Exploring the memory advantage for moving scenes

William J. Matthews

University of Essex, UK

Luciano G. Buratto

Koen Lamberts

University of Warwick, UK

Running Head: Memory for moving scenes

Keywords: scene memory; moving scenes; dynamic superiority effect;

attention; image reversal; similarity

Author note

Correspondence address: William Matthews, Department of Psychology, University of Essex, Colchester, CO4 3SQ, United Kingdom, E-mail will@essex.ac.uk

Abstract

Recognition memory is better for moving images than for static images (the *dynamic superiority effect*), and performance is best when the mode of presentation at test matches that at study (the *study-test congruence effect*). We investigated the basis for these effects. In Experiment 1, dividing attention during encoding reduced overall performance but had little effect on the dynamic superiority or study-test congruence effects. In addition, these effects were not limited to scenes depicting faces. In Experiment 2, movement improved both old-new recognition and scene orientation judgments. In Experiment 3, movement improved the recognition of studied scenes but also increased the spurious recognition of novel scenes depicting the same people as studied scenes, suggesting that movement improves the identification of individual objects or actors without necessarily improving the retrieval of associated information. We discuss the theoretical implications of these results and highlight directions for future investigation.

Exploring the memory advantage for moving scenes

The visual world is intrinsically dynamic and our visual and mnemonic systems have developed to encode and retrieve information about a moving environment (Gibson, 1979). In recent years there has been a surge of interest in the effects of movement on memory for faces (Knight & Johnston, 1997; Lander & Bruce, 2003; Lander & Chuang, 2005; O'Toole, Roark, & Abdi, 2002; Pike, Kemp, Towell, & Phillips, 1997) and objects (Balas & Sinha, 2009; Koban & Cook, 2009; Stone, 1998, 1999; Vuong & Tarr, 2004). These studies have shown that dynamic information can be stored in long-term memory; people encode the temporal sequence of views of a moving stimulus - its "spatiotemporal signature" – so that the information extracted from dynamic stimuli is more than just a set of static views, particularly when the motion is non-rigid (Knight & Johnston, 1997; Lander & Chuang, 2005; Stone, 1999).

To date, there has been surprisingly little research on memory for moving scenes – natural images depicting complex agglomerations of objects and people. In an early study, Goldstein, Chance, Hoisington, and Buescher (1982) had participants study film clips or still images taken from those clips, and administered a recognition memory test a few minutes later. They reported a *dynamic superiority effect*: Recognition was better for moving images than for static images, although Goldstein et al.'s comparison was not well controlled as the durations of moving and static stimuli were not equal. More recently, Matthews, Benjamin, and Osborne (2007) presented participants with moving and static scenes of equal duration drawn from a wide variety of sources, and tested recognition memory at intervals ranging from 3 days to 1 month. Recognition of moving scenes was better than that of static scenes, and this moving advantage was independent of chromaticity. Matthews et al. also examined performance in a "multistatic" condition in which single frames drawn from regular intervals in the moving clips were presented in succession. Memory for these stimuli was the same as for single static images, suggesting that the moving advantage is not wholly due to there being more static views in the dynamic scenes: Fluid motion seems to be important (Pike et al., 1997).

Buratto, Matthews, and Lamberts (2009) recently extended the results of Matthews et al. (2007). In a 3x3 factorial design, they crossed mode of presentation at study (static, multistatic, moving) with mode of presentation at test (static, multistatic, moving). Overall recognition was higher for scenes that were moving in the study phase; the main effect of movement in the test phase was less clear. There was also some evidence that multistatic scenes were better recognized than single still images. More importantly, Buratto et al. found a marked study-test congruence effect: Recognition was best when the scene was presented in the same mode at study and test,

indicating that people are sensitive to the match between the spatiotemporal properties of a studied scene and the test item (See Lander & Davies, 2007, for a similar result with faces). Note that in the stimuli used by Buratto et al. the camera was always stationary; these were scenes depicting moving objects rather than self-motion through an environment (cf. DeLucia & Maldia, 2006). This type of stimulus forms the focus of the current paper, and for brevity we refer to such stimuli as "moving scenes" throughout.

The study of memory for moving scenes is in its infancy and many basic questions are as yet unanswered. It is therefore important to establish the effects of simple manipulations and boundary conditions to motivate and constrain theoretical development. In three experiments, we examined the factors that influence memory for dynamic natural scenes. In Experiment 1 we ask whether dividing attention during the study phase modulates the dynamic superiority and study-test congruency effects; we also determine whether these effects are content-specific. In Experiment 2 we use an image-reversal paradigm to probe memory for visual details and to establish whether the superior recognition of moving scenes requires a precise perceptual match between the probe stimulus and the memory trace. In Experiment 3 we vary the similarity of novel scenes to studied scenes to determine whether movement helps or hinders the rejection of similar lure stimuli.

Experiment 1

In Experiment 1 we ask whether the dynamic superiority and study-test congruence effects are affected by an important encoding manipulation: The completion of a secondary task that divides attention (Gardiner & Parkin, 1990). One suggestion is that moving images are better remembered because they are better at engaging attention or are more deeply processed than static images (Lander & Bruce, 2003). A strong version of this idea would hold that the dynamic superiority effect results from a type or depth of encoding that is only possible when the observer fully attends to the stimulus, in which case dividing attention will eliminate the memory advantage for moving scenes. A more likely expectation is that dividing attention will reduce the effect, producing an interaction between movement and attention such that the moving advantage is larger under full attention. It is therefore important to ask (a) whether the dynamic superiority effect requires full attention, and (b) whether the magnitude of the effect depends upon the division of attention. Examining performance on the secondary task provides an alternative way to attack the issue: If moving images capture more attention, performance on the secondary task should be worse during the viewing of the moving scenes.

The effect of attention on the study-test congruence effect is also important. Dividing attention may render observers less able to encode the perceptual details of the studied scenes, including their movement status, reducing the study-test congruency effect reported by Buratto et al. (2009). On the other hand, if the encoding of movement status is automatic and relatively effortless, dividing attention will not modulate the effect.

As an additional manipulation, we varied the content of the stimuli. Many of the scenes used by Matthews et al. (2007), and all of those used by Buratto et al. (2009), depicted people. As noted above, there has been considerable research into the role of movement in face recognition, and the question arises: Are the dynamic superiority and study-test congruence effects reported in our previous research specific to facial stimuli? To establish the generality of these effects, we tested participants with two types of stimuli: *face* clips, in which people's faces were clearly visible, and *hand* clips, which depicted hands performing some action with no faces visible.

Method

Participants. Sixty University of Warwick students (31 males; age: *M* = 20.6, *SD* = 3.2) participated in the study, 30 in each encoding condition (full vs. divided attention). Each was paid £10.

Materials. In all experiments, stimuli were film clips drawn from the same collection as used by Buratto et al. (2009). Clips were black and white, silent, with no cuts and no camera movement. All clips showed men and women from a broad range of ages, epochs and countries, talking, moving, and interacting with each other. In this experiment, the stimuli were 300 short video clips extracted from this set. The clips either showed people interacting with each other (e.g., a man handing a package to another man) or hands performing some action (e.g., a hand giving a package to another hand). In the *face* clips, there were always two people and their faces could be seen; in the *hand* clips, there were one or several hands manipulating objects, and faces could not be seen. Examples of the stimuli from all three experiments are available from the authors.

Each clip was presented for 3 s in one of two modes, *moving* or *static*. Moving clips contained 75 frames shown at 25 frames per second; static clips consisted of a single frame shown for 3 s. The static frames were carefully selected to provide a fair representation of the gist of the scene from which they were taken. All clips measured 560 x 420 pixels and were presented on a 19" monitor (1152 x 864 pixels) viewed from approximately 40 cm.

The video clips were divided into 12 sets of 25 clips (6 sets of face clips and 6 sets of hand clips). The sets were matched for average complexity and visual properties. Complexity was evaluated in terms of file size (*png* and *jpg* formats), compression ratio (*bmp* file size relative to *jpg*

size) and edge density (the percentage of edge pixels) (Rosenholtz, Li, & Nakano, 2007; Szekely et al., 2005). The sets of 25 clips were then randomly allocated to conditions and presentation modes.

Design and procedure. The experiment employed a 2 (clip type: faces vs. hands) × 2 (study mode: moving vs. static) × 2 (test mode: moving vs. static) × 2 (attention: full vs. divided) mixed design with clip type, study mode and test mode manipulated within participants and attention manipulated between participants.

