Volatile visual representations: Failing to detect changes in recently processed information

MARK W. BECKER and HAROLD PASHLER
University of California, San Diego, La Jolla, California

Research documenting people's inability to detect large changes in visual scenes suggests that visual representations may be sparse and volatile, providing no cumulative record of the attended items in a scene. However, these studies have failed to control for attention. Thus, the visual system may construct a cumulative record of all attended stimuli and still miss such changes, because they involve items that were never attended to. In two experiments, subjects saw 12-digit arrays and identified either the highest digit in the array (Experiment 1) or the lowest digit *not* in the array (Experiment 2). Subsequent change-detection tasks revealed that subjects often failed to detect changes that involved the same digits they had previously identified to perform the digit tasks successfully. This provides additional evidence that our usable visual representations are relatively impoverished and volatile.

Our subjective visual experience would seem to suggest that our minds construct a rich and detailed visual representation of any scene that we apprehend. However, the neural representation created when an observer looks at a scene seems at odds with this assumption. For example, due to the optics of the eye, neural convergence, and cortical magnification, at any instant only the foveal portion of a scene is crisp and detailed (Anstis, 1998). There are two ways that one might reconcile the subjective experience of a rich visual representation with the apparent impoverishment of the representation created during a single fixation. One view suggests that information is gathered over time, and that this accumulation of information sums to create a detailed mental representation of a visual scene (henceforth, *mental representation*). An alternate view rejects the accuracy of the subjective experience, and is based on the assumption that mental representations are sparse and volatile, never attaining the complexity that the subjective experience suggests.

According to the view that people generate a detailed mental representation, a mechanism for compiling information is required. An intuitive approach to this problem is that saccadic eye movements bring different parts of the scene to the high-acuity fovea. If an observer attends to a foveated object, then a relatively stable, detailed representation of that object may be added to the mental representation of the scene. Over successive fixations, the number of items that have been attended to increases and,

thus, the complexity of the mental representation increases. This intuitive approach suggests that mental representations contain a cumulative record (CR) of many of the aspects of a scene that have been attended to.[1] The CR hypothesis is so intuitively appealing that, as O'Regan (O'Regan, 1992; O'Regan & Noë, 2001) has pointed out, it is an implicit assumption of many experiments on visual perception. However, as O'Regan noted, this hypothesis enjoys little empirical support. In fact, a number of recent findings suggest that mental representations are limited, supporting the alternative view that they never approach the level of complexity that our subjective experience suggests.

As a result of such recent findings, a number of authors (Dennett, 1991; O'Regan, 1992; O'Regan, Rensink, & Clark, 1999; Rensink, 2000a; Simons, 2000; Simons & Levin, 1997; Wolfe, 1998) have rejected both the CR hypothesis and the accuracy of the subjective visual experience. Instead, they have advocated an alternative, which we term the *volatile visual representation* (VVR) hypothesis.

According to the VVR hypothesis, when a scene is viewed, observers create two representations. One contains the gist, or meaning, of a scene, and the other represents the visual details of a small portion of the scene. The gist representation is thought to be relatively stable, and more conceptual than visual in nature (Wolfe, 1998). By contrast, the representation of visual details is thought to be volatile and fleeting. At any instant, observers represent only the small portion of the visual environment to which they are currently attending (Pashler, 1988; Rensink, 2000b; Rensink, O'Regan, & Clark, 1997). Most attempts to quantify the number of items contained in this representation suggest that it can hold only about three to five items with enough detail to support explicit change detection (Pashler, 1988; Pylyshyn & Storm, 1988; Rensink, 2000a, 2000b). This limit is consistent with the pro-

posal that the mental representation of a scene is limited to the items held in visual short-term memory (VSTM) (Irwin & Andrews, 1996; Rensink, 2000a, 2000b). Presumably, as an observer's attention shifts, so do the contents of VSTM, leaving no memory or representation of the previously attended visual stimuli (Wolfe, Klempen, & Dahlen, 2000).

