Spontaneous allocation of visual attention:
Dominant role of uniqueness

HAROLD PASHLER and CHRISTINE R. HARRIS
University of California, San Diego, La Jolla, California

Common sense suggests that unique or abruptly changing (transient) elements in a visual scene often draw attention involuntarily. Visual search studies paint a seemingly different picture, however: Unique items usually draw attention involuntarily only when observers seek a unique target. One type of transient—abrupt onsets—draws attention involuntarily, but only when the observer is seeking an onset (Folk, Remington, & Johnston, 1992). One way of reconciling common-sense with these findings is to suppose that when people view a scene with no specific goal or task, they adopt a *default set,* which might prioritize novelty and/or transients. In two experiments, 336 subjects saw a single display of six items for 900 msec, expecting to have to describe it (Experiment 1) or make an aesthetic judgment about it (Experiment 2). One item in the display was either uniquely flashing (surrounded by static items) or uniquely static (surrounded by flashing items). In both studies, the unique item, even if static, was more often reported than the non-unique item, with flashing items enjoying an additional advantage.

Common sense suggests that unique and/or abruptly changing stimuli sometimes seem to grab our attention "automatically." Not surprisingly, this idea was commonly asserted in early phenomenologically based writings on attention. Titchener (1908), for example, went so far as to claim that any sudden change would distract a person who was trying to concentrate on something else (p. 192), and James (1890/1950) made similar comments. Pillsbury (1908) said "an object in motion in any part of the field of vision will at once attract the attention, and will hold it as long as it continues to move" (p. 48).

In recent years, attention researchers have examined the effects of introducing unique items and various kinds of transients in visual search tasks. The results are not completely congenial to the commonsense views just described. Consider, for example, what happens when a stimulus is unique in a simple physical dimension, such as color (e.g., a red bar among a field of green bars). It is true that such stimuli can usually be detected with a parallel search (Green & Anderson, 1956; Treisman & Gelade, 1980), and for this reason it is often loosely said that they *pop out* (suggesting an involuntary occurrence). Indeed, when people look for an item that is unique in a specified dimension other than color (e.g., a "/" among a field of

"\"s) and one item (not the target) is unique in color, this *singleton* often disrupts search, presumably because it grabs attention (Pashler, 1988). However, when people search for something not itself a singleton, rather different results are found. For example, if the subject searches for prespecified target letter, a uniquely colored nontarget may have no detectable effect (Folk & Annett, 1994). Todd and Kramer (1994) found the same with small display set sizes, although with large display set sizes they found some impairment from the irrelevant singleton. In general, though, it seems that when uniqueness is not relevant to the task at hand, singletons have, at most, a very weak tendency to draw attention involuntarily (see Bacon & Egeth, 1997; Ghirardelli & Egeth, 1998; Yantis & Egeth, 1999).

What about visual transients—that is, motion, change, or the abrupt appearance or disappearance of objects? Many studies using visual search tasks confirm that one particular type of transient—the abrupt onset of a new object—can trigger a shift of visual attention to the object even when the shift interferes with performance of the instructed search task (Yantis, 1994, 2000). Thus, when displays contain both stimuli that "fade in" and those that appear suddenly, the targets that appear suddenly are apparently detected more quickly (even when sudden appearance does not predict that something is a target). Similarly, nonpredictive cues presented prior to a display speed up responses to targets in the same locations as the cues (Yantis, 1994; Yantis & Hillstrom, 1994). Remington, Johnston, and Yantis (1992) found attention shifts even in blocks of trials in which the target never appeared in the position of the onset cue, thus providing a maximum incentive to ignore it.

On the other hand, sudden offsets of a stimulus or changes in the color of a stimulus do not generally trig-

ger involuntary shifts detectable in search paradigms (see Yantis, 2000, for an overview, and Chastain & Cheal, 1999, for a challenge to this conclusion on offsets). Moreover, even the shifts of attention triggered by onsets in particular seem to be contingent on what task set a person has chosen to adopt and, therefore, are not automatic in the strongest sense. Folk, Remington, and Johnston (1992) had subjects make a speeded discrimination on the basis of shape. When the target display consisted of just a single item whose shape the subject was required to judge, with the position of that item varying from trial to trial, a preceding *onset cue* (briefly flashed spots surrounding the position occupied by one of the items) produced involuntary orienting. However, when the target display consisted of four shapes (three green and one red), with the subject responding only to the red one, the onset cue had no such effect. By contrast, a precue consisting of red dots in one location and green dots in the other locations had precisely the opposite effect, triggering involuntary orienting in the color task (but not in the single-item task).

