

Do images involuntarily trigger search? A test of Pillsbury's hypothesis

HAROLD PASHLER *University of
California, San Diego, LaJolla, California*

and

LING-PO SHIU *Chinese University
of Hong Kong, Hong Kong, China*

Pillsbury (1908) suggested that deciding to search for something in a scene consists of nothing more than forming a visual image of the target. If so, imaging should trigger search even when it would be more advantageous *not* to search. Subjects were cued to form an image of a specified object (e.g., *tiger*) and to press a key when they had done so. This initiated the presentation of a sequence of pictures, with a single target digit interspersed; the subject's task was to report this digit. The sequence contained a picture of the same type of object that the subject had just imaged (e.g., a tiger), either before or after the target digit. If this picture was detected involuntarily, an *attentional blink* should have impaired digit detection when the picture preceded the digit. This was confirmed in two experiments, even when instructions specifically encouraged subjects to discard the image and to avoid searching for it. The results support Pillsbury's hypothesis.

What is the relationship between central mental processes (thinking, planning, decision making) and the control of visual attention? One approach to this issue involves analyzing people's ability to perform visual search while carrying out other, concurrent tasks. Research of this kind suggests that people are able to search a visual display while concurrently performing central mental operations, such as selecting and producing responses to auditory stimuli (Pashler, 1989; Rollins & Hendricks, 1980). On the other hand, when a concurrent task is performed continuously over many seconds, visual search performance deteriorates (Pashler & Johnston, 1994). Furthermore, merely holding onto concurrent memory loads impairs search performance to some fairly modest degree, as compared with a single-task control (Logan, 1978). One way to understand these observations is to suppose that, although the central mental machinery required for decision making and response planning need not actually carry out the search, it participates in maintaining the state of preparation necessary to allow the search to take place (hence, the deterioration in performance with a protracted concurrent task).

The question considered in the present article is *how* central mechanisms prepare or control the process of search and what this process of preparation actually con-

sists in. Although these issues have scarcely been considered in contemporary literature, psychologists of the late 19th and early 20th centuries offered several intuitively plausible suggestions. Several writers noted that attention seems to be drawn to inputs that are related in some way to an ongoing train of thought. Thus, W. B. Pillsbury (1908) proposed that the act of preparing to search might consist, not in some mental operation specific to search per se, but merely in forming a mental image of the target to be searched for. He wrote:

searching for anything consists ordinarily of nothing more than walking about the place where the object is supposed to be, with the idea of the object kept prominently in mind, and thereby standing ready to facilitate the entrance of the perception when it offers itself. It is for that reason too that it is much easier to find an object again after finding it once, because you can look with a more definite image and can apply it more accurately.

Another theorist of the same period, Oswald Kulpe (e.g., 1909), suggested a broader principle along similar lines. According to Kiilpe, "impressions which repeat or resemble ideas already present in consciousness are especially liable to attract the attention" (p. 439). The mental set to search for something might, on this account, simply be the state of having "in consciousness" representations that are conceptually related to what is being searched for. It is obvious that such a wired-in connection between central mental activity and the control of visual attention might have functional advantages (see Pashler, 1997, for a discussion).

The suggestions of Pillsbury (1908) and Kulpe (1909) were based on introspection, of course, and introspective

The authors thank Marvin Chun, John Duncan, Ron Kinchla, Allen Osman, and an anonymous reviewer for useful discussion, and Matt Remuzzi for assistance with data collection. Correspondence concerning this article should be addressed to H. Pashler, Department of Psychology 0109, University of California, San Diego, La Jolla, CA 92094 (e-mail: hpashler@ucsd.edu).

access to mental operations may not be trustworthy. Furthermore, even if people do customarily form an image of whatever target they hope to find, as Pillsbury contended, imaging might be an initial step in retrieving specific information about the target, rather than constituting the state of preparation *per se*. This preparation might involve additional and quite different mental operations.¹ Several contemporary findings seem broadly consistent with Pillsbury's conjecture and also with the alternative interpretation just described. In searching a single briefly presented picture, subjects are more accurate when they see a picture of the target beforehand, rather than just its name (Pachella, 1975). Reaction times to search an array of drawings are faster when the observer has seen a picture of the target, rather than just a name (Paivio & Begg, 1974). These results are in line with Pillsbury's suggestion, but they do not argue very strongly that forming an image is either a necessary or a sufficient condition for carrying out a search.² As described in the General Discussion section, below, a recent article by Dark, Vochatzer, and VanVoorhis (1996) reports additional findings consistent with Kulpe's suggestions.

