Journal of Experimental Psychology:
Human Perception and Performance
1993, Vol. 19, No. 3, 676-681

0096-1523/93/$3.00

OBSERVATIONS

Stimulus-Driven Attentional Capture and Attentional Control Settings

Steven Yantis

Jonides and Yantis (1988) found that abrupt-onset singletons capture attention in visual search when onset is orthogonal to the target's defining and reported attributes and that color and brightness singletons do not. They concluded that abrupt onset may be unique in capturing visual attention. Folk, Remington, and Johnston (1992) challenge this conclusion and argue that (a) the occurrence of attentional capture is contingent on the adoption of an appropriate attentional control setting by the observer and (b) properties other than onset (in particular, color) can capture attention involuntarily. In this article, each of these claims is critically evaluated, and it is argued that the results reported by Folk et al., though important, do not definitively corroborate either one. The available evidence concerning stimulus-driven attentional capture is summarized, and 3 empirical generalizations that characterize the evidence are advanced.

Selection of information from visual displays is widely believed to be controlled in at least two distinct ways. *Goal-directed selection* refers to the observer's ability to control what regions or objects in the visual field are selected for further visual processing given a set of goals and beliefs about the current task; this is sometimes called *top-down* or *endogenous* control over the locus of attention. *Stimulus-driven selection* refers to the fact that certain properties of the stimulus may capture attention independently of the observer's goals and beliefs; this is sometimes referred to as *bottom-up* or *exogenous* control over the locus of attention. Either one of these mechanisms, or some combination of them, may determine how attention is distributed.

Although goal-directed selection has been a major research question for several decades, stimulus-driven selection became a significant focus of investigation relatively recently.[1] The approach has been to determine the conditions under which the observer's distribution of attention is determined by attributes of the stimulus and not by the observer's goals or intentions (see Yantis, in press, for a review).

The experiments reported by Folk, Remington, and Johnston (1992) fall into an intermediate category; they are concerned with the interaction between the intentions of the observer and stimulus-driven attentional capture. The experiments provide clear evidence that when observers are prepared to identify a display element that is defined by some prespecified featural singleton (e.g., name the red target object that is displayed among white nontarget objects), then a preceding featural singleton (a to-be-ignored "cue") in that same dimension cannot be ignored. They further show that when subjects are in the state of readiness induced by these instructions, irrelevant variation in another dimension (e.g., an onset singleton) does not capture attention. This is an important finding and adds significantly to our previous understanding of attentional capture. In particular, the central point made by Folk et al. that the bottom-up control of attention by stimuli interacts with the observer's state of attentional readiness provides an important foundation for further developments in attentional theory.

However, Folk et al. make two related claims that are not fully justified by their experiments or by their analysis of the literature. The first claim (which I call the *contingent-capture hypothesis*) is that attentional capture by any attribute (including abrupt onset) is contingent on the observer's adoption of an appropriate attentional set in advance; this claim is incompatible with the finding of Jonides and Yantis (1988) that attention is captured by abrupt onset when observers are in a "neutral" state of attentional readiness with respect to abrupt onset. The second claim (which I call the *broadened-scope hypothesis*) is that attentional capture occurs involuntarily for attributes other than abrupt onset; this claim is incompatible with the finding of Jonides and Yantis (1988) that abrupt onsets may be unique in this regard.

In this article, I critically examine both of these claims. My goal in undertaking this examination is to shed further light

Steven Yantis, Department of Psychology, Johns Hopkins University.

The research discussed in this article has been supported by Grant R01-MH43924 from the National Institute of Mental Health.

I thank Bill Bacon, Howard Egeth, Chip Folk, Anne Hillstrom, Doug Johnson, and Ray Klein for valuable comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Steven Yantis, Department of Psychology, Johns Hopkins University, Baltimore, Maryland 21218. Electronic mail may be sent to yantis@jhuvms.hcf.jhu.edu.

