BRIEF REPORT

Haptic experiences influence visually acquired memories: Reference frames during multimodal spatial learning

**Jonathan W. Kelly · Marios N. Avraamides ·
Nicholas A. Giudice**

Abstract In two experiments, we investigated whether reference frames acquired through touch could influence memories for locations learned through vision. Participants learned two objects through touch, and haptic egocentric (Experiment 1) and environmental (Experiment 2) cues encouraged selection of a specific reference frame. Participants later learned eight new objects through vision. Haptic cues were manipulated, whereas visual learning was held constant in order to observe any potential influence of the haptically experienced reference frame on memories for visually learned locations. When the haptically experienced reference frame was defined primarily by egocentric cues, cue manipulation had no effect on memories for objects learned through vision. Instead, visually learned locations were remembered using a reference frame selected from the visual study perspective. When the haptically experienced reference frame was defined by both egocentric and environmental cues, visually learned objects were remembered in the context of the haptically experienced reference frame. These findings support the common reference frame hypothesis, which proposes that locations learned through different sensory modalities are represented within a common reference frame.

J. W. Kelly (✉)
Department of Psychology, Iowa State University,
W112 Lagomarcino Hall,
Ames, IA 50011-3180, USA
e-mail: jonkelly@iastate.edu

M. N. Avraamides
Department of Psychology, University of Cyprus,
Nicosia, Cyprus

N. A. Giudice
Department of Spatial Information Science and Engineering,
University of Maine,
Orono, ME, USA

Spatial tasks, such as remembering where different seeds are planted in a vegetable garden or finding an alternative route home from work, rely on accurate spatial memory retrieval. Research on the organizational properties of spatial memory indicates that locations are typically remembered in the context of a spatial reference system. Retrieval tasks such as imagined perspective taking (Shelton & McNamara, 2001) and map drawing (Kelly, 2011) typically reveal preferred access to spatial memories from one or two perspectives, which is considered evidence for a reference frame parallel to the facilitated perspective(s) (Klatzky, 1998). Reference frame selection is influenced by available cues during learning, including egocentric cues such as experienced perspectives (Kelly, Avraamides, & Loomis, 2007; Shelton & McNamara, 1997) and environmental cues such as room axes (Shelton & McNamara, 2001). According to the reference frame theory of spatial memory (Shelton & McNamara, 2001), reference frames are selected from the first experienced perspective and are updated only if a subsequent perspective provides better information (e.g., the first perspective is misaligned with environmental axes, and the second perspective is aligned with the axes).

Although most research on reference frames has focused on how visual cues influence reference frame selection (Greenauer & Waller, 2008; Kelly & McNamara, 2008; Mou & McNamara, 2002; Shelton & McNamara, 2001), reference frames also characterize spatial memories acquired

through touch (Giudice, Betty, & Loomis, 2011; Newell, Woods, Mernagh, & Bülthoff, 2005), hearing (Yamamoto & Shelton, 2009), and language (Avraamides & Kelly, 2010; Mou, Zhang, & McNamara, 2004). The present study was designed to understand how reference frames are selected when learning a layout through both touch and vision.

Research on multimodal spatial learning has shown that locations learned through different modalities can be organized by a common reference frame. Giudice, Klatzky, and Loomis (2009) had participants learn three objects through vision and three through touch. Participants later performed judgments of relative direction (JRD) in which they imagined standing at the learning position facing one object and pointed to another object from the imagined perspective. JRDs were constructed using objects learned through the same or different modalities. When learning modality overlapped temporally, pointing performance was similar for trials testing within-modality and across-modality object pairs. If the layouts were stored in separate reference frames, between-layout JRDs would have required mentally aligning the reference frames, resulting in transformation costs. The lack of transformation costs for between-layout judgments led the authors to conclude that locations learned through different modalities were organized within a common reference frame.