Participants first took part in a study phase. They were presented with a sequence of 200 clips (50 moving and 50 static face clips, 50 moving and 50 static hand clips) in random order. In the *full attention* condition, participants were simply told to look at each clip and to respond "yes" on a keypad if they could see faces in the scene and to respond "no" otherwise. In the *divided attention* condition, participants looked at the clips and, at the same time, listened to tones played over headphones. They were to respond "yes" if two out of three 800-ms tones played during the presentation of a clip had the same frequency, and to respond "no" otherwise. The tones were presented at four frequencies: 360, 400, 440 and 480 Hz. Inter-tone intervals varied between 200–400 ms, so that each tone-monitoring trial lasted 3 s (i.e., the same duration as the clip). The first two tones always had different frequencies so that participants had to hold both tones in working memory in order to judge whether the third tone was equal to one of the previous two tones. The correct response was "yes" on half of the tone-monitoring trials.

In both encoding conditions, participants were only allowed to enter the yes/no responses after image offset so that they were exposed to moving and static pictures for the same amount of time (3 s). Participants were aware of the subsequent memory test and were told to be as accurate as possible on the tone-monitoring task while also trying to pay attention to the studied scenes. Participants were not told about the congruency manipulation (i.e., that a moving clip at study might be shown as a still frame at test or vice-versa).

Participants returned for the test phase 24 h after the study phase. The test phase comprised 300 trials: 200 *old* clips (which had been presented in the study phase) and 100 *new* clips (which had not been seen in the study phase). Of the 200 old clips, there were 100 face clips (50 moving and 50 static) and 100 hand clips (50 moving and 50 static). For both face and hand clips, 25 of the old clips were *congruent moving* clips (studied moving/tested moving), 25 were *congruent static* clips (studied static/tested static) and 50 were *incongruent* (25 moving/static, 25 static/moving). Of the 100 new clips, 50 were new face clips (25 moving, 25 static) and 50 were new hand clips (25 moving, 25 static). Allocation of face and hand clips to conditions (i.e., old/new, moving/static, congruent/incongruent) was randomised across participants.

Participants judged each test clip on a 6-point scale ranging from "definitely old" to "definitely new". It was emphasized that presentation mode was irrelevant for their decision: If, for example, they recognised a static frame from a moving study clip, they should respond "old". No feedback on responses was provided.

The responses on the 6-point scale were used to produce receiver operating characteristic (ROC) curves for each condition. A maximum-likelihood, unequal-variance, signal-detection model was then fitted to the ROC curves (RscorePlus: Harvey, 2001). The model parameters were used to estimate d_a, a multi-point discriminability measure that corresponds to the difference between the mean familiarity of targets and lures normalised by the average variance of target and lure distributions (MacMillan & Creelman, 2005). Unlike d' (the standard single-point sensitivity measure), d_a estimates do not assume equal variance of target and lure familiarity distributions. The equal-variance assumption is generally not borne out by empirical data (e.g., Mickes, Wixted, & Wais, 2007).

Results

Memory sensitivity (d_a). The sensitivity data are shown in Figure 1 and were subjected to a 2 (study mode: moving, static) × 2 (test mode: moving, static) × 2 (clip type: hands, faces) × 2 (attention: full, divided) mixed-design ANOVA. (Throughout this article, alpha = .05 and all tests are two-tailed). The results showed that dividing attention lowered performance, $F(1,58) = 10.37$, $MSE = 1.36$, $p = .002$ $\eta_p^2 = .15$, but there was no difference between clips depicting hands and clips depicting faces, $F < 1$, $p = .34$. Thus, the secondary task successfully diverted attention during encoding, and the face and hand clips were well-matched for overall memorability.

As in our previous work, memory was better for items that were moving in the study phase. $F(1,58) = 137.83$, $MSE = 0.10$, $p < .001$, $\eta_p^2 = .70$. However, there was no main effect of movement in the test phase, $F < 1$, $p = .57$. Buratto et al. (2009) found this pattern of results in their Experiment 2, but in their Experiment 1 there was a main effect of movement in both study and test phases. It is not clear why this difference arises. As in our previous work, there was a congruency effect – an interaction between study and test mode, $F(1,58) = 74.15$, $MSE = 0.09$, $p < .001$, $\eta_p^2 = .56$, indicating that sensitivity was highest when study and test modes matched. This effect is clearly visible in every panel of Figure 1.

Do the advantage for moving stimuli and the study-test congruency effect depend upon the presence of faces in the scene? Although there was a significant interaction between study mode and clip type, $F(1,58) = 7.15$, $MSE = 0.07$, $p = .01$, $\eta_p^2 = .11$, it was such that the advantage accrued

from studying moving pictures was higher for face clips than for hand clips [Faces: d_a(moving) = 1.07; d_a (static) = 0.80; Hands: d_a(moving) = 1.18; d_a (static) = 0.78]. Moreover, the clip content × study mode × test mode interaction was not significant, $F < 1$, $p = .39$, indicating that clip content made no difference to the study-test congruency effect.

Do the advantage for moving stimuli and study-test congruency effect depend upon full attention at encoding? The interaction between study mode and attention was not significant, but it approached significance, $F(1,58) = 3.61$, $MSE = 0.10$, $p = .06$, $\eta_p^2 = .06$, with the advantage for scenes that were moving in the study phase being slightly higher in the full attention condition than in the divided attention condition. However, when the data from just the divided attention condition were analysed, there remained a significant effect of study-phase movement, $F (1,29) = 72.68$, $MSE = .07$, $p < .001$, $\eta_p^2 = .72$. Thus, dividing attention may slightly reduce the advantage for stimuli which are moving in the study phase, but it does not eliminate the effect. Similarly, the study mode × test mode × attention interaction was not significant, $F(1,58) = 1.40$, $MSE = .09$, p = .241, $\eta_p^2 = .02$, indicating that the study-test congruence effect was unaffected by attention. None of the other effects were significant, Fs < 1.68, ps > .20.

Tone monitoring (divided attention condition). We also examined performance in the secondary tone-monitoring task used to divide attention. The proportion of correct responses in the tone-monitoring task did not differ between static pictures ($M = .87$, $SE = .02$) and moving pictures ($M = .86$, $SE = .02$), $F(1,29) = 1.03$, $MSE = 0.00$, $p = .32$, $\eta_p^2 = .02$. There was also no difference in accuracy between face ($M = .87$, $SE = 0.02$) and hand clips ($M = .86$, $SE = .02$), $F < 1$, $p = .42$, and no interaction between clip content and movement, $F(1, 29) = 3.05$, $MSE = 0.00$, $p = .09$, $\eta_p^2 = .10$. However, response times in the tone-monitoring task were affected by clip type. (Recall that participants only made a response once the last tone was played – that is, after the end of the clip.) Responses were faster after static clips ($M = 1007$ ms, $SE = 51$) than after moving clips ($M = 1085$ ms, $SE = 63$), $F(1,29) = 7.23$, $MSE = 25678$, $p = .01$, $\eta_p^2 = .20$, and were faster after faces ($M = 1019$ ms, $SE = 54$) than after hands ($M = 1073$ ms, $SE = 60$), $F(1,29) = 5.08$, $MSE = 16839$, $p = .03$, $\eta_p^2 = .15$. Clip content and movement did not interact, $F < 1$, $p = .57$.

The finding that responses in the secondary task were slower after viewing moving scenes suggests that the superior memory for these scenes may result from greater attentional engagement with their content. In other words, participants may have paid more attention to moving items in the study phase at the expense of slower RTs in the tone-monitoring task. However, such a trade-off is unlikely to account for the dynamic superiority and study-test congruence effects: Restricting our

analyses to participants whose accuracy in the tone-monitoring task was at least as high for moving stimuli as for static stimuli (N = 11) did not change the effects of movement on the old-new discrimination at test [i.e., there was a significant main effect of study mode, $F(1,10)$ = 15.76, MSE = 0.09, p = .003, η_p^2 = .61, a significant study mode × test mode interaction, $F(1,10)$ = 5.16, MSE = 0.08, p = .046, η_p^2 = .34, and no effect of test mode, $F(1,10)$ = 1.13, MSE = .18, p = .31].

Discussion

The results of Experiment 1 may be simply summarized. Scenes which were moving in the study phase were remembered better than scenes which were static, and performance was improved when the mode of presentation at test was the same as at study. These results were not dependent on the presence of a face in the scene; the dynamic superiority and study-test congruency effects reported by Matthews et al. (2007) and Buratto et al. (2009) are general phenomena which are not restricted to scenes containing moving faces. (Note that this experiment was not concerned with comparing the absolute memory for "hands" and "face" stimuli, only with showing that the same patterns of results hold for both types.) Finally, the results show that, although dividing attention at study lowers overall performance, it has little effect on either the superior memory for moving scenes or the study-test congruence effect. We discuss these results further below.

