Eye movements serialize memory for objects in scenes

GREGORY J. ZELINSKY
State University of New York, Stony Brook, New York

and

LESTER C. LOSCHKY
Kansas State University, Manhattan, Kansas

A gaze-contingent short-term memory paradigm was used to obtain forgetting functions for realistic objects in scenes. Experiment 1 had observers freely view nine-item scenes. After observers' gaze left a predetermined target, they could fixate from 1–7 intervening nontargets before the scene was replaced by a spatial probe at the target location. The task was then to select the target from four alternatives. A steep recency benefit was found over the 1–2 intervening object range that declined into an above-chance prerecency asymptote over the remainder of the forgetting function. In Experiment 2, we used sequential presentation and variable delays to explore the contributions of decay and extrafoveal processes to these behaviors. We conclude that memory for objects in scenes, when serialized by fixation sequence, shows recency and prerecency effects that are similar to isolated objects presented sequentially over time. We discuss these patterns in the context of the serial order memory literature and object file theory.

The ability to represent identity and location information about objects in scenes requires the visual system to solve a sort of perceptual puzzle. Object information is extracted during *fixations*, brief periods in which gaze is held stationary over a scene, but about three times each second our high-resolution fovea are repositioned over a scene by rapid movements of gaze called *saccades*. Despite their obvious importance in accumulating visual information, these saccades also serve to fracture our visual world and make impossible any simple binding of object identity to location. An object viewed during one fixation will change its retinal location and many aspects of its visual appearance following a saccade to another location in a scene. The task of representing an object in a scene therefore requires mapping these changing retinal descriptions onto a more stable representation that maintains its continuity across movements of gaze (for a review, see Bridgeman, Van der Heijden, & Velichkovsky,

1994). The fact that we experience a visual world in which objects appear not to jump and change appearance with each eye movement attests to the skill of our visual system in solving this puzzle.

How can object representations in a scene remain stable in the face of a constantly changing visual input? An early theory attempting to answer this question took the puzzle metaphor literally and hypothesized the existence of an integrative visual buffer in which foveal snapshots of individuated objects could be accumulated and arranged in a scene-based reference frame (McConkie & Rayner, 1976). Subsequent work, however, questioned this buffer conception of object representation by documenting clear cases of information not being integrated across saccades (Bridgeman & Mayer, 1983; Irwin, Brown, & Sun, 1988; McConkie & Zola, 1979; Rayner & Pollatsek, 1983). Also difficult to reconcile with an integrative visual buffer is Irwin's (1991, 1992) work on transsaccadic memory, showing that the act of making an eye movement effectively erases the sensory representation of a previously fixated item. More recent extensions of this work have clarified the object properties that can survive an eye movement and have found these properties to be highly abstracted with regard to their visual detail (e.g., general color and shape, approximate location, categorical definition), not unlike the properties often discussed in the context of visual short-term memory (STM; Carlson- Radvansky, 1999; Carlson-Radvansky & Irwin, 1995; Henderson, 1997; Henderson & Siefert, 2001; Irwin & Gordon, 1998).

To account for the survival of object properties across changes in gaze, Irwin and colleagues proposed transsaccadic object file theory (OFT), a modification to tradi-

tional OFT that makes explicit the relationship between object memory and eye movements (Irwin, 1996; Irwin & Andrews, 1996). Like traditional OFT (Kahneman & Treisman, 1984; Kahneman, Treisman, & Burkell, 1983; Kahneman, Treisman, & Gibbs, 1992; Treisman, 1988), transsaccadic OFT states that object identity–location bindings exist in the form of a limited number of spatially indexed property lists known as *object files*. However, transsaccadic OFT adds to this framework the assumption that visual STM capacity can accommodate only 3–4 object files across changes in eye position. If an object in a scene corresponds to one of these 3–4 available object files, the object properties associated with its file can be retrieved and used in a working memory task.

This assumption of transsaccadic OFT has important implications for the relationship between eye movements and object representation in scenes. Suppose a person is faced with the task of remembering the objects in an unfamiliar room. A reasonable plan might be to move gaze from object to object, pausing long enough on each to encode its identity and location into memory. Assuming that fixations on new objects are likely to be accompanied by the opening of an object file, transsaccadic OFT predicts good memory for the 3–4 objects visited most recently by gaze. Of course, objects viewed earlier in the sequence of fixations might also be remembered if some of their properties were encoded using a more enduring representation, but the existence of object file representations should still result in a memory benefit for the most recently fixated objects.

The recency benefit predicted by transsaccadic OFT is reminiscent of the recency effects reported for over a century in the memory literature on serial position effects (see Baddeley, 1986, Greene, 1986, and Neath, 1998, for historical treatments of this topic). Moreover, there is at least a surface similarity between the sequence of gaze shifts made while encoding the objects of a scene and the serial presentation conditions that are central to the expression of serial memory phenomena. Given these similarities, one might expect the serial memory literature to inform the question of scene-based object representation, but the exchange of ideas between these two research communities has been minimal. Slowing the flow of information are significant methodological differences between these two classes of studies. On the one hand, scene-based object memory studies typically present objects simultaneously as part of a scene, and then ask observers to make some judgment about one of the scene objects (for a review, see Henderson & Hollingworth, 1999a). On the other hand, studies designed to investigate serial position effects on visual memory typically present a series of individual patterns (e.g., histoforms, kaleidoscope patterns, snowflakes) at a central display location, and then ask observers whether a test pattern appeared in this study series (Broadbent & Broadbent, 1981; Neath, 1993; Phillips, 1983; Wright, 1998).[1] Two methodological concerns, therefore, prevent the gener-

alization of findings from the serial memory literature to the question of object representation in scenes. First, no equivalency has been demonstrated between the sequential presentation paradigm and the order in which objects are viewed in a scene. Are the objects fixated in a scene serialized in memory in the same way as they would be if they were presented individually one after the other? Second, because serial memory studies typically present stimuli at the same location in space, these studies provide no information about the representation of objects scattered over many locations in a scene. In short, many serial memory studies lack a spatial memory component and therefore cannot be used to study the identity–location binding that is central to OFT and, more generally, the representation of objects in scenes.

The extensive literature on spatial working memory suffers from the opposite problem with regard to informing research on scene-based object memory. In a typical spatial memory study, stimuli are presented serially at different display locations and observers are tested in one of three ways: (1) they are either asked to recall the locations of all the items from the memory set without regard for order (e.g., Hale, Myerson, Rhee, Weiss, & Abrams, 1996), (2) they are presented with a spatial probe and asked whether the probe location corresponds to one of the item locations from the memory set (e.g., Awh, Jonides, & Reuter-Lorenz, 1998; Courtney, Petit, Maisog, Ungerleider, & Haxby, 1998), or (3) they are asked to reproduce the entire spatiotemporal sequence in which the study items were presented (e.g., Jones, Farrand, Stuart, & Morris, 1995; Smyth & Scholey, 1996). Given this taxonomy of spatial memory studies, several methodological differences again prevent the generalization of findings to the question of object–location binding in scenes. First, many spatial memory studies are unconcerned with how the stimulus presentation order might affect one's spatial memory ability and therefore do not include order of stimulus presentation in their analyses (e.g., Awh et al., 1998; Courtney et al., 1998; Hale et al., 1996). Second, the studies that do take presentation order into account are primarily concerned with memory for temporal order per se, not the effect of presentation order on memory for object representations (e.g., Jones et al., 1995; Smyth & Scholey, 1996). Third, in many spatial memory studies, object identity is irrelevant to the task. Popular spatial memory paradigms, such as the Corsi blocks task or the 7/24 spatial recall test, do not even manipulate object identity, instead using simple dots or checker shapes as stimuli (e.g., Rao, Hammeke, McQuillen, Kharti, & Lloyd, 1984; Shimozaki et al., 2003). Thus, the majority of existing spatial memory studies, while important, shed little light on real-world scene representation.

To our knowledge, only Walker and colleagues (Walker, Hitch, Doyle, & Porter, 1994; Walker, Hitch, & Duroe, 1993) manipulated both the identity and location of objects for the purpose of studying how serial presentation order affects memory. Walker et al. (1993) presented a sequence of four geometric patterns at four horizontally

arranged display positions, followed by the presentation of a test stimulus at a neutral display location. The observer's task was to indicate the previous location of this test pattern in the study array. When plotted as a serial position function, the data from this study revealed a pronounced recency effect that was largely limited to the last item presented in the study sequence. This finding is important because it demonstrates serial position effects on the basis of the conjunction of spatial and identity information about an object, thereby potentially informing the scene representation literature. However, it is also important not to generalize too quickly from this finding to object memory in scenes. The patterns used by Walker et al. (1993) were highly confusable nonsense shapes, whereas those appearing in scenes are typically real-world objects. The serial position effects reported by these authors may therefore be influenced by this choice of stimuli. Moreover, the placement of patterns in the Walker et al. (1993) study was highly predictable and limited to only four locations in a horizontal array; the spatial distribution of objects in scenes is far less constrained. Finally, and most important to the present discussion, is the fact that Walker et al. (1993) still used a sequential presentation paradigm. In the real world, observers typically shift their gaze from object to object as they encode patterns into memory. The relationship between order of stimulus presentation and the eye movements to objects appearing in scenes remains a largely unaddressed question in the serial memory community.