Contingent Involuntary Orienting

To explain this pattern, Folk et al. (1992) proposed what they termed the contingent involuntary orienting (CIO) hypothesis. According to CIO, what has been referred to as involuntary orienting is actually a consequence of observers' adopting a task set designed to optimize performance. This set causes attention to be shifted in response to the cues, as well as in response to the search display itself. If the relevant item appears suddenly in an unpredictable location, as in the single-item displays just described, people adopt a set to orient to onsets, because this assists in performing the task itself. Apparently, this set cannot be in place by the time the display appears without having been in place several hundred milliseconds earlier. Consequently, an abrupt-onset cue will cause an attention shift even when this shift is always disadvantageous. Presumably, the set is essential to performing the primary task or, at least, has benefits on primary task performance that outweigh the costs of shifts induced by the cues.

The CIO account receives support from other results reported by Folk and Remington (1999) and Gibson and Kelsey (1998).[1] The theory makes a further, rather counterintuitive prediction: If people search a crowded display and try to ignore interspersed distractors, performance should not be impaired by having the distractors include rapid onsets or transients. However, CIO makes this prediction only if subjects have no reason to set themselves for onsets, as they would in a conventional visual search where they prepare for the display to appear suddenly. In a recent experiment (Pashler, in press), this problem was circumvented by allowing the subject to preview the search displays before being told what target to search for. Displays consisted of numerous red and green items spatially intermixed. The subject previewed the display, and then the computer played a spoken message that indicated which item should be searched for (among the

red items only). Performance was no worse when the green distractors flashed or were constantly replaced with new distractors, as compared with when these distractors remained static; in fact, having the distractors flash actually facilitated search (presumably, by providing subjects with an additional cue to segregate targets from distractors). The results strongly support CIO and indicate that, when onset-hood is irrelevant (or opposed) to the subject's task set, onsets are impotent to draw attention.

Reconciling Common Sense and Experimental Results

Why does it so often seem that transients—not only abrupt onsets, but also flashing, moving, and jumping stimuli—attract our attention involuntarily, given that the results of the studies described above suggest that such effects are set dependent (and largely restricted to abrupt onsets as against other kinds of transients)? Were Titchener and Pillsbury misled in their comments quoted at the beginning of this article? Are advertisers wasting their money putting flashing lights on signs, for example? Perhaps not. One possible way of reconciling these conflicting perspectives is to suppose that when viewing a scene with no particular goal or task beyond seeing what is there, one may adopt a *default* attentional set. This default set might trigger orienting to onsets in particular or, perhaps, to visual transients in general (a possibility noted by Folk et al., 1992). Alternatively, the default set might incline the visual system to orient to whatever is unusual in the scene. In most visual scenes, onsets are relatively infrequent and, hence, might enjoy a benefit that is intuitively noticeable. On this account, transients in general (and onsets in particular) are not inherently special, except with respect to the statistics of typical visual environments.

EXPERIMENT 1

To examine the nature of a possible default attentional set empirically, one must use an experimental design in which observers view displays without having been instructed to search for anything in particular. To disentangle the role of transients from the role of uniqueness requires contriving scenes that violate the general rule that transients are unusual. In the present experiment, we tried to achieve these goals by exposing a brief CRT display just once to a large number of observers given no information whatsoever about what they were to see (except, unavoidably, that they would see some sort of video display). The subjects were told merely to gaze at a fixation point and to be ready to describe what they had seen. The observers saw only a single trial in order to stop them from developing expectancies about the display or strategies for deploying their attention with these types of displays.

In its relative lack of constraints, the present design is unconventional by present-day standards. However, attention researchers in the early 20th century sometimes examined the spontaneous deployment of attention (e.g., Friedline & Dallenbach, 1929). More recently, in a study by

Lewis (1975), subjects simplv described 15-sec film clips as they viewed them; the film clips contained a variety of objects, stationary and moving, human and nonhuman. The results were somewhat ambiguous with regard to motion, however: Moving objects were more frequently mentioned than were stationary objects, but there was no tendency to report them earlier in the 15-sec period. The methods used in the present research also have some similarities to the *inattentional blindness* studies of Mack and Rock (1998), although the present tasks are less constrained.

Method

Subjects. One hundred sixty-eight University of California, San Diego (UCSD) undergraduates participated in partial fulfillment of a course requirement.