[1] The earliest examples include work by Posner and Jonides (e.g., see Jonides, 1981; Posner, 1980). With the exception of Jonides (1981), however, stimulus-driven attentional capture was not usually the focus of this work; instead, peripheral cues were used as a tool to capture attention in the pursuit of other theoretical or empirical goals.

on the important issues raised by Folk et al. I will reinterpret some of their results in the context of other recent findings to provide a more complete understanding of the mechanisms responsible for stimulus-driven visual selection.

For the discussion that follows, it will be useful to invoke a distinction suggested by Duncan (1985) between the defining attribute and the reported attribute of a target in visual search. The defining attribute of a target is the attribute that makes that target relevant for a given visual-search task; it is what the observer is "looking for" during search. The reported attribute is what the observer uses in his or her response. For example, in a task requiring subjects to name the red letter that is displayed among one or more green letters, the color red is the defining attribute, and the letter name is the reported attribute. In a task requiring observers to report the color of the X in a background of other colored letters, these roles are reversed. Finally, in a task requiring observers to determine whether an X was present anywhere in a multicolored display, letter name (or shape) is both the defining and the reported attribute (and variation in color is irrelevant or orthogonal to the task). I take it as an uncontroversial assumption that the defining attribute determines the observer's selective readiness, or what Folk et al. call their *attentional control setting*, and that once the object possessing the defining attribute is selected, its reported attribute is computed (if necessary) and reported. I define *stimulus-driven attentional capture* as attentional capture by an attribute that is independent of either the defining or the reported attribute of the target.

Contingent-Capture Hypothesis

Folk et al. claim that attention is captured only when the observer has actively adopted an appropriate attentional set. The claim is that involuntary capture of attention is contingent on the adoption of the appropriate set and will not occur if such a set is not adopted. This claim is inconsistent with the finding of Jonides and Yantis (1988) that abrupt onsets capture attention in the absence of any deliberate selective readiness for onset. The central point of contention here is whether there can be stimulus-driven capture of attention in the absence of a deliberate state of attentional readiness on the part of the observer. In this section, I review the Jonides and Yantis experiments and evaluate their implications for the contingent-capture hypothesis.

In Jonides and Yantis (1988), the defining and reported attribute of the stimulus was letter shape: Observers were required to determine whether a prespecified letter was present in a visual search array. The purpose of the experiment was to assess the effect (if any) of an irrelevant singleton in color, brightness, or onset. In the color condition, one of the letters was red and the rest were green (or vice versa); in the brightness condition, one letter was bright and the rest were dim; in the onset condition, one of the letters had an abrupt onset and the others did not (they were no-onset stimuli revealed by the removal of irrelevant camouflage at display onset). In each case, the existence of a featural singleton (in color, brightness, or onset) was orthogonal to the observer's task. Observers were encouraged to ignore these singletons if they could because they provided no information about the location of the upcoming target: in particular, the singleton was no more likely to be a target than was any other element in the display.

We found in the color and brightness conditions that the time to detect the target increased linearly with display size and did not depend on whether the target was a singleton. In the onset condition, the results were quite different. The time to detect the target when it was the onset singleton did not depend on display size (i.e., the display-size function was "flat," a signature of attentional capture), whereas search time when the target was one of the no-onset elements increased linearly with display size. This demonstrates that when observers are in a "neutral" state with respect to color, brightness, or onset singletons (i.e., when those dimensions are orthogonal to both the defining and the reported attributes of the target), then only onset singletons capture attention.

Folk et al. suggest that capture was observed in the onset condition of Jonides and Yantis (1988) but not in the color or brightness conditions, because the onset condition required observers to "monitor" for luminance onset singletons (i.e., adopt an attentional control setting for onset), whereas the color and brightness conditions did not require observers to monitor for color or brightness singletons (i.e., no attentional control setting for color or brightness, respectively, was required). They state that in the onset condition, both targets and distractors were signaled by luminance changes and that capture "could have resulted from subjects' use of luminance changes to locate the target" (p. 1031). They contrast this with the color condition of the Jonides and Yantis (1988) experiments, "in which an entire search set was simultaneously presented and one element of that set appeared in a unique color, [so] locating the target did not require monitoring for a static discontinuity" (p. 1031); the term *static discontinuity* refers to the color singleton. The implication of this passage is that the onset condition did require monitoring for a *dynamic discontinuity*, that is, an onset.