Many studies in the visual cognition literature have used eye movement measures to assess memory in multiobject displays (Hayhoe, Bensinger, & Ballard, 1998; Henderson & Hollingworth, 1999b; Henderson, McClure, Pierce, & Schrock, 1997; Henderson, Pollatsek, & Rayner, 1987, 1989; Henderson & Siefert, 1999, 2001; Hollingworth, Schrock, & Henderson, 2001; Hollingworth, Williams, & Henderson, 2001; Zelinsky, 2001), but here, too, there have been few attempts to specify how memory is affected by the order of fixations on objects. Two recent studies have addressed this relationship directly. Irwin and Zelinsky (2002) allowed observers to view a seven-object display for either 1, 3, 5, 9, or 15 fixations. Following this gaze-dependent viewing period, a spatial probe cued one of the object locations in the study display, and the observer was given a seven alternative forced choice (AFC) recognition test for the cued target. When Irwin and Zelinsky analyzed their gaze data in terms of recency of object fixation (Figure 8, p. 889), a clear relationship emerged between recognition accuracy and when the target was fixated during study. When the target was fixated immediately prior to probe onset, recognition accuracy was over 90%. However, when the target was fixated 1 or 2 fixations earlier, accuracy dropped to about 80%. Accuracy dropped further to 65% for targets fixated even farther back in the eye movement sequence (3, 4, or 5 fixations prior to the spatial probe) but then remained constant at this above-chance level. Irwin and Zelinsky interpreted this recency benefit for objects viewed within

the last 3 fixations as support for transsaccadic OFT and its assumption of a 3–4 object memory capacity.

In another recent study, Hollingworth and Henderson (2002) performed a similar post hoc analysis of their eye position data and reached a very different conclusion. These authors made type or token changes to objects in realistic scenes as observers inspected these scenes in anticipation of a memory test. Observers were also asked to indicate whether a scene change was noticed during viewing. Among their many interesting results, Hollingworth and Henderson failed to find a reliable recency effect in their data. Unlike Irwin and Zelinsky (2002), detection accuracy in their task did not vary systematically with the number of fixations since the target was last viewed. However, consistent with Irwin and Zelinsky (2002; cf. Zelinsky & Loschky, 1998), Hollingworth and Henderson did find an above-chance prerecency level of accuracy extending several items back in the viewing sequence, which they interpreted as evidence for an LTM representation of object properties.

As evidenced by the discrepancy between Irwin and Zelinsky (2002) and Hollingworth and Henderson (2002), the literature does not yet have a clear picture of how serial fixation order affects memory for objects in scenes, nor is it even clear what theoretical framework will best describe these data patterns once they have stabilized. Although transsaccadic OFT's prediction of a 3–4 object recency effect is well supported by Irwin and Zelinsky and much of the previous transsaccadic memory literature (Irwin, 1992; Irwin & Andrews, 1996; Irwin & Gordon, 1998), this prediction is greater than the one-object recency effect often observed in studies using a sequential presentation paradigm (Phillips & Christie, 1977a, 1977b; Wright, Santiago, Sands, Kendrick, & Cook, 1985; for reviews, see Phillips, 1983, and Wright, 1998) and the zero-object recency effect reported by Hollingworth and Henderson. The present study was conducted to address these discrepant estimates of recency in the scene-based object memory literature, as well as to explore factors potentially contributing to above-chance levels of prerecency accuracy reported by Irwin and Zelinsky and elsewhere (Hollingworth & Henderson, 2002; Kerr, Avons, & Ward, 1999; Phillips, 1983; Phillips & Christie, 1977a; Walker et al., 1994; Walker et al., 1993; Wright, 1998; Wright et al., 1985). By outlining a method of serializing the encoding of scene objects using eye movements, it was also our hope to build a bridge between the scene-based object memory literature and the many studies using a sequential presentation paradigm to describe serial order effects on object memory.

EXPERIMENT 1

One possible reason for the discrepancy between the findings of Irwin and Zelinsky (2002) and Hollingworth and Henderson (2002) may be that different methods of quantifying serial position were used in the two studies. Hollingworth and Henderson plotted percent accuracy

for a target object as a function of the number of fixations made after gaze left the target. The forgetting function derived by Irwin and Zelinsky (Figure 8, p. 889) was somewhat different, plotting target accuracy as a function of the number of objects inspected following fixation on the target. This distinction is important because the number of fixations made during scene viewing does not typically map perfectly onto the number of fixated objects. Often, multiple fixations are devoted to a single object, and sometimes two neighboring objects are inspected with only a single fixation (Zelinsky, Rao, Hayhoe, & Ballard, 1997). Although Irwin and Zelinsky (2002) and Hollingworth and Henderson (2002) both plotted memory accuracy as a function of viewing order, it is possible that their forgetting functions were not capturing the same serial position information and consequently could not be directly compared.

Experiment 1 describes a purely object-based forgetting function for items viewed serially in simple scenes. As in the Irwin and Zelinsky (2002) study, we monitor the eye movements of observers as they freely inspect a multiobject scene, and we make the duration of this study period contingent on their oculomotor behavior. However, rather than terminating the study display after a set number of fixations, we make testing contingent on the number of different objects fixated by observers after their gaze leaves a predesignated target. This online monitoring of the number of posttarget objects fixated by observers allows us to parametrically manipulate the number of objects viewed between study and test. Rather than relying on post hoc fixation analyses to derive recency and prerecency serial memory functions, our manipulation of intervening objects as an independent variable means that our experimental design will have sufficient power to discern the memory patterns in question. We believe that this more powerful design, when combined with an object-based gaze-contingent methodology, will provide a clearer measure of object memory in scenes.

Method

Stimuli. The stimuli consisted of nine common real-world objects (toy, tool, or food items) arranged on an appropriate background sur-
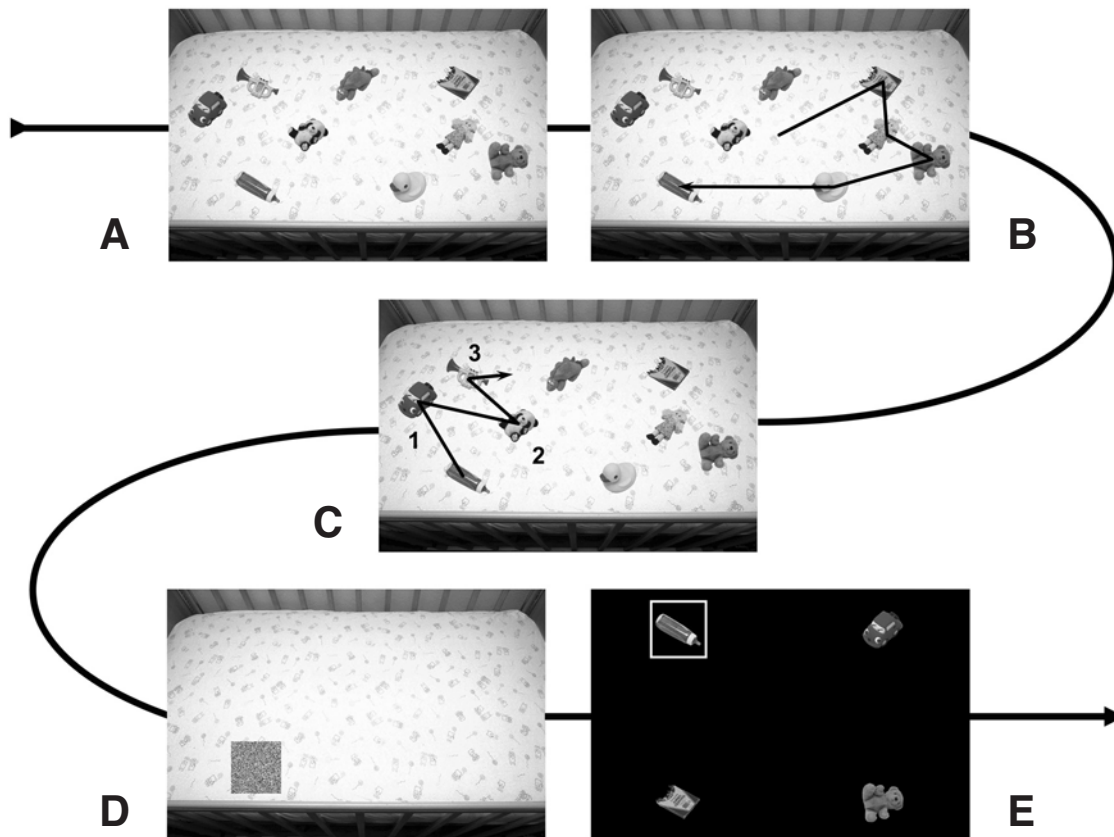

Figure 1. Events comprising a typical trial in Experiment 1. (A) A nine-object scene was presented to observers. Their task was to study these objects in anticipation of a memory test. (B) Observers freely viewed the objects with the intention of remembering their identities and locations, eventually fixating the target object (in this case, the baby bottle). (C) Unknown to the observer, the display program would then count the number of objects fixated after gaze left the target. (D) When the criterion number of different posttarget objects (in this case, three) was fixated, the multiobject display was replaced by a spatial probe at the target's location. (E) The observer then had to select the target object from among four alternatives.

face (a crib, workbench, or dining table). Each simple scene subtended $18° \times 11.6°$ of visual angle and was of near-photographic quality (16-bit color).[2] Individual objects were scaled to fit inside a 2.4° bounding box, and their locations in the scene were constrained to 18 positions, creating what appeared to be a haphazard arrangement of items lying on a surface. As a result of these placement constraints, scene objects had a minimum and maximum center-to-center separation of 2.4° and 14.5°, respectively. Multiple trials for a given scene type were created by randomly pairing the objects to locations (i.e., the same nine objects would occupy different locations in each scene), with the constraints that no two display configurations were repeated and that the memory target would appear equally often in each of the 18 allowable positions.