EVIDENCE FOR CHANGE DETECTION

Much of the evidence supporting the VVR hypothesis comes from experiments investigating people's ability to detect prominent changes in visual scenes. Under normal viewing conditions, people are very good at detecting actual changes; however, these changes are presumably detected on the basis of low-level transient or motion signals, which attract the observer's attention to the location of the change (Phillips & Singer, 1974; Stelmach, Bourassa, & di Lollo, 1984). Change detection is poor when the change occurs during a blink, a saccade, or a blank interstimulus interval (ISI). Presumably, this is because these situations create changes that are not signaled by motion transients. This forces the observer to rely on a representation of the scene to detect a change (for a review, see Simons & Levin, 1997). Thus, under these conditions, people are poor at detecting even rather large changes in arrays (Pashler, 1988) or photos of real-world scenes (O'Regan, Deubel, Clark, & Rensink, 2000; O'Regan et al., 1999; Rensink, O'Regan, & Clark, 2000).

These "change blindness" results have been interpreted as evidence in favor of the VVR hypothesis. However, change blindness may be consistent with the CR hypothesis as well. The CR hypothesis claims that mental representations contain a cumulative record of the items that have been attended to, but items that have not received attention are not included in this representation. Thus, it is possible that observers have a cumulative record of the items to which they have attended but in change blindness studies they fail to detect changes because they have not attended to the changing items prior to the change.

The critical difference between the VVR and the CR hypotheses concerns the fate of objects that were once the focus of attention, but are no longer. According to the VVR hypothesis, these items are no longer represented in VSTM and, thus, observers should not notice changes involving these items. According to the CR hypothesis, these items will continue to be part of an observer's internal representation of the scene, and so changes in these items should be detected.

Advocates of the VVR hypothesis have assumed that change blindness occurs even if observers attend to the to-be-changed item prior to the change. However, the evidence for this claim is indirect and less than compelling. Rensink et al. (2000) found that observers missed changes even if they were allowed to view the original image for 8 sec prior to the change. They reasoned that this should

have been ample time for completion of visual processing and, thus, the failure to detect changes was not due to a failure to ever attend to the information from the first scene. This conclusion assumes, of course, that observers used the additional exposure time to attend to and identify more details of the scene, including the aspect of the scene that subsequently changed. Since many researchers used pictures of realistic scenes, observers could analyze an almost indefinite number of potential attributes, rendering this assumption tenuous. It is possible that the observers achieved as complete a representation as they believed they needed, and carried out no further processing. The observers may have failed to detect changes because the to-be-changed items were of marginal interest and were therefore never processed. Indeed, Rensink et al. (2000) found that change blindness is much more likely to arise for items of marginal interest than for those of focal interest.

In other studies, Simons and Levin have claimed that observers miss changes to attended items. They found that, in both videotaped vignettes (Levin & Simons, 1997) and real-life interactions (Simons & Levin, 1998), the main actor in the scene could be switched with another actor (during a camera cut or an occluding event) without observers' noticing. The authors argue that the actor was the focus of the interaction, and had received observers' attention. Although this may be so, the specific attributes (e.g., facial features, hair style) that would allow an observer to distinguish the original actor from his or her replacement may not have been the focus of attention (Pani, 2000).

Indeed, Simons and Levin (1998) found that observers who were the same age as the actors were fairly good at detecting the switch, whereas age-peer observers often failed to see the change. This age-related difference disappeared when the actors dressed as construction workers (i.e., most same-age observers now failed to notice the change). The authors suggest that observers attend to only enough detail to differentiate the actor from others. When the actor is a peer of the observer, he or she is not easily differentiated from other peers; thus, observers attend to more detailed characteristics of the actor. Having done so, they retain a more complete representation of the actor and are thus good at detecting the switch. This interpretation suggests that observers do notice changes that occur in aspects of a scene to which they have attended, and it could be taken as a challenge to the VVR hypothesis.