The goal of the two experiments reported here was to determine whether forming an image is sufficient to trigger search, by asking whether image formation makes it obligatory to search for an object like the one imaged. To demonstrate that mental operation B occurs as an obligatory consequence of operation A, one needs a situation in which operation A occurs but carrying out B would not be to the subject's advantage—that is, where it would harm rather than help performance. How could visual search harm performance? In the experiments described below, a situation was contrived in which detecting a particular target could only interfere with another task the subject is attempting to perform.

The subjects were given a word or a phrase, such as *fish* or *swimming pool*, and were instructed to form a complete, clear mental image of it. Having done so, they were immediately called on to perform a different task that was, as far as they could tell, completely unrelated to the image: viewing a rapid serial visual presentation (RSVP) of line drawings, with the goal of detecting and identifying a target digit located midway through the sequence. A picture of the type of object they had just imaged (the *critical picture*) was present somewhere in the sequence. If this picture was involuntarily detected, this detection should interfere with detection of the digit target through the *attentional blink* effect first described by Broadbent and Broadbent (1987). This term describes the tendency for detection of an initial target (T1) in an RSVP display to reduce the ability to detect or report a subsequent target (T2) arriving close on the heels of T1. If the subjects involuntarily detected any drawing of the type they had just imaged, detection of digits presented shortly after the critical picture should be depressed by the blink effect.

The digit was always presented in position five of a sequence, with four pictures preceding and four pictures

following the digit. The critical picture appeared in either position three or position seven; on the basis of previous research with the attentional blink (e.g., Shapiro, Raymond, & Arnell, 1994), position three would be expected to produce as large an attentional blink as one can observe, should the effect occur. Notice that the subject's task and the position of the digit are equated across the two conditions.

EXPERIMENT 1

Method

Subjects. Twelve undergraduates from the University of California, San Diego, volunteered to participate and provided informed consent. All had normal or corrected-to-normal vision.

Stimuli. The pictorial stimuli consisted of line drawings of common objects taken from a pool of 142 images. The drawings depicted relatively familiar objects (e.g., elephant, sun, flower, tie, iron, airplane), and measured approximately 7.5 X 7.5 cm; digits measured 2.5 X 2.5 cm. Pictures and digits were drawn with black lines against a white background. On each trial, nine objects were presented in a sequence: four line drawings followed by a digit followed by four additional line drawings; thus, the digit was always in position five.

Design. The practice phase consisted of two blocks of 16 trials each, during which the subject merely identified the target digit. In addition to providing practice, this made it possible to estimate an appropriate frame rate with which to begin the subsequent test phase. This duration was selected with the goal of allowing the subject to identify the digit at approximately 75% accuracy.

The test phase consisted of eight blocks of 16 trials. Again, each trial involved presentation of a sequence of eight pictures with a single digit embedded in the sequence. In this phase, however, an *image cue* (a word or a phrase) was presented, telling the subject to form an image of a particular type of object (e.g., *elephant*), and a picture of a corresponding object (an elephant) appeared in either the third or the seventh position of the sequence (the digit and its mask were counted as both occupying position five).

In each test block, there were eight trials of each of two conditions: critical picture in position three versus critical picture in position seven. The order of conditions was independently randomized.

Procedure. During the practice phase, the task was to search for a digit target and report its identity, without any speed pressure. The subject initiated the display sequence by pushing the space bar. Once the display sequence was finished, the subject responded by pressing a key on the numeric keypad, to indicate his or her best guess about what number had been presented.

The procedure was the same during the test phase, with one exception. Before the display was presented, an image cue was displayed, telling the subject to form a mental image. The subjects were told to take as long as they needed to make as clear an image as possible. Once they had done this, they were to push the space bar, initiating the display sequence. At this point, the instructions made it clear that they should focus exclusively on searching for and accurately reporting the digit. The subjects were told that they could close their eyes while forming the image, if that helped in forming a clear image, but that, naturally, their eyes should be open when they began the visual sequence. At this point, they were led to understand, there was no need to maintain the image. The subjects were not told that the image might correspond to any of the pictures in the sequence. Feedback was provided after each trial, consisting of the correct digit response.

At the outset of practice, the digits were initially presented at a rate of 120 msec per item, and during practice, this was adjusted in-

dividually on the basis of each subject's responses, as follows. When the subjects were correct three times in a row, the digit presentation time was reduced by 15 msec, and when they were incorrect, the digit presentation time was increased by 15 msec. During the experiment itself, the exposure duration was set on the basis of the practice-phase adjustment, and this was only altered when detection accuracy fell outside the range of between 75% and 87%.