Procedure. The experimental procedure is shown in Figure 1. A scene appeared at the start of a trial (Figure 1A), and the observers' task was to remember the identity and location of every object—in this case, toys in a baby's crib (there were also tools on a workbench and food items on a dining table). Given this formidable memory task, observers invariably made eye movements to the individual objects, a response that was anticipated and crucial to the current paradigm but not explicitly mentioned in the experiment instructions. Eye position was monitored every millisecond and analyzed online to determine the object in the scene being fixated by the observer. Unknown to the observer, one object in the display was predesignated as the memory target for that particular trial (e.g., the baby bottle in Figure 1). As the observer freely scanned the scene, gaze would eventually be directed to this target object (Figure 1B). This fixation event was detected by the program controlling the experiment, which then started to count the number of different objects fixated after gaze left the target. This "count" constituted the termination criterion for the display, meaning that if the count was preset to three, the observer would be allowed to fixate exactly three objects after the target (Figure 1C). As gaze moved away from the third posttarget object, the study scene was replaced by a 1-sec duration probe (a colored noise mask) appearing at the target's location on an "emptied" background surface (Figure 1D). Following the probe, a display showing four objects was presented, and the observer had to indicate which of them appeared at the probed location (Figure 1E).[3] One of these objects was always the target; the other three were randomly selected from the study scene. Because observers were on a bite-bar and could not easily speak, they responded by looking at the desired object, causing a white box to be drawn around the item, then pressing a hand-held button when satisfied with their selection. Observers were asked to respond as accurately as possible without regard for time.

The independent variable of interest in this study was intervening objects, the number of different objects looked at after the target. There were seven intervening object conditions (1–7). For example, in the one-intervening-object condition, the study display terminated during the saccade away from the first posttarget object (i.e., gaze was not allowed to land on a second nontarget item after leaving the target). Likewise, in the seven-intervening-object condition, the observer fixated exactly seven different nontarget objects following target fixation. The amount of time that observers viewed the study scene therefore varied, it was determined by when they first fixated the target and the duration of their gaze on each object, as well as the intervening object criterion set for that particular trial. The intervening object conditions were randomly interleaved throughout the experiment, resulting in an unpredictable study scene duration. Postexperiment questioning revealed that the participants were uniformly unaware that scene presentation time depended on their own pattern of eye movements, attributing the variable interval, instead, to some random schedule of presentation durations.

Participants. Six experimentally naive observers from the University of Illinois at Urbana-Champaign each participated in 378 trials: 54 in each of the seven intervening object conditions. Eye position was monitored throughout each trial, using a Generation V dual-Purkinje-image (DPI) eyetracker sampling at 1000 Hz. The spatial precision of this tracker was better than 3 min of visual angle at a viewing distance of 62 cm. Saccadic eye movements were determined online using a velocity-based algorithm implementing a roughly 12.5°/sec detection threshold. The experiment required two 1.5-h sessions conducted on separate days, and the participants were paid $24 upon completion.

Results and Discussion

Our gaze-contingent memory paradigm allows us to plot response accuracy as a function of intervening objects for trials in which the target was fixated only once during viewing (i.e., no target refixations). This analysis, shown in Figure 2A by the solid markers, reveals two clear patterns of behavior. First, memory declined precipitously between the 1 and 3 intervening object conditions, dropping linearly from an 87% level of accuracy to only 65% [$F(6,30) = 6.53$, $MS_e = 0.044$, $p < .001$]. This pattern suggests that the act of fixating the first three objects after the target dramatically interfered with memory for the target object—much like the interference observed when objects are presented one after another in a sequential memory paradigm. Second, this drop in accuracy reached an abrupt asymptote at three intervening objects, meaning that item fixations over the 4–7 intervening object range resulted in no additional object-related interference. Moreover, because chance response in this 4AFC task should produce a 25% level of accuracy, the 65% level described by this asymptote indicates a clearly above-chance and relatively good memory for an object, irrespective of the number of items fixated after it during study.[4] We refer to the region of declining accuracy in this intervening object function as the *recency* memory component and the above-chance asymptotic region of accuracy as the *prerecency* memory component.

Our clear evidence for distinct recency and prerecency behavior, although consistent with the data from Irwin and Zelinsky (2002), is inconsistent with Hollingworth and Henderson's (2002) failure to observe a recency effect. To determine whether this discrepancy is due to our quantification of serial order in terms of intervening objects rather than intervening fixations, we reanalyzed our data using Hollingworth and Henderson's intervening fixation method (the open markers in Figure 2A). This comparison revealed considerable agreement between the two functions, with mean differences in accuracy amounting to only 2.5%.[5] This high degree of similarity would be expected if observers tended to devote only a single fixation to each object during study. More important, accuracy again declined with the number of fixations occurring after gaze left the target [$F(6,30) = 4.93$, $MS_e = 77.38$, $p < .001$]. Note, however, that this recency effect is less pronounced than that in the intervening object data. We attribute this pattern to a special role that objects play in producing recency behavior. Differences between the intervening object and intervening fixation functions are caused by the occasional multiple fixation of individual objects during study. Given that these fix-
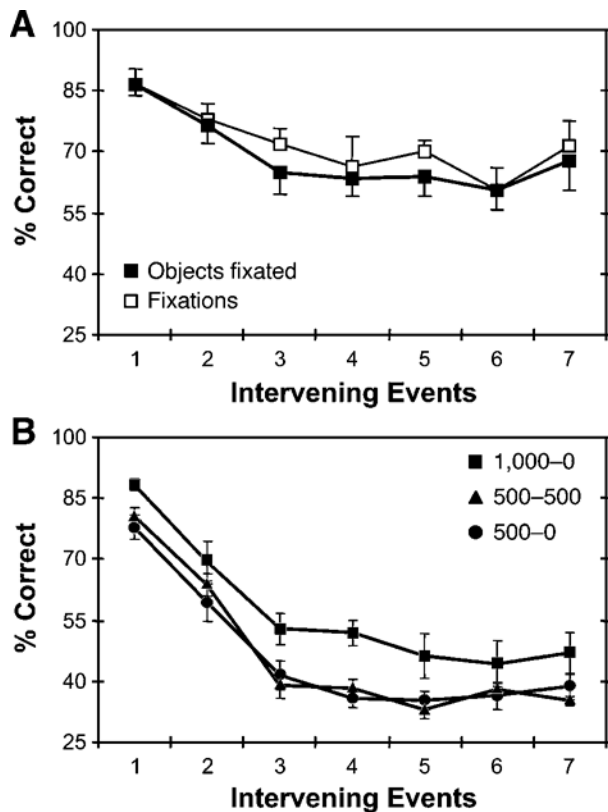
Figure 2. Memory forgetting functions relating accuracy to the number of different events following viewing of the target. Lines extending from the markers indicate one *SEM*. (A) Data from Experiment 1, in which a single nine-object study scene was presented on each trial and memory accuracy was defined in terms of either the number of intervening objects fixated between target and test or the number of intervening fixations. (B) Data from Experiment 2, in which each trial consisted of nine one-object study displays presented sequentially over time. Square markers indicate data from a 1,000–0 condition in which individual objects were presented for 1 sec each without an interstimulus interval (ISI). Data from the 500–500 (500-msec presentation, 500-msec ISI) and 500–0 (500-msec presentation, 0-msec ISI) conditions are shown by the triangle and circle markers, respectively.

ations on the same object were less detrimental to memory than were fixations on different objects (as indicated by the slightly better accuracy in the intervening fixation data), it appears that objects, not fixations, are primarily responsible for the accuracy decline that defines recency in our task. Moreover, if one were to replot such an object-based recency effect as a function of intervening fixations, the result should be a stretching of the object-based recency effect over more intervening events (i.e., fixations), which is the exact pattern appearing in Figure 2A. This dilution of recency in the intervening fixation data suggests one factor contributing to Hollingworth and Henderson's (2002) failure to find a recency effect in their study. As recency is stretched over more of the intervening event function, the slope becomes shallower and the effect more difficult to observe. For this reason, we be-

lieve that the preferred method of quantifying recency during scene viewing is to serialize the data by objects rather than by fixations.

