2008

Spatial memories of virtual environments: How egocentric experience, intrinsic structure, and extrinsic structure interact

Jonathan W. Kelly, *Vanderbilt University*
Timothy P. McNamara, *Vanderbilt University*

Spatial memories of virtual environments: How egocentric experience, intrinsic structure, and extrinsic structure interact

Jonathan W. Kelly and **Timothy P. McNamara**
Vanderbilt University

Abstract

Previous research has uncovered three primary cues that influence spatial memory organization: egocentric experience, intrinsic structure (object-defined), and extrinsic structure (environment-defined). The current experiments assessed the relative importance of these cues when all three are available during learning. Participants learned layouts from two perspectives in immersive virtual reality. In Experiment 1, axes defined by intrinsic and extrinsic structures were in conflict and learning occurred from two perspectives, each aligned with either the intrinsic or extrinsic structure. Spatial memories were organized around a reference direction selected from the first perspective, regardless of its alignment with intrinsic or extrinsic structures. In Experiment 2, axes defined by intrinsic and extrinsic structures were congruent, and spatial memories were organized around reference axes defined by those congruent structures rather than the initially experienced view. Findings are discussed in the context of spatial memory theory as it relates to real and virtual environments.

Memories for learned environments play a critical role in many spatial tasks, from selecting an alternative route home during rush hour to finding the kitchen in the dark. Much of the relevant experimental work has focused on spatial memory organization, and a nearly ubiquitous finding in this research is that spatial memories are represented with respect to one or two reference directions (see McNamara, 2003 for an overview). These reference directions are selected through a combination of egocentric experience within the learned environment and properties of the environmental structure (e.g., Shelton & McNamara, 2001). Both can both be considered cues which influence the reference system used to represent spatial memories. Three broad categories of these cues have emerged from this research: egocentric experience, extrinsic structure, and intrinsic structure.

Because the body is typically used to act on the remembered environment, selecting the primary reference axis from the learning perspective seems prima facie to be a natural way to organize spatial memories. In fact, early results indicated that egocentric experience is dominant under certain conditions. After learning object locations from multiple perspectives, subsequent recall and recognition is best for the learned perspectives and degrades with increasing angular distance from those learned perspectives (Diwadkar & McNamara, 1997; Shelton & McNamara, 1997). While some experiments report that spatial memories are organized around multiple reference axes based on multiple experienced views, others have found that a single reference direction is established from the initial learning perspective, and this organization persists even after extensive experience with other perspectives (Avraamides & Kelly, 2005; Kelly, Avraamides & Loomis, in press).

Address correspondence: Jonathan W. Kelly, Department of Psychology, Vanderbilt University, 111 21st Ave South, Nashville, TN 37203, Email: E-mail: jonathan.kelly@vanderbilt.edu.

Other work has demonstrated an important role for extrinsic structure, whereby salient environmental features influence reference frame selection. Werner and Schmidt (1999; also see Montello, 1991) found that residents of a German city were best at recalling relative object locations from imagined perspectives parallel to major city streets compared with oblique perspectives. Other environmental features such as lakes and nearby buildings can have a similar impact on spatial memories (McNamara, Rump & Werner, 2003). Furthermore, extrinsic structure has been shown to interact with egocentric experience in predictable ways. Shelton and McNamara (2001) had participants learn object locations from two perspectives within a rectangular room. One perspective was aligned and one was misaligned with the extrinsic axes redundantly provided by the room walls and a square mat on the floor. The sequence of experienced perspectives was varied such that half of the participants learned the misaligned view first and the aligned view second, and the other half received the reverse viewing order. Subsequent retrieval of relative object locations was best when imagining the aligned learning perspective. Performance when imagining the misaligned learning perspective was no better than for novel perspectives that were never experienced, and this held true regardless of the order in which the aligned and misaligned perspectives were experienced during learning. However, when the extrinsic structures defined by the room and mat were misaligned with one another by 45°, performance was best when imagining the initially experienced view, regardless of the learning perspective order. Thus, egocentric experience appears to dominate in the absence of a single coherent extrinsic structure. When a reliable extrinsic axis is present and learning occurs from a perspective aligned with that axis, spatial memories are influenced by the extrinsic axis rather than egocentric experience. Interestingly, our daily environments commonly have multiple potential extrinsic structures misaligned with one another, such as streets, tree lines, waterfronts and buildings. If these separate environmental structures compete with each other, then egocentric cues may predominate.