Thus, there would appear to be no conclusive evidence establishing that people regularly fail to detect changes involving recently attended attributes or objects. Such evidence would strengthen the case for VVR. On the other hand, contrary evidence would suggest that visual representations might, after all, include a cumulative record of the attributes which have been attended to, thus favoring the CR hypothesis.

To test directly whether people fail to notice a change involving an aspect of a scene to which they have recently

attended, we conducted two experiments, each of which combined two tasks. The first task was to name the highest digit in an array of digits (Experiment 1), or to name the lowest digit that did *not* appear in an array of digits (Experiment 2). Both of these tasks forced subjects to attend to and identify some of the digits in the scene. The second task was a change-detection task that determined whether or not people had detected a change to an aspect of the scene that they had just made use of for the first task, namely the value of one of the digits.

EXPERIMENT 1

Subjects identified the highest digit among an array of 12 numbers. This task presumably required observers to attend to the values of the individual digits (Pashler & Badgio, 1985). After the subjects had completed this task, one of the digits in the scene changed, and change-detection performance was assessed.

Method

Subjects. Thirty-eight undergraduate students, who reported normal or corrected-to-normal vision, participated for class credit.

Apparatus and Stimuli. The experiment was programmed in Macromedia Director and was run on a PowerMac 6400 with a 15-in. Sony monitor running at 66 Hz.

All displays consisted of a circular array of numbers, like a clock face with 12 single-digit numbers. Each digit was presented in black on a white background and subtended ~1° X 1.5° of visual angle. The radius of the clock face was 6°, and all stimuli were legible while the center of the display was fixated; however, observers were not required to maintain fixation at the center of the display.

To construct the displays, the computer randomly chose a number between 6 and 9 to be the highest digit in the display. An additional 11 numbers were randomly selected, ranging from 1 to the highest digit minus 1. All numbers were randomly placed in one of the 12 positions on the screen. On change trials, 1 position was chosen at random and the digit at that location was switched, during a blank (white) ISI, to a different randomly selected number ranging from 1 to 9.

Design. The experiment had a within-subjects design. Each subject participated in two blocks of trials: a dual-task experimental block, and a single-task control block. In the dual-task block, observers had two tasks. First, they named aloud the highest number in the original display, and second, they reported whether or not a change had occurred during the ISI. In the control block, the observers needed only to say whether or not a change had occurred. Each block consisted of 60 trials: 40 in which a change occurred, and 20 catch trials in which no change occurred. For all trials, the subjects' responses on a change/no-change decision were recorded. In the dual-task block, the subjects' highest-digit responses were also recorded.

Procedure. The subjects were randomly assigned to receive either the control block first or the experimental block first. On the basis of this assignment, the experimenter gave verbal instructions for only the first block of trials. The subject faced the monitor, and the experimenter sat at the computer keyboard behind the subject and out of his or her view. Fifteen practice trials familiarized each subject with the task. Data were not collected for these trials.

Each trial began with a fixation cross displayed at the center of the screen for 500 msec. The initial display of numbers then appeared around the cross, and it remained on the screen for 2 sec. The display disappeared for 150 msec, and reappeared for 500 msec either with no change or with one digit changed to a different number.

In the experimental block, the subjects named the highest digit aloud as soon as they had identified it. In almost all instances, this response occurred while the first scene was still being displayed. After the blank ISI, the subjects responded by saying "change" or "no change" to indicate whether or not they had detected a change in the scene. In the control block, the subjects only made the "change"/"no-change" response.

Verbal responses were used so that the subjects would not need to look away from the screen to respond. After each trial, the experimenter recorded the subjects' responses by typing them on the computer keyboard. These data were stored and analyzed later.

At the end of the first block, the subjects were given a 3-min break, and then the experimenter gave the instructions for the second block. Another 15 practice trials were performed, followed by the second block of trials.

