No pause for a brief disruption: Failures of visual awareness during ongoing events.

Daniel T. Levin

and

Donald A. Varakin

Vanderbilt University

Address correspondence to the first author at:
Department of Psychology and Human Development
Vanderbilt University
Peabody College # 512
230 Appleton Place
Nashville, TN. 37203-5701
615-322-1518
daniel.t.levin@vanderbilt.edu

Abstract

Past research has repeatedly documented the close relationship between visual attention and awareness. Most recently, research exploring change blindness, inattentional blindness, repetition blindness, and the attentional blink has converged on the conclusion that attention to one aspect of a scene or event may lead to a highly circumscribed awareness of only the specific information attended, while other information, even that which is spatially or temporally nearby can go completely unnoticed. In the present report, we extend these observations to the dynamic allocation of attention during a well-structured meaningful event. In two experiments, subjects viewed brief videos of simple events and were told to pay close attention to them. During the events, an unexpected disruption consisting of a brief low spatial frequency motion field occurred. Despite intensive questioning and opportunities for recognition, the majority of subjects reported no awareness of 200, 400, or 600 ms disruptions. In a second experiment, blank-screen disruptions were added, and these resulted in no increase in detection. We conclude that visual attention may result in far more transitory awareness of visual information than previously appreciated.

When researchers describe visual attention, they find metaphor almost irresistible. Visual attention has been described as a spotlight, a zoom lens, a filter, and even a hand (see for example, Erikson & St James, 1986; Deutsch & Deutsch, 1963; Rensink, 2000; Cave & Bichot, 1999). One interesting property of almost all of these metaphors is that they imply that visual attention, selective though it may be, is either in continuous contact with the visual world, or at least maintains contact over some period of time. So, whether we are attending to a few objects, or only one aspect of one object, the presumption is that we are generally attending to *something* unless our eyes are closed. A related assumption can also be seen in philosophical discussions of vision which imply that even outside the focus of attention, we are globally aware of the continued existence of visual information, even if that awareness is of only “visual stuff” or “more of the same” (Dennett, 1991; O’Regan, 1992). However, one problem with both of these conclusions is that they might be based on the common constraint imposed by almost all introspections and lab tasks used to study visual attention and awareness. In lab tasks subjects are usually required to continuously monitor visual displays for specific visual information. Whether subjects are searching displays for a specific target, trying to detect a visual change, or trying to compare one thing with another, they are usually required to focus attention on some concrete aspect of the visual world. Introspections are similarly constrained. As a number of authors have pointed out, the act of attending to visual information (or even thinking about it more abstractly) brings it into awareness, and it is therefore very difficult to conceptualize what, exactly, goes on outside of awareness (see Blackmore, 2002).

Although the traditional methods are good ways of studying the process of seeing, they may not reflect more dynamic real-world contexts. In many situations, it is probably not necessary to continually focus attention on, and be aware of visual information. For example, a variety of findings suggest that people tend to parse ongoing visual action into discrete events that may serve to structure visual attention temporally (for review see Zacks & Tversky, 2001). Accordingly, visual awareness may only be necessary during moments that have the potential to specify the meaning of events, and be less critical while those events unfold in expected patterns. For example, Newtson and Engquist (1976) asked subjects to search for deletions of up to 500 ms in films of simple events and observed better detection when the deletions occurred during between-event transitions than when the occurred within events. More recently, Baldwin, Pederson, Craven, Andersson, and Bjork (in prep; see also Saylor & Baldwin, in press) describe experiments in which subjects viewed brief videos, and searched for single-frame (33 ms) disruptions in which the picture became sepia-toned, or was blanked. Most subjects detected the disruptions, but were slower to do so when they occurred during an event.

Research documenting the attentional blink converges to suggest that visual awareness can be transitory by demonstrating that subjects miss targets that follow another target. Most important is the hypothesis that the attentional blink occurs because people have difficulty becoming aware of one stimulus while they are doing the elaboration necessary for awareness of another (Chun & Potter, 1995). If this is true, then the efficient allocation of visual attention might require weighing the benefits of awareness against the cost of elaborating on a continuous stream of aspects that may not be particularly important. Therefore, there may be a strong ecological demand to default to a limited awareness during only the specific moments when visual information is most valuable. This would allow effective information pick-up while reducing the possibility that important information cannot be incorporated into our plans because it occurred while we are processing some irrelevant aspect of our environment.