In addition to environmental features, the layout of the objects themselves can provide a salient intrinsic structure. Mou and McNamara (2002) had participants learn a regular arrangement of objects organized in rows and columns, viewed from a single learning perspective oblique to the intrinsic axes. The object layout consisted of two primary orthogonal axes defined by the object locations. During learning, the experimenter emphasized the intrinsic structure of the layout through verbal instruction, and subsequent testing showed that the objects were represented along reference axes defined by the intrinsic rows and columns. This occurred despite the fact that observers never actually experienced a perspective aligned with the intrinsic axes.

The experiments reported here are aimed at understanding how cues provided by egocentric experience, extrinsic structure, and intrinsic structure interact when learning a spatial layout. Although previous experiments have highlighted the individual importance of each cue, it is unknown how spatial memories are formed when all three are available during learning. This represents an important step toward understanding the organization of spatial memories for our daily environments, which commonly contain numerous potential reference frames. In these experiments, spatial layouts were learned in immersive virtual reality (VR), where movement through the virtual environment was achieved by physically turning and walking, just like in a real environment. VR has previously been used to replicate and extend real world findings on spatial memory retrieval (e.g., Kelly et al., in press; Williams, Narasimham, Westerman, Rieser, & Bodenheimer, in press). However, the use of VR in studying spatial memory encoding has produced mixed results. Some studies report substantial differences between spatial memories acquired through real world navigation and those learned through desktop VR, where movement through the virtual environment is controlled by joystick or keyboard (Richardson, Montello & Hegarty, 1999). But these findings may be due to the unnatural movement through desktop VR, compared to the physical rotations and translations that accompany real world exploration. Indeed, other work indicates that complex foraging

tasks requiring accurate spatial memories can benefit greatly from physical movement, and that physically walking while navigating through a virtual environment (Ruddle & Lessels, 2006) produces performance on par with performance in real environments (Lessels & Ruddle, 2005). Further work using immersive VR has indicated the importance of egocentric experience when objects are learned in a virtual environment devoid of strong environmental axes (Kelly et al., in press). This finding suggests that spatial memories of virtual environments may be structured similarly to memories for real environments.

Experiment 1

Experiment 1 was designed to understand how egocentric experience, extrinsic structure, and intrinsic structure affect reference frame selection when all three cues are available during learning. Participants learned a layout of objects in a virtual environment that contained two sets of environment-defined axes: one defined by the columnar organization of the virtual objects (intrinsic) and the other defined by the square walls of the virtual room (extrinsic). These two sets of axes were misaligned with one another by 45°, and this environment is referred to as the incongruent environment (Figure 1, left panel). All participants studied the objects from two perspectives, one that was aligned with the intrinsic structure but misaligned with the extrinsic structure (the 0° view in Figure 1, left) and one that was aligned with the extrinsic structure but misaligned with the intrinsic structure (the 135° view in Figure 1, left), and viewing order was manipulated. If either the intrinsic or extrinsic structure is more salient than the other, then recall should be best for perspectives aligned with that salient reference frame, regardless of the viewing order. If, on the other hand, the conflicting intrinsic and extrinsic structures negate the influence one another, then egocentric experience should dominate and the preferred reference axis should correspond to the initially experienced view.

Additionally, Experiment 1 was designed to provide supporting evidence for current spatial memory theories using virtual environments. In order for VR to be a useful tool to study complex spatial behaviors like navigation and route planning, spatial memories for simple, well-controlled virtual environments should be organized similarly to spatial memories for real environments. To our knowledge, this is the first thorough investigation of reference frames in spatial memories for immersive virtual environments.

Method

Participants—Twenty-four participants (12 males) from the Nashville community participated in exchange for monetary compensation.

Stimuli and Design—Stimuli were presented through an nVisor SX (from NVIS, Reston, VA) head mounted display (HMD). The HMD presented stereoscopic images at 1280 × 1024 resolution, refreshed at 60 Hz. The HMD field of view was 47° horizontal by 38° vertical. Graphics were rendered by a 3.0 GHz Pentium 4 processor with a GeForce 6800 GS graphics card using Vizard software (WorldViz, Santa Barbara, CA). Head orientation was tracked using a three-axis orientation sensor (InertiaCube2 from Intersense, Bedford, MA) and head position was tracked using a passive optical tracking system (Precision Position Tracker, PPT X4 from WorldViz, Santa Barbara, CA). Graphics displayed in the HMD were updated based on sensed position and orientation of the participant's head.