Experiment 2

In Experiment 2 we examined memory for the perceptual details of moving and static scenes by investigating the effects of image reversal on recognition. Changing the left-right orientation of a picture between study and test is a classic way to probe memory for perceptual details rather than the semantic meaning or "gist" of the scene (e.g., Dallett, Wilcox, & D'Andrea, 1968). We employed a paradigm in which participants study a set of pictures and are then shown test items comprising studied items, mirror-reversed copies of studied items, and completely novel images . Participants were told to respond "old" to pictures seen at study, regardless of orientation. For pictures identified as old, participants indicated whether the picture has the same orientation as during the study phase (Bartlett, Gernsbacher, & Till, 1987; Klatzky & Forrest, 1984). This approach allows us to answer two questions: (1) does the movement advantage found in previous experiments require a precise perceptual match between the probe and the memory trace, or does it extend to mirror-reversed copies of the studied scenes? And (2) does movement improve memory for the perceptual details of the scene (its left-right orientation), or only for the gist? In order to see whether these

effects changed over time, participants completed study sessions 7 days, 1 day and 1.5 hours prior to the single test session (Bartlett et al., 1987).

As an additional manipulation, we varied the movement congruence (i.e., the match between a scene's motion status at study and test.) Both our Experiment 1 and Buratto et al. (2009) demonstrated a congruency effect, and we sought to establish whether this effect varies with retention interval and image reversal. It may be that the dynamic features of moving scenes are lost over time, such that dynamic representations gradually reduce towards static representations (for a similar argument regarding image detail and colour representations see Homa & Viera, 1988; Wichmann, Sharpe, & Gegenfurtner, 2002). If this is the case, the memory advantage in the congruent moving condition (study moving / test moving) relative to the incongruent moving condition (study moving / test static) should gradually diminish with increasing retention interval.

Method

Participants. Thirty University of Warwick students (16 males; age: *M* = 21, *SD* = 2.6) participated in the study. Participants took part in all retention interval conditions (90 min, 1 day and 7 days) and were paid £15 each.

Materials. Stimuli consisted of 600 video clips from the same collection of films used in Experiment 1. For each clip, a mirror-reversed version was produced by reversing all 75 frames in the clip. Static scenes were produced by selecting a representative frame from each clip. The original 600 scenes were split into 30 groups of 20 scenes, matched for complexity and low-level visual properties. The scenes were then counterbalanced across participants using a Latin square, yielding 30 assignments of scene groups to conditions, one different assignment for each participant.

Design and Procedure. Study mode (moving vs. static), test mode (moving vs. static), retention interval (90 min vs. 1 day vs. 7 days) and scene type (same vs. reversed vs. unrelated lure) were all manipulated within participants. The experiment consisted of four sessions: Three study sessions and one test session. Participants returned six days after the first study phase for the second study phase, and again on the following day for the third study phase. Participants then returned 90 minutes after the third study phase to take part in the test phase.

In each study session, participants were presented with 160 clips (80 moving and 80 static, intermixed). For each clip, participants reported whether they could see a female in the scene by entering "yes" or "no" on a joypad (this search task was designed to ensure engagement with the stimuli). They were also informed about the subsequent memory test. Importantly, participants were not told about the left-right orientation manipulation (i.e., no attention was drawn to orientation during the study phase). The recognition memory test consisted of 600 trials: 240 targets

(120 moving, 120 static), 240 mirror-reversed lures (120 moving, 120 static) and 120 unrelated lures (60 moving, 60 static). One third of the targets had been studied 7 days before the test, one third had been studied 1 day before the test, and the last third were clips studied 90 minutes before the test. Similarly, one third of the mirror-reversed scenes had been studied in their original left-right orientation 7 days before the test, one third had been studied 1 day before the test, and the last third were studied 90 minutes before the test. Half of the studied clips were presented in the same mode at test (i.e., study moving/test moving or study static/test static) and half were presented in the other mode (i.e., study moving/test static or study static/test moving). There were 20 trials per study mode × test mode × retention interval × orientation combination.

At test, participants completed an old-new recognition task in which they decided whether each clip had been presented in the study phase. They were instructed to respond "old" if they recognised the clip from the study phase, even if it was a reversed version of the original. If they responded "new", the next test scene was presented. If they responded "old", they were asked to decide whether the scene was the same as the one previously studied, or whether it was a mirror-reversed version. Participants were told that half of the old clips were identical to the studied ones and half were mirror-reversed. Feedback on responses was not provided. Participants completed a short practice session to make sure they understood the instructions.

Results

We first describe the results of the old-new recognition task before going on to examine performance in the orientation judgment task.

Old-new recognition. We first examine false alarm rates, to check whether movement influences response bias. We then examine hit rates from all conditions. Finally, we restrict our attention to the congruent conditions, which will form the basis for our exploration of orientation judgment performance.

Responses to unrelated lures. The proportion of "old" responses given to unrelated lures (that is, the false alarm rate) provides a measure of response bias. A paired-sample *t*-test revealed that the proportion of false alarms for moving scenes (*M* = .26, *SE* = .03) was no different from that for static scenes (*M* = .25, *SE* = .02, t < *1, p*=.88), indicating no effect of movement on response bias.

Responses to target items. The proportions of "old" responses for each type of target stimulus are shown in Table 1, and were subjected to a 2 (study mode: moving vs. static) × 2 (test mode: moving vs. static) × 2 (orientation: same vs. reversed) × 3 (retention interval: 90 minutes, 1 day, 7 days) repeated measures ANOVA. The proportion of "old" responses was higher for items which were moving in the study phase, $F(1,29) = 54.29$, *MSE* = 0.02, $p < .001$, $\eta_p^2 = .65$, and for

scenes which were moving in the test phase, $F(1,29) = 4.67$, $MSE = 0.01$, $p = .04$, $\eta_p^2 = .14$. However,

as in Experiment 1, these main effects were modulated by a study × test interaction, $F(1,29) = 79.14$,

$MSE = 0.01$, $p < .001$, $\eta_p^2 = .73$, such that participants responded "old" to same and reversed scenes

more often in congruent than in incongruent conditions (moving/moving: $M = .65$, $SE = .03$;

moving/static: $M = .55$, $SE = .03$; static/moving: $M = .50$, $SE = .03$; static/static: $M = .56$, $SE = .02$).

The proportion of "old" responses was higher for same scenes than for reversed scenes,

$F(1,29) = 35.62$, $MSE = 0.01$, $p < .001$, $\eta_p^2 = .55$, and for items encountered more recently, $F(1.42,$

$41.22) = 11.27$, $MSE = 0.03$, $p = .001$, $\eta_p^2 = .28$, with more "old" responses to scenes studied 90 min

and 1 day before the test than to scenes studied 7 days before the test ($ps \leq .001$) but no difference

between the 90-min and 1-day retention intervals ($p = .10$; post-hoc Least Significant Difference

test). Retention interval and orientation did not interact with the other variables, Fs < 1, ps > .33,

and the four-way interaction was not significant $F < 1$, $p > .73$.

The fact that the study-test congruency effect was not modulated by retention interval (i.e.,

the null study × test × retention interaction) suggests that dynamic representations stored in

memory were not reduced to a common static representation over the course of 7 days. We

included the congruency manipulation to see whether the study-test interaction found in previous

experiments depends on retention interval. Having established that it does not (at least between 90

minutes and 7 days), we henceforth restrict our analyses to the congruent conditions (i.e., study

static-test static, and study moving-test moving). This not only simplifies matters, it also improves

comparability with Experiment 3 (where only congruent conditions are used) and with Bartlett et al.

(1987). First, we consider the proportion of "old" responses in the old-new recognition task (that is,

we repeat the foregoing analysis using only the data from the congruent conditions). We then go on

to consider the classification of items as "same" or "reversed".

Results for congruent conditions. We conducted a 2 (mode: moving, static) × 2 (orientation:

same, reversed) × 3 (retention interval: 90 min, 1 day, 7 days) repeated-measures ANOVA on the

proportion of "old" responses to congruent moving and congruent static scenes. The proportion of

"old" responses was higher for moving scenes than static scenes, $F(1,29) = 38.70$, $MSE = 0.02$, $p <$

$.001$, $\eta_p^2 = .57$, higher for scenes presented in the same orientation as at study than for reversed

scenes, $F(1,29) = 17.73$, $MSE = 0.01$, $p < .001$, $\eta_p^2 = .38$, and higher for scenes that were studied more

recently, $F(1.50, 43.39) = 7.34$, $MSE = 0.02$, $p = .001$, $\eta_p^2 = .20$ (specifically, there were significantly

more "old" responses to scenes studied 90 min and 1 day before the test than to scenes studied 7

days before the test, $ps < .01$, but no difference between the 90-min and 1-day scenes, $p = .35$; post-

hoc Least Significant Difference test). None of the interactions were significant (all Fs < 1, ps > .33). The results for the congruent condition thus replicate the results from the previous section.