The distinction drawn here is illusory: The defining attribute of the target in all conditions of the experiment was letter shape. There is no sense in which the onset singleton (or any other luminance change) defined or was even correlated with the position of the target, just as neither the color singleton in the color condition nor the brightness singleton in the brightness condition defined or was correlated with the position of the target.[2] All of the elements in the onset condition of the experiment were accompanied by a luminance change: The onset element produced a luminance increment, and the no-onset elements produced a (smaller) luminance decrement. But this cannot be characterized as an attentional

[2] One might argue that even though none of the singletons was correlated in position with the target, observers used the onset singletons to guide search anyway because they were subjectively more salient than the color or brightness singletons. In debriefing, however, we have found that observers rarely notice the existence of onsets in these displays; it usually takes a fair amount of explaining to convey what we even mean by *onsets*. Color and brightness singletons are much more subjectively salient than onset singletons.

control setting for onset, and it does not explain why the onset element alone enjoyed an attentional advantage.

Folk et al. therefore acknowledge that there may be a "default" attentional set for luminance onsets: "When there is little motivation to configure the system for any other property, abrupt luminance change . . . may be instantiated as the 'default' setting" (p. 1042). But they stress that this does not necessarily make onsets "special."

This latter claim, however, is misleading. There is a fundamental difference between the circumstances in which abrupt onsets capture attention and those in which other stimulus attributes (such as color) capture attention: Capture by abrupt onset can occur in the absence of a deliberate top-down intention to detect onsets, whereas capture by other attributes occurs only in the presence of a deliberate top-down intention to detect those attributes. In just this sense, abrupt onset is apparently unique.

Of course, to claim that onsets have a special status is not to claim that they have absolute control over attention. In their experiments, Folk et al. provide clearcut evidence that when observers are set for a color singleton, onset singletons do not capture attention. Similarly, Yantis and Jonides (1990) found that when observers focus their attention at a spatial location in anticipation of a target event there, an abrupt onset elsewhere does not capture attention. These results reveal that capture by abrupt onset can be overridden by top-down control. This makes adaptive sense: If an organism has some reason to attend to an attribute or location, then onsets elsewhere should not be allowed to arbitrarily distract from it. This ability to override the effects of abrupt onset is presumably just what Folk et al. have in mind as the function of an attentional control setting. Even though their influence is not absolute, however, I argue that onsets play a special role in vision.

Broadened-Scope Hypothesis

Another claim made by Folk et al. is that stimulus attributes other than abrupt onset (for example, color) can capture attention involuntarily. This claim conflicts with the conclusion of Jonides and Yantis (1988) that onsets may be unique in this regard. The point of contention here concerns how the term *involuntary* should be construed.

To assess this point, it is useful to view the Folk et al. experiments in terms of the defining and reported attributes of the task (see earlier discussion of Duncan, 1985). In all conditions of the experiments, the reported attribute was target shape (X vs. =). In the color-target condition, the defining attribute was color, whereas in the onset-target condition, the defining attribute was onset.[3] Folk et al. found that when a preceding to-be-ignored featural singleton (the "cue") appeared that matched the defining attribute of the target, it captured attention (as evidenced by prolonged reaction times to identify the target). The claim is that the cue involuntarily captured attention when it matched the defining attribute of the target and that this capture was contingent on the attentional set of the observer. That this occurred for color as well as for onset is the basis for the claim that the scope of involuntary attentional capture should be broadened beyond abrupt onset to include (at least) color.