To explore the dynamics of visual attention during natural events, we asked subjects to view short videos with the intention of answering unspecified questions about them. During the videos, 200-600 ms disruptions unexpectedly occurred. These disruptions were either low spatial frequency motion fields consistent with the direction of on-screen object motion, or (in Experiment 2) blank screens, and we sought to determine whether subjects would detect them. In doing so, we follow the Newtson & Engquist (1976) and Saylor & Baldwin (in press) work described above. However, in those experiments subjects intentionally searched for disruptions, and missed potentially subtle discontinuities in the action, or very brief periods during which the tonal qualities of the image changed. In contrast, the disruptions tested here were much longer, and constituted periods during which the action was off-screen entirely. Failures to detect these disruptions would be consistent with the hypothesis that although even sophisticated visual processing may be ongoing, it may not be characterized by continuous attention to and/or awareness of any particular aspect of an attended stream of visual information.

Experiment 1

Method

Participants. Seventy-six Kent State University undergraduates participated for course credit. All participants reported normal or corrected-to-normal vision.

Equipment. Videos were shot with a Panasonic S-VHS camera (model #: AG456UP) and edited on a Power Macintosh 9600/200 computer using Media100-le hardware and software. The videos were transferred to a standard VHS cassette tape using a Panasonic S-VHS VCR (model #: AG1980P). Videos were presented on a 19" television screen using a commercial VHS VCR.

Stimuli. One of two short events was depicted in each video, which we refer to as the *Accident event* and the *Newspaper event*. The Accident event opened with a shot of a man carrying a stack of boxes up some stairs, and cut to a different man walking down a hallway while reading a newspaper. The third and fourth shots depicted the two men colliding as the man with the boxes rounded a corner at the top of the stairs. The disruption occurred as the boxes fell. The first four shots of the Newspaper event depicted a man entering a room, picking up a pad of paper and then sitting down. The fifth shot was a close-up of the man sitting down, opening the paper backwards and flipping the page, writing on the pad of paper, and flipping the page a second time. The disruption was overlaid on the second page flip. The Accident and Newspaper events were 27 and 65 seconds in length respectively.

The 200ms disruption was 6 frames of a low spatial frequency dark blur moving horizontally across the screen. The 400ms and 600ms disruptions consisted of the same six frames being overlaid two and three times consecutively (see Figure 1). The midpoint was used as a point of reference for lengthening the disruption. Thus, the onset of the 400ms disruption was 3 frames earlier than the onset of the 200ms disruption and so forth. The disruptions were overlaid on the video and therefore did not disturb its temporal continuity. For each event, 200ms, 400ms and 600ms disruptions were overlaid that were consistent or inconsistent with the motion depicted in the critical event, yielding a total of 12 experimental videos.

Insert Figure 1 about here

Procedure. Groups of one to four participants viewed one of the twelve videos subsequent to participating in an unrelated experiment. Viewing distance ranged from about 3-6 feet. Participants were told to pay close attention to the video because they would be asked a few questions about it, but were not told specifically about the disruptions. Once the video clip ended, the tape was paused and an experimenter asked three questions to determine whether subjects recalled the disruption: 1) if they noticed anything unusual, 2) if they noticed any unusual disruptions of the video, and 3) if they noticed and dropouts, skipped frames or other momentary disruptions of the camera's view, such as a person walking quickly in front of the camera. Next, the participants were shown the disruption in isolation and asked if they had seen it in the video. If the participants answered "yes" to any question, they were asked to indicate what they saw and when it occurred. All responses were recorded on answer sheets provided by the experimenter.

Results

In our initial analysis, two raters coded subjects as having seen the disruption if they answered "y" to any of the open ended detection questions or the recognition question, and followed that with any possibly correct explanation of what the disruption was. The raters agreed on 88% of the open-ended responses, and 96% of the recognition responses. Disagreements were resolved through discussion. Explanations were coded because some subjects indicated they had seen a disruption, but clearly described another event. For example, some subjects gave an affirmative response, but described a moment when the initial actor in the accident video briefly disappeared behind a staircase. To avoid an overly restrictive criterion for detection, it was emphasized to raters that they should be liberal in accepting an open-ended specification, and that they should only reject it if they could clearly recognize that the description referred to something other than the actual disruption. Based on this measure, of the 76 subjects who viewed the video, only 27.6% detected the disruption (based on open ended responses), and there was no effect of disruption duration (detection rates: 200ms = 21.4%, 400ms = 41.7%, 600ms = 20.8%, X^2 (1, N = 76) = 3.458, p = .177) or direction (detection rates: consistent = 28.9%, inconsistent = 26.3%, X^2 (1, N = 76) = .798, p = .798). Results from the recognition test were similar. Only 15.8% of subjects recognized the disruption, and again there was no effect of duration (detection rates: 200ms = 3.6%, 400ms = 20.8%, 600ms = 25%, X^2 (1, N = 76) = 5.134, p = .077) or direction (detection rates: consistent = 15.8%, inconsistent = 15.8%; see Figure 2).