Orientation judgments. We next looked at the proportion of correct "same" vs. "different" classifications, restricting our attention to congruent conditions. For compatibility with Bartlett et al. (1987), we considered both unconditioned responses (obtained by dividing the number of correct orientation judgments by the number of trials in the condition) and conditioned responses (obtained by dividing the number of correct orientation judgments by the number of correct "old" responses). The unconditioned results are shown in the left panel of Figure 2. A 2 (mode: moving, static) × 2 (orientation: same, reversed) × 3 (retention interval: 90 min, 1 day, 7 days) repeated-measures ANOVA showed that orientation judgments were more accurate for moving stimuli, $F(1,29) = 37.37$, $MSE = 0.01$, $p < .001$, $\eta_p^2 = .56$, and for stimuli encountered more recently, $F(1.52, 43.98) = 19.00$, $MSE = 0.03$, $p < .001$, $\eta_p^2 = .40$. There were more correct classifications of same scenes than reversed scenes, $F(1, 29) = 115.34$, MSE = 0.06, p < .001, $\eta_p^2 = .80$, mimicking the pattern seen in previous work (Klatzky & Forrest, 1984; McKelvie, 1983). Crucially, movement and orientation interacted, $F(1, 29) = 8.67$, MSE = 0.01, $p = .01$, $\eta_p^2 = .23$. As can be seen in Figure 2, movement produced a much bigger improvement in the orientation classification of same scenes than reversed scenes. No other interactions were significant: RI × movement $F(2, 58) = .40$, $MSE = 0.01$, $p = .67$, $\eta_p^2 = .01$; RI × orientation, $F(2,58) = 2.83$, $MSE = 0.11$, $p = .07$, $\eta_p^2 = .09$; RI × movement × orientation $F(2, 58) = 3.02$, $MSE = 0.01$, $p = .057$, $\eta_p^2 = .094$.

The conditioned results are shown in the right panel of Figure 2. The pattern is identical except that the main effect of movement just missed significance, $F(1,29) = 4.05$, $MSE = 0.02$, $p = .053$, $\eta_p^2 = .12$ [movement improved orientation judgments for same scenes, $F(1, 29) = 9.63$, $MSE = 0.02$, $p = .004$, $\eta_p^2 = .25$, but not reversed scenes, $F < 1$] and there was an orientation × retention interval interaction such that correct "reversed" responses were more negatively affected by retention interval than correct "same" responses, $F(2,58) = 7.34$, $MSE = 0.02$, $p = .001$, $\eta_p^2 = .20$.

Orientation bias. Finally, we asked whether participants preferred to respond "same" regardless of scene orientation by looking at participants' "same" / "reversed" responses to unrelated lures which had been erroneously identified as "old" in the old-new discrimination. That is, we split the false alarms in the old-new task by the response made in the subsequent same-reversed discrimination. For moving lures, the proportion of false "same" responses (*M* = .183, *SE* = .021) was higher than the proportion of false "reversed" responses (*M* = .073, *SE* = .021; summing up

to the overall false alarm probability of .256 for moving lures). Thus, participants chose to respond "same" more often than "reversed" when the test scenes were unrelated foils. Similarly, for static lures, the proportion of false "same" responses ($M = .166$, $SE = .018$) was higher than the proportion of false "reversed" responses ($M = .088$, $SE = .013$; summing up to .254). This replicates the "same"-"reversed" bias described by Bartlett et al. (1987). A 2 (orientation response) × 2 (mode) repeated measures ANOVA confirmed these trends: There was a main effect of orientation response, $F(1,29) = 26.07$, $MSE = 0.01$, $p < .001$, $\eta_p^2 = .47$, but no main effect of mode, $F < 1$, $p = .91$. (This latter result simply reiterates the finding of no net difference in false alarm rates for moving and static stimuli reported above). Importantly, orientation and mode did not interact, $F(1,29) = 3.53$, $MSE = 0.00$, $p = .07$, $\eta_p^2 = .11$; the bias to respond "same" was not affected by scene movement. Thus, the results of the preceding sections cannot be attributed to a greater bias to respond "same" or "reversed" to moving images. (See Bartlett et al., 1987, p. 32, for a fuller discussion of this issue.)

Discussion

Experiment 2 produced the following results. First, old-new recognition is improved by movement at study and test, and these main effects are modulated by an interaction such that performance is best when the study and test modes match. That is, we replicated the dynamic superiority and congruency effects reported previously. Second, old-new performance declines over time and is impaired by mirror-reversal of the stimuli, but these effects do not interact with the dynamic superiority and study-test congruency effects. Third, orientation judgments are more accurate for same scenes than for reversed scenes, are better for moving scenes than for static scenes, and are better for more recently encountered items. Fourth, movement improves the accuracy of same classifications more than reversed classifications. Fifth, when orientation judgment accuracy is made conditional on having given an "old" response, the effect of movement is diminished, and there is a retention interval × orientation interaction such that the effect of retention interval is greater for correct "reversed" responses. And sixth, these results are unlikely to be due to response bias: Movement does not affect the tendency to judge novel scenes "old", and although there is a bias towards responding "same orientation" to novel scenes, this is not significantly dependent on movement and is not a necessary prerequisite for the asymmetry in same/different classification accuracy (Bartlett et al., 1987).

The finding that correct "same" orientation judgments are more common than correct "reversed" responses, and that a manipulation which improves old-new recognition (in this case movement) produces a larger improvement in correct "same" classifications than correct "reversed" classifications, replicates several previous studies. Bartlett et al. (1987) provide a detailed framework

for understanding this pattern, and we do not consider it further. For our purposes, the key results are that movement improved the old-new recognition of both same and reversed scenes, and that movement improved overall orientation judgment accuracy. We discuss these results in more detail below.

Experiment 3

Experiment 2 found that movement improved the old-new recognition of scenes which were very similar to studied items (mirror-reversed copies). In Experiment 3 we examined the effects of target-lure similarity in more detail by choosing novel clips which shared differing amounts of content (people, settings, actions) with target items.

Method

Participants. One hundred and twenty University of Warwick students (50 males; age: M = 20.8, SD = 3.9) participated in the study, 40 in each retention interval condition (90 min, 1 day or 14 days). Each was paid £12.

Materials. The stimuli were 480 video clips drawn from drawn from the set used in Experiment 2. Of the 480 clips, a fixed group of 240 was presented at study. Half of the study clips were shown in the moving mode and half in the static mode. Moving and static clips were matched for average complexity and visual properties, and counterbalanced across participants. Another set of 120 clips was presented in the test phase as *similar* lures (60 clips) and *related* lures (60 clips). A clip was defined as *similar* if it depicted the same people and background as a clip presented in the study phase but differed in the action performed in the scene (e.g., a man in a room stirring tea with a spoon vs. the same man in the same room drinking the tea). A clip was defined as *related* if it featured some of the people present in a studied clip but differed in background and action (e.g., two people having dinner in a restaurant vs. one of them walking along a street). The remaining 120 clips (two from each film) were presented in the test phase as unrelated lures. A clip was *unrelated* if it depicted neither the same people nor the same background as a studied clip. (The terms *similar*, *related*, and *unrelated* are somewhat arbitrary, but the stimuli reflect a clear gradient of similarity between the lures and the studied item.) Of the 120 clips, 60 were presented in the moving mode and 60 in the static mode. They were matched for complexity, and the assignment of clips to modes was counterbalanced across participants.

Design and Procedure. Presentation mode (moving vs. static) and scene type at test (target vs. similar lure vs. related lure vs. unrelated lure) were manipulated within participants. Retention interval (90 min vs. 1 day vs. 14 days) was manipulated between participants.

In the study phase, participants were presented with 240 clips (120 moving and 120 static, intermixed). For each clip, participants decided whether there was a female present in the scene by entering a "yes" or "no" response on a joypad. They were also informed about the subsequent memory test.

Participants returned for the test phase 90 min, 1 day or 14 days after the study phase. The test phase consisted of 360 trials. One third of the test trials consisted of old clips (i.e., clips presented in the study phase in the same presentation mode: 60 moving and 60 static) and two thirds consisted of new clips (i.e., clips not presented in the study phase). Of the 240 new clips, 60 were similar, 60 were related and 120 were unrelated to studied items. Similar and related new clips were presented in the same mode as the corresponding studied clip (30 similar moving, 30 related moving, 30 similar static, 30 related static). Unrelated clips had no connection to studied clips and were thus randomly assigned to the moving (60 clips) or the static modes (60 clips). For each test clip, participants had to decide on a 6-point scale (from "definitely old" to "definitely new") whether or not the clip had been presented in the study phase.