Using a different paradigm, Kelly and Avraamides (2011) also found that locations learned through different modalities were organized by a common reference frame. Participants viewed two objects on a table, and parallel stripes drawn on the table encouraged selection of a reference frame aligned with the stripes. Participants subsequently moved to another perspective and learned seven additional objects through touch. JRD performance when recalling locations learned through touch indicated that touched objects were remembered using a reference frame aligned with visual environmental cues, and this cross-sensory reference frame transfer suggests that locations learned through different senses were organized by a common reference frame. In this context, cross-sensory reference frame transfer refers to the influence of a reference frame acquired through one sensory modality on spatial memories acquired through another modality. This can arise from recoding a modality-specific reference frame into another modality (recoding hypothesis), or creating an amodal spatial representation dissociated from any sensory system (amodal hypothesis). The term cross-sensory transfer is used for ease of exposition and is not intended to support the recoding or amodal hypothesis. Evidence for and against these hypotheses is presented in the General discussion section.

Because vision is the dominant sense for most humans, cross-sensory transfer might be specific to visually defined reference frames. Similar visual dominance, or "visual

capture," is evidenced by the ventriloquist effect, in which the perceived location of a sound is biased toward the location of a simultaneous visual stimulus (Warren, Welch, & McCarthy, 1981). Vision also influences perceived spatial properties of stimuli experienced through touch (Rock & Victor, 1964) and proprioception (Hay, Pick, & Ikeda, 1965; Pavani, Spence, & Driver, 2000). Although capture of visually perceived locations by nonvisual stimuli can also occur, this typically requires visual stimulus degradation (Alais & Burr, 2004). The present experiments were designed to test the bidirectionality of cross-sensory reference frame transfer by evaluating whether a reference frame acquired through touch could influence locations subsequently learned through vision.

Experiment 1

Participants studied two objects through touch, and the two objects formed a column that was oriented parallel to the haptic learning perspective. This was expected to result in selection of a reference frame parallel to the egocentric cues defined from the haptic study perspective. Participants subsequently learned eight new objects through vision. The haptic learning perspective was either aligned (touch–vision aligned condition) or misaligned (touch–vision misaligned condition) with the visual learning perspective to evaluate the influence of the haptic reference frame on memories acquired through vision. If the reference frame used to remember visual object locations depends on the haptically acquired reference frame, then this will provide evidence of cross-sensory transfer from touch to vision.

Method

Participants Forty-two undergraduates (20 men and 22 women) participated for course credit.

Stimuli and design Learning stimuli consisted of 10 objects on a round table (76-cm diameter), centered within a 3.9 × 4.1 m room (Fig. 1). Participants studied two objects through touch from 0° (touch–vision misaligned condition) or 315° (touch–vision aligned condition) and later studied the entire layout through vision from 315°. Touch–vision alignment was manipulated between participants. Haptic objects were selected from a subset of three objects (underlined objects, Fig. 1) so that the two haptic objects were aligned with the haptic learning perspective. The unused haptic object was added to the layout during subsequent visual haptic learning, such that all participants viewed the same 10 objects from 315°.

Spatial memories were later tested using JRD in which participants imagined standing at one object, facing a second object, and pointed to a third object from the
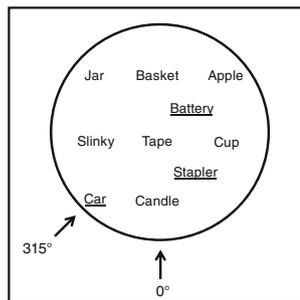
Fig. 1 Object layout used in Experiment 1. The *circle* represents the round table on which objects were placed, and the surrounding *rectangle* represents the room walls (not drawn to scale). Participants studied two objects through touch from 0° (touch–vision misaligned condition) or 315° (touch–vision aligned condition) before studying the same two objects plus eight new objects through vision from 315°. Haptic objects were selected from a subset of three objects (shown in *underlined text*) so that the two haptic objects were aligned with the haptic learning perspective (stapler and battery from 0°, car and battery from 315°). The unused haptic object was added to the layout during subsequent visual learning, so that all participants viewed the same 10 objects from 315°

imagined perspective. JRD included only objects studied exclusively through vision. Eight imagined perspectives were tested (spaced every 45°), and imagined perspective was manipulated within participants. For each perspective, six JRD trials were constructed requiring egocentric pointing responses of 45°, 90°, 135°, 225°, 270°, and 315°. Participants completed 48 JRD in a random sequence. Pointing error and latency were recorded.