Insert Figure 2 about here

Discussion

The majority of subjects in Experiment 1 failed to recall or recognize 200-, 400-, and 600 ms disruptions to an ongoing event. It is important to note that although the effect of disruption duration is nonsignificant, we do not conclude that duration does not impact detection, only that a substantial number of subjects miss even 600 ms disruptions.

Although this failure might be striking, there are several potential problems with this experiment. First, the motion-field disruptions were in some ways similar to what one might expect to see during the events themselves. For example, if the falling boxes were shown in a closeup, one might expect to see a brief "smearing" motion similar to the stimulus we used. Therefore, in Experiment 2 subjects experienced one of two kinds of disruption: either a motion-field as in Experiment 1, or a simple blank screen. In the blank-screen condition, the disruption was simply a series of blank grey frames that were adjusted to match mean luminance of the last pre-disruption frame and the first post-disruption frame. Second, the recognition test we used in Experiment 1 did not reinstate the context surrounding the disruption. This may have caused to disruption to appear different from how it looked in the original film, perhaps due to a lack of forward and backward masking from the surrounding visual events. Subjects might then have responded 'no' to the recognition test because they did see the disruption, but did not believe that the recognition stimulus was very similar to what they saw. In Experiment 2, we fixed this problem by including the surrounding event in the recognition test. Finally, it is possible that subjects did not see the disruption because they failed to attend to any of the information in that part of the video. Although it seems unlikely that 80% of our subjects had simply "spaced out" after 30-60 seconds of video (see Schooler, Ryan, & Halpern, in press), we tested this hypothesis by asking them a series of multiple choice questions about the events immediately surrounding the disruption. If a global breakdown in attention to the stimulus is responsible for low detection rates, then subjects who miss the disruption should be unable to answer these questions, while subjects who do see it should be much more successful.

Experiment 2

Method

Participants. Eighty-two Kent State University undergraduates participated for course credit. Data from five subjects were discarded because they had previously participated in this or a related experiment, leaving 77.

Equipment. The same equipment was used as in Experiment 1 with the following exceptions. Instead of standard VHS cassette tapes, videos were recorded from the computer onto SVHS tapes and played from a Panasonic SVHS VCR (model #: AG1980P).

Stimuli. Slightly shorter versions same events from Experiment 1 were used for Experiment 2. For the Newspaper event, the disruption was overlaid during the first page flip, and the second page flip was omitted, shortening the video to approximately 60 seconds. Also, the Accident event was shortened by about one second at the end. The motion disruption was the same as in Experiment 1, but only 200ms and 600ms disruptions that were consistent with the action motion were used. In addition, 200ms and 600ms blank disruptions were also overlaid the critical action. The blank disruptions were homogeneous and grayish in color, and approximated the luminance of the frames just before and after their onset. There were eight experimental videos.

Two additional videos were also used for the control experiment. The control videos were identical to the originals, except they ended prior to the critical event surrounding the disruption. The Newspaper event ended as the man sat at the table (about 7 seconds before the disruption), and was about 34 seconds long. Therefore, subjects did not see him set the paper down, turn the pages, and write in his notepad. The Accident event ended after the shot of the man carrying the paper (about 5 seconds before the disruption) and was about 18 seconds. In this case, no shots had both the men in the same view.

Procedure. The procedure was nearly identical to Experiment 1 except as follows. After asking the first three detection questions, four additional questions about the content of the videos were asked. Participants who viewed the Accident event were asked 1) about the direction in which the boxes fell, 2) the side of the body with which the man with the boxes bumped the other man, 3) whether the man dropped all of the boxes and 4) which man was taller. Participants who viewed the Newspaper event were asked 1) how many times the newspaper page was flipped, 2) whether the paper was opened from the front page or back page, 3) whether the man wrote on a pad of paper or the newspaper and 3) whether the man was left handed or right handed. After the content questions, a context recognition test was given. Participants viewed the critical event and were told when the disruption would occur. After reporting whether they noticed the disruption in the second viewing, they were asked if the disruption was also present the first time they viewed the event.