Presentation modes were always matched at study and test (i.e., moving scenes at study were always tested with moving targets or their corresponding moving lures and static scenes at study were always tested with static targets or their corresponding static lures). Participants were told that only one third of the clips were old and that two thirds were new (either because they were similar to studied clips or because they were different from studied clips). Examples of similar, related, and unrelated clips were given at test to clarify the instructions. No feedback on responses was provided. Stimuli were yoked across retention intervals, so that each set of three participants saw exactly the same study and test stimuli in the same order [e.g, participants 1 (90-min interval condition), 41 (1-day condition) and 81 (14-day condition) saw the same clips in the same order]. Clip order and assignment of clips to presentation modes was counterbalanced across each group of three participants.

Results

The proportions of "old" responses to each type of stimulus are shown in Figure 3. The top panels show the results grouped by retention interval; the bottom panel shows the data collapsed over retention interval. We conducted a 2 (mode: moving, static) × 4 (scene type: target, similar, related, unrelated) × 3 (retention interval: 90 min, 1 day, 14 days) mixed-design ANOVA on the

proportion of "old" responses. The results showed that there were significantly more "old" responses to moving than to static scenes, $F(1,117) = 41.50$, $MSE = 0.01$, $p < .001$, $\eta_p^2 = .26$. There was also a main effect of scene type, $F(2.13, 248.78) = 848.93$, $MSE = 0.01$, $p < .001$, $\eta_p^2 = .88$, with more "old" responses to targets than to similar lures, more "old" responses to similar lures than to related lures, and more "old" responses to related lures than to unrelated lures (all Bonferroni-adjusted pairwise comparisons were significant, ps < .001).

The two-way interactions were all significant [mode × retention interval: $F(2,117) = 7.06$, $MSE = 0.01$, $p = .001$, $\eta_p^2 = .11$; scene type × retention interval: $F(4.25, 248.78) = 22.53$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = .28$; mode × scene type: $F(2.89, 337.74) = 11.70$, $MSE = 0.01$, $p < .001$, $\eta_p^2 = .09$]. There was, however, no main effect of retention interval, $F < 1$, $p = .46$, and no three-way interaction, $F < 1$, $p = .84$.

To deconstruct the significant two-way interactions, we conducted separate paired-sample t-tests or one-way ANOVAs on the proportion of "old" responses. First, we examined the mode × scene type interaction by comparing moving and static conditions separately for each of the four scene types. There were significantly more "old" responses to moving than to static scenes for target, similar and related scenes, ts(119) > 2.91, ps < .01, but not for unrelated scenes, $t(119) = .96$, $p = .34$. These results, illustrated in Figure 3, indicate that similar and related lures were incorrectly recognised as "old" to a larger extent when the scenes were moving than when they were static. Importantly, the fact that false alarms to unrelated lures did not differ between moving and static scenes suggests that there was no overall bias to say "old" more often to moving than to static scenes. We then dissected the mode × retention interval interaction by comparing moving and static conditions for each retention interval. There were significantly more "old" responses to moving than to static scenes 1 day and 14 days after study, ts(39) > 2.86, ps < .01, but not after 90 minutes, $t(39) = 1.62$, $p = .11$.

Finally, we examined the scene type × retention interval interaction with four one-way ANOVAs to determine the effect of retention interval for each scene type. Increasing the retention interval lowered the hit rate and increased false alarms to unrelated lures, but did not affect false alarms to similar and related lures. [For hits, $F(2,117) = 19.49$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = .25$; for unrelated lures, $F(2,117) = 4.00$, $MSE = 0.01$, $p = .02$, $\eta_p^2 = .06$; for similar lures: $F(2,117) = 1.28$, $MSE = 0.02$, $p = .28$, $\eta_p^2 = .02$; for related lures: $F(2,117) = 1.93$, $MSE = 0.02$, $p = .15$, $\eta_p^2 = .03$.]

As an additional analysis, we assessed the influence of lure similarity on the ability to discriminate targets from lures by calculating separate d_a estimates for each lure type (i.e., target vs.

similar lure, target vs. related lure and target vs. unrelated lure). For the sake of brevity we do not report the full results here, but they were as one would expect from the preceding analyses: Target-lure discrimination was better for moving scenes than static scenes, but this superiority was reduced when the lures were similar to the targets.

Discussion

The key result of Experiment 3 is that although movement improved the successful recognition of studied scenes, it also increased the false alarm rate for novel scenes which were similar to studied items. This is not due to response bias: As in our previous studies, the false alarm rates for completely novel scenes were the same for moving and static images.

General Discussion

These experiments provide considerable information about the representation and recognition of moving scenes. In Experiment 1 we found that the memory advantage for moving scenes and the study-test congruence effects reported by Matthews et al. (2007) and Buratto et al. (2009) are not limited to scenes depicting faces. More importantly, dividing attention during encoding lowered overall performance but had little effect upon the dynamic superiority or study-test congruence effects. In Experiment 2 movement improved both old-new and scene orientation judgments. This experiment also showed that the dynamic superiority and study-test congruency effects were unaffected by retention interval. Finally, in Experiment 3 movement improved the successful recognition of studied scenes but also increased the spurious recognition of novel scenes which were similar to studied items. This was not due to a response bias: The false alarm rate for completely novel scenes did not differ between moving and static images in any of the experiments.

What do these results tell us about the representation and recognition of moving and static scenes? In what follows, we discuss the implications of our findings, focussing on the general conclusions that can be drawn and on the constraints that they place on accounts of scene memory. We then go on to identify key directions for future research.

Attention, content, and time

The fact that the memory advantage for moving scenes is unaffected by dividing attention during encoding argues against a strong version of the hypothesis that the moving advantage reflects greater engagement with moving stimuli (Lander & Bruce, 2003). The same conclusion is urged by the fact that even those participants for whom tone-monitoring was better during the

static stimuli showed the dynamic superiority effect. Moreover, if moving images elicit greater attentional engagement, accuracy on the secondary task should have been worse in the moving condition, but this was not found (although there was some effect of movement on response time after stimulus offset, perhaps indicating that participants took longer to disengage from watching a moving clip in order to press the response key). This is not to say that there is no differential attention to moving stimuli, only that this is unlikely to be the sole difference in how the stimuli are processed and encoded. In addition, the fact that dividing attention made no difference to the study-test congruency effect suggests that participants need not pay full attention to an image in order to encode its movement status. This complements the finding by Buratto et al. (2009) that intentional encoding is not a prerequisite for the study-test congruence effect; it seems that the movement status of an image is encoded automatically, without much effort.

Experiment 1 also established that the dynamic superiority and study-test congruence effects are not limited to scenes depicting faces. It seems likely that the kinds of "characteristic motion" used to recognize moving faces (Lander & Davies, 2007) apply to other classes of stimuli. In addition, the fact that the dynamic superiority and study-test congruence effects did not interact with retention interval in Experiment 2 argues against the idea that the dynamic features of moving scenes are lost over time, with dynamic representations gradually reducing towards static representations (see also Homa & Viera, 1988; Wichmann et al., 2002). Information about the movement status of a studied scene remains apparently undiminished for intervals ranging from 90 minutes to 7 days.

The moving advantage is robust, but movement does not always lead to more accurate memory. For static images, people typically mis-remember scenes as more extended than they really were , a phenomenon known as boundary extension (Intraub & Richardson, 1989). DeLucia and Maldia (2006) found no difference in boundary extension between scenes which were moving and static at study: Movement did not improve accuracy. There were many procedural differences between this work and our own – for example, the stimuli depicted self-motion in depth rather than moving objects with a fixed camera position – but it illustrates that stimulus motion does not improve memory for every aspect of a scene. This point is echoed in Experiment 3 of the current study.

The representation of moving and static scenes: What is remembered?

Historically, a distinction has been drawn between memory for the "gist" of a scene – loosely, its semantic meaning – and memory for the perceptual details (e.g., Brainerd, Reyna, & Mojardin, 1999). In Experiment 2, movement improved same-reversed orientation judgments,

indicating that stimulus motion improves memory for perceptual details. Movement also improved the old-new recognition of reversed scenes. One interpretation of this latter result is that movement enhances memory for orientation-independent features such as gist or semantic meaning, but it might equally be that the observer has a more complete perceptual record and that the memory system is flexible enough to detect the correspondence between reflected copies (Standing, Conezio, & Haber, 1970). In either case, Experiment 2 shows that movement improves memory for orientation, and that the movement recognition-boost does not require precise perceptual overlap between the probe and the memory record.