As a working interpretation of our data, we hypothesize that retroactive interference (RI) is a likely cause of the observed recency effect. Consistent with a classic RI explanation for STM forgetting (Waugh & Norman, 1965; see also Intraub, 1984, for RI applied to scenes), each object fixated after the target in our task might have interfered with memory for the target, eventually driving recognition performance into the prerecency baseline. However, interference might be proactive as well as retroactive (Keppel & Underwood, 1962; Wickens, Born, & Allen, 1963), and given that a relatively small set of objects was being reused in each trial of this experiment, the potential for proactive interference (PI) influencing the data would seem high. Previous research using a serial presentation paradigm has found that between-trial PI effects on picture recognition memory are greatest early in an experiment rather than late (Jitsumori, Wright, & Shyan, 1989; Wright, 1998), particularly under conditions of immediate testing (e.g., a 1-sec retention interval). It might then be the case that observers were able to accurately remember the target over a large intervening object range early in the experiment but, as a result of PI accumulating over trials, this high level of accuracy gradually declined in the later trials to chance performance. If the forgetting function was changing during the course of the experiment, the recency and prerecency components in Figure 2A might be mere artifacts of averaging over trials. To explore this possibility, we evenly divided the 378 trials into "early" and "late" conditions. If PI was building up throughout the experiment, prerecency accuracy in the early-trial condition should be higher than in the late-trial condition. However, as can be seen from Figure 3, this was clearly not the case. If anything, accuracy actually improved somewhat over the 3–7 intervening object range in the late trials. Rather than suffering from a buildup of PI, observers benefited slightly from their growing familiarity with the set of stimulus objects.

Although between-trial PI appears not to have played a role in producing the Figure 2A intervening object function, there still exists the possibility that another form of PI might have influenced our data. Given our free-viewing paradigm, the number of nontarget objects fixated before gaze first shifted to the target necessarily varied with the observers' own idiosyncratic scanning preferences. If observers happened to inspect more pretarget objects in the 3–7 intervening object conditions, and if the fixation of these objects interfered with target memory, the buildup of PI within a trial might result in the lower accuracy observed in the prerecency component.

To evaluate this potential for within-trial PI, we first tested the premise that pretarget fixations were less frequent in the 1–2 intervening object conditions compared with the 3–7 intervening object conditions. The results from this analysis are shown in Table 1. As expected on the basis of our random selection of targets, the number
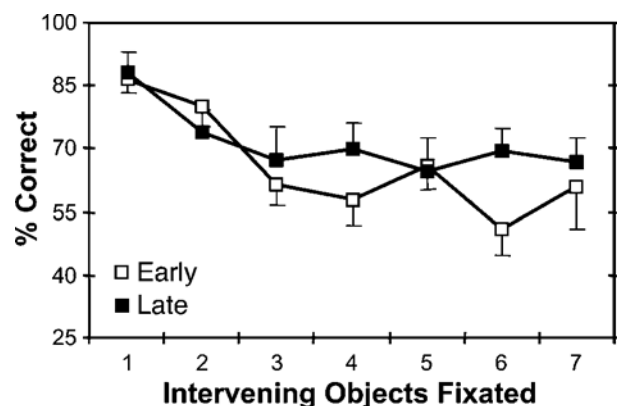
Figure 3. Intervening object functions derived from an early-versus-late segregation of the Experiment 1 data. Open markers show data from the first 189 trials of the experiment; closed markers show data from the final 189 trials.

of distractor objects fixated before the target in a trial did not meaningfully differ between the 1–2 and the 3–7 intervening object conditions [Pearson χ^2 (8) = 3.25, p = .918; N = 1,457]. Our failure to find an overrepresentation of pretarget distractor fixations in the 3–7 intervening object trials therefore weakens the argument that within-trial PI could account for our recency data. We also directly looked for a within-trial PI effect by comparing trials in which the target was fixated very early in the sequence of fixations (with 0, 1, or 2 objects fixated before the target) with trials in which the target was first fixated later (three or more objects fixated before the target). If within-trial PI played a role in determining the recency effect, accuracy should be higher when there were fewer pretarget fixations. Table 2 indicates that this, too, was not the case. Trials in which the target was fixated very early (i.e., few pretarget fixations) were no more accurate than trials in which the target was fixated much later (i.e., several pretarget fixations). We can therefore conclude that pretarget fixations, and whatever within-trial PI they may have exerted, cannot explain the recency effects reported in Figure 2A.

Unresolved Questions

Since we dispatched PI as a factor shaping the Figure 2A forgetting functions, an explanation appealing to RI becomes more attractive. However, two remaining alternative explanations cannot be ruled out, based on the Experiment 1 data. First, it is possible that the recency behavior reported in Figure 2A reflects a decay process, rather than RI from objects fixated after the target. Decay-based explanations for STM forgetting enjoyed considerable popularity in the early days of modern memory research (e.g., Peterson & Peterson, 1959; see also Altmann & Gray, 2002, for a recent formulation of this view) but were largely replaced by interference-based explanations after several studies demonstrated little or no forgetting over time when item interference was controlled (Loess & Waugh, 1967; Waugh & Norman, 1965; Wickens et al., 1963). Although the latter studies certainly reduce the plausibility that decay was responsible for the currently observed recency behavior, this factor cannot be ruled out for the specific presentation and testing conditions used in Experiment 1. A second factor potentially affecting the Figure 2A forgetting data is specific to the simultaneous presentation paradigm. Our use of eye movements to serialize object processing assumes that processing is restricted to only the currently fixated object during free viewing. Several studies, however, have also demonstrated processing benefits for extrafoveally viewed objects (Henderson & Anes, 1994; Pollatsek, Rayner, & Henderson, 1990). If multiple objects were being processed during each display fixation, this opportunity for extrafoveal processing of the target might explain the above-chance level of accuracy observed in the Figure 2A prerecency data. Although such extrafoveal processing benefits have been reported only for the intended target of a saccade and not every object in a display (Henderson, Pollatsek, & Rayner, 1989), the potential for a more generalized form of extrafoveal processing benefit affecting our data should be considered.

EXPERIMENT 2

Experiment 2 directly addressed the possibilities that decay or extrafoveal processes may have influenced the Experiment 1 data. The question of whether decay or interference underlies the Figure 2A recency effect was resolved by holding constant the number of intervening objects appearing between study and test but varying the retention interval. If decay was responsible for the observed decline in accuracy, lengthening the retention interval should accelerate the accuracy decline and constrict the expression of recency on the forgetting function.

Table 1
Percentage of Trials With a Given Number of Pretarget Object Fixations
(0–8) as a Function of Two Intervening Object Groups (1–2 vs. 3–7)

	Number of Pretarget Objects Fixated								
Number of Intervening Objects	0	1	2	3	4	5	6	7	8
1–2	11.2	13.4	11.2	12.5	9.0	11.2	10.3	10.6	10.6
3–7	10.3	11.3	12.4	12.0	10.2	10.4	10.9	10.6	12.0

Note—Data exclude trials in which the target object was refixated. n = 554 and 903 cases for the 1–2 and 3–7 intervening object groupings, respectively.

Table 2
Memory Accuracy as a Function of the Number
of Objects Fixated Before the Target

Objects Fixated	% Correct	SEM
0	70.3	8.05
1	71.2	5.68
2	72.7	4.34
3+	70.8	3.83

Note—Data in the 0 row indicate cases in which the target was fixated by the first saccade following display onset. SEM = standard error of the mean.

However, if our recency effect was caused by interference, varying the retention interval should have no meaningful effect on accuracy. To address the question of whether extrafoveal processing was contributing to prerecency performance in Experiment 1, we adopted in Experiment 2 a standard sequential presentation methodology. If objects are presented one after the other in time rather than simultaneously as scenes, there could be no opportunity for extrafoveal processing to affect memory performance. By comparing prerecency accuracy between Experiments 1 and 2, we can therefore assess any extrafoveal influences.