The virtual environment consisted of seven virtual objects (shark, ball, train, bug, cup, lamp, plant) placed on identical green pillars that were 60 cm tall. Each object was scaled to fit within a 20 cm cube. Objects were arranged along a rectilinear grid, evenly spaced every 75 cm (see Figure 1). Objects were arranged in three columns, the left and middle column containing three objects and the right column containing one object. In addition, a square (3 m) virtual room surrounded the scene. The room was rotated about its vertical axis to be misaligned by 45°

with the intrinsic structure. Room walls were covered with a brick pattern containing strong texture gradient depth cues, and the floor was covered with a random carpet texture.

All participants learned the object locations from two perspectives, 0° and 135° (corresponding to two unique viewing locations; see Figure 1), and viewing order was manipulated. Figure 2 shows the participant's perspective of the incongruent environment from 135°. After learning, participants were removed from the virtual environment and performed judgments of relative direction (JRD) in which they were asked to imagine standing at the location of one object, facing a second object, and then point to a third object from that imagined perspective. JRD trials were presented on a 19″ computer monitor and pointing responses were made using a joystick (Freedom 2.4 by Logitech, Freemont, CA) placed on the desk in front of the participant.

The independent variables were viewing order (0° then 135° or 135° then 0°) and imagined perspective in the JRD task. Viewing order was manipulated between participants and imagined perspective was manipulated within participants. JRD stimuli were selected to test eight different imagined perspectives, spaced every 45° from 0° to 315°. Objects appeared as standing, orienting, and pointing objects with equal frequency. For each imagined perspective, six trials were constructed to create correct egocentric pointing directions of 45°, 90°, 135°, 225°, 270°and 315°, resulting in 48 total trials.

The dependent measures were pointing error (the absolute angular error of the pointing response) and pointing latency (the latency between presentation of the three object names for a given trial and completion of the pointing response).

Procedure

Learning phase: The experimenter met the participant in front of the learning room. After providing informed consent, the participant donned the HMD while standing in the hallway. Once the HMD was in place, the participant was led into the lab and positioned for the first view (either 0° or 135°). During transport, the blank HMD display served as a blindfold, blocking any view of the physical lab structure. Additionally, the lab lights were turned off to prevent viewing the lab around the edges of the HMD.

Once the participant was properly positioned at the first viewing location and the HMD was turned on, the experimenter named each of the test objects in a random sequence. To direct the participant's attention to a specific location, the experimenter pressed a key to temporarily change the color of the pillar underneath the object from green to white. After all of the objects were named, the participant was instructed to study the layout for 30 s. During learning, participants were told not to move from the study location but were free to turn their heads to look around, and participants regularly did so in order to study all objects. Following this study interval, the objects were hidden but the pillars on which they rested remained. The experimenter then indicated each pillar in a random sequence and asked the participant to name the corresponding object from that location. This learning sequence was repeated until the participant successfully named all objects twice. Typically this occurred after two or three learning intervals.

After learning from the first location, the display was turned off (serving as a blindfold during transport) and the participant walked, guided by the experimenter, to the second viewing location (approximately 3 m from the first viewing location). When the display was turned back on, the participant's view of the virtual environment reflected the physical movements that they underwent during transport between the first and second viewing locations. This procedure simulates those of Shelton and McNamara (2001), where participants walked blindfolded from the first to the second viewing location. After completion of the same learning

procedures from the second perspective, the display was turned off again and the participant was led back out into the hallway and the HMD was subsequently removed.

Testing phase: After learning, the participant was led to another room on the same floor of the building to complete the JRD task. The participant was seated in front of a computer monitor and completed four practice JRD trials using either buildings on the Vanderbilt campus or U.S. cities, depending on participant familiarity. After practice, the participant completed 48 trials composed of previously studied objects. Each trial was initiated when the participant pressed a button on the joystick, at which point the imagined standing object, facing object, and target object were simultaneously displayed in text (e.g., "Imagine standing at the train, facing the shark. Point to the plant."). Trials were pseudo-randomized with the constraint that the same perspective was never tested twice in a row. The participant responded by moving the joystick in the direction of the target object from the imagined perspective. Each response was recorded when the joystick was deflected by 30° from vertical.