In Experiment 3, movement increased successful recognition of studied scenes but also boosted false alarms to scenes which depicted the same actor as a studied scene performing a different action, either against the same backdrop (similar lures) or in a different setting (related lures). How are we to explain this result? The increased false alarm rate means that movement has improved the recognition of the familiar components of the scene. This follows because there was no effect of movement on the tendency to respond "old" to completely novel scenes: The increased false alarms to similar lures must result from better recognition of the familiar components of the scene (the central actor or actors). However, this improvement is not matched by a corresponding enhancement of memory for precisely where the familiar person was seen before and what they were doing. This follows because participants were instructed to respond "old" only when probe stimuli exactly matched items from the study phase. To take a concrete example, suppose the participant studies a scene in which a woman drinks tea in a restaurant, and is probed with a scene showing the same woman walking along the street. The increased false alarm rate in the moving condition implies greater recognition of the woman without a concomitant retrieval of the details of where or when she was seen before (or, indeed, of the action she was performing).

One way to conceptualize this finding is that movement differentially affects item recognition (in this case, the person in the studied scene) and associative recognition (in this case, the links between the actor, his or her actions, and the background context) (see, for example, Yonelinas, 2002). A related conceptualization emphasizes the use of a *recall to reject* strategy. When people discriminate studied word-pairs from novel pairs in which the words are presented in new associations, the probe cues recall of the original pairing which in turn allows rejection of the test pair as novel (Rotello & Heit, 2000). That is, recall of the original association leads to confident rejection of the lure, even if its component words are highly familiar. The results of Experiment 3 indicate that movement impairs the use of recall to reject novel scenes which contain familiar elements. This suggestion generates tangible predictions regarding the receiver operating

characteristic (Rotello, MacMillan, & Van Tassel, 2000) and time-course of recognition (Rotello & Heit, 2000).

There is superficial tension between the image-reversal results of Experiment 2 (which indicate that movement improves memory for perceptual details) and the increased false alarm rate in Experiment 3 (which indicates a failure to retrieve the details of the encoding event). However, the type of detail memory probed in Experiment 2 (left-right orientation) is different from that required to reject similar lures in Experiment 3. Moreover, the key result from Experiment 3 is not that movement has impaired memory for visual details: it is that the increased tendency to judge a whole moving scene as "old" on the basis of the familiar components is not offset by a corresponding increase in the ability to recall precise information about where those familiar components were seen before. Movement may well have produced superior memory for some details of the studied scenes (including orientation) in Experiment 3, but the overall impression that a similar lure had been encountered in the study phase outweighed this.

These results have potentially important practical implications. Legal psychologists have been interested in the identification of individuals from video footage (Bruce et al., 1999) and how doctored film evidence can influence eyewitness testimony (Wade, Green, & Nash, in press). Our results suggest that movement will lead witnesses to show superior recognition of replayed crime scenes, but that they may falsely recognize novel scenes that depict the same actors. One possibility, with marked implications, is that movement improves people's ability to recognize an individual without necessarily remembering where the person was seen or what they were doing at the time. It is important to note that in the moving condition of Experiment 3, the actor depicted in both the studied scene and the similar/related lure scenes was always moving at both study and test. Moreover, participants only ever judged whether they recognized the whole scene, rather than judging the old/new status of specific scene components. The results of systematically varying both of these factors (see below) will be an important determinant of the real-world implications of our results.

Future directions

Rather than speculate further, we conclude by identifying three key questions that we believe should be addressed in future research.

Which aspects of moving scenes are remembered? An important progression from the current work will be the investigation of memory for individual components of dynamic scenes. One possibility is that moving objects are prioritized for fixation and encoding (Abrams & Christ, 2006; Brockmole & Henderson, 2005; Matsukura, Brockmole, & Henderson, 2009), leading to better

recognition of those items at the expense of the static components or background of the scene. A simple starting point would be the comparison of memory for foreground objects and background context from moving or static versions of the same scenes (Kensinger, Garoff-Eaton, & Schacter, 2007). A related question concerns how movement itself is remembered: How far do people remember the perceptual details of object motion, and how far do they rely on the "gist" of the action? Finally, some previous studies have examined the effects of movement on the extraction of spatial layout (e.g., Garsoffky, Schwan, & Hesse, 2002; Garsoffky, Schwan, & Huff, 2009; Munger et al., 2006), and it will be important to integrate these results with work on recognition of individual scene components.

Where do people look in moving scenes? For static scenes, fixations on individual objects produce detailed representations which are indexed to particular spatial locations in long-term memory (Hollingworth & Henderson, 2002; Irwin, 1992), providing information which can support accurate recognition even after lengthy retention intervals (Hollingworth, 2005). A straightforward extension is that detailed object representations are indexed to locations which change over time (Kahneman, Treisman, & Gibbs, 1992; Matthews et al., 2007). Determining how the visual system tracks and records these spatiotemporal patterns will require the application of eye-tracking techniques to dynamic stimuli (e.g., Cristino & Baddeley, 2009; Hirose, Kennedy, & Tatler, 2010).

Which processes underlie recognition? Many researchers argue that recognition judgments are based on two distinct processes: A general feeling of stimulus familiarity, or the recollection of specific details of the encoding episode (Atkinson & Juola, 1973; Jacoby, 1991; Mandler, 1980; Yonelinas, 2002). There are tools for evaluating the contribution of recollection and familiarity which could be used to determine whether the phenomenology of recognition is different for moving and static scenes (e.g., Tulving, 1985). A related question concerns the time course of recognition (e.g., Boldini, Russo, Punia, & Avons, 2007). We required participants to view an entire clip before responding, but it will be important to establish how quickly moving scenes (or scene components) are recognized, and how this compares with performance for static stimuli (Pike et al., 1997). Similarly, deeper understanding of the mental representation of dynamic scenes will require investigation of how movement parameters influence the memory advantage. For example, how do optical velocity, number of moving objects, and presence or absence of self-motion (DeLucia & Maldia, 2006) influence the encoding of dynamic scenes?

Conclusion

The dynamic superiority and study-test congruence effects are robust phenomena which are unaffected by large changes in retention interval, are not specific to face stimuli, and do not require

full attention during encoding. Movement improves recognition of mirror-reversed copies of studied scenes, and also improves the ability to detect image reversal. Finally, movement boosts the false alarm rate for scenes which are similar to studied items, consistent with an increase in memory for item information (for example, the people present in a scene) without a corresponding increase in memory for associative information (their background environment and the actions they performed). We hope that these results will motivate and constrain future study of memory for dynamic natural scenes.

References

Abrams, R. A., & Christ, S. E. (2006). Motion onset captures attention: A rejoinder to Franconeri and Simons (2005). *Perception & Psychophysics, 68*(6), 114-117.

Atkinson, R. C., & Juola, J. F. (1973). Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 583-612). New York: Academic Press.

Balas, B., & Sinha, P. (2009). A speed-dependent inversion effect in dynamic object matching. *Journal of Vision, 9*(2), 1-13.

Bartlett, J. C., Gernsbacher, M. A., & Till, R. E. (1987). Remembering left-right orientation of pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*(1), 27-35.

Boldini, A., Russo, R., Punia, S., & Avons, S. E. (2007). Reversin the picture superiority effect: A speed-accuracy trade-off study of recognition memory. *Memory & Cognition, 35*(1), 113-123.

Brainerd, C. J., Reyna, V. F., & Mojardin, A. H. (1999). Conjoint recognition. *Psychological Review, 106*(1), 160-179.

Brockmole, J. R., & Henderson, J. M. (2005). Prioritization of new objects in real-world scenes: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 31*(5), 857-868.

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P. J. B., Burton, A. M., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied, 5*(4), 339-360.

Buratto, L. G., Matthews, W. J., & Lamberts, K. (2009). When are moving images remembered better? Study–test congruence and the dynamic superiority effect. *Quarterly Journal of Experimental Psychology, 62*(10), 1896-1903.

Cristino, F., & Baddeley, R. (2009). The nature of visual representations involved in eye movements when walking down the street. *Visual Cognition, 17*(6/7), 880-903.

Dallett, K., Wilcox, S. G., & D'Andrea, L. (1968). Picture memory experiments. *Journal of Experimental Psychology, 76*(2), 312-320.

DeLucia, P. R., & Maldia, M. M. (2006). Visual memory for moving scenes. *Quarterly Journal of Experimental Psychology, 59*(2), 340-360.

Gardiner, J. M., & Parkin, A. J. (1990). Attention and recollective experience in recognition memory. *Memory & Cognition, 18*(6), 579-583.

Garsoffky, B., Schwan, S., & Hesse, F. W. (2002). Viewpoint dependency in the recognition of dynamic scenes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(6), 1035-1050.

Garsoffky, B., Schwan, S., & Huff, M. (2009). Canonical views of dynamic scenes. *Journal of Experimental Psychology: Human Perception and Performance, 35*(1), 17-27.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton-Mifflin.