Results and Discussion

Figure 1 presents mean performance on the highest-digit task and the change-detection task for both the dual-task block and the change-only block. Since the change-detection task was a simple change/no-change decision, there was a high probability that the subjects would correctly guess that there was a change when none was actually detected. Therefore, the change detection data presented in the figure have been adjusted to eliminate trials in which a correct "change" response was identified as a guess. To make this adjustment, we solved the following formula for x (the portion of trials in which the subject actually detected the change):

Proportion of correct change detection (PCCD) =

$$x + (1 - x) * \text{Proportion of false alarms.}$$

The change-only block and the dual-task block had raw change detection rates of 47.96% and 44.9%, respectively, with false alarm rates of 19.1% and 16.6%, respectively. Thus, the change-detection performance, corrected for guessing, was 35.7% for the change-only block and 34% for the dual-task block.
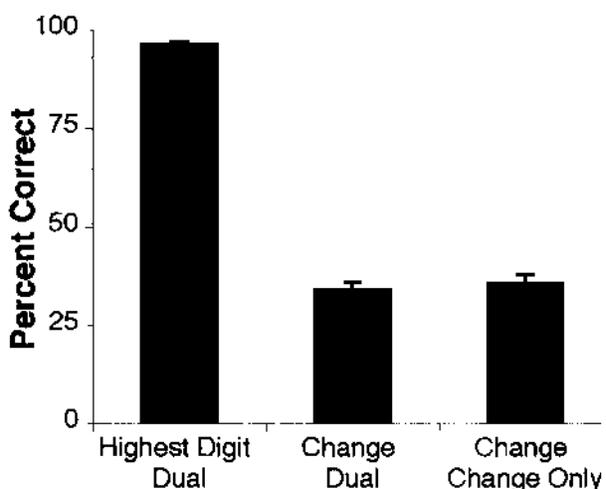

Figure 1. Mean (n = 38) performance for the highest-digit-naming task and for the change-detection task, both when each was the only task and when observers had to simultaneously perform the highest-digit task. Change-detection data are corrected for guessing. Observers performed almost perfectly on the highest-digit task, yet still performed poorly at change detection.

Figure 1 clearly shows that, during dual-task block, observers were excellent (96.4% correct) at identifying the highest digit. If we assume that correctly identifying the highest digit requires an observer to identify and process virtually every digit in the display, then this high level of performance provides evidence that subjects were attending to each item in the dual-task block. Even so, the change-detection performance in this block was fairly poor (34%). The change-only block produced a similar level of performance (35.7%), suggesting that the poor performance in the dual-task block was not caused by subjects' having to perform two tasks.

Although the CR hypothesis does not make a precise numerical prediction about change-detection performance, change-detection performance is far worse than one would expect under that hypothesis. Even in the dual-task block, when observers attended to the identity of each digit during the highest-digit task, they detected a change in these identities only about one third of the time, suggesting that these digits were not still part of their mental representation of the visual scene.

On the other hand, the VVR hypothesis makes a more specific prediction of change-detection performance. Under the VVR hypothesis, the factor limiting an observer's ability to detect a change is the number of items that can be stored in VSTM with enough detail to support explicit change detection. The PCCD should be equal to the number of items held in VSTM (n) divided by the set size (12). If we take previous work (Pashler, 1988) as the basis for our estimate of how many digits should be held in VSTM, our estimate would be about four items.[2] Thus, the VVR hypothesis predicts that about 33% of the changes should be detected. This value is very close to the observed values in both the dual-task and the change-only blocks.

Alternatively, we could use the observed PCCD (corrected for guessing) to estimate the number of items that our subjects held in VSTM. Doing so demonstrates that our subjects held 4.3 items in the change-only block and 4.1 items in the dual-task block. Both of these values are consistent with earlier estimates of the size of VSTM (Pashler, 1988). Thus, it seems that the data from the first experiment strongly support the VVR hypothesis.

The data from the dual-task block yield a second finding that supports the VVR hypothesis. It is apparent that people are excellent (87.19% correct) at detecting the change when it occurs to the highest digit.[3] This is what one would expect under the VVR hypothesis if attention was focused on the highest digit at the completion of the highest-digit task and, thus, there was a detailed representation of the highest digit in VSTM.

