtained when stimuli are shown briefly and masked, with accuracy the dependent variable. Further, her studies suggest that an apparent exception to the hypothesis (Joseph & Optican, 1996) can be explained on the assumption that search for an L among Ts is carried out in singleton detection mode.

Folk, Leber, and Egeth (submitted) have subsequently extended the application of the contingent capture hypothesis to the realm of RSVP tasks. As several investigators have shown, the RSVP task is attentionally demanding. More specifically, Joseph, Chun, and Nakayama (1997) have shown that when attention is focused at the center of the field the seemingly preattentive task of detecting an orientation singleton in the periphery becomes difficult. Nakama and Egeth (1999) have shown that the task is demanding even when the first target is absent from the stream. This may challenge theories that claim the blink is due solely to processing of the first target, but it does nothing to detract from the idea that identifying a letter in an RSVP stream is an attentionally demanding task. These results set the stage for determining whether spatially focused attention can be disrupted by a peripheral distractor that shares the defining feature of a centrally displayed target. Such disruption was observed. The form of this disruption is much like that of the attentional blink, but with the role of the traditional first target played here by an irrelevant peripheral distractor. The conditions under which the disruption occurred follow closely what would be expected on the basis of the contingent capture hypothesis. When the target was, say, a red letter in a stream of variously colored letters, only the red distractor interfered with performance. However, when the target was a red letter in a stream of gray letters, then both a red and a green distractor interfered with performance. This suggests subjects perform this latter task by adopting singleton detection mode. Leber and Egeth (2001) have begun to systematically explore what determines whether subjects adopt the feature-search or singleton-detection mode of processing.

References

Bacon, W.F., & Egeth, H.E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics, 55*, 485-496.

Braun, J. (1998). Vision and attention: the role of training. *Nature, 393*, 424-425.

Braun, J. & Sagi, D. (1990). Vision outside the focus of attention. *Perception & Psychophysics, 48*, 45-58.

Braun, J. & Sagi, D. (1991). Texture-based tasks are little affected by second tasks requiring peripheral or central attentive fixation. *Perception, 20*, 483-500.

Breitmeyer, B. G. (1984). *Visual masking: An integrative approach*. New York, Oxford.

Broadbent, D. E., & Broadbent, M. H. (1987). From detection to identification: Response to multiple targets in rapid serial visual presentation. *Perception & Psychophysics, 42*, 105-113.

Chun, M.M., & Potter, M.C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 109-127.

Egeth, H.E., & Dagenbach, D. (1991). Parallel versus serial processing invisual search: Further evidence from subadditive effects of visual quality. *Journal of Experimental Psychology: Human Perception and Performance*, 17(2), 551-560.

Egeth, H.E., & Nakama, T. (1999). Pop-out line detection affected by string length of concurrent RSVP task and "attentional capture" by initial item [Abstract] *Investigative Ophthalmology & Visual Science, 40*, S808.

Folk, C.L., Leber, A.B., & Egeth, H.E. (submitted). Made you blink! Contingent Attentional Capture Produces a Spatial Blink.

Folk, C.L., & Remington, R. (1998). Selectivity in distraction by irrelevant featural singletons: Evidence for two forms of attentional capture. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 847-858.

Folk, C.L., Remington, R.W., & Johnston, J.C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception & Performance, 18*, 1030-1044.

Francolini, C. M., & Egeth, H. E. (1979). Perceptual selectivity is task-dependent: The pop-out effect poops out. *Perception & Psychophysics, 25*, 99-110.

Hendel, S.K. (1999). Attentional shifts to orientation differences are contingent on task demands. (Doctoral dissertation, Johns Hopkins University, 1998). *Dissertation Abstracts International, 60-02B,* 849.

Jolicoeur, P. (1998). Modulation of the attentional blink by on-line response selection: Evidence from speeded and unspeeded task decisions. *Memory & Cognition, 26,* 1014-1032.

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics, 43*, 346-354.

Joseph, J. S., Chun, M. M., & Nakayama, K. (1997). Attentional requirements in a "preattentive" feature search task. *Nature, 387,* 805-807.