Results and Discussion

The average exposure duration was 55 msec, and accuracy in reporting digits in any condition correctly ranged between 69% and 89%. Digit report accuracy averaged 77.7% when the critical picture was in position three and 85.6% when the critical picture was in position seven. The difference was in the direction that would be expected if the picture was detected on at least some proportion of the trials and if this caused an attentional blink for the detection of the digit. The difference held for 11 of 12 subjects and was significant by test [$t(11) = 4.86, p < .05$].

The finding is consistent with the view that a person's having formed a mental image involuntarily triggers detection of the type of object just imaged, even when the person attempts to search for something else and lacks any incentive to maintain the image, much less to search for a new exemplar of the same type.

Is it nonetheless possible that some *demand characteristic* of the experiment inadvertently encouraged the subjects to search for what they had just imaged? Even though the data suggest that searching hindered rather than helped performance, the subjects may still have noticed exemplars of the type they imaged and concluded that it was advantageous for them to search for these exemplars.

EXPERIMENT 2

The purpose of the second experiment was to deliberately counteract any such demand characteristics that might have led to voluntary, as against involuntary, searching.

Method

Subjects. Nineteen subjects from the same population as that in Experiment 1 volunteered and gave informed consent.

Stimuli, Procedure, and Design. This experiment was the same as Experiment 1, except for the following changes. First, the instructions stated that it would require several seconds, at least, to form a good, complete mental image. This was done to discourage the subjects from simply identifying the word and leaving it at that. Second, the instructions emphasized that the image should not, under any circumstances, be maintained during the subsequent search for the digit. To encourage compliance, the instructions stated that forming good mental images would *prime* the cognitive search processes for a search, but they also stated that, although this priming effect would be helpful in digit detection, it would not work unless the subject *got rid of* the image before the display appeared and avoided searching for anything related to the *image word*.

Results and Discussion

The average exposure duration was 65 msec. When the critical picture was in position three, the subjects re-

ported the digit 81.8% of the time; when it was in position seven, they reported it 85.1% of the time. The difference was again significant [$t(18) = 2.3, p < .05$]. Fourteen subjects showed superior performance in the position seven condition. The results replicate those of Experiment 1, despite the additional precautions taken to discourage voluntary searching. The size of the apparent blink effect in this experiment was significantly smaller than in the previous experiment. [$F(1,29) = 4.34, p < .05$], perhaps because fewer subjects maintained the image while the pictures were presented.

GENERAL DISCUSSION

Two experiments required subjects to form an image of a particular type of object and then to search for a digit in a stream of pictures. When the critical picture (i.e., a picture of the same type of object as that which the subject had just imaged) preceded the digit, detection of the digit was impaired, relative to performance when the picture was presented after the digit. Evidently, the critical picture was detected on many or all trials, inducing an attentional blink effect that reduced digit detection accuracy.

Was the picture searched for and detected *involuntarily*? Plainly, both searching for the picture and detecting it could only harm the subjects' performance in their instructed task—namely, digit detection. This is true not only because of the attentional blink, but also because searching for the digit target would be hindered by trying to search for any other target at the same time, whether or not the other target was detected (Horlitz, Johnston, & Remington, 1992). In Experiment 2, there was an attempt to introduce additional demand characteristics directly opposing any tendency to search, and the results were basically unchanged. It would seem, then, that search occurs even when the subject has an incentive not to try to find the critical picture; therefore, the search is evidently involuntary. Of course, one could still insist that subjects are immune to these incentives and voluntarily search out of sheer perversity, but such a contention could be raised regarding any behavior normally viewed as involuntary, including the knee jerk reflex.

The size of the effect observed here appears to be quite modest, by comparison with what is typically found in experiments in which subjects voluntarily search for a blink-inducing target (sometimes 15%-30%, depending on the task and exposure conditions). Does this mean that Pillsbury's (1908) hypothesis is only partly correct and that involuntary search for the imaged object is an occasional or weak process, as compared with voluntary search? Perhaps. Alternatively, however, the effect might be smaller simply because the subjects had no incentive to maintain the image while they searched for the digit and, therefore, allowed it to weaken. Maintaining an image might produce a full-strength blink effect even in the absence of a voluntary intention to search.³