Procedure Participants were blindfolded and guided into the learning environment, which contained two objects on a round table. Participants in the touch–vision aligned condition were led to the 315° perspective, and participants in the touch–vision misaligned condition were led to the 0° perspective. The participant was seated, and the experimenter guided the participant's right hand to the two objects. Participants studied the objects by freely moving their right hands for 20 sec; then, they pointed to each object. This study-then-point sequence was repeated until participants correctly pointed to each object twice. Participants in the touch–vision misaligned condition were then led to the 315° perspective, whereas participants in the touch–vision aligned condition remained at 315°. Eight new objects were then added to the layout, and participants lifted their blindfolds. The experimenter named and pointed to all 10 objects: first the two previously experienced haptic objects and then the eight new objects in a random sequence. Participants visually studied for 60s before replacing the blindfold and pointing to each object in a random order determined by the experimenter. Learning ended when participants correctly pointed to all objects twice, after which they were blindfolded and led to another room for testing.

JRD were presented as sentences on a monitor ("Imagine standing at the battery, facing the apple. Point to the cup."). Participants pointed by deflecting the joystick in the intended direction.

Results

There was no indication of speed–accuracy trade-off. Within-participants correlations between pointing latency and error averaged 0.15 ($SD = .38$), significantly above zero, $t(41) = 2.50$, $p = .016$. In the interest of brevity, we focus on errors, but all significant ANOVA results were also found in latencies.

Absolute pointing error (Fig. 2) was analyzed in a mixed-model ANOVA with terms for touch–vision alignment and imagined perspective. Manipulation of prior haptic experience had no effect on the reference frame used to remember visually learned locations: Regardless of touch–vision alignment, JRD performance was best when imagining the 315° perspective. Statistical analyses supported this conclusion. The main effect of imagined perspective was significant, $F(7, 280) = 9.64$, $p < .001$, $\eta_p^2 = .19$. Performance was better when imagining the 315° perspective as compared with all other perspectives, $F(1, 40) = 34.16$, $p < .001$, $\eta_p^2 = .46$. The main effect of touch–vision alignment was marginally significant, $F(1, 40) = 3.13$, $p = .085$, $\eta_p^2 = .07$, indicating superior performance in the touch–vision misaligned condition ($M = 29.15°$, $SE = 3.84°$) than in the touch–vision aligned condition ($M = 39.15°$, $SE = 4.15°$). The interaction between perspective and condition was not significant, $F(7, 280) = 1.20$, nor was the interaction contrast directed at the 0° and 315° perspectives, $F(1, 40) = 0.73$.
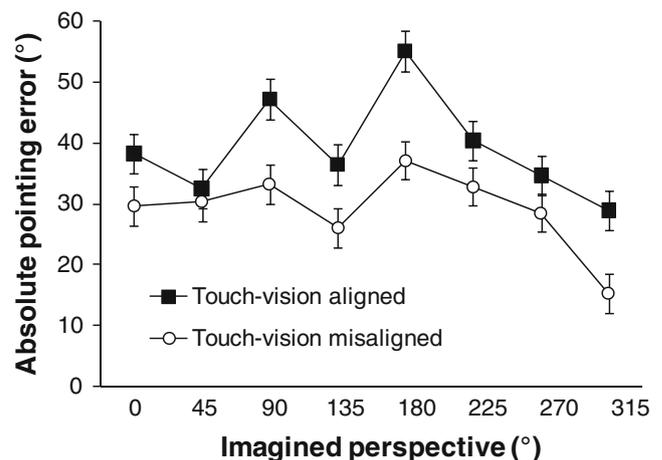

Fig. 2 Absolute pointing error as a function of imagined perspective and touch–vision alignment in Experiment 1. *Error bars* represent standard errors estimated from the ANOVA
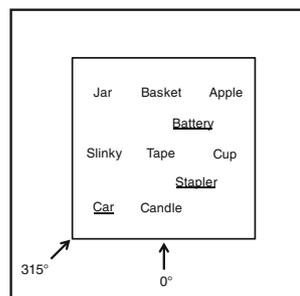
Discussion

The reference frame used to remember locations learned through vision was selected from the visual study perspective and was unaffected by manipulation of the prior haptic learning perspective. This led to superior pointing performance when imagining the visual study perspective as compared with all other perspectives, which is inconsistent with cross-sensory reference frame transfer from touch to vision.

