Binding actions and scenes in visual long-term memory

Zhisen Jiang Urgolites & Justin N. Wood

Springer

BRIEF REPORT

Binding actions and scenes in visual long-term memory

Zhisen Jiang Urgolites · Justin N. Wood

Abstract How does visual long-term memory store representations of different entities (e.g., objects, actions, and scenes) that are present in the same visual event? Are the different entities stored as an integrated representation in memory, or are they stored separately? To address this question, we asked observers to view a large number of events; in each event, an action was performed within a scene. Afterward, the participants were shown pairs of action–scene sets and indicated which of the two they had seen. When the task required recognizing the individual actions and scenes, performance was high (80 %). Conversely, when the task required remembering which actions had occurred within which scenes, performance was significantly lower (59 %). We observed this dissociation between memory for individual entities and memory for entity bindings across multiple testing conditions and presentation durations. These experiments indicate that visual long-term memory stores information about actions and information about scenes separately from one another, even when an action and scene were observed together in the same visual event. These findings also highlight an important limitation of human memory: Situations that require remembering actions and scenes as integrated events (e.g., eyewitness testimony) may be particularly vulnerable to memory errors.

Keywords Visual long-term memory · Action memory · Scene memory · Event memory · Binding

Using visual information to guide behavior requires a long-term memory mechanism, known as *visual long-term memory* (VLTM), that maintains the information over time. Previous studies have shown that VLTM has a large and detailed storage capacity for the visual features of objects (Brady, Konkle, Alvarez, & Oliva, 2008; Standing, 1973), actions (Urgolites & Wood, 2013), and scenes (Konkle, Brady, Alvarez, & Oliva, 2010). To date, however, little is known about how VLTM stores these different entities when building a representation of a larger visual event. Are objects, actions, and scenes that are observed together in the same visual event stored as an integrated representation in VLTM? Or does long-term memory store entities from the same visual event separately from one another?

Behavioral and neurophysiological studies have provided suggestive evidence that VLTM maintains different visual entities separately from one another. For example, on the behavioral level, studies of eyewitness testimony have shown that visual memories are susceptible to agent–action binding errors, in which observers mistakenly remember an agent and action as having been part of the same event, when in reality they were seen in different events (Earles, Kersten, Curtayne, & Perle, 2008; Kersten & Earles, 2010; Kersten, Earles, Curtayne, & Lane, 2008; Loftus, 1976; Perfect & Harris, 2003; Ross, Ceci, Dunning, & Toglia, 1994). Similarly, unconscious transference errors—in which someone seen in a noncriminal context is mistakenly identified as a perpetrator in a criminal context—occur frequently in eyewitness testimony. For example, mere exposure to mug shots of a suspect significantly increases the likelihood that in a later lineup, the witness will identify that suspect as a perpetrator (e.g., Brown, Deffenbacher, & Sturgill, 1977; Deffenbacher, Bornstein, & Penrod, 2006; Deffenbacher, Carr, & Leu, 1981; Kassin, Tubb, Hosch, & Memon, 2001; Perfect & Harris, 2003), and the presence of a familiar but innocent bystander in a lineup significantly increases the rate of false identification (Brown et al., 1977; Buckhout, 1974; Loftus, 1976). These studies have shown that in many situations, we remember people, actions, and scenes, but fail to accurately bind these different entities into integrated representations in long-term memory. More generally, these results suggest that objects, actions, and scenes that are present in the same visual event are stored separately from one another in VLTM.

Neurophysiological studies have provided further support for this hypothesis, with the discovery that visual cognition

Z. J. Urgolites (✉) · J. N. Wood (✉)
Department of Psychology, University of Southern California,
3620 South McClintock Avenue, Los Angeles, CA 90089, USA
e-mail: zhisenji@usc.edu
e-mail: justin.wood@usc.edu

depends on at least four dissociable cortical pathways (reviewed by Kravitz, Saleem, Baker, & Mishkin, 2011). Specifically, the ventral pathway supports object recognition, and the dorsal pathway contains three more specialized pathways, including one with projections to the prefrontal cortex that supports the ability to track the movements and locations of objects, a second with projections to the medial temporal lobe that maintains scene-based information for visual navigation, and a third with projections to the premotor cortex that supports visually guided action. Together, these findings provide evidence for separate cortical pathways for processing information about objects, actions, and scenes. Since long-term memories appear to be incorporated within the information-processing functions of specialized brain areas (e.g., Eichenbaum & Cohen, 2001; McTighe, Cowell, Winters, Bussey, & Saksida, 2010), memories of objects, actions, and scenes are likely stored in at least partially distinct parts of the cortex. If this is so, some additional process would be needed to bind information about objects, actions, and scenes stored in distinct cortical structures.