To test subjects' accuracy on the content questions, we compared them with another group of 31 subjects who were given the same instructions but saw the control videos without the critical portion surrounding the disruption. The control group was asked to answer the same content questions as the other subjects. They were told to answer the questions to the best of their ability, based on what they had seen.

Results

Similar to Experiment 1, open-ended responses (e.g. recall) were again coded by two raters. Overall agreement was 89.6%, and disagreements were resolved through discussion.

As in Experiment 1, many subjects failed to detect the disruption, both when it was a motion field, and when it was a blank (see Figure 3a). Blank disruptions were not detected more often than the motion-field disruptions (detection rates: based on recall, blank = 48.8%, motion field = 47.2%, $\underline{X}^2(1, N = 77) = .019$, $p = .891$; based on recognition, blank = 53.7%, motion field = 55.6%, $\underline{X}^2(1, N = 77) = .028$, $p = .868$). In addition, duration did not significantly affect detection, and a large proportion of subjects failed to detect both the 200ms and 600ms disruptions (detection rates: based on recall, 200ms = 40.5%, 600ms = 55%, $\underline{X}^2(1, N = 77) = 1.610$, $p = .205$; based on recognition, 200ms = 48.6%, 600ms = 60.0%, $\underline{X}^2(1, N = 77) = .999$, $p = .318$). As before, results from the recall and context-recognition tests were similar. 51.9% of subjects failed to report the disruption outright, and 45.5% failed to recognize it.

Responses to the content questions were not closely associated with awareness of the disruption. Subjects who failed to recall the disruption were correct on an average of

74.4% (n = 40) of the four questions, and subjects who did recall the disruption were correct on 81.8% (n = 37) of questions (t(105) = 1.44, MSE =.801, p > .15). Similarly, subjects who did not recognize the disruption got 75.7%(n = 35) correct, compared with 79.8% (n = 42) correct for subjects who did recognize it (t(105) = .786, MSE =.813, p > .43). Subjects in the control condition were correct on 48.4% of questions (n = 31), which was significantly less accurate than subjects who failed to recognize the disruptions (t(105) = 4.91, MSE =.813, p < .001) and those who failed to recall them (t(105) = 4.85, MSE =.801, p < .001).

Insert Figure 3 about here

Discussion

In summary, we found that a substantial proportion of subjects failed to report disruptions to videotaped events that lasted between 200 and 600 ms. These disruptions went unnoticed whether they were low spatial frequency motion fields or simple blanks, and subjects failed to report them both on recall tests and recognition tests. Further, subjects who missed the disruptions showed good recognition for specific meaningful aspects of the events immediately surrounding the disruptions, and were nearly as accurate as subjects who detected the disruptions.

One of the most potentially problematic aspects of these findings is that it is not clear how aware the subjects were of the event when it actually occurred. It is, for example, possible that subjects did briefly perceive the disruption, but forgot about in the seconds intervening between the disruption and test (Moore, 2001; Wolfe, 1999). This is particularly important here because of the mundane nature of the stimulus combined with the incidental nature of the task. In contrast, in previous research the unexpected event was either highly unusual (e.g. the sudden substitution of one person for another or the appearance of a very distinctive object; Levin & Simons, 1997; Simons & Levin, 1998), or a target the subject was expecting (e.g. a second member of a target category as in the repetition blindness and inattentional blink paradigms). In these cases, the novelty of the event or the fact that it was expected might lead the subject to elaborate upon it once it was perceived, making the hypothesis that the subject was fully aware of it, then forgot it, less probable. However, other cases are more similar to that tested here. For example, in the inattentional blindness paradigm, subjects are told to perform a difficult discrimination task, and while they are doing this, an unexpected stimulus appears for 200 ms (Mack & Rock, 1998). Because the stimulus is simply a small black box, it is not particularly unusual, or memorable.

Although some variant of the inattentional amnesia hypothesis is plausible, emerging analyses of visual awareness suggest that temporary awareness of visual information should be distinguished from a more durable, and reflective awareness that is necessary to remember visual information, to incorporate it into ongoing plans, and to communicate it to others. For example, several recent discussions of visual consciousness focus on the fact that we may be aware of a large amount of information that we nonetheless fail to reflect upon (Schooler, Reichle, & Halpern, in press), or attend to (Lamme, 2003). In both cases, information may be processed and even identified, but still cannot be used for many important purposes. On this hypothesis, it is, indeed, worthwhile to ask whether unreported visual stimuli are identified, but an equally important category of experience includes a reflective or elaborative awareness that may be associated with memorable conscious experience.