Results

Pointing error was more responsive to the independent variables than was pointing latency, which was largely unaffected[1]. In the interest of brevity, we focus on the error data. Pointing error (plotted in Figure 3 as a function of imagined perspective) was analyzed in a 2 (gender) $\times$ 2 (viewing order) $\times$ 8 (imagined perspective) mixed-model ANOVA. The significant main effect of imagined perspective [$F(7,140)=2.77$, $p=0.01$, $\eta_p^2=.12$] was qualified by a significant interaction between imagined perspective and viewing order [$F(7,140)=2.18$, $p=0.04$, $\eta_p^2=.10$]. No other main effects or interactions were significant. Participants were more accurate on the imagined perspective corresponding to the first learning perspective (0° for the 0°-135° learning order and 135° for the 135°-0° learning order) than the second learning perspective, as indicated by the interaction contrast [$F(1,20)=6.03$, $p=0.023$, $\eta_p^2=.23$]. There was no overall benefit for imagined perspectives aligned with either the intrinsic (i.e., the 0°, 90°, 180° and 270° perspectives) or the extrinsic (i.e., the 45°, 135°, 225° and 315° perspectives) axes.

Discussion

In Experiment 1, the incongruent intrinsic and extrinsic environmental structures resulted in spatial memories with a preferred reference direction selected from the initially experienced perspective. Although one learning perspective was aligned with the intrinsic structure and one was aligned with the extrinsic structure, neither had a consistent impact across the two viewing orders. These results parallel findings by Shelton and McNamara (2001), who manipulated learning perspective order in the presence of two incongruent extrinsic structures defined by the room walls and a mat on the floor. Under these conditions, they found that the primary reference direction used to structure spatial memories was selected from the initially viewed perspective, regardless of the viewing order. This provides initial evidence that intrinsic and extrinsic structures exert similar influences on spatial memory, and that neither is preferentially used when they are placed in competition.

By replicating and extending a primary finding on reference frame congruency, Experiment 1 also serves to further validate the use of immersive VR technology for studying spatial memory. This finding provides further support for the use of VR, which has been previously used demonstrate the role of egocentric experience (Kelly et al., in press) but not environmental structure.

---

[1]Latency did not vary as a function of imagined perspective, and was slightly but non-significantly greater for the 0°-135° than the 135°-0° learning order.

Experiment 2

Experiment 2 was designed to investigate the roles of egocentric experience, intrinsic structure, and extrinsic structure when the latter two are congruent with one another. To achieve this, the axes defined by the object layout and the room walls were placed in alignment, and this environment is referred to as the congruent environment (see Figure 1, right). Results from Experiment 1 indicate that intrinsic and extrinsic structures interact with each other in a manner functionally equivalent to two extrinsic structures. As such, we anticipated that congruent intrinsic and extrinsic structures would produce spatial memories organized around their congruent axes, regardless of viewing order. This prediction is based on previous results using environments with congruent extrinsic structures (Shelton & McNamara, 2001).

Methods

Participants—Twenty-four participants (12 males) from the Nashville community participated in exchange for monetary compensation. One female participant was replaced due to failure to follow experimenter instructions.

Stimuli, Design and Procedures—Design and procedures were identical to Experiment 1. Stimuli were modified so that intrinsic and extrinsic structures defined congruent axes (see Figure 1, right). The 0° learning view was aligned with both structures and the 135° view was misaligned with both structures. Viewing order was manipulated between participants and imagined perspective was manipulated within participants.

Results

Pointing error (plotted in Figure 4) was analyzed in a 2 (gender) $\times$ 2 (viewing order) $\times$ 8 (imagined perspective) mixed-model ANOVA. This analysis revealed only a main effect of imagined perspective [$F(7,140)=9.71$, $p<0.001$, $\eta_p^2=.33$], where performance was best on perspectives aligned with the intrinsic and extrinsic structures (0°, 90°, 180°and 270°) compared to misaligned perspectives (45°, 135°, 225°and 315°) [$F(1,20)=53.17$, $p<0.001$, $\eta_p^2=.73$]. No other main effects or interactions were significant.

Discussion

When axes defined by intrinsic and extrinsic structures were congruent with one another, spatial memories were organized around the orthogonal reference axes redundantly defined by the environment. Performance on the 135° perspective was no better than other novel misaligned perspectives, and performance on the 0° perspective was no better than other novel aligned perspectives, indicating that memories were organized around environmentally-defined axes and not based strictly on egocentric experience. This replicates previous findings with congruent extrinsic structures (Shelton & McNamara, 2001), and serves to further validate immersive VR as a tool for studying spatial memory.