One question for future research is whether the involuntary attention capture suggested by these findings causes storage in short-term memory of visual details of the task-irrelevant target, as one would expect if the image instigates a full-scale search that runs to its natural completion. Another question for future research is whether involuntary detection of targets and the consequent attentional blink effect is elicited only by visual images or whether it can be triggered by any activation of nonvisual representations. Along the lines of Kulpe's (1909) suggestions, the correct generalization may be that attention is drawn to anything even associatively related to the contents of consciousness (cf. Pashler, 1997). Dark et al. (1996) described a *semantic priming effect* that seems consistent with that idea. Subjects saw a single prime word, which they were asked to read but not respond to; this was followed by two 100-msec masked target words that the subject attempted to report. When the prime was related to one of the target words, that word was reported more accurately, and the other, unrelated member of the pair was reported slightly less accurately. This interesting result might reflect attention shifts to the semantic associate. Alternatively, semantic priming might enhance the identifiability of the target word, thereby depressing report of the other word, possibly by lessening output competition.

The general topic explored in the present report and in that of Dark et al. (1996)—namely, conceptually based attentional control occurring as a by-product of other cognitive events—has been neglected in recent research on attention, although it greatly interested early attention theorists. The topic seems to deserve further investigation, especially because, in ordinary life, attention is normally subject to indirect control, rather than to explicit direction. Indeed, viewing a scene with the specific goal of finding a certain target, as in the great majority of attention experiments, is probably the exception, rather than the rule, outside of the laboratory.

REFERENCES

- BROADBENT, D. E., & BROADBENT, M. H. P. (1987). From detection to identification: Response to multiple targets in rapid serial visual presentation. *Perception & Psychophysics*, 42, 105-113.
- DARK, V. J., VOCHATZER, K. G., & VANVOORHIS, B. A. (1996). Semantic and spatial components of selective attention. *Journal of Experimental Psychology: Human Perception & Performance*, 22, 63-81.
- FODOR, J. A. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*. Cambridge, MA: MIT Press.
- HORLITZ, K. L., JOHNSTON, J. C., & REMINGTON, R. W. (1992, November). *What makes it difficult to process multiple targets in rapid serial visual presentation?* Paper presented at the 33rd Annual Meeting of the Psychonomic Society, St. Louis.
- KULPE, O. (1909). *Outlines of psychology* (E. Titchener, Trans.). London: Swan Sonnenschein.
- LOGAN, G. D. (1978). Attention in character classification tasks: Evidence for the automaticity of component stages. *Journal of Experimental Psychology: General*, 107, 32-63.
- PACHELLA, R. G. (1975). The effect of set on the tachistoscopic recognition of pictures. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance K* (pp. 136-156). New York: Academic Press.
- PAIVIO, A., & BEGG, I. (1974). Pictures and words in visual search. *Memory & Cognition*, 2, 515-521.
- PASHLER, H. (1989). Dissociations and dependencies between speed and accuracy: Evidence for a two-component theory of divided attention in simple tasks. *Cognitive Psychology*, 21, 469-514.
- PASHLER, H. (1997). *The psychology of attention*. Cambridge, MA: MIT Press.
- PASHLER, H., & JOHNSTON, J. C. (1994). *Continuous task performance and dual-task interference: Chronometric studies*. Manuscript submitted for publication.
- PILLSBURY, W. B. (1908). *Attention*. New York: Macmillan.
- ROLLINS, H. A., & HENDRICKS, R. (1980). Processing of words presented simultaneously to eye and ear. *Journal of Experimental Psychology: Human Perception & Performance*, 6, 99-109.
- SHAPIRO, K., RAYMOND, J., & ARNELL, J. (1994). Attention to visual pattern information produces the attentional blink in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception & Performance*, 20, 357-371.

NOTES

1. This point is reminiscent in some ways of contemporary criticism of British Empiricist philosophers' tendency to equate imagery with thought. As Fodor (1987) and others have pointed out, introspections may correctly indicate that imagery often accompanies thinking, but grave difficulties arise when one tries to suppose that the images actually *are* the thoughts. These difficulties would not apply to Pillsbury's hypothesis, however.
2. Pillsbury's (1908) remarks suggested to the present authors that he viewed the image as sufficient, but perhaps not necessary. One reviewer of an earlier version of this article did not share this reading, however, and suggested that the hypothesis tested here should be associated with Kulpe (1909) exclusively.
3. In these experiments, instructions to maintain the image were avoided, of course, because they would seem to introduce demand characteristics favoring search for the imaged object.

(Manuscript received February 19, 1998;
revision accepted for publication July 10, 1998.)