Apparatus and Stimuli. Experiments were controlled by Pentium II computers controlling 15-in. Sony Trinitron Multiscan 100GS SVGA monitors. The display consisted of six different words four letters in length (black capital letters against a white background) arrayed in an ellipse about the center of fixation. In each display, one word was unique with respect to whether it was transient. In the unique-flashing condition, the unique item was flashing, whereas the other five words were static; in the unique-static condition, it was static, and the others were flashing. Flashing words were presented for 100 msec on, 50 msec off, 100 msec on, and so forth. Each word measured 0.5 cm (height) X 1.8 cm (width) (.41° X 1.47° of visual angle, based on a typical viewing distance of 70 cm). The overall extent of the display was 11.5 cm (height) X 11 cm (width).

Design. Each subject saw 1 of 12 different displays. All 12 displays consisted of the same six words presented in the same six positions; what varied was the position of unique item (1-6) and whether its uniqueness consisted in being flashing (among static items) or being static (among flashing items). This generated a total of 6 X 2, or 12 possibilities. Equal numbers of subjects saw each of the 12 displays.

Procedure. All instructions were provided to the subject by the computer. The instructions stated the following: "We are going to show you a brief display and then we want you to describe what you saw by typing on the keyboard. Before the display is presented, we want you to look directly at a plus sign that we will put up in the middle of the screen. Before we show you the display, we need to have you looking right in the center of the screen. We will have you rest your index finger of your left hand on the space bar (please rest it there now). Then we will put up a plus sign, and when you are ready to proceed, just press the space bar with your finger. The display will then appear briefly, and that will be replaced with a text box in which you can type a description of what you saw. Just type everything you remember seeing and when you are done, hit enter." The subjects were then asked if they had any questions, and if they did not, the trial began, and the plus sign was presented. As soon as the subject pressed the space bar, the display was presented for 900 msec. Immediately afterward, a text box appeared in which the subject could type his or her responses, taking as long as he or she wished.

Results and Discussion

Some subjects provided an elaborate narration of the events, whereas most just typed in one or more individual words. For each subject, we determined the proportion of the unique and the nonunique items mentioned in the report. Only exact matches were counted (e.g., CART for CAT was considered an error). Figure 1 shows the mean proportion of items reported for the unique and the non-unique items, depending on whether the item was flashing or static. The results suggest that unique items are far more often reported than nonunique items, with flashing- unique items enjoying a far greater advantage than did static-unique items.

To confirm this impression, an analysis of variance was performed, with uniqueness (unique vs. nonunique) as a within-subjects variable and the feature that was unique in the display (unique-flashing vs. unique-static) as a
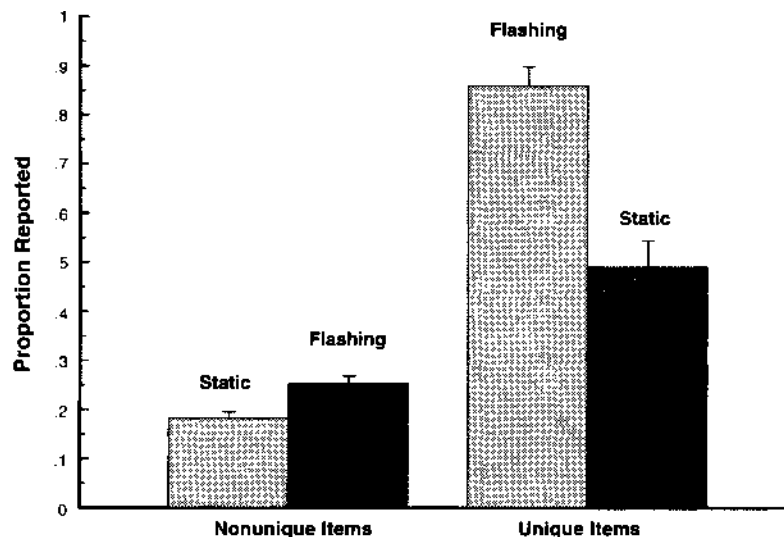

Figure 1. Mean proportion of words recalled in Experiment 1 as a function of whether the item was unique in the display and whether the item was flashing or static (the standard error is shown in the error bar). A given trial contained either flashing-unique items plus static—nonunique items (contributing data represented by light bars) or static-unique items plus flashing-nonunique items (contributing data represented by dark bars).

between-subjects variable. The results showed a significant effect of uniqueness [$F(1,166) = 147.1$, $MS_e = 0.0868$, $p < .001$] and unique feature [flashing vs. static; $F(1,166) = 21.4$, $MS_e = 0.1189$, $p < .001$], as well as a sig-nificant interaction of the two [$F(1,166) = 34.3$, $MS_e = 0.1189$, $p < .001$].