General Discussion

The primary purpose of these experiments was to evaluate the interaction between egocentric experience, intrinsic structure, and extrinsic structure in a spatial memory task. Previous work has indicated that the effect of multiple extrinsic structures (such as the room walls and a mat on the floor) depends on their congruency with one another. Shelton and McNamara (2001) found that when extrinsic structures were incongruent, spatial memories were organized around a reference direction selected from the initially experienced perspective. When those same extrinsic structures were congruent, spatial memories were organized around the redundantly defined axes rather than the initially experienced view. Likewise, incongruent intrinsic and extrinsic structures in Experiment 1 resulted in egocentric selection of a reference direction

from the initially learned perspective. Neither intrinsic nor extrinsic structure was preferentially used when learning the incongruent environment. Instead, incongruency between multiple types of environmental structures resulted in egocentric spatial memories. Because environments often contain multiple potential reference frames such as roads, trees, and hills, it is possible that egocentric experience predominates spatial memories of our daily environments. However, further research is needed to determine how differences in spatial memory organization might lead to differences in navigation performance within those environments. Additionally, the intrinsic structure provided by the columnar object organization in these experiments is rather artificial, and it remains to be seen how more naturalistic structures influence spatial memories.

In Experiment 2, when intrinsic and extrinsic structures were congruent with one another, spatial memories were organized around the congruent environmental axes rather than egocentric experience. We hypothesize that when learning first occurred from a perspective misaligned with the congruent intrinsic and extrinsic structures (the 135° learning view in Experiment 2), participants selected an egocentric reference direction from that misaligned perspective. However, subsequent learning from an aligned perspective (the 0° learning view in Experiment 2) resulted in restructuring of the memory around the redundantly provided intrinsic and extrinsic axes. Recent experiments have shown that this restructuring occurs within 10 to 30 s of learning from the aligned view (Valiquette, McNamara & Lebrecque, 2007).

Previous experiments using immersive VR have found that spatial memories for virtual environments lacking strong extrinsic or intrinsic structure are heavily influenced by egocentric experience (Kelly et al., in press), an important replication of similar findings in real environments (Diwadkar & McNamara, 1997; Shelton & McNamara, 1997). The current experiments extend this work to more complex virtual environments, containing multiple environmental structures. Based on the current evidence, it appears that spatial memories for virtual environments are organized similarly to memories for real environments, and are similarly responsive to environment structure and egocentric experience. VR holds obvious advantages for studying spatial memory due to the ease with which novel environments can be created, from room-sized to city-sized environments. The current experiments, however, only address room-sized environments. Other studies using joystick-based locomotion have shown substantial differences in spatial memories for real and virtual large-scale environments (Richardson et al., 1999). It is unclear if these divergent findings result from differences in environmental scale or differences in exploration mode, and further work is needed to disentangle these factors. Furthermore, learning in these experiments only occurred from two perspectives, whereas natural exploration typically involves learning from many perspectives. It remains to be seen how these results extend to more complex learning conditions.

Acknowledgments

References

Avraamides, MN.; Kelly, JW. Imagined perspective-changing within and across novel environments. In: Freksa, C.; Nebel, B.; Knauff, M.; Krieg-Bruckner, B., editors. Lecture Notes in Artificial Intelligence: Spatial Cognition IV. Berlin: Springer-Verlag; 2005. p. 245-258.

Diwadkar VA, McNamara TP. Viewpoint dependence in scene recognition. Psychological Science 1997;8(4):302–307.

Kelly JW, Avraamides MN, Loomis JM. Sensorimotor alignment effects in the learning environment and in novel environments. Journal of Experimental Psychology: Learning, Memory, and Cognition. in press

Lessels S, Ruddle RA. Movement around real and virtual cluttered environments. Presence: Teleoperators and Virtual Environments 2005;14:580–596.

McNamara, TP. How are the locations of objects in the environment represented in memory?. In: Freksa, C.; Brauer, W.; Habel, C.; Wender, KF., editors. Lecture Notes in Artificial Intelligence: Spatial cognition III. Berlin: Springer-Verlag; 2003. p. 174-191.

McNamara TP, Rump B, Werner S. Egocentric and geocentric frames of reference in memory of large-scale space. Psychonomic Bulletin & Review 2003;10(3):589–595. [PubMed: 14620351]

Montello DR. Spatial orientation and the angularity of urban routes: A field study. Environment and Behavior 1991;23(1):47–69.