In the present study, we provide a direct empirical test of whether VLTM maintains actions and scenes separately from one another. To do so, we used the two-alternative forced choice method that has previously revealed high-fidelity VLTM representations of actions (Endress & Wood, 2011; Urgolites & Wood, 2013) and scenes (Konkle et al., 2010). In the study phase, participants viewed a large number of action–scene sets, in which an action was displayed within a scene. In the test phase, participants were then shown pairs of action–scene sets and indicated which of the two sets they had seen. Critically, we tested memory for actions and scenes in contexts that either did or did not require memory for action–scene binding. If actions and scenes are necessarily bound together in VLTM, then participants should be able to remember which actions had occurred within which scenes. In contrast, if actions and scenes are stored separately in VLTM, then participants should have accurate memory for individual actions and individual scenes, but impaired memory for integrated action–scene sets.

Experiment 1

Each participant was tested in either a binding or an entity memory task. In the binding memory task, the participants needed to remember which actions had been presented within which scenes. In the entity memory task, they needed to remember both the individual actions and the individual scenes, but did not need to bind the actions and scenes together in memory.

Method

Participants A group of 40 participants (ten males, 30 females; mean age=20 years, SD=1.35) with normal or corrected-to-normal vision participated in order to receive credit toward a course requirement.[1] Informed consent was obtained. Eight of the participants were randomly assigned to each of the five conditions described below.

Stimuli The stimuli consisted of 80 animated actions and 80 images of scenes. The actions were selected from a previous study that had examined the fidelity of the action representations stored in VLTM (Urgolites & Wood, 2013), and the scenes were selected from a previous study that had examined the fidelity of the scene representations stored in VLTM (Konkle et al., 2010). Each action–scene set displayed one action being performed within one scene (see Fig. 1). The action and scene presented in each set were randomly selected. The actions appeared in the middle of the scene. The acting agent subtended 2.3º×6.0º of viewing angle, and the scenes subtended 17.5º×14.6º of viewing angle in the center of a 17-in. computer screen. All of the actions had the same duration (1,000 ms) and frame rate (30 frames per second), were performed by the same agent, and were displayed from the same viewpoint. Each action started and ended at the same neutral standing position and reached its maximal deviation from the neutral position at the 15th frame. The action stimulus set included jumps, turns, bends, arm raises, and crouches. The scene stimuli set included images of oceans, forests, cities, and mountains.

Procedure The experiment consisted of a study phase and a test phase. In the study phase, the participants viewed 40 randomly selected action–scene sets. Each trial began with a 1,000-ms black screen, followed by an action–scene set, which was in turn followed by another 1,000-ms black screen. Each action–scene set was displayed for 1,000 ms, and each action was performed once. To maintain their attention, participants performed a repeat-detection task during the study phase. Specifically, ten action–scene sets were presented on two separate trials in the study phase, such that zero, one, two, three, or four trials intervened between the first set and the repeat set. After the presentation of each stimulus set, the phrase "Old or New" appeared, prompting participants to indicate whether the action–scene set was one of the previous sets or a new set. The participants responded without time pressure.