Joseph, J.S., & Optican, LM. (1996). Involuntary attentional shifts due to orientation differences. *Perception & Psychophysics, 58*, 651-665.

LaBerge, D. (1983). Spatial extent of attention to letters and words. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 371-

379.

Leber, A. B., & Egeth, H. E. (2001, May). *Exploring mode selection in visual search.* Poster session presented at the annual meeting of the Vision Sciences Society, Sarasota, FL.

Mounts, J.R.W. (2000). Evidence for suppressive mechanisms in attentional selection: Feature singletons produce inhibitory surrounds. *Perception & Psychophysics, 62*, 969-983.

Mueller, H.J., & Rabbitt, P.M. (1989). Spatial cueing and the relation between the accuracy of "where" and "what" decisions in visual search. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 41,* 747-773.

Nakama, T., & Egeth, H.E. (1999). The dependence of attentional blink on the temporal position of the target in RSVP [Abstract]. *Investigative Ophthalmology & Visual Science, 40*, S49.

O'Regan, J.K., Rensink, R.A. & Clark, J.J. (1999). Change-blindness as a result of "mudsplashes". *Nature, 398*, 34.

Pashler, H. (1988). Cross-dimensional interaction and texture segregation. *Perception & Psychophysics, 43*, 307-318.

Posner, M.I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3-25.

Posner, M. I., & Cohen, Y. A. (1984). Components of visual orienting. In H. Bouma and D. G. Bouwhuis, Eds. *Attention and Performance*. Hillsdale, NJ: Erlbaum. 10: 531-554.

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,* 849-860.

Remington, R.W., Johnston, J.C., & Yantis, S. (1992). Involuntary attentional capture by abrupt onsets. *Perception & Psychophysics, 51*, 279-290.

Sagi, D. & Julesz, B. (1985). "Where" and "what" is vision. *Science, 228*, 1217-1219.

Teichner, W.J. & Krebs, M.J. (1974). Laws of visual choice reaction time. *Psychological Review, 81*, 75-98.

Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception & Psychophysics, 50,* 184-193.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics, 51*, 599-606.

Theeuwes, J. (1994). Stimulus-driven capture and attentional set: Selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human Perception & Performance, 20,* 799-806.

Theeuwes, J. (1996). Perceptual selectivity for color and form: On the nature of the interference effect. In A.F. Kramer, M. Coles, & G. Logan (Eds.), *Converging operations in the study of visual selective attention* (pp. 297-314). Washington, DC: American Psychological Association.

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

Treisman, A., Kahneman, D., & Burkell, J. (1983). Perceptual objects and the cost of filtering. *Perception & Psychophysics, 33*, 527-532.

Turatto, M. & Galfano, G. (2001). Attentional capture by color without any relevant attentional set. *Perception and Psychophysics, 63*, 286-297.

Visser, R. A. W., Bischof, W.F., & Di-Lillo, V. (1999). Attentional switching in spatial and nonspatial domains: Evidence from the attentional blink. *Psychological Bulletin, 125*, 458-469.

Yantis, S. (1993). Stimulus-driven attentional capture. *Current Directions in Psychological Science, 2*, 156-161.

Yantis, S., & Egeth, H. E. (1999). On the distinction between visual salience and stimulus-driven attentional capture. *Journal of Experimental Psychology: Human Perception & Performance, 25*¸661-676.

Yantis, S., & Hillstrom, A.P. (1994). Stimulus-driven attentional capture: Evidence from equiluminant visual objects. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 95-107.

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: voluntary vs. automatic allocation. *Journal of Experimental Psychology: Human Perception & Performance, 16*¸121-134.

Author Note

The research reported in this chapter was supported in part by grants to H.E. Egeth from NIMH (R01MH57388) and the FAA (2001-G-020) and by a grant from NSF to C.L. Folk (BCS-9817673). The authors would like to thank Vince DiLollo, Anne Hillstrom, Jim Johnston, Julian Joseph, Peter Jusczyk, Hal Pashler, Matt Peterson, Eric Ruthruff, Jeremy Wolfe, and Steve Yantis for helpful comments on various parts of this work.