These findings show that people are able to access virtually all of the semantic information in a display when they perform a highest-digit task, but, having just done so, they retain only enough information to report a change for the few items that can be adequately stored in VSTM. Thus, they often fail to notice a change in the very information they have just accessed and used. This seems to rule out the idea that poor performance in change detection occurs because observers never identify the to-be-changed item in the original scene. Instead, it supports the idea of a VVR, in which attention rapidly shifts around the scene, leaving little or no residual information about what has just been attended to.

These conclusions are based on the assumption that the observer had to identify all the numbers in order to detect the highest digit. However, it is possible that the highest digit can be detected on the basis of feature detection, without necessarily identifying all of the items in the scene. For instance, an observer may have relied on a strategy of searching first for a 9, then for an 8, and so on, until he or she detected a digit. Once the highest digit was detected, the subject could reliably name it. During this type of search, an observer may be able to reject a potential digit without fully identifying it. Although we doubt that this occurred, if it did, it would challenge the assertion that subjects had identified each digit during the highest-digit task, thus weakening the interpretation of Experiment 1.

To eliminate this potential alternative account, we ran a second experiment, which required subjects to identify a set of numbers conclusively.

EXPERIMENT 2

In this experiment, the highest-digit task was replaced with the task of detecting the lowest digit that *did not* occur in the display. This task differs from the highest-digit task in important ways. First, the subject is no longer searching for the presence of an item, but for the absence of an item. Earlier research investigating the search for simple features suggests that a search for the presence of a simple feature may be done in parallel, whereas a search for the absence of the same type of simple feature is much less efficient (Treisman & Souther, 1985). If people are able to detect the highest digit in a scene by some form of simple-feature detection, they should not be able to detect the absence of a digit using the same procedure. Thus, this potential explanation of the results would be eliminated.

Second, given that searching for the absence of a stimulus is a more difficult type of task, we believe that it is much more likely to force subjects to adopt a strategy such as the one mentioned above: Search first for a 0, then a 1, and so on until, in this case, they fail to find a digit that matches the target of that particular iteration of the search. This type of strategy would be open to some testable predictions. First, reaction time (RT) measures should show that the time required to identify the missing digit should increase monotonically as the missing digit's value increases, Second, for an observer to know, for example, that 3 is the lowest missing digit (LMD), he/she must identify a 0, a 1, and a 2 in the display. Thus, in this task we can infer that subjects have identified the

items in the display that are lower than the correctly identified LMD. Furthermore, if the visual representation contains a cumulative record of all the items in a scene which have been identified, this representation must include the digits that are below a correctly identified LMD. As a result, the CR hypothesis would predict that subjects would be fairly good at detecting changes to digits that are lower than the missing digit. On the other hand, if the possible criticism of Experiment 1 is valid—namely, that while searching for a specific number, observers can eliminate distractors without identifying them—then digits higher than the LMD should rarely have been identified. Thus, according to the CR hypothesis, these digits should be subject to change blindness.

Method

Subjects. Nineteen undergraduate students participated for class credit. All reported having normal or corrected-to-normal vision.

Stimuli. The physical characteristics of the displays were the same as those used in Experiment 1. In Experiment 2, the program randomly selected a digit between 0 and 9 to be the LMD. The program then selected every number below this LMD to be in the display, and randomly assigned each one to a position on the screen. The remaining locations were filled with random digits between 0 and 9, with the provision that none could be the LMD.

In this experiment, we wanted to ensure that changes did not result in the new display's containing a different LMD than that in the original display, so the change involved randomly selecting two of the digits in the display and having them switch locations. Thus, the display always comprised the same digits before and after the change, with two of the items in opposite locations.