The test phase began approximately 1 min after the study phase. In the test phase, participants performed a two-alternative forced choice test. Two action–scene sets were presented on the screen: one previously seen "old" action–scene set and one new foil action–scene set. The participants received 40 test trials, each of which began with a 1,000-ms black screen, followed by an action–scene set on the left half

---

[1] Across all experiments, four additional participants were excluded from the analyses because they failed to follow the instructions.
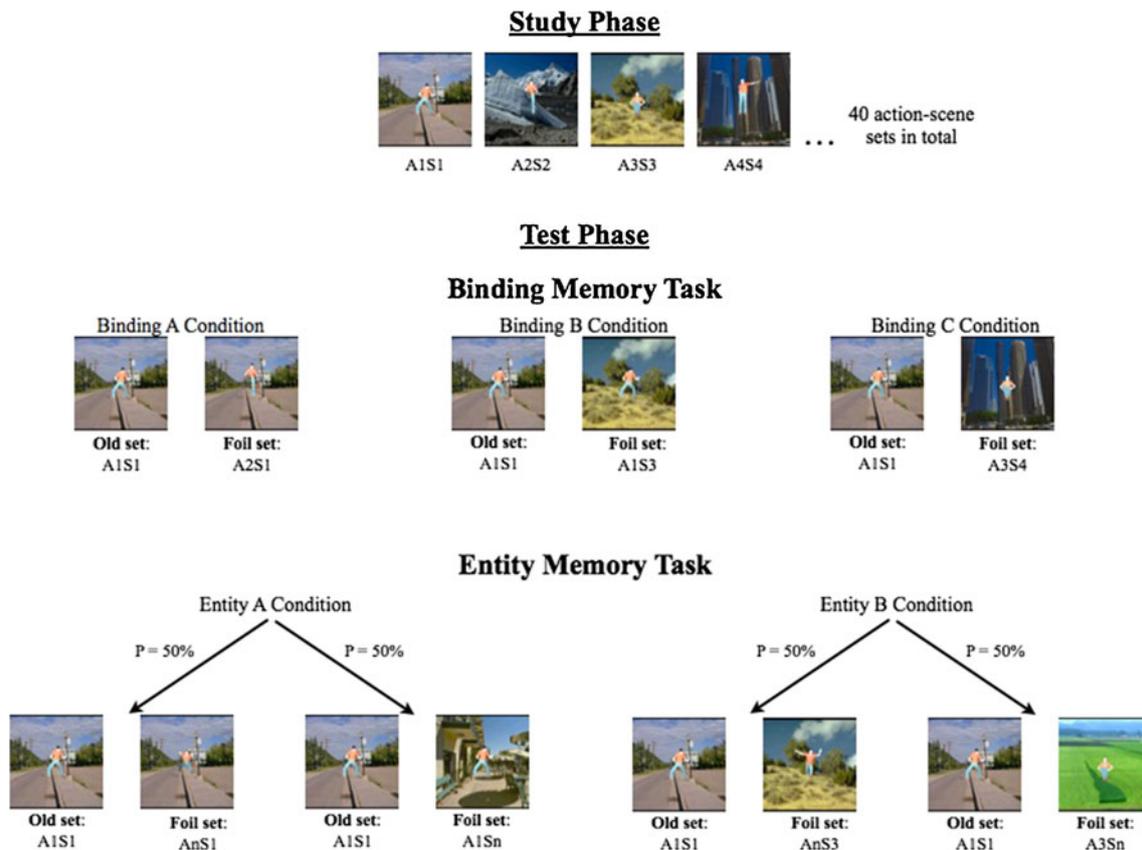
Fig. 1 Depiction of the design for each condition. In the binding memory task, participants needed to remember which actions had been displayed within which scenes. In the entity memory task, they needed to remember the individual actions and scenes, but did not need to bind the actions and scenes together in memory. The images show the actions' maximal deviations from the neutral position. "A1S1" indicates that the action–scene set contained Action 1 within Scene 1 from the study phase, "A2S2" indicates that the action–scene set contained Action 2 within Scene 2 from the study phase, and so forth. Likewise, "A1Sn" indicates that the action–scene set contained Action 1 from the study phase within a novel scene, "AnS1" indicates that the action–scene set contained a novel action within Scene 1 from the study phase, and so forth

of the screen, which after an interval of 1,000 ms was replaced by another action–scene set on the right half of the screen. Both the old set and the foil set were displayed only one time. We counterbalanced the side of the previously seen old set, such that the old set was presented equal numbers of times on the left and right sides of the screen across the test phase. After viewing the second set, participants indicated which set (left or right) had previously been seen in the study phase, by pressing one of two keys on the keyboard.