Goldstein, A. G., Chance, J. E., Hoisington, M., & Buescher, K. (1982). Recognition memory for pictures: Dynamic vs. static stimuli. *Bulletin of the Psychonomic Society, 20*(1), 37-40.

Harvey, L. O., Jr. (2001). *Parameter estimation of signal detection models: RscorePlus user's manual.* Boulder, CO: Author.

Hirose, Y., Kennedy, A., & Tatler, B. W. (2010). Perception and memory across viewpoint changes in moving images. *Journal of Vision, 10*(4), 1-19.

Hollingworth, A. (2005). The relationship between online visual representation of a scene and long-term scene memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(3), 396-411.

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance, 28*(3), 113-136.

Homa, D., & Viera, C. (1988). Long-term memory for pictures under conditions of thematically related foils. *Memory & Cognition, 16*(5), 411-421.

Intraub, H., & Richardson, M. (1989). Wide-angle memories of close-up scenes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(2), 179-187.

Irwin, D. E. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(2), 307-317.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*(5), 513-541.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*(2), 175-219.

Kensinger, E. A., Garoff-Eaton, R. J., & Schacter, D. L. (2007). Effects of emotion on memory specificity: Memory trade-offs elicited by negative visually arousing stimuli. *Journal of Memory and Language, 56*(4), 575-591.

Klatzky, R. L., & Forrest, F. H. (1984). Recognizing familiar and unfamiliar faces. *Memory & Cognition, 12*(1), 60-70.

Knight, B., & Johnston, A. (1997). The role of movement in face recognition. *Visual Cognition, 4*(3), 265-273.

Koban, A., & Cook, R. (2009). Rotational Object Discrimination by Pigeons. *Journal of Experimental Psychology-Animal Behavior Processes, 35*(2), 250-265.

Lander, K., & Bruce, V. (2003). The role of motion in learning new faces. *Visual Cognition, 10*(8), 897-912.

Lander, K., & Chuang, L. (2005). Why are moving faces easier to recognize? *Visual Cognition, 12*(3), 429-442.

Lander, K., & Davies, R. (2007). Exploring the role of characteristic motion when learning new faces. *Quarterly Journal of Experimental Psychology, 60*(4), 519-526.

MacMillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide* (2nd ed.). Mahwah, N.J.: Erlbaum.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87*(3), 252-271.

Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology, 57*(3), 203-220.

Matsukura, M., Brockmole, J. R., & Henderson, J. M. (2009). Overt attentional prioritization of new objects and feature changes during real-world scene viewing. *Visual Cognition, 17*(6/7), 835-855.

Matthews, W. J., Benjamin, C., & Osborne, C. (2007). Memory for moving and static images. *Psychonomic Bulletin & Review, 14*(5), 989-993.

McKelvie, S. J. (1983). Effects of lateral reversal on recognition memory for photographs of faces. *British Journal of Psychology, 74*(3), 391-407.

Mickes, L., Wixted, J. T., & Wais, P. E. (2007). A direct test of the unequal-variance signal detection model of recognition memory. *Psychonomic Bulletin & Review, 14*(5), 858-865.

Munger, M. P., Dellinger, M. C., Lloyd, T. G., Johnson-Reid, K., Tonelli, N. J., Wolf, K., et al. (2006). Representation momentum in scenes: Learning spatial layout. *Memory & Cognition, 34*(7), 1557-1568.

O'Toole, A. J., Roark, D. A., & Abdi, H. (2002). Recognizing moving faces: a psychological and neural synthesis. *Trends in Cognitive Sciences, 6*(6), 261-266.

Pike, G. E., Kemp, R. I., Towell, N. A., & Phillips, K. C. (1997). Recognizing moving faces: The relative contribution of motion and perspective view information. *Visual Cognition, 4*(4), 409-437.

Rosenholtz, R., Li, Y., & Nakano, L. (2007). Measuring visual clutter. *Journal of Vision, 7*(2), 1-22.

Rotello, C. M., & Heit, E. (2000). Associative recognition: A case of recall-to-reject processing. *Memory & Cognition, 28*(6), 907-922.

Rotello, C. M., MacMillan, N. A., & Van Tassel, G. (2000). Recall-to-reject in recognition: Evidence from ROC curves. *Journal of Memory and Language, 43*(1), 67-88.

Standing, L., Conezio, J., & Haber, R. N. (1970). Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. *Psychonomic Science, 19*(2), 73-74.

Stone, J. V. (1998). Object recognition using spatiotemporal signatures. *Vision Research, 38*(7), 947-951.

Stone, J. V. (1999). Object recognition: view specificity and motion specificity. *Vision Research, 39*(24), 4032-4044.

Szekely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., et al. (2005). Timed action and object naming. *Cortex, 41*, 7-25.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology, 26*(1), 1-12.

Vuong, Q. C., & Tarr, M. J. (2004). Rotation direction affects object recognition. *Vision Research, 44*(14), 1717-1730.

Wade, K. A., Green, S. L., & Nash, R. A. (in press). Can fabricated evidence induce false eyewitness testimony? *Applied Cognitive Psychology*.

Wichmann, F. A., Sharpe, L. T., & Gegenfurtner, K. R. (2002). The contributions of color to recognition memory for natural scenes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(3), 509-520.

Yonelinas, A. P. (2002). The nature of recollection and familiarity. *Journal of Memory and Language, 46*(3), 441-517.

Table 1. Proportion of "old" responses for each condition of the old-new recognition task of Experiment 2.

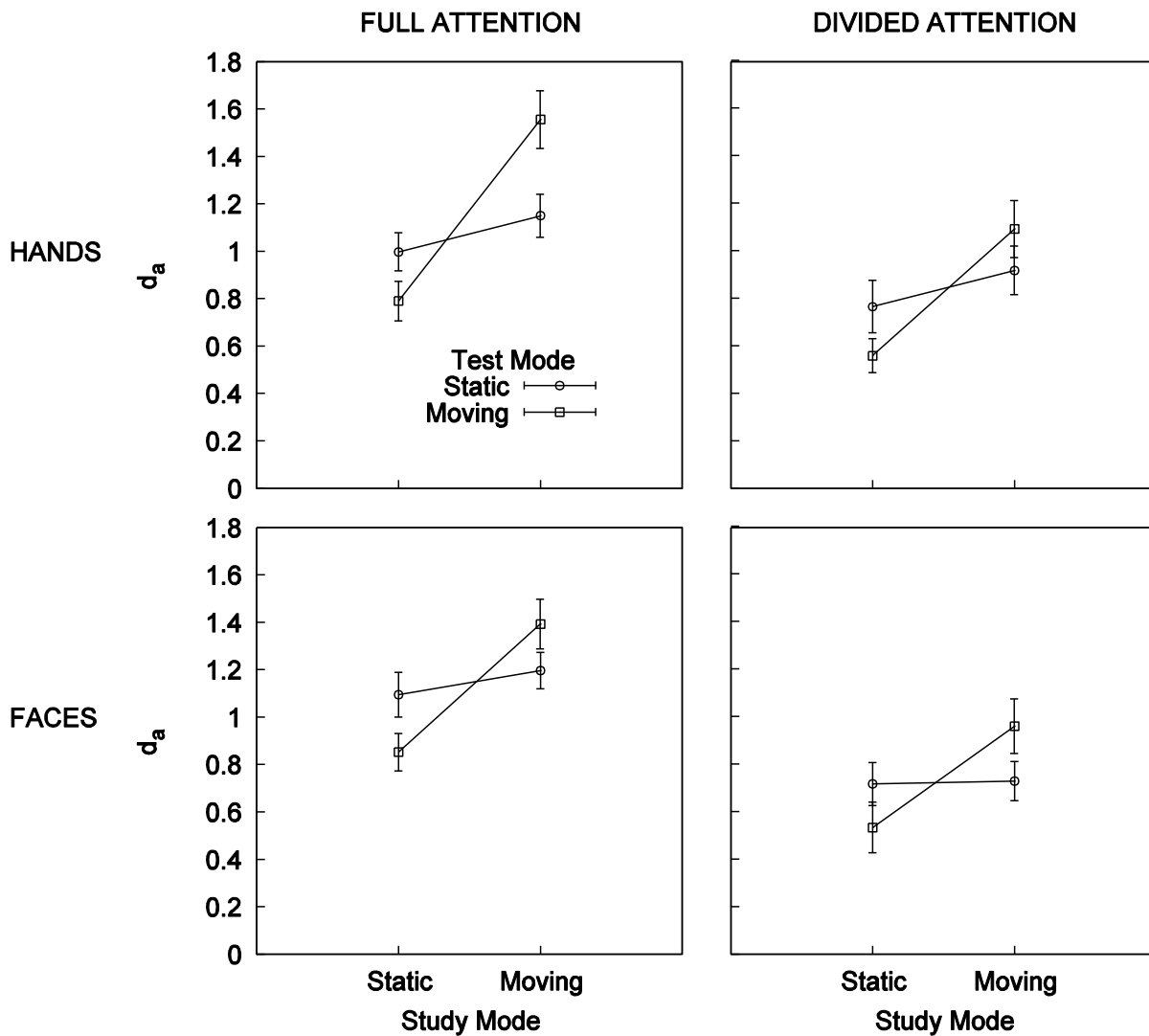| | | Retention Interval | | | | | |
		90 min		1 day		7 days	
	Study/Test Mode	*M*	*SE*	*M*	*SE*	*M*	*SE*
Same scenes	Static/Static	.61	.04	.61	.04	.54	.03
	Static/Moving	.55	.04	.54	.03	.51	.04
	Moving/Static	.62	.04	.58	.03	.51	.03
	Moving/Moving	.70	.04	.66	.04	.64	.03
Reversed scenes	Static/Static	.56	.04	.55	.04	.50	.03
	Static/Moving	.51	.04	.47	.04	.42	.03
	Moving/Static	.58	.04	.53	.04	.49	.03
	Moving/Moving	.65	.05	.65	.03	.58	.03