The results of Experiment 1 diverge from those reported by Kelly and Avraamides (2011), in which locations learned through touch were remembered using a visually acquired reference frame. One possible explanation is that cross-modal reference frame transfer is not bidirectional. Alternatively, the haptic cues in Experiment 1 may have been insufficient to induce transfer. Kelly and Avraamides used a combination of egocentric and environmental cues to induce reference frame selection during visual learning. In contrast, reference frame selection during haptic learning in Experiment 1 relied primarily on egocentric haptic cues defined by the study perspective and relatively weak environmental haptic cues defined by the alignment of the two study objects with the haptic learning perspective. Experiment 2 was designed to determine whether a combination of egocentric and environmental cues during haptic learning would be sufficient to cause cross-sensory reference frame transfer from touch to vision. This was accomplished by substituting a square table for the round table used in Experiment 1.

Experiment 2

In order to supplement the haptic environmental cues to reference frame selection, objects were placed on a square table. The 0° haptic study perspective was aligned with the table edges, and the 315° haptic study perspective was misaligned with the table (Fig. 3). Furthermore, participants were instructed to feel the table edges during haptic learning. It was expected that haptic learning from 0° would result in selection of a reference frame aligned with the 0° study perspective, and that haptic learning from the

315° perspective would result in selection of a reference frame aligned with the 315° study perspective. These predictions are based on the reference frame theory of spatial memory (Shelton & McNamara, 2001), whereby reference frame selection occurs from the initial study perspective unless another experienced perspective provides better access to environmental cues (e.g., alignment with environmental structure). Subsequent visual learning occurred from 315° for all participants. If the reference frame used to remember visually acquired object locations depends on the haptically acquired reference frame, then this will provide evidence of cross-sensory reference frame transfer from touch to vision.

Method

Participants Forty-four undergraduates (18 men and 26 women, evenly distributed across condition) participated for course credit.

Stimuli, design, and procedure Learning stimuli from Experiment 1 were modified by replacing the circular table with a 61 × 61 cm square table. Table edges were parallel to the 0°–180° and 90°–270° axes of the layout. The stimuli, design, and procedure were otherwise identical to those in Experiment 1.

Results

There was no indication of speed–accuracy trade-off. Within-participants correlations between pointing latency and error averaged 0.17 (SD = .41), significantly above zero, $t(43)$ = 2.77, p = .008. For brevity, we focus on angular errors.

Fig. 3 Object layout used in Experiment 2. The inner *square* represents the edges of the square table on which objects were placed, and the surrounding *rectangle* represents the room walls (not drawn to scale)

Jar   Basket   Apple
Battery
Slinky   Tape   Cup
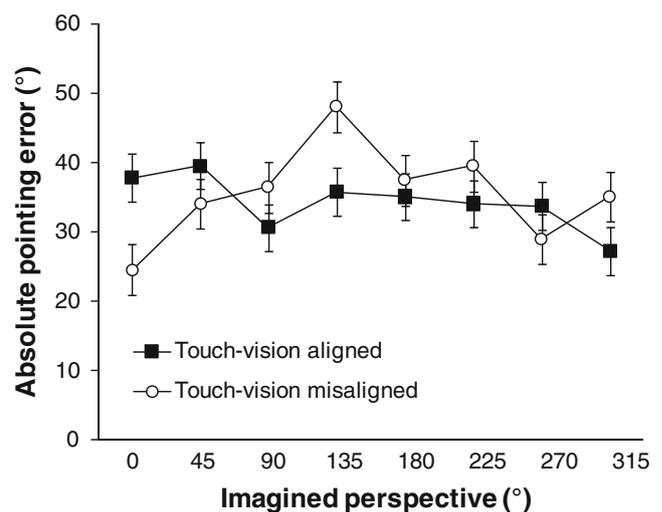Stapler
Car   Candle

315°   0°

Fig. 4 Absolute pointing error as a function of imagined perspective and touch–vision alignment in Experiment 2. *Error bars* represent standard errors estimated from the ANOVA