Clearly, these initial findings suggest a whole series of additional questions. Most important, if we assume that people are normally aware only of the information necessary to specify events, then we would expect disruption-blindness to be more likely within events than between events as was the case with Saylor and Baldwin (in press). In an initial experiment, we have, indeed observed this. We found that 65% of subjects failed to recall 200 ms motion-field disruptions when they occurred during an action (e.g. as the man turned the page in the newspaper video), but that only 12% failed to recall the same disruptions when they were placed outside actions (e.g. when the man was reading the paper). At this point it is, however, not clear whether increased detectability of between-action disruptions is more a consequence of the perceptual conflict between the disruption and the still frame that is usually induced by pauses between actions, or more the result of a meaning-driven parsing of continuous perceptual events.

Finally, it is interesting to consider commonalities between our phenomenon and related situations where absent perceptual information is “filled in” or extrapolated based on the perceptual surround. The most obvious spatial analogs to disruption misses are illusory contours. In this case, aligned edges or terminators are used as visual evidence that a contour exists in the word even though it may not be entirely visible as a continuous break in luminance. In our task, it may be that the middle of events are filled in based on the beginnings and ends of events. Although this is a tempting analogy, it may not be entirely appropriate because the kind of filling-in we would be hypothesizing is conceptually more problematic than that implicated in illusory contours. In the case of illusory contours, there is evidence that filling-in occurs early in the visual system – some neurons in area V2 will show the same orientation-tuned response to complete edges and illusory edges (see von der Heydt, 1995 for review). In contrast, there is no particular reason to believe that subjects perceive a completed middle event in our experiments. Instead, they may be aware only of the event as an abstract entity while having no particular commitments about the totality of perceptual information constituting the event over time (Noë, Pessoa, & Thompson, 2000; Levin & Simons, 2000). Therefore, we are hesitant to argue that subjects are filling-in the events, and instead suggest that subjects perceive an event based on a few time-limited samples of visual information, and that inferences following from that are normally cognitive in nature and do not require "reperceiving" the middle of the events (for discussion see Dennett, 1991).

Summary and Conclusions

The two experiments reported here demonstrate the possibility that brief disruptions to attended visual events often go unnoticed. We conclude that visual awareness, and perhaps visual attention, may be more transitory than previously appreciated. It is important to notice the hedge in the previous sentence – we leave open the possibility that attentional processes remain focused on, and are guided by, some continuous contact with visual information. For example, continuous tracking of object locations is necessary to plan eye/hand movements, and if this requires visual attention then the disruptions that our subjects have missed may, in some sense, have been detected, and may even have interfered with some important perceptual process. On this account these failures may actually illustrate a dissociation between visual attention and visual awareness. The important thing to keep in mind is that, although visual attention is important, visual awareness is as well. If something does not reach a reflective and durable awareness, it will not be represented in our plans, we will not be able to describe it to others, and we will not remember it. As such, in a very important way, it wasn't seen.

References

Baldwin, Pederson, Craven, Andersson, & Bjork (in preparation). Change detection as a window on processing of dynamic, intentional action. University of Oregon, Eugene, OR.

Blackmore, S. (2002). There is no stream of consciousness. Journal of Consciousness Studies, 9, 5-6, 17-28.

Cave, K. R., & Bichot, N.P. (1999). Visuospatial attention: Beyond the spotlight metaphor. Psychonomic Bulletin and Review, 6, 204-223.

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.

Dennett, D.C. (1991). Consciousness Explained. New York: Littleton, Brown & Company.

Deutsch, J.A., & Deutsch, D. (1963). Attention: some theoretical considerations. Psychological Review, 70, 51-61.

Erikson, C.W., & St. James, J.D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. Perception and Psychophysics, 70, 225-240.

Lamme, V.A.F. (2003). Why visual attention and awareness are different. Trends in Cognitive Science, 7, 12-18.

Levin, D.T., & Simons, D.J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin and Review*, **4**, 501-506.

Levin, D.T., & Simons, D.J. (2000). Perceiving stability in a changing world: Combining shots and integrating views in motion pictures and the real world. Media Psychology, 2, 357-380.

Mack, A., & Rock, I. (1998). Inattentional Blindness. Cambridge, MA: MIT Press.