But in this experiment, the observer was in a state of attentional readiness for red singletons, so one cannot claim that attention was captured involuntarily by the red singleton. On the contrary, the attentional response was entirely consistent with the observer's goals concerning the defining attribute of the target. The limitation revealed by the Folk et al. experiments is more properly characterized as a purely temporal one: Observers evidently cannot efficiently select the object satisfying the conjunction of color with a small temporal window. That is, the task Folk et al. set for their subjects was to name the shape of the red object appearing in Temporal Position 2 of the sequence, where Temporal Position 1 is occupied by the cue and Temporal Position 2 is occupied by the target display (the cue was presented for 50 ms, followed by a 100-ms blank interval, and then the target display for 50 ms). That the observers were set in advance for a color singleton undermines the claim that attentional capture by a color singleton was involuntary.[4] Indeed, observers were actively seeking a color singleton. What the Folk et al. experiments reveal is that this attentional set cannot be engaged or switched over very small time scales (e.g., on the order of a few hundred ms). This is an important finding, but it differs significantly from the claim advanced by Folk et al.

In fact, there is evidence that observers cannot ignore irrelevant singletons even if they are in a different dimension than the relevant one (Pashler, 1988, Experiment 6; Theeuwes, 1991a, 1992). This only occurs, however, if the defining attribute of the target is a singleton (i.e., if the observer has adopted an attentional set for a singleton). For example, in Theeuwes (1991a), the reported attribute of the target was orientation (horizontal or vertical line segment), and the defining attribute was the color or brightness of a circle surrounding each of the display locations. When the defining attribute was a color singleton (i.e., report the orientation of the line segment inside the red circle embedded among green circles), then on some trials another location contained an (irrelevant) brightness singleton (i.e., a bright green circle). The presence of the irrelevant singleton disrupted perfor-

[3] This is not strictly correct. Although it is true that the target in the onset-target conditions was the only element in the visual field with an abrupt onset (because it was the only element in the visual field at all), it was also the only element in the visual field with any attribute (including color, shape, etc.). The task set for the subject was to name the shape of the single element that appeared in the visual field. Folk et al. assume that the implicit defining feature (and hence the attentional control setting) in the onset–target condition was onset. However, one could as easily assert that the defining attribute was color or shape.

[4] Folk et al.'s finding in their Experiment 4 that a green singleton captures attention when the target was defined as a red singleton is an important additional part of this story. It reveals that the particular attribute that is the target of search is not the only singleton that will capture attention; instead, when observers are prepared for a red singleton, singletons of other colors will also capture attention. A similar conclusion has been advanced by Theeuwes (1991a, 1992); see discussion in the next paragraph.

Table 1
Effects of Static and Onset Singletons on Attention as a Function of the Observer's State of Selective Readiness

Selective state	Singleton	
	Static	Onset
Attentional set for static singleton	Capture (Folk et al., 1992; Pashler, 1988; Theeuwes, 1991a, 1992)	Do not capture (Folk et al., 1992)
No relevant attentional set	Do not capture (Folk, 1990; Jonides & Yantis, 1988; Martin & Benson, 1991)	Capture (Jonides & Yantis, 1988; Lambert, Spencer, & Mohindra, 1987; Theeuwes, 1991b; Yantis & Jonides, 1984, 1990)
Attentional focus on spatial location	Do not capture (assumed but not yet known)	Do not capture (Theeuwes, 1991b; Yantis & Jonides, 1990)

Note. A static singleton as defined by Folk, Remington, and Johnston (1992) is a singleton in a feature dimension that varies over space and not time. It includes at least the following: color, form, orientation, brightness.

mance, even though it was in a different dimension from the target's defining feature.

The general conclusion that can be drawn from these experiments, together with those of Folk et al., is that when observers adopt a perceptual set for a specific featural singleton, then both relevant and irrelevant singletons (if they are salient enough) tend to capture attention. This lends support to Folk et al.'s claim that singletons other than onset can capture attention but only if the task requires observers to deliberately monitor for singletons. There is no evidence that any attribute other than onset will capture attention in the absence of a deliberate attentional set for singletons. In other words, the only stimulus attribute that has been shown to exhibit stimulus-driven visual capture is abrupt visual onset.