Absolute pointing errors (Fig. 4) were analyzed in a mixed-model ANOVA with terms for touch–vision alignment and imagined perspective. Manipulation of prior haptic experience influenced the reference frame used to remember visually learned locations: JRD performance when recalling visually learned locations was best from perspectives aligned with the haptically acquired reference frame. In the touch–vision misaligned condition, JRD performance was best when imagining perspectives aligned with the 0° haptic learning perspective. In the touch–vision aligned condition, JRD performance was best when imagining the 315° haptic and visual learning perspective. Statistical analyses supported these conclusions. A significant main effect of imagined perspective, $F(7, 294) = 2.31$, $p = .026$, $\eta_p^2 = .05$, was qualified by a significant interaction, $F(7, 294) = 2.86$, $p = .007$, $\eta_p^2 = .06$. The interaction contrast between imagined perspective (0° and 315°) and touch–vision alignment was significant, $F(1, 42) = 5.06$, $p = .03$, $\eta_p^2 = .11$. Participants in the touch–vision aligned condition performed best when imagining the 315° perspective as compared with all others, $F(1, 21) = 4.61$, $p = .044$, $\eta_p^2 = .18$. Participants in the touch–vision misaligned condition performed best when imagining the 0° perspective as compared with all others, $F(1, 21) = 9.64$, $p = .005$, $\eta_p^2 = .32$.

Discussion

Manipulation of haptic cues influenced the reference frame used to remember locations learned through vision, demonstrating cross-sensory reference frame transfer from touch to vision. Participants in the touch–vision misaligned condition established a reference frame through touch from the 0° perspective and interpreted new visual objects studied from the 315° perspective in the context of the haptically defined reference frame. Participants in the touch–vision aligned condition established a reference frame through touch from the 315° perspective and maintained that organization during visual learning from 315°.

The primary difference between Experiments 1 and 2 was the presence of the square table during learning in Experiment 2. Therefore, the environmental structure defined by the square table appears to have played a critical role in establishing cross-sensory reference frame transfer. The results of Experiment 2 are analogous to those reported by Kelly and Avraamides (2011), in which manipulation of a visually defined reference frame established through both experiential and environmental cues influenced memories for locations subsequently learned through touch.

General discussion

The present study was designed to evaluate whether a reference frame acquired through touch could influence locations subsequently learned through vision. In two experiments, participants learned two objects through touch before learning eight objects through vision. Haptic cues were manipulated in order to observe any influence of haptically acquired reference frames on memories for visually learned locations. In Experiment 1, haptic egocentric cues defined by the experienced haptic perspective did not result in cross-sensory reference frame transfer. Instead, participants remembered visually learned objects using a reference frame aligned with the visual study perspective, regardless of prior experiences during haptic learning. This led to superior retrieval of visually learned objects when imagining perspectives aligned with the visual study perspective. The results of Experiment 1 contrast with previous work demonstrating cross-sensory reference frame transfer from vision to touch (Kelly & Avraamides, 2011) and also with research indicating that locations learned through different modalities are stored in a common reference frame (Giudice et al., 2009).

In Experiment 2, haptic environmental cues were added to address methodological differences between Experiment 1 and past research (Kelly & Avraamides, 2011). This was accomplished by placing all objects on a square table and instructing participants to feel the edges of the table during haptic learning. The resulting haptic egocentric and environmental cues were sufficient to cause cross-sensory reference frame transfer from touch to vision, so that participants remembered visually learned objects using a reference frame aligned with the haptic cues. This led to superior retrieval of visually learned objects when imagining perspectives aligned with the haptic study perspective.

Taken together, the results from the present studies indicate that cross-sensory reference frame transfer can occur from touch to vision, so that locations learned through both modalities are stored within a common reference frame. Furthermore, cross-sensory transfer from touch to vision required haptic environmental cues. However, experiments reported by Giudice et al. (2009) show that environmental cues are not always necessary. In that study, participants learned some objects through touch and others through vision in a setting devoid of environmental cues. Despite the lack of environmental cues, participants remembered locations within a common reference frame as long as learning modality overlapped temporally. In light of this evidence, the contrasting results from Experiments 1 and 2 might not be due to environmental cues but could instead be explained by cue salience. In Experiment 2, the combined egocentric and environmental cues experienced through touch were sufficiently salient to establish and maintain a reference frame from the haptic study perspective, but other methods of increasing haptic cue salience (increased study time, verbal instructions, etc.) might have a similar effect.