Participants were tested in either a binding or an entity memory task. In the binding memory task, the participants were told that they would see a large number of action–scene sets and would then be tested on their ability to remember which actions had occurred within which scenes. In the test phase, an old action–scene set was paired with a foil action–scene set that consisted of a new combination of an old action and an old scene that during the study phase had been presented in different action–scene sets. Thus, to succeed, participants needed to remember which actions had occurred within which scenes. To ensure that the results

would generalize across different testing contexts, we probed binding memory in multiple ways (see Fig. 1). Specifically, our design included three binding conditions: In the Binding A condition, on each test trial, both action–scene sets had the same scene and different actions. In the Binding B condition, both action–scene sets had the same action and different scenes. And in the Binding C condition, both action–scene sets had different actions and different scenes. Testing binding memory across these different conditions allowed us to explore how participants make memory judgments about previously experienced events. For example, binding memory might be enhanced when the same action or the same scene appear in both alternatives. If so, performance should be better in the Binding A and Binding B conditions than in the Binding C condition.

In the entity memory task, participants were told that they would see a large number of action–scene sets and then be tested on their ability to remember which actions and scenes had been present in the study phase. In the test phase, an old action–scene set was paired with a foil action–scene set that

contained either a new action or a new scene that had not been present in the study phase. On one-half of the trials, the foil set contained a new action displayed within an old scene, and on the other half, the foil set contained an old action displayed within a new scene. Thus, in order to succeed, participants needed to remember the individual actions and scenes. Critically, since each foil set could contain either a new action or a new scene, participants needed to remember both the actions and the scenes that had been presented in the study phase. We also probed entity memory in multiple ways in order to ensure that the results would generalize across different testing contexts (see Fig. 1). Specifically, our design included two entity conditions. In the Entity A condition, the old action–scene set and the foil action–scene set had the same old scene on one-half of the trials (thus, the foil set had a new action) and the same old action on the other half of the trials (thus, the foil set had a new scene). In the Entity B condition, both action–scene sets had different actions and different scenes (i.e., no shared action or scene). As with the binding memory task, testing entity memory across these different conditions allowed us to explore how participants make judgments about whether they have seen a particular visual entity. For example, memory for visual entities might be enhanced when the same action or the same scene appears in both alternatives. If so, performance should be better in the Entity A than in the Entity B condition.

For both the binding and entity memory tasks, verbal instructions were provided before the experiment. Written instructions were also provided on the screen before each phase of the experiment.

Results

Performance on the repeat-detection task in the study phase was high for all conditions, with accuracy rates of 84 % or higher across the conditions (M=86 %, SEM=1 %). Performance in the test phase was similar for action–scene sets that had and had not been repeated in the study phase: entity memory task, $t(23)$=0.25, p=.81; binding memory task, $t(15)$=0.08, p=.94. Thus, the data for repeated and nonrepeated action–scene sets were collapsed for the remaining analyses.

The results are depicted in Fig. 2a. In the binding memory task, performance was surprisingly low: Participants correctly recognized the old action–scene set on 59 % (SEM=2 %) of the trials in the Binding A condition, 57 % (SEM=3 %) of the trials in the Binding B condition, and 60 % (SEM=3 %) of the trials in the Binding C condition. These accuracy rates were, however, higher than chance level (50 %): $t(7)$=3.99, p<.01 (Binding A condition); $t(7)$=2.54, p=.04 (Binding B condition); $t(7)$=2.91, p=.02 (Binding C condition). Performance did not differ significantly across these conditions, $F(2, 21)$= 0.37, p=.61, η_p^2=.03. This indicates that performance was similar whether or not the two test alternatives shared a common action or a common scene. Because performance

did not differ across the conditions, we pooled the data for the analyses presented below.