Tukey contrast tests verified that in displays where the unique item was flashing, the unique item was better reported than the nonunique items $(p < .01)$ and that in those in which the unique item was static, again the unique item was better reported than the nonunique item $(p < .01)$. When only unique items are considered, flashing was associated with better report than was being static $(p < .01)$. However, when only nonunique items are considered, there was no significant difference in report of flashing versus static items.

The mean total number of words reported in each condition was very similar (1.76 in displays in which the flashing word was unique, 1.75 in displays in which the static word was unique), indicating a near-perfect tradeoff between performances on the unique and the nonunique items. This suggests that the effects of flashing reflect differences in allocation of attention almost exclusively, rather than sensory effects of flashing per se (this is not unexpected, given the results of Kowler & Sperling, 1983).

EXPERIMENT 2

Although the subjects were not told what to expect in the previous experiment, nonetheless the instructions did indicate that the subjects would need to describe what they saw. Thus, the findings of Experiment 1 might be specific to a rather general set to remember and describe, not reflective of any true default set. To investigate this possibility, the subjects in Experiment 2 were told that they would be making an aesthetic rating of a display about whose contents they knew nothing. Only after the display had been presented were the subjects told to describe what they had seen.

Method

Subjects. One hundred sixty-eight UCSD undergraduates (not subjects in Experiment 1) participated in partial fulfillment of a course requirement.

Apparatus and Stimuli. The display consisted of six line drawings arrayed in a circle about the center of fixation. Six line drawings were selected from a large normed digitized set (Snodgrass & Vanderwart, 1980), choosing some of the most easily identifiable drawings that were, in the norming study, virtually always named with the same single word (book, chair, key, lemon, sock, sun). As in Experiment 1, the composition of the array was held constant across subjects, and one item was unique with respect to whether it was flashing or not. Each picture measured 3.5 cm (height) X 3.5 cm (width; 2.9° X 2.9° visual angle based on a typical viewing distance of 70 cm). The outer extent of the display was 13 cm (height) X 13 cm (width).

Design. The design was identical to that in Experiment 1.

Procedure. The instructions stated that the subjects should make an aesthetic judgment about each array as a whole: "We are going to show you a brief display and we want you to rate how aesthetically pleasing you find it," using a 1-10 scale. As soon as the subject finished reading the instructions and pressed the space bar, the display was presented for 900 msec. Approximately 2.0 sec after the
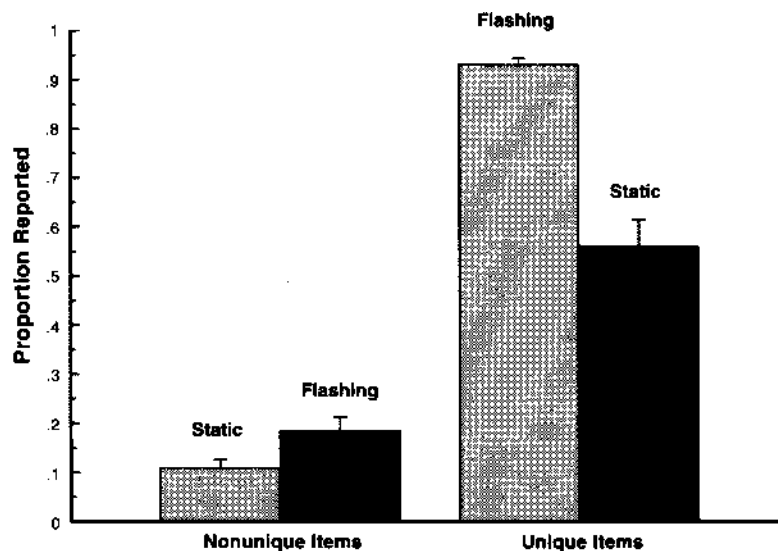

Figure 2. Mean proportion of pictures correctly named in Experiment 2 as a function of whether the item was unique in the display and whether the item was flashing or static (the standard error is shown in the error bar). A given trial contained either flashing-unique items plus static-nonunique items (contributing data represented by light bars) or static-unique items plus flashing-nonunique items (contributing data represented by dark bars).

offset of the display, a message was played through the computer speakers, saying "What objects did you see? Please type them in." A text box then appeared, allowing the subject to type in his or her responses. The subject was then given the opportunity to make the expected aesthetic judgment.