This point is perhaps the most important one I wish to make. In my view, the key distinction between stimulus-driven attentional capture and goal-directed attentional effects is as follows: (a) Goal-directed attention allocation depends on what observers are intending to do given task instructions, their expectations about what they will see, and other similar goal-related factors. (b) Stimulus-driven attentional capture can be said to occur only when the attribute that elicits it is independent of the defining and reported attributes of the target; it depends solely on properties of the visual system that are insensitive to the kinds of goal-related factors delineated in (a).[5] According to this view, onsets capture attention in a stimulus-driven fashion, but other attributes do not.

Summary

Folk et al. made two claims about the properties of stimulus-driven attentional capture. First, they claimed that singletons in general (and onset singletons in particular) can capture attention only when observers have adopted in advance an attentional set for that attribute, and not otherwise (the contingent-capture hypothesis). I have argued that in the experiments of Jonides and Yantis (1988), there was no deliberate attentional set for onset or color or brightness (the defining and reported attributes of the target were shape) and that onset singletons nevertheless captured attention, whereas color and brightness singletons did not. If this argument is correct, attentional capture by onset singletons is not contingent on the adoption of an appropriate attentional set.

Second, Folk et al. claimed that stimulus attributes other than onset (e.g., color) may capture attention involuntarily (the broadened-scope hypothesis). I have argued that this holds only if observers adopt a deliberate perceptual set in which the target's defining attribute is a singleton and that this cannot readily be interpreted as involuntary with respect to that attribute.

Survey of Related Evidence

Many recent studies have examined the conditions under which attention is captured in a stimulus-driven fashion. The following statements (summarized in Table 1) are supported by the available evidence.

Statement 1. When the observer's task requires an attentional set for a featural singleton (i.e., the defining attribute of the target is a singleton), then almost any to-be-ignored singleton will capture attention, including the following: (a) an irrelevant matching singleton (e.g., a red singleton when the target is defined as a red singleton [see Folk et al., 1992, Experiment 2; Theeuwes, 1991a], or a motion singleton when the target is defined as a motion singleton [see Folk & Wright, 1992]); (b) an irrelevant singleton with a different value in the same dimension (e.g., a green singleton when the target is defined as a red singleton; see Folk et al., 1992, Experiment 4); (c) an irrelevant singleton in a different dimension (e.g., a simultaneous color singleton when the target is defined as a form singleton [Pashler, 1988, Experiment 6; Theeuwes, 1992], or a simultaneous brightness singleton when the target is defined as a color singleton [Theeuwes, 1991a]).

[5] Except insofar as goal-related factors can override stimulus-driven capture, as shown in Folk et al.'s (1992) Experiment 1 and in Yantis and Jonides (1990).

Statement 2. When the observer's task requires an attentional set for a static featural singleton (e.g., a color singleton), then an onset singleton does not capture attention (Folk et al., 1992, Experiment 1).

Statement 3. When the observer's task does not require a deliberate attentional set for a featural singleton, then abrupt-onset singletons capture attention (e.g., Jonides & Yantis, 1988; Lambert, Spencer, & Mohindra, 1987; Theeuwes, 1991b; Yantis & Jonides, 1984, 1990).

Statement 4. When the observer's task does not require a deliberate attentional set for a featural singleton, then singletons in the following dimensions do not capture attention: color (Folk, 1990; Jonides & Yantis, 1988; Martin & Benson, 1991); brightness (Jonides & Yantis, 1988); and motion (Folk & Wright, 1992; Hillstrom & Yantis, 1992).

Statement 5. When the observer focuses attention on a spatial location in advance, then an onset singleton elsewhere does not capture attention (Theeuwes, 1991b; Yantis & Jonides, 1990).

Statement 6. It follows from Statements 3 and 5 (but has not been demonstrated directly) that if attention is focused on a spatial location in advance, then no singleton of any kind will capture attention.