Method

Unlike Experiment 1, in which we inferred serial encoding from the observers' pattern of eye movements to the objects, Experiment 2 more closely replicated past object memory studies by using a sequential presentation paradigm. Although we used the same real-world objects and backgrounds in Experiment 2 as in Experiment 1, a sequential paradigm required that each of these objects now be presented separately over time (Figure 4). Instead of showing observers a single multiobject display, we decomposed each Experiment 1 display and presented the component objects singly one after the other. Because a nine-object scene was always presented in each Experiment 1 trial, for each trial in Experiment 2 we presented a sequence of nine single-object displays. The identity and location of these objects were identical to a counterpart scene from a trial in Experiment 1. Thus, if the nine Experiment 2 displays for a given trial were superimposed, the corresponding Experiment 1 scene would be obtained. The presentation duration of each single-object display, and the interstimulus interval (ISI) between the offset of one object and the onset of the next, were independent variables manipulated in three conditions. In the 500–0 condition, each object was presented for 500 msec, followed immediately (0-msec ISI) by the next object in the sequence; the 500–500 condition showed each object for 500 msec, followed by a 500-msec ISI showing only the "empty" background scene surface; and the 1,000–0 condition showed each object for a full second without an accompanying ISI.[6] Thirty-six observers from Stony Brook University, 12 in each of the above three conditions, participated for
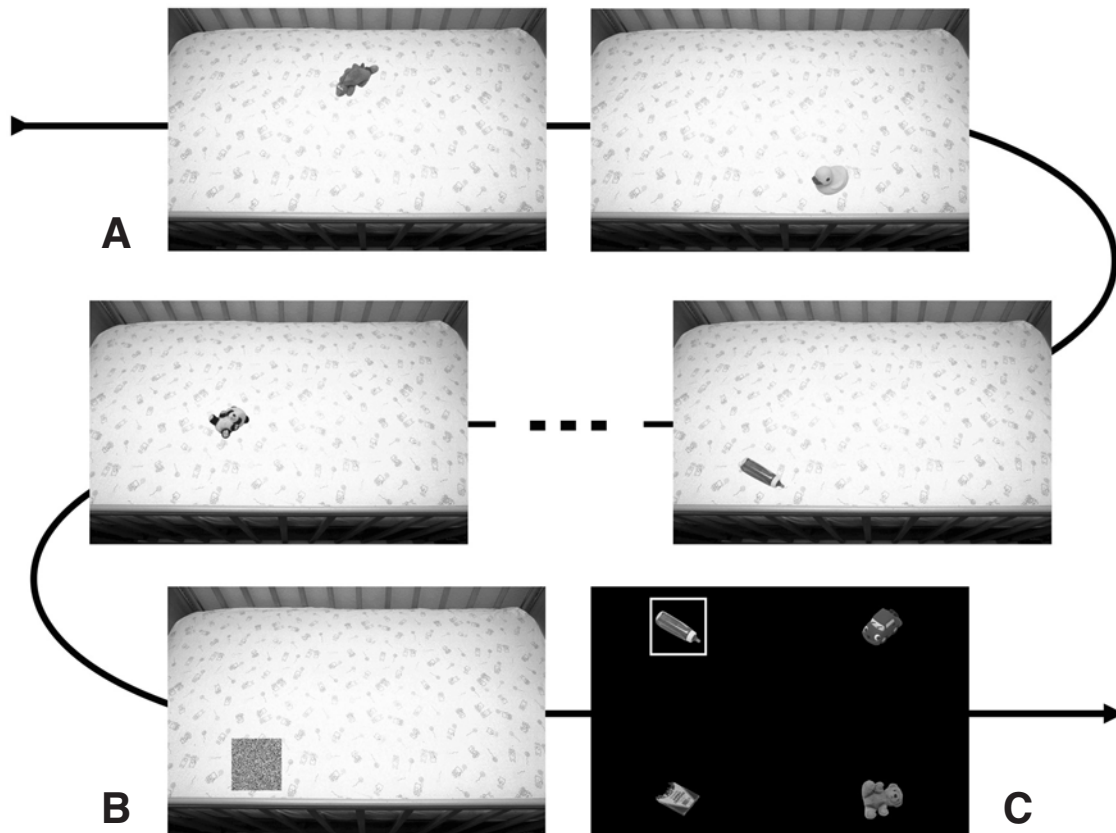

Figure 4. Events composing a typical trial in Experiment 2. (A) A sequence of nine one-object scenes was presented to observers. (B) Following the presentation of this sequence, a spatial probe appeared at the target object's location. (C) The observer then had to select the target from among four alternatives.

course credit. There were 378 trials per observer (as in Experiment 1), which were completed in one 2.5-h session. Because eye movements were not monitored in this experiment, the observers now registered their selection of a memory target by moving a computer mouse cursor to the desired object in the response grid (Figure 4C) and clicking a button.[7] All other aspects of the design and procedure, including the presentation of the memory probe and the identities and locations of objects in the response grid, were identical to those in Experiment 1.

Results and Discussion

The data from Experiment 2 are shown in Figure 2B. Despite dramatically different presentation methods, the recency effects for the three Experiment 2 conditions were remarkably similar to the pattern observed in Experiment 1. Accuracy in the 500–0 condition was 78% after one intervening object, 59% after two, and 41% after three; accuracy in the 500–500 condition was 81%, 64%, and 39%, respectively. The prerecency level of accuracy in both the 500–0 and 500–500 conditions was approximately 36%. A somewhat different pattern was found in the 1,000–0 data. Accuracy after one intervening object was 88%, with this initially high level of performance declining sharply over the second (69%) and third (53%) intervening objects. Prerecency accuracy in the 1,000–0 condition was 47%, 18% lower than what was found in Experiment 1 and roughly 11% higher than in the other Experiment 2 conditions. To quantify these patterns, we performed a two-way mixed design ANOVA with number of intervening objects as a within-subjects factor and presentation condition (500–0, 500–500, and 1,000–0) as a between-subjects factor. As expected, we found a highly significant main effect of intervening objects [$F(4.28, 141.26$, Greenhouse–Geisser adjusted) = 175.75, $p < .001$], and a significant main effect of presentation condition [$F(2,33) = 5.00, p = .013$], but these variables did not significantly interact [$F(8.56, 141.26$, Greenhouse–Geisser adjusted) = 1.03, $p = .418$]. The effect of presentation condition was driven by significant differences (Sidak corrected) between the 1,000–0 condition and the 500–0 [$t(22) = 2.79, p = .026$] and 500–500 [$t(22) = 2.69, p = .033$] conditions. The two 500-msec presentation conditions were not significantly different [$t(22) = 0.105, p = .999$].

Is the recency component a result of decay or of intervening objects? We used the dissociation logic described by Waugh and Norman (1965) to tease apart decay and interference factors in Experiment 2, varying the presentation rate of consecutive objects while holding constant the number of items appearing in each intervening object condition. Objects in the 500–0 condition appeared at a rate of two per second; objects in the 500–500 condition appeared at a rate of one per second. These conditions therefore allow a straightforward test of the decay versus interference hypotheses over the timeframes relevant to object memory during free viewing. Assuming that a decay process was responsible for the observed recency effects and that the rate of this decay

was constant, the recency component in the 500–500 data should be twice as steep as that in the 500–0 data. This prediction follows from the fact that a decay process in the 500–500 condition would have had twice as long to degrade object representations compared with the 500–0 condition.[8] However, if the observed recency effects were due to interference introduced by the objects appearing between the target and probe, accuracy should not differ between these two conditions. Figure 2B provides clear support for the interference hypothesis. The data patterns in the 500–0 and 500–500 conditions overlap and suggest no interaction between retention interval and intervening objects. Consistent with Waugh and Norman, we therefore conclude that a decay process was not a meaningful contributor to the reported recency effects, even at the relatively brief exposure durations used in this study. Rather, the recency component is better described by an interference process that accumulates with the addition of each intervening object to the stimulus sequence.

Do extrafoveal processes contribute to prerecency memory? The sequential presentation paradigm adopted in Experiment 2 provides a straightforward test of whether extrafoveal processing contributed to the above-chance level of prerecency memory reported in Experiment 1. According to this argument, if observers in Experiment 1 were extracting information from multiple objects while fixated on an individual item, the target (and many of the nontargets) would have received considerably more processing than what was previously assumed—processing that might translate into an above-chance prerecency component. This hypothesis therefore predicts a significantly higher prerecency component for Experiment 1, compared with the three Experiment 2 viewing conditions in which extrafoveal processing was precluded by sequential presentation.

To evaluate this hypothesis, we analyzed the prerecency data from Experiments 1 and 2 and found a significant difference [$t(40) = 4.75, p < .001$] between the gaze-contingent Experiment 1 data ($M = 64.17, SD = 9.47$) and the data from Experiment 2 collapsed across the three presentation conditions ($M = 40.82, SD = 11.38$). Multiple comparisons (Sidak corrected) further revealed significant differences between the Experiment 1 prerecency level of memory and each of the three prerecency components from Experiment 2 [$t(16) \geq 2.31, p \leq .022$]. Experiment 2 therefore provides partial support for the existence of extrafoveal benefits shaping performance in our task. Consistent with this proposal, prerecency accuracy in all three of the Experiment 2 conditions was well below the level observed in Experiment 1, meaning that the high prerecency component from Experiment 1 may have been due in part to extrafoveal processes benefiting target encoding. However, such extrafoveal processes cannot explain why prerecency accuracy was higher in the 1,000–0 condition relative to the 500–0 and 500–500 conditions. Because a se-

quential presentation paradigm was used in all three of the Experiment 2 conditions, those differences cannot be explained in terms of an extrafoveal processing advantage.

Consistent with previous work (Hollingworth & Henderson, 2002; Phillips & Christie, 1977a), we believe that our Experiment 2 data suggest that longer viewing times lead to better memory for the target, regardless of the number of intervening items. Such a relationship between viewing duration and target memory might explain, in part, the high prerecency level of accuracy observed in Experiment 1. The average first-pass target gaze duration in Experiment 1 was 683 msec over the 3–7 intervening item range, intermediate to the 500-msec and 1,000-msec presentation durations used in Experiment 2. Given that the target objects in Experiment 1 were processed an average of 183 msec longer than those in the short-duration Experiment 2 conditions, it is reasonable to speculate that viewing duration may have contributed to the Experiment 1 prerecency component. However, because the average Experiment 1 target viewing time was considerably shorter than in the 1,000–0 condition from Experiment 2, we must also conclude that viewing duration is not the sole determinant of prerecency performance. Additional work will be needed to weight the contribution of these factors, but for now it appears that both viewing duration and extrafoveal processing may have played roles in determining prerecency accuracy in Experiment 1.[9]

GENERAL DISCUSSION

Two clear data patterns emerged from this study. First, recency effects exist for objects presented in scenes, with this memory advantage declining steadily with fixation of the first three intervening objects (Experiment 1). As for the cause of this recency advantage, we were able to rule out both between- and within-trial PI. The forgetting functions for early and late trials did not differ, and cases in which the target was fixated early in a trial produced no accuracy advantage over those in which the target was fixated later in a trial. By shifting to a sequential presentation paradigm and varying the retention interval (Experiment 2), we were also able to exclude decay as a possible cause of this recency effect. The serial buildup of RI from fixations on intervening objects appears to be the only plausible explanation. Regardless of whether this serial order of processing is imposed on the memory set by the experimenter (as in a sequential presentation paradigm) or imposed by the observer's own eye movements (as in our simultaneous presentation paradigm), the recency memory advantage rapidly declines with the number of objects viewed between target and test.