In the entity memory task, performance was significantly better. In the Entity A condition, participants correctly recognized the old action–scene set on 79 % (SEM=3 %) of the trials: They were correct on 76 % (SEM=5 %) of the trials when the foil set contained a new action, and on 81 % (SEM= 2 %) of the trials when the foil set contained a new scene. In the Entity B condition, participants correctly recognized the old action–scene sets on 80 % (SEM=2 %) of the trials: They were correct on 73 % (SEM=3 %) of the trials when the foil set contained a new action, and on 88 % (SEM=2 %) of the trials when the foil set contained a new scene. These accuracy rates were significantly higher than chance level – the Entity A condition: $t(7)$=5.70, p=.001 (actions), and $t(7)$=13.79, p<.001 (scenes); the Entity B condition: $t(7)$=8.19, p<.001 (actions), and $t(7)$=22.91, p<.001 (scenes). Performance did not differ significantly between the Entity A and Entity B conditions, $F(1, 14)$=0.18, p=.68, η_p^2=.01. This indicates that performance was similar whether or not the two test alternatives shared a common action or a common scene. Because performance did not differ across the conditions, we pooled the data for the analyses presented below.

In the entity memory task, accuracy rates were higher when distinguishing old scenes from new scenes than when distinguishing old actions from new actions (see Fig. 2a), $t(15)$=2.58, p=.02. Since accurate memory of a bound action–scene set requires representing both the action and the scene, accuracy should be limited by the more difficult entity. Thus, to compare entity memory with binding memory, we used the lowest of the two accuracy rates in the entity memory task—that is, the accuracy rates for remembering actions. This analysis comparing entity memory to binding memory revealed a highly significant difference, $F(1, 38)$=29.05, p<.001, η_p^2=.43; performance was substantially lower in the binding than in the entity memory task.

Discussion

In this experiment, we tested observers' ability to remember individual actions and scenes and their ability to bind actions and scenes together in memory. We found that memory for individual actions and scenes was remarkably high, whereas memory for action–scene binding was much lower. Despite being explicitly instructed to remember which actions had occurred within which scenes in the binding memory task, participants had considerable difficulty distinguishing previously seen action–scene sets from novel action–scene sets. This suggests that actions and scenes experienced together in the same visual event are not necessarily stored as integrated representations in VLTM.

The high performance in the entity memory task indicates that the errors in the binding memory task were not due to

Fig. 2 Accuracy rates for recognizing the action–scene sets. (a) In Experiment 1, each action–scene set was displayed once (1,000-ms total duration). (b) In Experiment 2, each action–scene set was displayed three times (3,000-ms total duration). Error bars indicate standard errors.
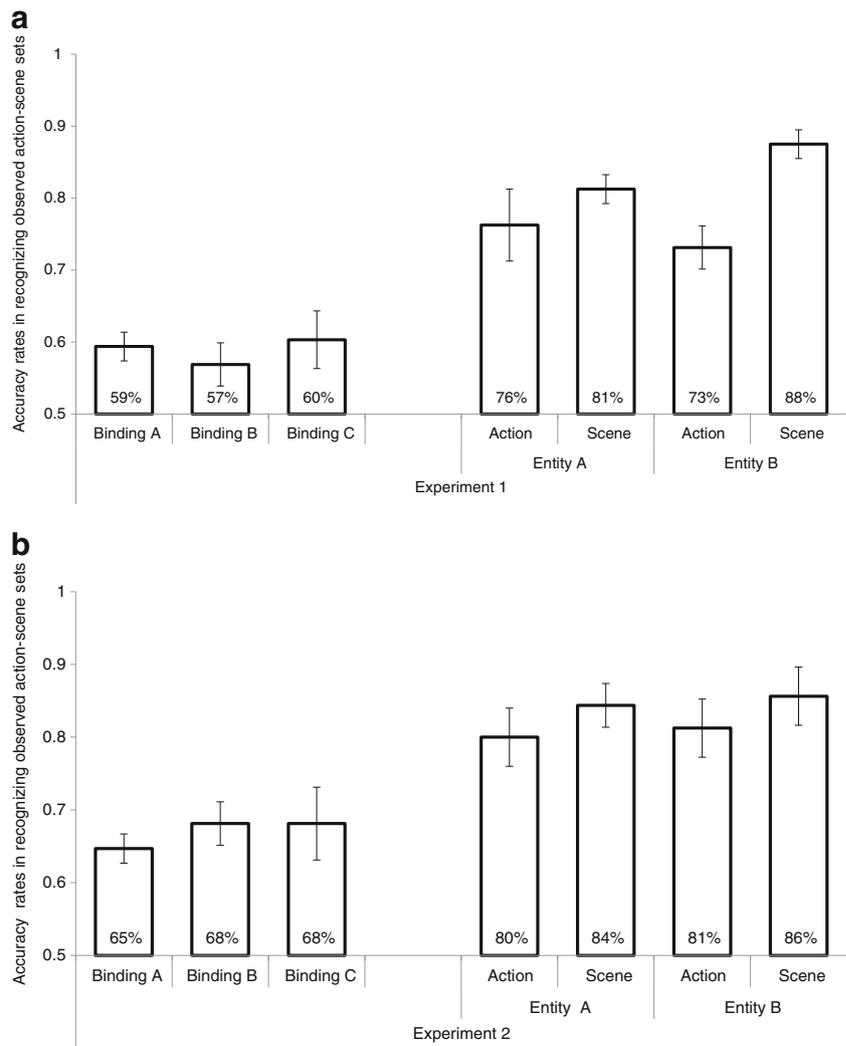
limitations in processes such as perceiving the actions and scenes and/or encoding them into memory, since observers were able to perceive and encode both the individual actions and the individual scenes in the entity memory task. Thus, the errors likely reflect a specific difficulty with binding actions and scenes together in memory.