Procedure. The procedure was similar to that of Experiment 1, with the following changes. Each trial began when the subjects moved the cursor over the word READY and clicked the mouse. A fixation cross appeared in the center of the display for 500 msec. The initial display then appeared, and it remained on the screen until the subjects clicked the mouse to indicate that they had identified the LMD. The subjects were informed that this portion of the task was an RT task, and to respond as quickly as possible while maintaining accuracy. One second after the mouse click, the display flickered off for 150 msec and then came back on. In this experiment, subjects were informed that every trial contained a change, and they were asked to indicate the location of the change by rolling the mouse over the changed digit and clicking. The subjects were instructed that if they noticed more than one change, they should select either of the changed items. If they did not notice any change, they were asked to guess. Feedback was given via tones that indicated whether the response was correct or not. After clicking a digit to indicate the location that had changed, the display disappeared and the query "LMD?" appeared on the screen. At this point, the subjects used the keyboard to input the display's LMD, and auditory feedback was given.

In single-task blocks, the first display was presented for 2 sec. The subjects had to identify changes, but they did not perform an LMD task.

Results and Discussion

Figure 2 plots the RT data from the dual-condition blocks. As predicted, RT increased in step with the value of the LMD ($r^2 = .998$). This strikingly linear increase is what would be expected if the subjects had performed multiple iterations of a search (searching first for a 0,
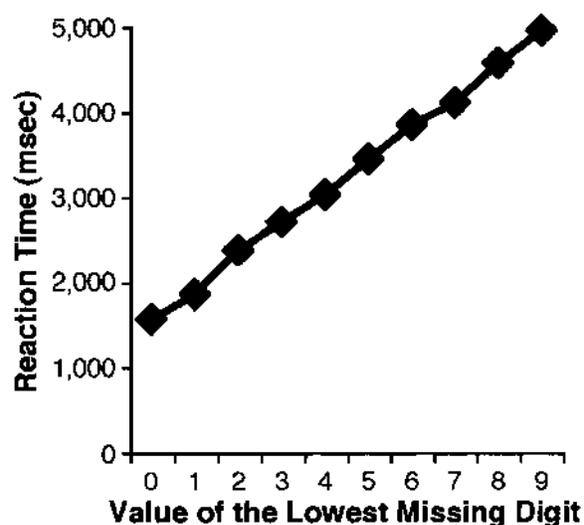

Figure 2. The reaction time (RT) required to find the lowest missing digit (LMD) is plotted as a function of the value of the LMD. RT increases linearly as the LMD's value increases, suggesting that observers use a strategy of first looking for a 0, then for a 1, then for a 2, and so forth, until one of the iterations of a search fails to find a matching digit.

then for a 1, and so on). This finding supports the assertion that people identify each of the digits lower than the LMD.

Each trial in the dual-task block was categorized as either a change in which at least one of the changing digits was lower than the LMD, or a change in which both digits that changed were above the LMD. Both categories of changes produced similar performance [52.8% and 52.3%, respectively].

If the alternative explanation of the results of Experiment 1 were true, then subjects in the present experiment should have identified virtually every number lower than the LMD, and they should have failed to identify numbers above the LMD. That is, the CR hypothesis predicts that changes involving numbers lower than the LMD should be more easily detected than changes involving numbers higher than the LMD. The present results are clearly inconsistent with this prediction, providing additional evidence against the CR hypothesis.

Whereas the CR hypothesis clearly predicts a difference in performance for these two categories of changes, the VVR does not make a strong prediction. One might assume that digits below the LMD would be more likely to have detailed representations within VSTM, and would thus yield higher change detection rates, but this does not have to be the case. It is possible that the contents of VSTM are replaced each time the subject searches for a higher digit; thus, the VVR hypothesis may be consistent with the finding that there is no difference between these types of changes.