The second salient pattern emerging from our data is the constant and relatively high level of accuracy over the 3–7 intervening object range. Multiple factors likely contribute to this above-chance prerecency level of accuracy. One such factor is the amount of time that observers have to view the display objects. Whereas viewing dura-

tion was necessarily uncontrolled in Experiment 1 and confounded with simultaneous presentation, Experiment 2 suggested that this factor may have contributed to prerecency memory in our task, with observers exploiting longer viewing times to improve accuracy over the prerecency component of the forgetting function. We also believe that the extrafoveal availability of objects in Experiment 1 may have played a role in elevating this prerecency level of performance. We base this assertion on the fact that differences in presentation duration in Experiment 2 could not account for all of the variation in prerecency performance between the simultaneous and sequential presentation conditions.

At this point, however, the nature of this extrafoveal contribution remains unclear. One possibility is that observers were able to process multiple items during each object fixation in Experiment 1, and they then used this extrafoveal information to improve their prerecency accuracy. However, given Henderson et al.'s (1989) finding of extrafoveal processing benefits being limited to the target of a saccadic eye movement, we would expect such an influence to be relatively small. A potentially more important form of extrafoveal influence may involve the formation of spatial codes relating one object to another. When an observer is free to control the order in which items are encoded into memory, as would be the case in Experiment 1, memory for this scanning order may enable the target's identity to be better retrieved upon presentation of the spatial probe. Future work will vary the orderliness with which objects are sequentially presented in the study scene in an attempt to isolate and weight the contribution of this factor to prerecency performance.

Relating the present data to the object memory literature. Both the recency and prerecency behavior observed in this study are in near-perfect agreement with the patterns reported by Irwin and Zelinsky (2002). Consistent with Irwin and Zelinsky's post hoc analysis of recency behavior, we found a three-object recency benefit (extending over the target and the first two intervening objects) using a completely object-based gaze-contingent methodology. Irwin and Zelinsky also reported a high and relatively constant 65% level of accuracy when testing occurred farther back in the viewing sequence, a pattern again very similar to the above-chance prerecency component that we found in our study. The consistency between these two studies is notable, given their very different display-termination criteria, and suggests that the reported data patterns are indeed real and quite robust.

The present data agree less well with the nonscene object memory literature that typically reports recency benefits for only the final pattern presented in a series (for a review, see Phillips, 1983). However, one should be careful when generalizing these one-object recency estimates to object memory in scenes. Objects in scenes appear in spatially diverse configurations, and this spatial information can be used to help access and individuate objects in memory as assumed by OFT. Most object memory studies, however, present objects one after the other in the

same spatial location, thereby preventing the use of spatial cues to assist memory and potentially underestimating our memory ability. The representation of identity and spatial information for objects in scenes is also a blend of visual and semantic codes that depends critically on the stimulus, task, and the observer's ability to chunk object features into meaningful units (Avons & Phillips, 1987; Cowan, 2001; Gordon & Irwin, 2000; Logan, 1995). The object memory literature, to the extent that it has relied on abstract visual patterns, minimizes the contribution of these nonvisual codes and therefore again potentially underestimates our ability to remember objects in scenes. Although teasing apart visual, verbal, and semantic contributions to object memory is a worthwhile topic for future research, when it comes to representing the identity and location of real-world objects in simple scenes, it is probably safe to say that our three-item recency estimate provides a better indication of human memory ability.

The present data also stand in sharp contrast to the absence of a recency effect reported by Hollingworth and Henderson (2002). We attribute this puzzling discrepancy in part to the fact that these authors quantified their forgetting function in terms of intervening fixations rather than intervening objects. If observers in their study tended to devote more than one fixation per visit to an object, as may well have been the case given that they were engaged in a change-detection task, any object-based recency effect would be diluted over a wider range of the forgetting function and might therefore become more difficult to discern. It might also have been the case that their post hoc analysis lacked an adequate number of trials per intervening fixation condition to observe a recency effect. If so, our manipulation of intervening objects as an independent variable may simply have provided our study with the power needed to observe recency behavior.

Unlike recency memory, which may be sensitive to stimulus type and methodological factors, there is no disagreement regarding our above-chance level of prerecency memory ability. With Hollingworth and Henderson (2002) and Irwin and Zelinsky (2002), three studies have now shown that memory exists for objects in scenes independent of when these objects were fixated in the viewing sequence. This finding has an important implication for scene representation in intentional memory tasks. Although we may have very good memory for only the last three objects that we view in a scene, we nevertheless have a reasonably good memory for many more scene objects.

The implications of recency and prerecency memory benefits may also extend beyond standard intentional working memory tasks and are potentially relevant to any visuocognitive task involving the free viewing of multi-item displays. For example, consider our often profoundly impaired ability to detect changes to objects in scenes (for reviews, see Rensink, 2002, Simons, 2000, and Simons & Levin, 1997). Change detection difficulty

has been traditionally attributed to sparse representation arising from attentional limitations (Grimes, 1996; Levin & Simons, 1997; O'Regan, Rensink, & Clark, 1999; Rensink, 2000a, 2000b; Rensink, O'Regan, & Clark, 1997). According to this view, limits on selective attention prevent the representation of one or both of the pre- or postchange objects, thereby resulting in a change detection failure. The present data suggest a different perspective from which to view this literature, one focused on working memory constraints rather than attentional limitations. The stimuli used in change detection studies range from realistic scenes (O'Regan et al., 1999; Rensink et al., 1997) to video clips (Levin & Simons, 1997) to naturalistic events (Simons & Levin, 1998), and in all of these stimuli there were potentially dozens of objects that observers might have inspected with dozens of fixations. On the basis of the forgetting function reported here and elsewhere (Irwin & Zelinsky, 2002; Zelinsky & Loschky, 1998), newly fixated objects in the scene should interfere with memory for previously fixated scene objects. The best detection performance should therefore be observed when the changed object was one of the last three objects fixated in the scene. If the change target was fixated farther back in the viewing sequence, detection of this change would have to rely on a less accurate prerecency memory (Hollingworth & Henderson, 2002), resulting in a higher frequency of change detection failure. Rather than an extremely sparse representation (O'Regan, 1992), change detection might therefore be served by a fallible but relatively dense representation extending over the last seven objects fixated during viewing, and possibly many more. In this sense, our memory-constrained view is consistent with recent explanations of change detection suggesting that information exists for many objects in a scene, but that this information is impoverished and not sufficient to affect performance on every memory task (Angelone, Levin, & Simons, 2003; Simons, Chabris, Schnur, & Levin, 2002; Zelinsky, 2003). If observers fail to fixate both the pre- and postchange objects during scene viewing, even prerecency memory would be unavailable to the task, and the probability of detection should approach chance (Henderson & Hollingworth, 1999b; Hollingworth, Schrock, & Henderson, 2001; but see Zelinsky, 2001).

Reconciling the intervening object function with object memory theory. Can OFT be reconciled with the clear patterns of recency and prerecency memory reported in this study? According to transsaccadic OFT, our STM for objects in a scene is limited to only 3–4 property lists at any given moment in time (Irwin, 1992, 1996; Irwin & Andrews, 1996; Irwin & Gordon, 1998; Irwin & Zelinsky, 2002; see also Cowan, 2001, for a perspective from the working memory literature). Assuming that each object fixated in our task resulted in the creation of a new object file, transsaccadic OFT would predict a 3–4 item recency benefit very much like the one observed in the present study. Note, however, that

OFT's suggestion of a capacity limitation makes its STM function analogous to the operation of a fixed-capacity conveyor belt (Zelinsky, 2001). Once the belt is filled with objects, each new file linked to an object must be accompanied by the deallocation of an existing object file. OFT therefore predicts a straightforward pattern of object displacement from memory; good memory should exist for the 3–4 most recently viewed objects still on the belt, and poorer memory should exist for those objects that have fallen off the belt as a result of being fixated farther back in the viewing sequence.

To account for above-chance prerecency memory, object file theorists have had to amend their theories to include mechanisms capable of longer term object representation. For example, in addition to a short-term object file representation, Irwin (1996) posited the existence of a semantic network containing information about object properties. Similarly, Hollingworth and Henderson (2002) recently suggested a dual-store modification to OFT. These authors proposed that short-term object file representations can consolidate into a more permanent record of an object's properties. This LTM representation is thought to remain available following the withdrawal of attention and the dissolution of the object's spatially indexed property list (see Henderson, 1994, 1997, Henderson & Anes, 1994, Henderson & Siefert, 2001, Irwin & Zelinsky, 2002, and Pollatsek et al., 1990, for related thoughts regarding the relationship between object files and long-term representation). Such a structural division of theoretical labor is attractive in that it offers a clear-cut explanation for both recency and prerecency patterns of object memory. Much like the theories proposed for object memory during sequential presentation (Kerr et al., 1999; Phillips, 1983; Phillips & Christie, 1977a), by assuming separate STM and LTM systems, each with their own processing limitations and capacity constraints, it becomes possible to explain recency and prerecency behavior simply by assigning each component of the forgetting function to a different memory system.