Results and Discussion

Report data were analyzed as in Experiment 1, with any response except the high-frequency response being considered an error. Figure 2 shows the mean proportion of pictures named for the unique and the nonunique items, depending on whether the item was flashing or static. As in Experiment 1, the subjects reported the unique item far more than the nonunique items, with flashing-unique items enjoying a greater advantage than did static-unique items.

To confirm this impression, an analysis of variance was performed, with uniqueness (unique vs. nonunique) as a within-subjects variable and the feature that was unique (unique-flashing vs. unique-static) as a between-subjects variable. The results showed a significant effect of uniqueness [$F(1,166) = 324.3$, $MS_e = 0.0859$, $p < .001$] and unique feature [flashing vs. static, $F(1,166) = 20.9$, $MS_e = 0.0859$, $p < .001$], as well as a significant interaction of the two [$F(1,166) = 44.8$, $MS_e = 0.0929$, $p < .001$].

Tukey contrast tests verified that in displays in which the unique item was flashing, the unique item was better reported than the nonunique items ($p < .01$) and that in displays in which the unique item was static, again the unique item was better reported than the nonunique item ($p < .01$). When only unique items are considered, flashing was associated with better report than was being static ($p < .01$). However, when only nonunique items are considered, there was no significant difference in report of flashing versus static items.

As in Experiment 1, the mean total number of pictures correctly named in each condition was very similar (1.46 in displays in which the flashing picture was unique, 1.47 in displays in which the static picture was unique), again suggesting that the experimental variables induced performance tradeoffs, consistent with an effect on attentional priority.

GENERAL DISCUSSION

In two rather unconventional attention tasks, observers were shown a single 900-msec exposure of six items (words in Experiment 1, pictures in Experiment 2). In Experiment 1, they expected to have to describe what they had seen, whereas in Experiment 2, they expected to make an aesthetic judgment, and the request to describe the display came as a surprise after the display presentation. In both experiments, there was a very strong tendency to attend to and remember the single unique item, whether it was unique in being a lone flashing item among static items or unique in being a lone static item among flashing items. However, flashing unique objects were more often reported than nonflashing unique objects. These differences almost certainly reflect attentional rather than sensory factors.

The results suggest that observers adopt a *default* set in preparation for seeing and interpreting an unknown scene. This set evidently generates a strong tendency to shift attention to unique elements, with some extra weight given to flashing items, rather than to nonflashing items. This conclusion supports a simple reconciliation between the commonsense view noted at the outset of this paper, according to which transients and unique items attract attention involuntarily, and the results of visual search studies described above, which show that certain types of transient stimuli (abrupt onsets) grab attention but that even that occurs only when a specific task set has been adopted. (For example, as was mentioned in the introduction, having distractor characters flash on and off, much like the objects in the present experiment, weakened, rather than amplified, their distracting power when subjects performed a color-search task; Pashler, in press.) The commonsense view, it would seem, correctly characterizes the attention shifts that occur when the *default set* is in place (as Folk et al., 1992, conjectured), but traditional writers (such as Pillsbury, 1908, quoted in the first paragraph of this article) incorrectly assumed that these tendencies are fixed. Attention shifts to transients and to unique items should not be described as *involuntary,* it seems, but rather as *contingently involuntary* (i.e., the tendency can be voluntarily "turned off" or suppressed when it is not necessary for performing a given task).

The results of these experiments raise several new and interesting questions. First, does it really make sense to refer, as we have here, to *"the* default set," suggesting that there is only one such state? Or do people perhaps have a variety of different default sets for processing scenes of different types? The resemblance between the results of the two experiments presented here, involving rather different task demands, supports the very tentative conjecture that there may indeed be a default set common to a fairly wide range of different mental states.

Second, does the default set exist prior to the onset of the display, or is it triggered by the recognition that the display contains some—at least one—flashing items and/or the fact that some items are unique in this respect? The present results do not answer this question, but one might investigate the issue by using displays briefer than 900 msec (probably sufficient time for significant changes in perceptual strategy to take place).