Figure 1. Results of Experiment 1, comparing performance for clips depicting hands (top panels) and human faces (bottom panels) under conditions of full attention (left panels) and divided attention (right panels) during the encoding phase. The overall advantage for stimuli that were moving at study, and the study-test congruence effect, are both apparent in all conditions. Here and throughout, error bars indicate plus or minus one standard error. Note that these provide no indication of the statistical significance of differences between means for within-subject designs (Masson & Loftus, 2003).
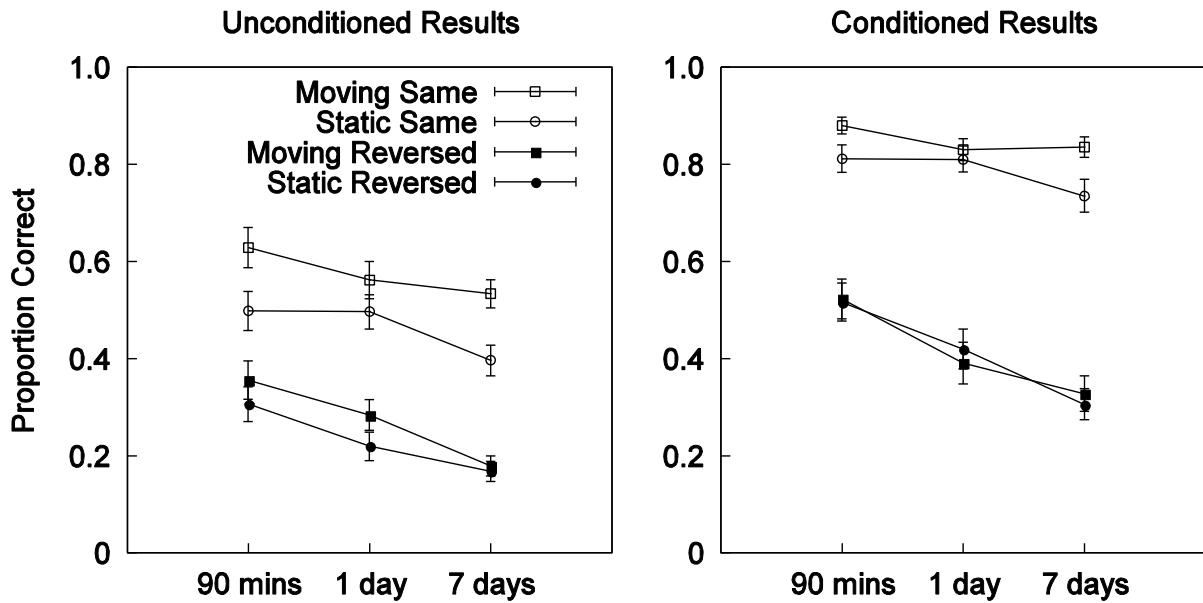
Figure 2. Results from the orientation judgment task of Experiment 2, using only the data from the congruent conditions. The left panel shows the unconditioned responses (that is, the proportion of all trials in a given condition for which a correct orientation judgment was made). Movement improved classification accuracy much more for same scenes than for reversed scenes, consistent with an increase in familiarity. The right panel shows the data conditional upon successful recognition of a studied scene as "old" (that is, the plot shows the proportion of stimuli judged "old" for which a correct orientation judgment was made).
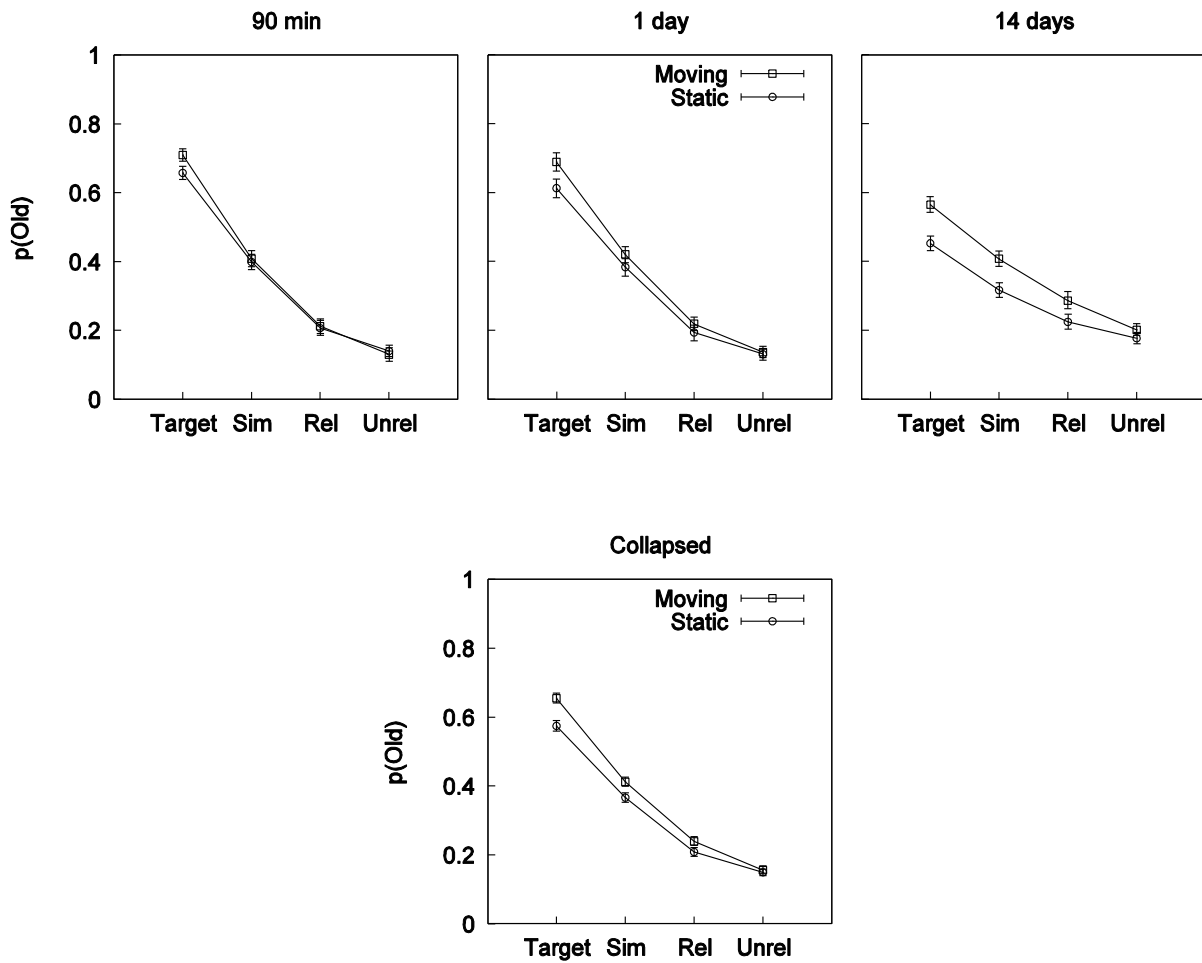
Figure 3. Results from Experiment 3. The top panels show the proportion of "old" responses to each type of stimulus at each retention interval. The bottom panel shows the results collapsed across retention interval. Targets were unchanged copies of scenes presented in the study phase; similar lures (Sim) depicted the same people and background as a studied scene, but with a different action being performed (and, correspondingly, an altered spatial layout); related lures (Rel) depicted one or more of the same people as a target, but against a different background; unrelated lures (Unrel) had neither background nor actors in common with any of the studied scenes.