However, despite the intuitive appeal of such a dual-store model of object memory, restraint should be exerted before adopting it. Given that a dual-store model has two free parameters, and that only two components of the forgetting function require explanation, applying such a model to the present data is as much a description as it is a theoretical advance. What is needed is the introduction of new, alternative theories of object memory in scenes, followed by critical evaluations of these theories to determine which offers the best and most parsimonious account of the data. The present study was intended to help bring about this crucial step in theory development. There currently exists a schism in the working memory community, with the visual cognition literature embracing capacity-limited models and dual-store theories (Hollingworth & Henderson, 2002; Irwin & Andrews, 1996; Luck & Vogel, 1997; Rensink, 2000a, 2000b) and the serial memory literature explaining performance limitations in terms of distinctiveness (Knoedler, Hellwig, & Neath, 1999; Neath, 1993), interference (Wright, 1998; Wright et al., 1985), and related forms of processing constraint. By using eye movements to serialize the encoding of scene objects into memory and then deriving from these data an intervening object forgetting function, it is our hope to bridge these communities and enable serial order memory theory to begin informing our memory for objects presented simultaneously in scenes.

REFERENCES

Altmann, E. M., & Gray, W. D. (2002). Forgetting to remember: The functional relationship of decay and interference. *Psychological Science*, **13**, 27-33.

Angelone, B., Levin, D., & Simons, D. (2003). The roles of representation and comparison failures in change blindness. *Perception*, **32**, 947-962.

Avons, S. E., & Phillips, W. A. (1987). Representation of matrix patterns in long- and short-term visual memory. *Acta Psychologica*, **65**, 227-246.

Awh, E., Jonides, J., & Reuter-Lorenz, P. A. (1998). Rehearsal in spatial working memory. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 780-790.

Baddeley, A. (1986). *Working memory* (Vol. 11). Oxford: Oxford University Press, Clarendon Press.

Bridgeman, B., & Mayer, M. (1983). Failure to integrate visual information from successive fixations. *Bulletin of the Psychonomic Society*, **21**, 285-286.

Bridgeman, B., Van der Heijden, A., & Velichkovsky, B. (1994). A theory of visual stability across saccadic eye movements. *Behavioral & Brain Sciences*, **17**, 247-292.

Broadbent, D. E., & Broadbent, M. H. (1981). Recency effects in visual memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, **33A**, 1-15.

Carlson-Radvansky, L. A. (1999). Memory for relational information across eye movements. *Perception & Psychophysics*, **61**, 919-934.

Carlson-Radvansky, L. A., & Irwin, D. E. (1995). Memory for structural information across eye movements. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 1441-1458.

Courtney, S. M., Petit, L., Maisog, J. M., Ungerleider, L. G., & Haxby, J. V. (1998). An area specialized for spatial working memory in human frontal cortex. *Science*, **279**, 1347-1351.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral & Brain Sciences*, **24**, 87-185.

Gordon, R. D., & Irwin, D. E. (2000). The role of physical and conceptual properties in preserving object continuity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 136-150.

Gottesman, C.V., & Intraub, H. (2002). Surface construal and the mental representation of scenes. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 1-11.

Greene, R. L. (1986). Sources of recency effects in free recall. *Psychological Bulletin*, **99**, 221-228.

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.

Hale, S., Myerson, J., Rhee, S. H., Weiss, C. S., & Abrams, R. A. (1996). Selective interference with the maintenance of location information in working memory. *Neuropsychology*, **10**, 228-240.

Hayhoe, M., Bensinger, D., & Ballard, D. (1998). Task constraints in visual working memory. *Vision Research*, **38**, 125-137.

Henderson, J. M. (1994). Two representational systems in dynamic visual representation. *Journal of Experimental Psychology: General*, **123**, 410-426.

Henderson, J. M. (1997). Transsaccadic memory and integration during real-world object perception. *Psychological Science*, **8**, 51-55.

HENDERSON, J. M., & ANES, M. D. (1994). Roles of object-file review and type priming in visual identification within and across eye fixations. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 826-839.

HENDERSON, J. M., & HOLLINGWORTH, A. (1999a). High-level scene perception. *Annual Review of Psychology*, **50**, 243-271.

HENDERSON, J. M., & HOLLINGWORTH, A. (1999b). The role of fixation position in detecting scene changes across saccades. *Psychological Science*, **10**, 438-443.

HENDERSON, J. M., McCLURE, K. K., PIERCE, S., & SCHROCK, G. (1997). Object identification without foveal vision: Evidence from an artificial scotoma paradigm. *Perception & Psychophysics*, **59**, 323-346.

HENDERSON, J. M., POLLATSEK, A., & RAYNER, K. (1987). Effects of foveal priming and extrafoveal preview on object identification. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 449-463.

HENDERSON, J. M., POLLATSEK, A., & RAYNER, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Perception & Psychophysics*, **45**, 196-208.

HENDERSON, J. M., & SIEFERT, A. B. [C.] (1999). The influence of enantiomorphic transformations on transsaccadic object integration. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 243-255.

HENDERSON, J. M., & SIEFERT, A. B. C. (2001). Types and tokens in transsaccadic object identification: Effects of spatial position and left–right orientation. *Psychonomic Bulletin & Review*, **8**, 753-760.

HOLLINGWORTH, A. (2004). Constructing visual representations of natural scenes: The roles of short- and long-term visual memory. *Journal of Experimental Psychology: Human Perception & Performance*, **30**, 519-537.

HOLLINGWORTH, A., & HENDERSON, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 113-136.

HOLLINGWORTH, A., SCHROCK, G., & HENDERSON, J. M. (2001). Change detection in the flicker paradigm: The role of fixation position within the scene. *Memory & Cognition*, **29**, 296-304.

HOLLINGWORTH, A., WILLIAMS, C. C., & HENDERSON, J. M. (2001). To see and remember: Visually specific information is retained in memory from previously attended objects in natural scenes. *Psychonomic Bulletin & Review*, **8**, 761-768.

INTRAUB, H. (1980). Presentation rate and the representation of briefly glimpsed pictures in memory. *Journal of Experimental Psychology: Human Learning & Memory*, **6**, 1-12.

INTRAUB, H. (1984). Conceptual masking: The effects of subsequent visual events on memory for pictures. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 115-125.

INTRAUB, H. (1999). Understanding and remembering briefly glimpsed pictures: Implications for visual scanning and memory. In V. Coltheart (Ed.), *Fleeting memories: Cognition of brief visual stimuli* (pp. 47-70). Cambridge, MA: MIT Press.

IRWIN, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology*, **23**, 420-456.

IRWIN, D. E. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 307-317.

IRWIN, D. E. (1996). Integrating information across saccadic eye movements. *Current Directions in Psychological Science*, **5**, 94-99.

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

IRWIN, D. E., BROWN, J., & SUN, J. (1988). Visual masking and visual integration across saccadic eye movements. *Journal of Experimental Psychology: General*, **117**, 276-287.

IRWIN, D. E., & GORDON, R. D. (1998). Eye movements, attention and trans-saccadic memory. *Visual Cognition*, **5**, 127-155.

IRWIN, D. E., & ZELINSKY, G. J. (2002). Eye movements and scene perception: Memory for things observed. *Perception & Psychophysics*, **64**, 882-895.

JITSUMORI, M., WRIGHT, A. A., & SHYAN, M. R. (1989). Buildup and release from proactive interference in a rhesus monkey. *Journal of Experimental Psychology: Animal Behavior Processes*, **15**, 329-337.

JONES, D., FARRAND, P., STUART, G., & MORRIS, N. (1995). The functional equivalence of verbal and spatial information in serial short-term memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 1008-1018.

KAHNEMAN, D., & TREISMAN, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 29-61). New York: Academic Press.

KAHNEMAN, D., TREISMAN, A., & BURKELL, J. (1983). The cost of visual filtering. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 510-522.

KAHNEMAN, D., TREISMAN, A., & GIBBS, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, **24**, 175-219.

KEPPEL, G., & UNDERWOOD, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning & Verbal Behavior*, **1**, 153-161.

KERR, J. R., AVONS, S. E., & WARD, G. (1999). The effect of retention interval on serial position curves for item recognition of visual patterns and faces. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 1475-1494.

KNOEDLER, A. J., HELLWIG, K. A., & NEATH, I. (1999). The shift from recency to primacy with increasing delay. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 474-487.

LAWRENCE, B. M., MYERSON, J., OONK, H. M., & ABRAMS, R. A. (2001). The effects of eye and limb movements on working memory. *Memory*, **9**, 433-444.

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

LOESS, H., & WAUGH, N. C. (1967). Short-term memory and inter-trial interval. *Journal of Verbal Learning & Verbal Behavior*, **6**, 455-460.

LOGAN, G. (1995). Linguistic and conceptual control of visual spatial attention. *Cognitive Psychology*, **28**, 103-174.

LUCK, S., & VOGEL, E. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, **390**, 279-280.

McCONKIE, G. W., & RAYNER, K. (1976). Identifying the span of the effective stimulus in reading: Literature review and theories of reading. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (2nd ed., pp. 137-162). Newark, DE: International Reading Association.

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

NEATH, I. (1993). Distinctiveness and serial position effects in recognition. *Memory & Cognition*, **21**, 689-698.

NEATH, I. (1998). *Human memory: An introduction to research, data, and theory*. New York: Brooks/Cole.