In Experiment 1, each action–scene set was displayed for a relatively short period of time (1,000 ms) in the study phase. To test whether these results would generalize to longer presentation times, we performed an additional experiment, increasing the display time to 3,000 ms. This provided 3 s of exposure to each action–scene set.

Experiment 2

Method

A group of 40 new participants (seven males, 33 females; mean age=20 years, SD=1.12) served in this experiment. It

was identical to Experiment 1, except that each action–scene set was displayed three times rather than one time on each trial in the study phase. A 1,000-ms black screen separated each of the three 1,000-ms repetitions of the action–scene sets, providing 3 s of total exposure to the event. In the test phase, the action–scene sets were displayed once.

Results and discussion

Performance on the repeat-detection task in the study phase was high for all conditions, with accuracy rates of 91 % or higher across the conditions (mean=93 %, SEM=1 %). Performance in the test phase was similar for action–scene sets that had and had not been repeated in the study phase: entity memory task, $t(23)$=0.12, p=.91; binding memory task, $t(15)$=0.91, p=.38. Thus, the data for repeated and nonrepeated action–scene sets were collapsed for the remaining analyses.

The results are depicted in Fig. 2b. In the binding memory task, participants correctly recognized the old action–scene set on 65 % (SEM=2 %) of the trials in the Binding A

condition, 68 % (*SEM*=3 %) of the trials in the Binding B condition, and 68 % (*SEM*=5 %) of the trials in the Binding C condition. These accuracy rates were significantly higher than chance level: $t(7)=6.42$, $p<.001$ (Binding A condition); $t(7)=5.28$, $p=.001$ (Binding B condition); $t(7)=3.51$, $p=.01$ (Binding C condition). These results did not differ significantly from one another, $F(2, 21)=0.27$, $p=.77$, $\eta_p^2=.03$, so the data were pooled for the further analyses. The accuracy rates in Experiment 2 were significantly higher than those in Experiment 1, $F(1, 46)=9.14$, $p<.01$, $\eta_p^2=.17$, which indicates that action–scene binding is facilitated when participants have longer exposure to each action–scene set.

In the entity memory task, participants correctly recognized the old action–scene set on 82 % (*SEM*=3 %) of the trials in the Entity A condition: They were correct on 80 % (*SEM*=4 %) of the trials when the foil set contained a new action, and 84 % (*SEM*=3 %) of the trials when the foil set contained a new scene. In the Entity B condition, participants correctly recognized the old action–scene set on 83 % (*SEM*=3 %) of the trials: They were correct on 81 % (*SEM*=4 %) of the trials when the foil set contained a new action, and 86 % (*SEM*=4 %) of the trials when the foil set contained a new scene. These accuracy rates were significantly higher than chance level – the Entity A condition: $t(7)=7.94$, $p<.001$ (actions), and $t(7)=9.92$, $p<.001$ (scenes); the Entity B condition: $t(7)=8.92$, $p<.001$ (actions), and $t(7)=9.30$, $p<.001$ (scenes). The results from the Entity A and Entity B conditions did not differ significantly from one another, $F(1, 14)=0.09$, $p=.78$, $\eta_p^2<.01$, so the data were pooled for the further analyses. Memory for scenes was higher than memory for actions (see Fig. 2b), although this difference did not reach statistical significance, $t(15)=1.58$, $p=.14$.