One notable difference between Experiment 2 and experiments reported by Kelly and Avraamides (2011) is the perceptual availability of environmental cues throughout learning. Kelly and Avraamides presented participants with visual environmental cues defined by stripes on a table aligned with surrounding room walls. Those visual cues were occluded by a blindfold during subsequent haptic learning. Participants in Experiment 2 of the present study experienced the square table through touch and viewed the square table during subsequent visual learning, albeit from a misaligned perspective. However, the visibility of the square table during visual learning does not undermine the conclusion of cross-sensory reference frame transfer, since the reference frame used to remember visually acquired objects was clearly influenced by manipulation of prior haptic experience. Instead, the table may have served as an external indicator of the haptic study perspective.

The results of the present studies support the common reference frame hypothesis, which proposes that locations learned through different senses are represented within a common reference frame (Kelly & Avraamides, 2011). However, the common reference frame hypothesis does not make claims about representational format. To that end, the recoding hypothesis and the amodal hypothesis present competing accounts of the sensory format of spatial representations acquired through multimodal learning. According to the recoding hypothesis, locations learned through touch are recoded into a visual format, or vice versa (Newell et al., 2005), so that the entire representation shares the same sensory format. The present results are consistent with the recoding hypothesis, since locations learned through vision or touch could have been recoded to support integration within a common reference frame. According to the amodal hypothesis, spatial locations learned through different senses are represented as an amodal "spatial image," dissociated from any sensory system (Bryant, 1997; Giudice et al., 2009; Loomis, Lippa, Klatzky, & Golledge, 2002). The present results are also consistent with the amodal hypothesis, since locations learned through vision and touch could have been stored in a common amodal representation.

The present results cannot adequately distinguish between the recoding and amodal hypotheses. However, behavioral research on functional equivalence and neuroscientific research on multimodal perception appear better suited to evaluate the competing accounts. Research on functional equivalence shows that locations learned through different senses can lead to functionally equivalent performance when recalling and acting on remembered locations. Numerous studies identify functional equivalence between locations learned through vision, hearing, touch, and spatial language (see Loomis, Klatzky, Avraamides, Lippa, & Golledge, 2007, for review). Functional equivalence should be relatively uncommon if object locations are represented in sensory-specific formats, but should be quite common if locations are represented in an amodal format. Further behavioral evidence favoring the amodal hypothesis comes from research demonstrating functional equivalence between blind and blindfolded sighted participants after learning locations through touch, audition, or spatial language (Giudice et al., 2011; Loomis et al., 2002).

Neuroscientific research showing the convergence of sensory-specific brain regions onto common multisensory regions (Macaluso & Driver, 2005) also supports the amodal hypothesis. Furthermore, brain regions showing greater activation during visual scene processing compared to visual object processing show similar patterns during haptic scene and object processing, and this holds true for blind and blindfolded sighted participants (Wolbers, Klatzky, Loomis, Wutte, & Giudice, 2011).

The present study advances our understanding of reference frames during multimodal learning. Previous work has shown that reference frames acquired through vision can be used to remember locations learned subsequently through touch (Kelly & Avraamides, 2011). The present experiments extend those findings by demonstrating that cross-sensory reference frame transfer can occur from touch to vision, and that such transfer may require environmental cues during reference frame acquisition. These findings are consistent with the common reference frame hypothesis, indicating that locations learned through different modalities can be represented in a common reference frame. Furthermore, the convergence of behavioral and neuroscientific evidence in favor of the amodal hypothesis suggests that participants in the present studies may have remembered all locations, whether acquired through vision or touch, within an amodal representation.

References

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology, 14,* 257–262.

Avraamides, M. N., & Kelly, J. W. (2010). Multiple systems of spatial memory: Evidence from described scenes. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 36,* 635–645.

Bryant, D. J. (1997). Representing space in language and perception. *Mind & Language, 12,* 239–264.