Mou W, McNamara TP. Intrinsic frames of reference in spatial memory. Journal of Experimental Psychology: Learning, Memory, and Cognition 2002;28(1):162–170.

Richardson AE, Montello DR, Hegarty M. Spatial knowledge acquisition from maps and from navigation in real and virtual environments. Memory & Cognition 1999;27(4):741–750.

Ruddle RA, Lessels S. For efficient navigational search humans require full physical movement but not a rich visual scene. Psychological Science 2006;17:460–465. [PubMed: 16771793]

Shelton AL, McNamara TP. Multiple views of spatial memory. Psychonomic Bulletin & Review 1997;4 (1):102–106.

Shelton AL, McNamara TP. Systems of spatial reference in human memory. Cognitive Psychology 2001;43(4):274–310. [PubMed: 11741344]

Valiquette CM, McNamara TP, Labrecque JS. Biased representations of the spatial structure of navigable environments. Psychological Research 2007;71:288–297. [PubMed: 16957954]

Werner S, Schmidt K. Environmental reference systems for large-scale spaces. Spatial Cognition and Computation 1999;1(4):447–473.
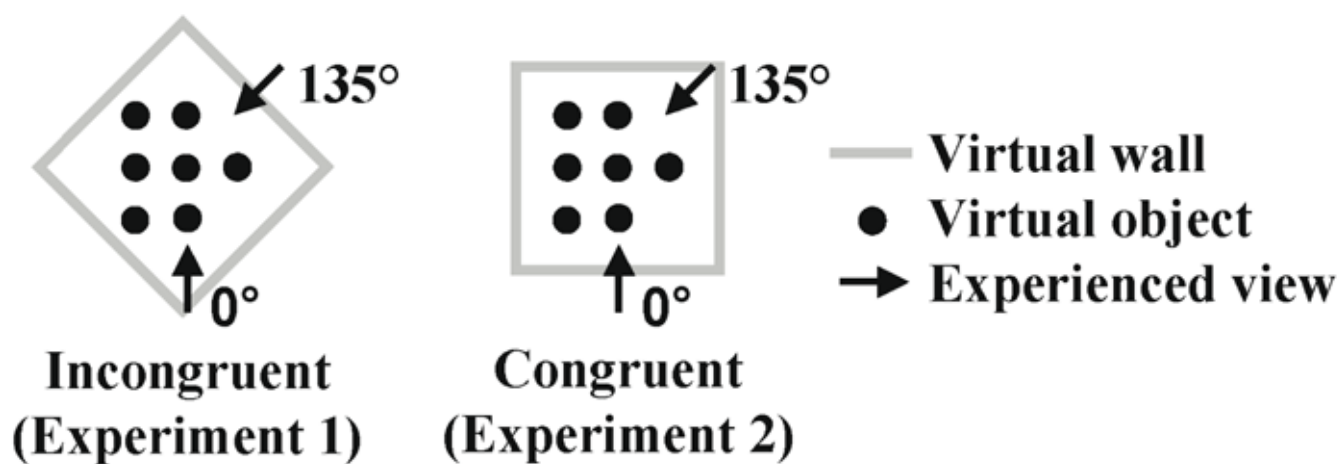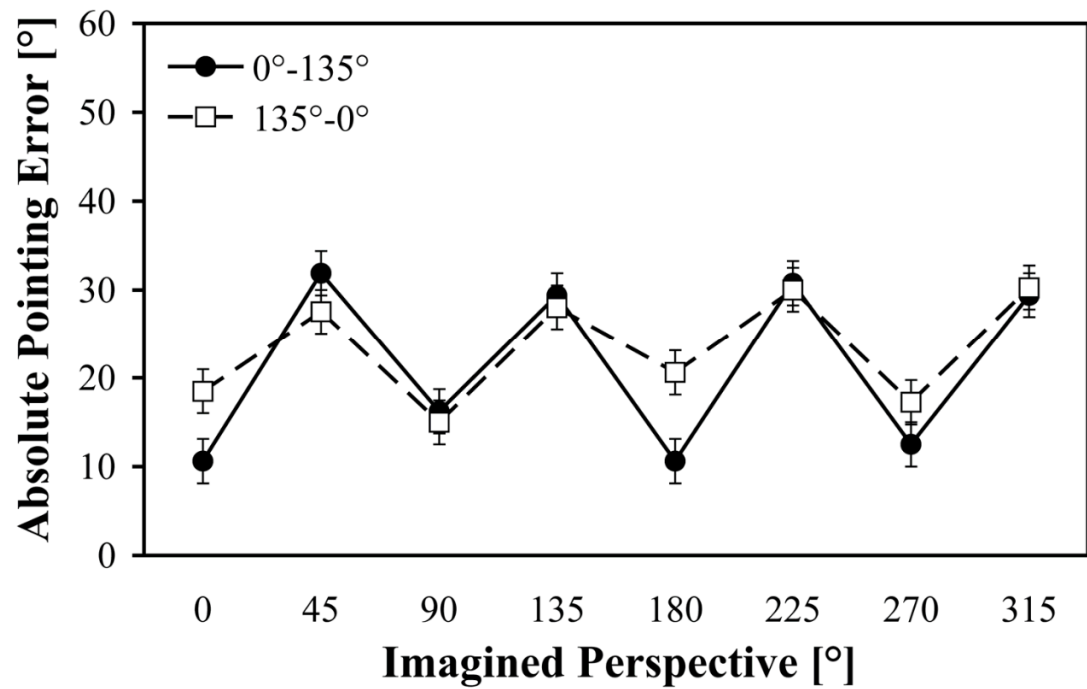
Figure 1.
Stimuli used in Experiment 1 (left panel) and Experiment 2 (right panel).