As in Experiment 1, we compared performance in the binding memory task to performance for the more difficult entity in the entity memory task (i.e., memory for actions). Performance was still significantly higher in the entity memory task than in the binding memory task, $F(1, 38)=16.96$, $p<.01$, $\eta_p^2=.31$. A two-way analysis of variance across Experiments 1 and 2 with the factors Display Duration (1,000 or 3,000 ms) and Memory Task (binding or entity memory task) revealed a significant main effect of display duration, $F(1, 76)=10.08$, $p<.01$, $\eta_p^2=.12$, and a significant main effect of memory task, $F(1, 76)=44.31$, $p<.001$, $\eta_p^2=.37$. The interaction was not significant, $F(1, 76)=0.24$, $p=.62$, $\eta_p^2<.01$.

Experiment 2 replicated the general pattern observed in Experiment 1: Although increasing the display time enhanced both binding memory and entity memory, performance in the binding memory task was still significantly lower than performance in the entity memory task. This provides additional support for the hypothesis that actions and scenes are not necessarily stored as integrated action–scene representations in VLTM.

General discussion

In this study, we investigated whether actions and scenes that are observed together in the same visual event are stored as integrated representations in visual long-term memory. To do so, we conducted two experiments that measured participants' ability to remember actions and scenes in contexts that either did or did not require action–scene binding. In the binding memory task, participants needed to remember which actions had occurred within which scenes. In the entity memory task, they needed to remember the individual actions and the individual scenes, but did not need to bind the actions and scenes together in memory. We tested memory across several different conditions and presentation times. In all cases, performance was significantly lower when the task required memory for action–scene bindings than when the task required memory for individual actions and scenes. This indicates that VLTM stores action information and scene information at least partially separately from one another, rather than as integrated action–scene representations.

Although these results show that action–scene binding in VLTM is difficult, they also show that it is possible. In both Experiments 1 and 2, participants were able to bind actions and scenes at above-chance levels. This pattern is consistent with a VLTM architecture in which actions and scenes are stored separately from one another and require some additional binding mechanism to integrate these different types of information. It would be interesting for future studies to combine the present method with neurophysiological measures to examine the processes that support action–scene binding. For example, researchers have proposed that gamma oscillations synchronize assemblies of neurons in different parts of the cortex that process different features of an event, thereby binding those features together in memory (reviewed by Singer, 2001). This gamma-binding proposal has been extended to binding both within and across sensory modalities (Engel & Singer, 2001), making it a potential candidate mechanism for supporting binding of action and scene information stored in different parts of the cerebral cortex.

Additional research will also be needed to determine whether contextual factors can enhance memory for action–scene binding. For example, Kersten (1998) showed that humans encode intrinsic movements (i.e., the motion of a body part with respect to the body as a whole) and extrinsic movements (i.e., the motion of a body part in relation to a larger reference frame) differently from one another. Kersten suggested that intrinsic movements are more strongly associated with "what" information (i.e., the ventral stream) than with "where" information (i.e., the dorsal stream). Since the ventral stream specializes in separating out and identifying the individual entities in an event, one might expect intrinsic

actions to be represented separately from information about the larger scene in which the actions take place. In contrast, a representation of an extrinsic action might be integrated with information about the scene, because extrinsic actions are defined in relation to the larger reference frame. For example, if an action involves entering a door or falling into a hole, these types of actions would contain contextual information linking the action to the scene, and thus might not require an additional binding mechanism to the same extent that intrinsic actions do.