Giudice, N. A., Betty, M. R., & Loomis, J. M. (2011). Functional equivalence of spatial images from touch and vision: Evidence from spatial updating in blind and sighted individuals. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 37,* 621–634.

Giudice, N. A., Klatzky, R. L., & Loomis, J. M. (2009). Evidence for amodal representations after bimodal learning: Integration of haptic-visual layouts into a common spatial image. *Spatial Cognition and Computation, 9,* 287–304.

Greenauer, N., & Waller, D. (2008). Intrinsic array structure is neither necessary nor sufficient for nonegocentric coding of spatial layouts. *Psychonomic Bulletin & Review, 15,* 1015–1021.

Hay, J. C., Pick, H. L., & Ikeda, K. (1965). Visual capture produced by prism spectacles. *Psychonomic Science, 2,* 215–216.

Kelly, J. W. (2011). Head for the hills: The influence of environmental slant on spatial memory organization. *Psychonomic Bulletin & Review, 18,* 774–780.

Kelly, J. W., & Avraamides, M. N. (2011). Cross-sensory transfer of reference frames in spatial memory. *Cognition, 118,* 444–450.

Kelly, J. W., Avraamides, M. N., & Loomis, J. M. (2007). Sensorimotor alignment effects in the learning environment and in novel environments. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 33,* 1092–1107.

Kelly, J. W., & McNamara, T. P. (2008). Spatial memories of virtual environments: How egocentric experience, intrinsic structure, and extrinsic structure interact. *Psychonomic Bulletin & Review, 15,* 322–327.

Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In C. Freksa, C. Habel, & K. F. Wender (Eds.), *Lecture notes in artificial intelligence: Spatial cognition* (pp. 1–17). Berlin, Germany: Springer-Verlag.

Loomis, J. M., Klatzky, R. L., Avraamides, M., Lippa, Y., & Golledge, R. G. (2007). Functional equivalence of spatial images produced by perception and spatial language. In F. Mast & L. Jäncke (Eds.), *Spatial processing in navigation, imagery, and perception* (pp. 29–48). New York, NY: Springer.

Loomis, J. M., Lippa, Y., Klatzky, R. L., & Golledge, R. G. (2002). Spatial updating of locations specified by 3-D sound and spatial language. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 28,* 335–345.

Macaluso, E., & Driver, J. (2005). Multisensory spatial interactions: A window onto functional integration in the human brain. *Trends in Neurosciences, 28,* 264–271.

Mou, W., & McNamara, T. P. (2002). Intrinsic frames of reference in spatial memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 28,* 162–170.

Mou, W., Zhang, K., & McNamara, T. P. (2004). Frames of reference in spatial memory acquired from language. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 30,* 171–180.

Newell, F. N., Woods, A. T., Mernagh, M., & Bulthoff, H. H. (2005). Visual, haptic and crossmodal recognition of scenes. *Experimental Brain Research, 161,* 233–242.

Pavani, F., Spence, C., & Driver, J. (2000). Visual capture of touch: Out of the body experiences with rubber gloves. *Psychological Science, 11,* 353–359.

Rock, I., & Victor, J. (1964). Vision and touch: An experimentally created conflict between the two senses. *Science, 143,* 594–596.

Shelton, A. L., & McNamara, T. P. (1997). Multiple views of spatial memory. *Psychonomic Bulletin & Review, 4,* 102–106.

Shelton, A. L., & McNamara, T. P. (2001). Systems of spatial reference in human memory. *Cognitive Psychology, 43,* 274–310.

Warren, D. H., Welch, R. B., & McCarthy, T. J. (1981). The role of visual-auditory "compellingness" in the ventriloquism effect: Implications for transitivity among the spatial senses. *Perception & Psychophysics, 30,* 557–564.

Wolbers, T., Klatzky, R. L., Loomis, J. M., Wutte, M. G., & Giudice, N. A. (2011). Modality-independent coding of spatial layout in the human brain. *Current Biology, 21,* 984–989.

Yamamoto, N., & Shelton, A. L. (2009). Orientation dependence of spatial memory acquired from auditory experience. *Psychonomic Bulletin & Review, 16,* 301–305.