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

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

PETERSON, L. R., & PETERSON, M. J. (1959). Short-term retention of individual items. *Journal of Experimental Psychology*, **61**, 12-21.

PHILLIPS, W. A. (1983). Short-term visual memory. *Philosophical Transactions of the Royal Society of London*: *Series B*, **302**, 295-309.

PHILLIPS, W. A., & CHRISTIE, D. F. M. (1977a). Components of visual memory. *Quarterly Journal of Experimental Psychology*, **29**, 117-133.

PHILLIPS, W. A., & CHRISTIE, D. F. M. (1977b). Interference with visualization. *Quarterly Journal of Experimental Psychology*, **29**, 637-650.

POLLATSEK, A., RAYNER, K., & HENDERSON, J. M. (1990). Role of spatial location in integration of pictorial information across saccades. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 199-210.

POTTER, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Perception & Performance*, **2**, 509-522.

POTTER, M. C., & LEVY, E. I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology*, **81**, 10-15.

POTTER, M. C., STAUB, A., RADO, J., & O'CONNOR, D. H. (2002). Recognition memory for briefly presented pictures: The time course of rapid forgetting. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 1163-1175.

RAO, S. M., HAMMEKE, T. A., McQUILLEN, M. P., KHARTI, B. O., & LLOYD, D. (1984). Memory disturbance in chronic progressive multiple sclerosis. *Archives of Neurology*, **41**, 625-631.

RAYNER, K., & POLLATSEK, A. (1983). Is visual information integrated across saccades? *Perception & Psychophysics*, **34**, 39-48.

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

RENSINK, R. (2000b). Visual search for change: A probe into the nature of attentional processing. *Visual Cognition*, **7**, 345-376.

RENSINK, R. (2002). Change detection. *Annual Review of Psychology*, **53**, 247-277.

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

SHIMOZAKI, S., HAYHOE, M., ZELINSKY, G., WEINSTEIN, A., MERIGAN, W., & BALLARD, D. (2003). Effect of parietal lobe lesions on saccade targeting and spatial memory in a naturalistic visual search task. *Neuropsychologia*, **41**, 1365-1386.

SIMONS, D. J. (2000). Current approaches to change blindness. *Visual Cognition*, **7**, 1-15.

SIMONS, D. J., CHABRIS, C. F., SCHNUR, T., & LEVIN, D. T. (2002). Evidence for preserved representations in change blindness. *Consciousness & Cognition*, **11**, 78-97.

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.

SMYTH, M. M., & SCHOLEY, K. A. (1996). Serial order in spatial immediate memory. *Quarterly Journal of Experimental Psychology*, **49A**, 159-177.

TREISMAN, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology*, **40A**, 201-237.

WALKER, P., HITCH, G. J., DOYLE, A., & PORTER, T. (1994). The development of short-term visual memory in young children. *International Journal of Behavioral Development*, **17**, 73-89.

WALKER, P., HITCH, G. J., & DUROE, S. (1993). The effect of visual similarity on short-term memory for spatial location: Implications for the capacity of visual short-term memory. *Acta Psychologica*, **83**, 203-224.

WAUGH, N. C., & NORMAN, D. A. (1965). Primary memory. *Psychological Review*, **72**, 89-104.

WICKENS, D. D., BORN, D. G., & ALLEN, C. K. (1963). Proactive inhibition and item similarity in short-term memory. *Journal of Verbal Learning & Verbal Behavior*, **2**, 440-445.

WRIGHT, A. A. (1998). Auditory and visual serial position functions obey different laws. *Psychonomic Bulletin & Review*, **5**, 564-584.

WRIGHT, A. A., SANTIAGO, H. C., SANDS, S. F., KENDRICK, D. F., & COOK, R. G. (1985). Memory processing of serial lists by pigeons, monkeys, and people. *Science*, **229**, 287-289.

ZELINSKY, G. J. (2001). Eye movements during change detection: Implications for search constraints, memory limitations, and scanning strategies. *Perception & Psychophysics*, **63**, 209-225.

ZELINSKY, G. J. (2003). Detecting changes between real-world objects using spatiochromatic filters. *Psychonomic Bulletin & Review*, **10**, 533-555.

ZELINSKY, G.J., & LOSCHKY, L. (1998). Toward a more realistic assessment of visual short-term memory [Abstract]. *Investigative Ophthalmology & Visual Science*, **39**, S224.

ZELINSKY, G.J., & LOSCHKY, L. (2003). Fuzzy object file theory: A framework for understanding recency effects for objects in scenes [Abstract]. *Journal of Vision*, **3**, 644a.

ZELINSKY, G.J., & MURPHY, G.L. (2000). Synchronizing visual and language processing: An effect of object name length on eye movements. *Psychological Science*, **11**, 125-131.

ZELINSKY, G.J., RAO, R.P., HAYHOE, M.M., & BALLARD, D.H. (1997). Eye movements reveal the spatiotemporal dynamics of visual search. *Psychological Science*, **8**, 448-453.

NOTES

1. Note that there is also a literature in which scene stimuli are presented in rapid sequence and observers have to indicate whether a particular test scene was included in that series (Intraub, 1980; Potter, 1976; Potter & Levy, 1969; Potter, Staub, Rado, & O'Connor, 2002; for a review, see Intraub, 1999). These studies deal primarily with the representations and processes required to differentiate one scene from another and not the representation of individual objects within a particular scene. In the context of the present investigation, we therefore see these studies as basically serial presentation memory studies that have substituted scene stimuli for isolated objects.

2. We define a scene as one or more objects obeying a lawful relationship to a background context, such as tools sitting on a workbench. We refer to our stimuli as simple scenes to acknowledge the fact that object placement in these displays was less relationally and contextually constrained than what is typical for fully realistic scenes. See Gottesman and Intraub (2002) for more discussion of what is, and is not, a scene.

3. Because OFT assumes that object properties can be indexed by their position in space (Kahneman et al., 1992), we adopted a task that encouraged observers to encode both the featural properties of an object, as well as the object's location. According to OFT, upon presentation of the spatial probe, observers should be able to use the probe position to access and retrieve any object properties on file at that location. In this sense, a spatial probe recognition task is ideally suited to evaluate the tenets of OFT. In contrast, predictions from OFT would be less clear in the case of a nonspatial recognition test (e.g., choosing, from among four alternatives, an object that appeared anywhere in the study scene). In the absence of a spatial probe, it would be difficult to know which of the existing object files was contributing information to the recognition judgment and how observers were accessing this information.

4. Note that our .25 estimate of chance in this 4AFC task assumes that neither the target nor any of the three lures were being remembered during test and that observers were truly guessing from among the four alternatives. Of course, if observers did remember one or more of the lures, estimates of chance would increase above .25, because observers would be able to exclude these lures from the recognition decision. However, one would have to assume a perfect memory for more than six (of the eight) nontargets per study display in order for exclusionary guessing to account for the .65 level of prerecency accuracy in the recognition test (Zelinsky & Loschky, 2003). Given that such a disproportionately high level of nontarget memory is unlikely, we believe that the prerecency level of accuracy obtained in Experiment 1 reflects memory for the target rather than guessing.

5. Note that fixations and fixated objects were highly correlated in our data ($r = .85$), meaning that the two data sets plotted in Figure 2A are not independent. For this reason, we could not use inferential statistics to test for significant differences between the intervening object and intervening fixation functions.

6. Our selection of the 500-msec and 1,000-msec presentation times used in Experiment 2 was intended to bracket the range of first-pass gaze durations on objects observed in Experiment 1.

7. Note that the hand movements associated with mouse usage occurred only at test, replacing the "eye cursor" used by observers to enter their selections in Experiment 1. Although it is reasonable to ask whether these different motor behaviors might differentially affect recency, we consider this an unlikely possibility. As argued by Lawrence, Myerson, Oonk, and Abrams (2001, p. 433), "all spatially directed movements appear to have similar effects on visuospatial working memory" regardless of whether they are performed by the eyes or hands (see also Hale et al., 1996). The present data would seem to support this claim. If memory was subject to motor interference in this study, it probably did not vary with the mode of response.

8. Because lengthening the retention interval also increases the opportunity for rehearsal, it is conceivable that a detrimental effect of decay might be offset by a rehearsal benefit, thereby resulting in no net difference between the 500–0 and 500–500 forgetting functions. We consider this possibility unlikely for two reasons. First, one would have to assume a very fortuitous set of conditions to achieve such a perfect offset between decay and rehearsal at each of our intervening object levels. Second, our 500-msec manipulation of retention interval offered only a minimal opportunity for rehearsal. Observers would likely be able to verbally recode and rehearse only one object, at most, during this interval (Zelinsky & Murphy, 2000), making it impossible for a rehearsal explanation to account for the similarity in accuracy over the range of intervening object levels reported in this study.

9. While the present study was under review, another study similar to our Experiment 2 came to press (Hollingworth, 2004). The author used a moving dot to impose a temporal sequence on the objects encoded from a scene. Consistent with our findings, the results revealed both recency and prerecency memory components and a lack of proactive interference from repeatedly presented objects and scenes. Note that the "follow the dot" paradigm, while important in offering converging evidence for serialization of object memory in scenes, differs from our Experiment 2 "one object at a time" paradigm in that all of the memory items remained visible during the dot sequence, thus leaving open the possibility that extrafoveal information influenced memory for the target object.