The VVR hypothesis does, however, make a clear prediction that overall change-detection performance should

be limited by the size of VSTM. Given that two digits changed in Experiment 2, estimating the number of items held in VSTM from our data becomes more complicated than it was for the case of Experiment 1. In Experiment 2, an observer could make a correct response if any of the following three scenarios occurred: (1) Altered digit A was stored in VSTM but altered digit B was not; (2) altered digit B was stored in VSTM but altered digit A was not; or (3) both altered digits were stored in VSTM. Thus, the total PCCD should be a sum of the outcomes of these three probabilities, and is given by the following formula, in which n is the number of items held in VSTM and k is the number of items in the display:

$$\text{PCCD} = \left(\frac{n}{k}\right)\left(\frac{k-n}{k-1}\right) + \left(\frac{n}{k}\right)\left(\frac{k-n}{k-1}\right) + \left(\frac{n}{k}\right)\left(\frac{n-1}{k-1}\right).$$

After simplifying the equation, we get

$$n^2 + n(1 - 2k) + \text{PCCD} * (k^2 - k) = 0.$$

Use of the quadratic equation to solve for n yields

$$n = \frac{-(1-2k) \pm \sqrt{(1-2k)^2 - \left[4 * \text{PCCD} * \left(k^2 - k\right)\right]}}{2}.$$

By substituting the set size of 12 for k and the observed percentages of correct detections for the dual-task block (52.7%) and the change-only block (63.9%), we can solve the formula for n. These substitutions yield estimates that subjects held 3.58 items in VSTM for the dual-task block and 4.58 items in VSTM for the change-only block. Both values are consistent with previous research on VSTM. Again, the data seem to be consistent with the prediction of the VVR hypothesis.

CONCLUSION

The main finding of these experiments is that observers failed to perceive changes in the identity of items that they had just attended to and identified. In Experiment 1, we found that, even after observers had successfully reported the highest digit in a display, they rarely retained sufficient information about the digits in the display to detect when one of them had changed. We suggest that the task of detecting the highest digit forces observers to identify all of the digits in the display. This result demonstrates that the failure to notice changes in typical change detection experiments is not due to the observers' never having processed the identity of the original items.

An alternative explanation of the findings from Experiment 1 would suggest that subjects can identify the highest digit without identifying other digits in the display. This might be accomplished if subjects search for a specific digit and are able to reject distractors without fully identifying them. Experiment 2 tested this alternative possibility by having observers search for the lowest digit that *did not* occur in the display. According to the alternative explanation of Experiment 1, observers in Experiment 2 should have identified all digits below the LMD and no digits above the LMD. If this were so, and

observers had detected changes to items to which they had attended, they should have detected many changes when the changes occurred to digits lower than the LMD, and few changes when they occurred to digits higher than the LMD. The results showed, however, that both of these types of change produced equally poor performance.

In summary, we believe that the two experiments presented here directly challenge the view that mental representations contain a cumulative record of all items to which an observer has attended. Instead, the results support the view that mental representations are far more volatile than ordinary intuition would suggest (Dennett, 1991; O'Regan, 1992; Rensink, 2000a; Simons & Levin, 1997; Wolfe, 1998). At any instant, little visual information is retained, and, over time and across fixations (Grimes, 1996; McConkie & Currie, 1996), the representation does not get more complete and complex. Instead, as attention shifts, so does the content of the mental representation.

REFERENCES

ANSTIS, S. (1998). Picturing peripheral acuity. *Perception, 27*, 817-825.

DENNETT, D. C. (1991). *Consciousness explained.* Boston: Little, Brown.

FERNANDEZ-DUQUE, D., & THORNTON, I. M. (2000). Change detection without awareness: Do explicit reports underestimate the representation of change in the visual system? *Visual Cognition, 7*, 323-344.

GRIMES, J. (1996). On the failure to detect changes in scenes across saccades. In K. Akins (Ed.), *Perception* (pp. 89-110). New York: Oxford University Press.

HAYHOE, M. M., BENSINGER, D. G., & BALLARD, D. H. (1998). Task constraints in visual working memory. *Vision Research, 38*, 125-137.

IRWIN, D. E., & ANDREWS, R. V. (1996). Integration and accumulation of information across saccadic eye movements. In T. Inui & J. L. McClelland (Eds.), *Attention and performance 16: Information integration in perception and communication* (pp. 125-155). Cambridge, MA: MIT Press.