Figure 2.
The participant's view of the incongruent environment from the 135° perspective. Some objects were enlarged for publication purposes.
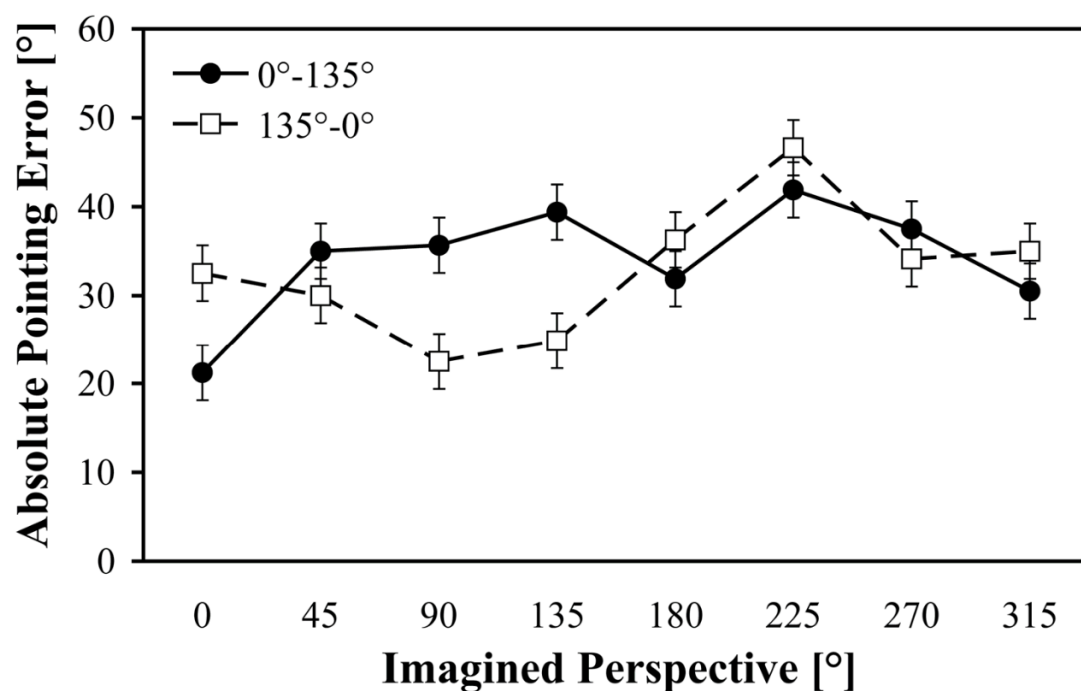
Figure 3.
Absolute pointing error in Experiment 1 after learning the incongruent environment, plotted separately for 0°-135° and 135°-0° viewing orders. Error bars are standard errors estimated from the ANOVA.

Figure 4.
Absolute pointing error in Experiment 2 after learning the congruent environment, plotted separately for 0°-135° and 135°-0° viewing orders. Error bars are standard errors estimated from the ANOVA.