These statements can be condensed into the following generalizations concerning stimulus-driven visual capture. First, when the observer's task requires a deliberate attentional set for a featural singleton, then both relevant and irrelevant singletons of any kind will capture attention.[6] Second, when the observer's task does not require a deliberate attentional set for a featural singleton, then onset singletons will capture attention, but other types will not. Finally, when the observer focuses attention on a spatial location, then no singleton appearing elsewhere will capture attention.

Conclusion

The evidence suggests, then, that onset singletons alone can capture attention in the absence of a deliberate attentional set for that attribute. Why are onsets different? One possible reason is that onset singletons usually accompany newly presented objects. Yantis and Hillstrom (in press) have found that new objects capture attention even in the absence of an accompanying luminance increment. Furthermore, they found that a luminance increment itself, without the appearance of a new object, is not sufficient to capture attention. These findings suggest that the creation of a new object file to represent the attributes of an object (e.g., Kahneman, Treisman, & Gibbs, 1992) triggers an attentional interrupt signal even in the absence of a relevant attentional set. Related ideas have been explored by Johnston, Hawley, Plewe, Elliott, and DeWitt (1990).

Although I have taken issue with some of the specific claims made by Folk et al. (1992), I must stress again that their findings provide important evidence concerning the interaction between goal-directed and stimulus-driven attentional selection. They contribute to a growing body of evidence concerning the conditions under which attention is directed in a stimulus-driven manner by highlighting the central role played by the observer's state of selective readiness in visual selection tasks. The similarities between the theoretical position they advance and the one I have defended in this article are both more numerous and more important than are the differences.

Nevertheless, the evidence reported by Folk et al. does not unconditionally corroborate the contingent-capture or the broadened-scope hypotheses as stated in their article. The qualifications I have outlined may contribute to a more complete understanding of the implications of their findings for theories of visual attention.

---

[6] This point is subject to the two caveats noted earlier: Folk et al. (1992, Experiment 1) showed that onset singletons will not capture if the observer is set for a color singleton; Theeuwes (1992) showed that a to-be-ignored singleton with low salience will not capture attention.

References

Duncan, J. (1985). Visual search and visual attention. In M. Posner & O. Marin (Eds.), *Attention and performance XI* (pp. 85–106). Hillsdale, NJ: Erlbaum.

Folk, C. L. (1990, March). *Spatial discontinuities and involuntary shifts of spatial attention.* Paper presented at the 61st Annual Meeting of the Eastern Psychological Association, Boston.

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

Folk, C. L., & Wright, J. H. (1992, April). *Does apparent motion capture attention?* Paper presented at the 63rd Annual Meeting of the Eastern Psychological Association, Boston.

Hillstrom, A. P., & Yantis, S. (1992, April). *Attentional capture by visual motion.* Paper presented at the 63rd Annual Meeting of the Eastern Psychological Association, Boston.

Johnston, W. A., Hawley, K. J., Plewe, S. H., Elliott, H. M. G., & DeWitt, M. J. (1990). Attention capture by novel stimuli. *Journal of Experimental Psychology: General, 119,* 397–411.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. B. Long & A. D. Baddeley (Eds.), *Attention and performance IX* (pp. 187–203). Hillsdale, NJ: Erlbaum.

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

Kahneman, D., Treisman, A., & Gibb, B. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24,* 175–219.

Lambert, A., Spencer, E., & Mohindra, N. (1987). Automaticity and the capture of attention by a peripheral display change. *Current Psychological Research and Reviews, 6,* 136–147.

Martin, D. W., & Benson, A. E. (1991). *Is there a color advantage in visual search?* Paper presented at the 32nd Annual Meeting of the Psychonomic Society, San Francisco.

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.

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

Theeuwes, J. (1991b). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception & Psychophysics, 49,* 83–90.

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

Yantis, S. (in press). Stimulus-driven attentional capture. *Current Directions in Psychological Science.*

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

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

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

Received March 24, 1992
Revision received July 13, 1992
Accepted July 13, 1992 ■