JONIDES, J., IRWIN, D. E., & YANTIS, S. (1982). Integrating visual information from successive fixations. *Science, 215*, 192-194.

JONIDES, J., IRWIN, D. E., & YANTIS, S. (1983). Failure to integrate information from successive fixations. *Science, 222*, 188.

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

MCCONKIE, G. W., & CURRIE, C. B. (1996). Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology: Human Perception & Performance, 22*, 563-581.

MCCONKIE, G. W., & ZOLA, D. (1979). Is visual information integrated across successive fixations in reading? *Perception & Psychophysics, 25*, 221-224.

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.

O'REGAN, J. K., DEUBEL, H., CLARK, J. J., & RENSINK, R. A. (2000). Picture changes during blinks: Looking without seeing and seeing without looking. *Visual Cognition, 1*, 191-211.

O'REGAN, J. K., & NOË, A. & (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*, 939-1031.

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

PANI, J. R. (2000). Cognitive description and change blindness. *Visual Cognition, 7*, 107-126.

PASHLER, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics, 44*, 369-378.

PASHLER, H., & BADGIO, P. C. (1985). Visual attention and stimulus

identification. *Journal of Experimental Psychology: Human Perception & Performance,* **11**, 105-121.

PHILLIPS, W. A., & SINGER, W. (1974). Function and interaction of on and off transients in vision. I. Psychophysics. *Experimental Brain Research,* **19**, 493-506.

PYLYSHYN, Z. W., & STORM, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision,* **3**, 179-197.

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

RENSINK, R. A. (2000b). Visual search for change: A probe into the nature of attentional processing. *Visual Cognition,* **1**, 345-376.

RENSINK, R. A., O'REGAN, J. K., & CLARK, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science,* **8**, 368-373.

RENSINK, R. A., O'REGAN, J. K., & CLARK, J. J. (2000). On the failure to detect changes in scenes across brief interruptions. *Visual Cognition,***7**, 127-147.

SIMONS, D. J. (2000). Current approaches to change blindness. *Visual Cognition,* **7**, 1-16.

SIMONS, D. J., & LEVIN, D. T. (1997). Change blindness. *Trends in Cognitive Sciences,* **1**, 261 -267.

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

STELMACH, L. B., BOURASSA, C. M., & DI LOLLO, V. (1984). Detection of stimulus change: The hypothetical roles of visual transient responses. *Perception & Psychophysics,* **35**, 245-255.

TREISMAN, A., & SOUTHER, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General,* **114**, 285-310.

WOLFE, J. M. (1998). Visual memory: What do you know about what you saw? *Current Biology,* **8**, R303-R304.

WOLFE, J. M., KLEMPEN, N., & DAHLEN, K. (2000). Postattentive vision. *Journal of Experimental Psychology: Human Perception & Performance,* 26, 693-716.

NOTES

1. This hypothesis should be distinguished from the suggestion that the visual system assembles a complete representation of both attended and unattended visual inputs in real-world coordinates (Jonides, Irwin, & Yantis, 1982), a view that has been challenged by a variety of recent studies (Jonides, Irwin, & Yantis, 1983; McConkie & Zola, 1979).

2. The previous studies that estimate VSTM capacity at about four items used tasks that require explicit reports. Recent research using implicit measures of change detection (Fernandez-Duque & Thornton, 2000; Hayhoe, Bensinger, & Ballard, 1998) has suggested that more visual information is represented than the explicit measures demonstrate. However, we feel that the VVR hypothesis is completely consistent with these implicit-measure studies. There may be residual information in VSTM about items that are no longer the focus of attention, but the fact that these representations are no longer detailed enough to support explicit change detection suggests that the information has been degraded, thus supporting the claim that these representations are volatile.

3. During change-only blocks, subjects who had not yet performed on a dual-task block should not have attributed any special import to the highest digit, and thus should not have detected changes to the highest digit so readily. The PCCD (not corrected for guessing) for changes to the highest digit in these trials was only 56%, suggesting that the highest digit did take on special import in the dual-task blocks.