Third, is the default set something that the visual system would relax into in the absence of a contravening set, or is it actively maintained, as are the task sets arising in conventional visual search experiments? To answer this question, one might determine whether flashing distractors disrupt search in a color-selection task if another task with central demands is performed concurrently?[2]

Fourth, can the default set coexist with the sorts of voluntary set evoked in ordinary visual search tasks? So, for example, when a person has a voluntary task set to search some part of a scene (as in the standard visual search task), is processing of inputs appearing in other parts of the visual field (or inputs in different modalities)[3] governed by

the default set, or is it instead governed by the same set (or perhaps an attenuated version of it)?

These are just a few of the kinds of questions that one might ask about the spontaneous deployment of attention. Contemporary research on attention has focused heavily on constrained tasks (see Mack & Rock, 1998, for a notable exception). As a result, we know much about tasks like visual search, but far less about the mechanisms and processes that govern the spontaneous deployment of attention outside of such tasks.

REFERENCES

BACON, W. F., & EGETH, H. E. (1997). Goal-directed guidance of attention: Evidence from conjunctive visual search. *Journal of Experimental Psychology: Human Perception & Performance, 23*, 948-961.

CHASTAIN, G., & CHEAL, M. (1999). Time course of attention effects with abrupt-onset and offset single- and multiple-element precues. *American Journal of Psychology, 112*, 411-436.

FOLK, C. L., & ANNETT, S. (1994). Do locally defined feature discontinuities capture attention? *Perception & Psychophysics, 56*, 277-287.

FOLK, C. L., & REMINGTON, R. (1999). Can new objects override atten-tional control settings? *Perception & Psychophysics, 61*, 727-739.

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.

FRIEDLINE, C. L., & DALLENBACH, K. M. (1929). Distance from point of fixation vs. intensity as a determinant of attention. *American Journal of Psychology, 41*, 464-468.

GHIRADELLI, T. G., & EGETH, H. (1998). Goal-directed and stimulus-drivien attention in cross-dimensional texture segregration. *Perception & Psychophysics, 60*, 826-838.

GIBSON, B. S., & KELSEY, E. M. (1998). Stimulus-driven attentional capture is contingent on attentional set for displaywide visual features. *Journal of Experimental Psychology: Human Perception & Performance, 24*, 699-706.

GREEN, B., & ANDERSON, L. K. (1956). Color coding in a visual search task. *Journal of Experimental Psychology, 51*, 19-24.

JAMES, W. (1950). *The principles of psychology* (Vol. 1). New York: Dover. (Original work published 1890)

JONIDES, J. (1980). 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.

KOWLER, E., & SPERLING, G. (1983). Abrupt onsets do not aid visual search. *Perception & Psychophysics, 34*, 307-313.

LEWIS, M. S. (1975). Determinants of visual attention in real-world scenes. *Perceptual & Motor Skills, 41*, 411-416.

MACK, A., & ROCK, I. (1998). *Inattentional blindness.* Cambridge, MA: MIT Press.

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

PASHLER, H. (in press). Involuntary orienting to flashing distractors in delayed search? In C. L. Folk & B. Gibson (Eds.), *Attraction, distraction, and action: Multiple perspectives on attentional capture.*

PILLSBURY, W. B. (1908). *Attention.* New York: SWAN Sonnenschein/ Macmillan.

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

SNODGRASS, J. G., & VANDERWART, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory, 6*, 174-215.

TITCHENER, E. B. (1908). *Lectures on the elementary psychology of feeling and attention.* New York: MacMillan.

TODD, S., & KRAMER, A. F. (1994). Attentional misguidance in visual search. *Perception & Psychophysics, 56*, 198-210.

TREISMAN, A., & GELADE, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12*, 97-136.

YANTIS, S. (1994). Stimulus-driven attentional capture. *Current Directions in Psychological Science, 2,* 156-161.

YANTIS, S. (2000). Goal-directed and stimulus-driven determinants of attentional control. In S. Monsell & J. Driver (Eds.), *Attention & performance XVlll* (pp. 73-103). Cambridge, MA: MIT Press.

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

NOTES

1. It should be noted, however, that when subjects search for a target specified by identity (e.g., a target letter), onsets appear to capture attention, whereas irrelevant color singletons do not.

2. Slightly reminiscent of this possibility is the finding by Jonides (1980) that a memory load reduced the effects of a central cue, but not a peripheral cue, in a visual search task.

3. It seems certain that subjects performing a visual search task would spontaneously notice a loud sound or a hand on their shoulder, even if their performance was immune to the effects of onset distractors in the visual displays (pointed out by W. A. Johnston, personal communication).