More generally, these findings have implications for eyewitness testimony, because they call into question the reliability of our ability to remember which actions occurred within which scenes. For instance, these results help explain why unconscious transference errors occur in eyewitness testimony. Since actions and scenes are not necessarily integrated in long-term memory, eyewitnesses make errors, often incorrectly stating that an action occurred within a scene when it did not. Future experiments could use the present method to further investigate the conditions under which action–scene binding errors do, and do not, occur. For example, this method could be adapted into a dual-task paradigm, allowing researchers to systematically test which processes are needed to bind actions and scenes together in long-term memory.

References

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences, 105,* 14325–14329. doi:10.1073/pnas.0803390105

Brown, E., Deffenbacher, K., & Sturgill, W. (1977). Memory for faces and the circumstances of encounter. *Journal of Applied Psychology, 62,* 311–318.

Buckhout, R. (1974). Eyewitness testimony. *Scientific American, 231,* 23–31.

Deffenbacher, K. A., Bornstein, B. H., & Penrod, S. D. (2006). Mugshot exposure effects: Retroactive interference, mugshot commitment, source confusion, and unconscious transference. *Law and Human Behavior, 30,* 287–307.

Deffenbacher, K. A., Carr, T. H., & Leu, J. R. (1981). Memory for words, pictures, and faces: Retroactive interference, forgetting, and reminiscence. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 299–305.

Earles, J. L., Kersten, A. W., Curtayne, E. S., & Perle, J. G. (2008). That's the man who did it, or was it a woman? Actor similarity and binding errors in event memory. *Psychonomic Bulletin & Review, 15,* 1185–1189. doi:10.3758/PBR.15.5.1185

Eichenbaum, H., & Cohen, N. (2001). *From conditioning to conscious recollection: Memory systems of the brain*. New York, NY: Oxford University Press.

Endress, A. D., & Wood, J. N. (2011). From movements to actions: Two mechanisms for learning action sequences. *Cognitive Psychology, 63,* 141–171.

Engel, A. K., & Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends in Cognitive Sciences, 5,* 16–25.

Kassin, S. M., Tubb, V. A., Hosch, H. M., & Memon, A. (2001). On the "general acceptance" of eyewitness testimony research: A new survey of the experts. *American Psychologist, 56,* 405–416. doi:10.1037/0003-066X.56.5.405

Kersten, A. W. (1998). A division of labor between nouns and verbs in the representation of motion. *Journal of Experimental Psychology: General, 127,* 34–54.

Kersten, A. W., & Earles, J. L. (2010). Effects of aging, distraction, and response pressure on the binding of actors and actions. *Psychology and Aging, 25,* 620–630.

Kersten, A. W., Earles, J. L., Curtayne, E. S., & Lane, J. C. (2008). Adult age differences in binding actors and actions in memory for events. *Memory & Cognition, 36,* 119–131. doi:10.3758/MC.36.1.119

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Scene memory is more detailed than you think: The role of categories in visual long-term memory. *Psychological Science, 21,* 1551–1556. doi:10.1177/0956797610385359

Kravitz, D. J., Saleem, K. S., Baker, C. I., & Mishkin, M. (2011). A new neural framework for visuospatial processing. *Nature Reviews Neuroscience, 12,* 217–230.

Loftus, E. F. (1976). Unconscious transference. *Law and Psychology Review, 2,* 93–98.

McTighe, S. M., Cowell, R. A., Winters, B. D., Bussey, T. J., & Saksida, L. M. (2010). Paradoxical false memory for objects after brain damage. *Science, 330,* 1408–1410.

Perfect, T. J., & Harris, L. J. (2003). Adult age differences in unconscious transference: Source confusion or identity blending? *Memory & Cognition, 31,* 570–580.

Ross, D. R., Ceci, S. J., Dunning, D., & Toglia, M. P. (1994). Unconscious transference and mistaken identity: When a witness misidentifies a familiar with innocent person. *Journal of Applied Psychology, 79,* 918–930.

Singer, W. (2001). Consciousness and the binding problem. *Annals of the New York Academy of Sciences, 929,* 123–146.

Standing, L. (1973). Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology, 25,* 207–222. doi:10.1080/14640747308400340

Urgolites, Z. J., & Wood, J. N. (2013). Visual long-term memory stores high-fidelity representations of observed actions. *Psychological Science, 24*(4), 403–411. doi:10.1177/0956797612457375.