Moore, C.M. (2001). Inattentional blindness: perception or memory and what does it matter? Psyche: An Inerdisciplinary Journal of Research on Consciousness, 7(2). NP, http://psyche.cs.monash.edu.au/v7/psyche-7-02-moore.html

Noë, A., Pessoa, V., & Thompson, E. (2000). Beyond the grand illusion: What change blindness really teaches us about vision. Visual Cognition, 7, 93-106.

Newtson, D., & Engquist, G. (1976). The perceptual organization of ongoing behavior. Journal of Experimental Social Psychology. 12, 436-450.

O'Regan, J.K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. Canadian Journal of Psychology, 46, 461-488.

Rensink, R.A. (2000). The dynamic representation of scenes. Visual Cognition, 7, 17-42.

Saylor, M, & Baldwin, D. (in press). Action Parsing in Infancy . In D. T. Levin (Ed), *Thinking and Seeing: Visual Metacognition in Adults and Children*. Cambridge, MA: MIT Press.

Simons, D.J., & Levin, D.T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin and Review*, **5**, 644-649.

Von der Heydt, R. (1995). Form analysis in visual cortex. In M. Gazzaniga (Ed.) The Cognitive Neurosciences. Cambridge, MA: MIT Press.

Wolfe, J.M. (1999). Inattentional Amnesia. In V. Coltheart (Ed.), *Fleeting Memories* (pp. 71-94). Cambridge, MA: MIT press.

Zacks, J.M. & Tversky, B. (2001). Event structure in perception and cognition. Psychological Bulletin, 127, 3-21.

Figure 1. Illustration of a 600 ms (18 video frames) disruption during he Accident event.

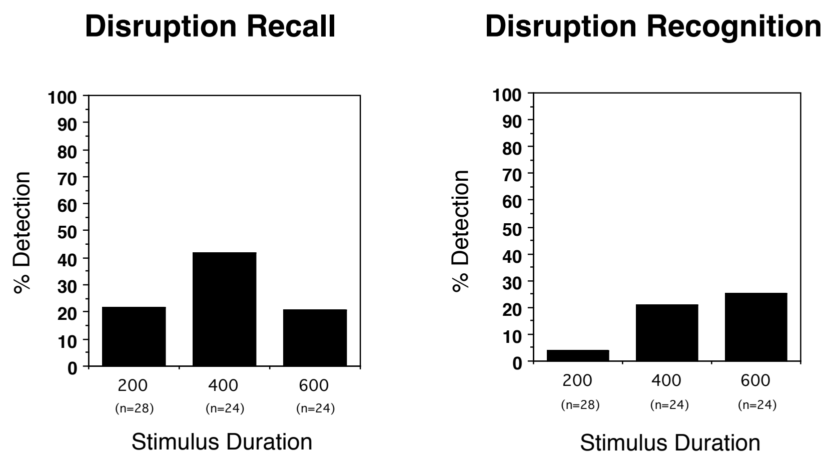

Figure 2. Recall and recognition of moving disruptions.

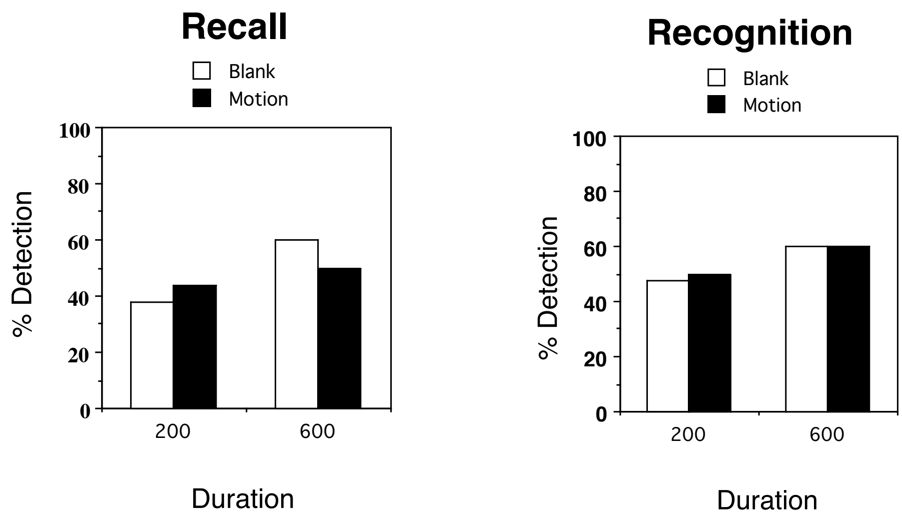

Figure 3. Recall and recognition of moving and blank-screen disruptions.