Psychological Research

Orientation-dependent spatial memories for scenes viewed on mobile devices

--Manuscript Draft--

Manuscript Number:	PRPF-D-18-00109
Full Title:	Orientation-dependent spatial memories for scenes viewed on mobile devices
Article Type:	Original Article
Corresponding Author:	Marios Avraamides University of Cyprus Nicosia, CYPRUS
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of Cyprus
Corresponding Author's Secondary Institution:	
First Author:	Savvas Avraam
First Author Secondary Information:	
Order of Authors:	Savvas Avraam
	Adamantini Hatzipanayioti
	Marios Avraamides
Order of Authors Secondary Information:	
Funding Information:	
Abstract:	We examined whether spatial representations for scenes experienced on the screens of mobile devices are orientation dependent and whether the type of movement (physical vs. simulated) during learning affects the encoding and the retrieval of spatial information. Participants studied a spatial layout depicted on a tablet and then carried out perspective taking trials in which they localized objects from imagined perspectives. Depending on condition, participants either rotated the tablet along with their body or remained stationary and swiped with their finger on the screen to change their viewpoint within the scene. Results showed that participants were faster and more accurate to point to objects from an imagined perspective that was aligned than misaligned to their initial physical orientation during learning, suggesting that they had formed an orientation-dependent representation. Although no differences were found between movement conditions during pointing, participants were faster to encode spatial information with physical than simulated movement.
Suggested Reviewers:	Jan Wiener Bournemouth University jwiener@bournemouth.ac.uk expert in spatial cognition
	Victor Schinazi ETH Zurich victor.schinazi@gess.ethz.ch expert in spatial cognition and technology
	Francesca Pazzaglia Universita degli Studi di Padova francesca.pazzaglia@unipd.it expert in spatial cognition
	Weimin Mou University of Alberta Department of Resource Economics and Environmental Sociology weimin.mou@ualberta.ca expert in spatial cognition, did similar work

Orientation-dependent spatial memories for scenes viewed on mobile devices

Savvas Avraam[1,2], Adamantini Hatzipanayioti[3], & Marios N. Avraamides[1,4]

[1] Department of Psychology, University of Cyprus

[2] Silversky3D Virtual Reality Technologies Ltd

[3] Max-Planck-Institute for Biological Cybernetics

[4] RISE

Author Note:

The research presented here was part of S.A’s MA thesis. We thank Maria Fotiou for help with data collection. Address Correspondence to Marios N. Avraamides, Department of Psychology, University of Cyprus, Email: mariosav@ucy.ac.cy, Telephone: +357 22892066.

Abstract

We examined whether spatial representations for scenes experienced on the screens of mobile devices are orientation dependent and whether the type of movement (physical vs. simulated) during learning affects the encoding and the retrieval of spatial information. Participants studied a spatial layout depicted on a tablet and then carried out perspective taking trials in which they localized objects from imagined perspectives. Depending on condition, participants either rotated the tablet along with their body or remained stationary and swiped with their finger on the screen to change their viewpoint within the scene. Results showed that participants were faster and more accurate to point to objects from an imagined perspective that was aligned than misaligned to their initial physical orientation during learning, suggesting that they had formed an orientation-dependent representation. Although no differences were found between movement conditions during pointing, participants were faster to encode spatial information with physical than simulated movement.

Orientation-dependent spatial memories for scenes viewed on mobile devices

Advances in modern technology have provided us with new ways to experience spatial environments. We can now view environments such as tourist sites, amusement parks, and surgery rooms, as immersive scenes within Virtual Reality (VR) head-mounted-displays, as mixed-reality scenes with Augmented Reality (AR) glasses, as 360° panoramas on the screens of our mobile devices, and so on. A question that arises about spatial cognition, is how similar the spatial memories created from such modern experiences are to those acquired through the direct experience of the physical environment.

Past research in spatial cognition suggests that the organizational structure of spatial representations constructed in VR and AR does not differ from those acquired by direct perception. For example, a study by Kelly, Avraamides, & Loomis (2007) showed that participants who have memorized objects in an immersive virtual environment from a particular learning orientation, pointed more accurately and/or faster to objects from imagined perspectives that were aligned than misaligned with the learning orientation. This finding, indicative of orientation-dependent spatial memories, is also reported by studies involving perceptual (Mou, McNamara, Valiquette, & Rump, 2004) and described objects (Avraamides & Kelly, 2008) within physical environments, as well as virtual objects embedded in real environments through AR (Mou, Biocca, Owen, Tang, Xiao, & Lim, 2004). In addition to this advantage for the learning orientation, the aforementioned studies have also shown an advantage for the orientation occupied during testing; that is, participants also pointed more efficiently when the imagined perspective coincided with their physical orientation during testing, suggesting that the constructed spatial representations in all cases were not only stored in a preferred orientation, but that they were also updated by movement following learning.

The fact that both direct perception and immersive VR or AR yield converging findings may not be so surprising: in both cases the observer is physically embedded in the spatial layout; given the similarity of these learning situations, any changes in the observer's orientation can support the successful updating of egocentric (i.e., self-to-object) relations.

The experience is very different though when experiencing distal scenes on external screens, where the user is decontextualized from the depicted environment. In these cases, movement in the scene is often simulated (e.g., effected by touch and swipe gestures on the screen of a mobile device) and thus decoupled from any physical changes in the actual location and orientation of the observer. In many instances, however, movement is linked to changes in orientation even when the observer is not physically embedded in the scene (e.g., when observers rotate their viewpoint in the scene by physically turning the device along with their body). To our knowledge, no study so far has compared directly the encoding of spatial relations in distal scenes resulting from simulated movement during learning with the encoding of the same spatial relations through physical user movement that is now possible with modern mobile devices.

Based on previous findings from research with real, virtual, and augmented reality environments (e.g., Kelly et al., 2007; Mou et al., 2004; Mou et al., 2004), the goal of the present study is twofold. First, it aims to extend past findings by investigating whether distal scenes encoded through mobile devices are also orientation dependent. Second, it aims to examine whether the type of movement employed when experiencing the scene on a mobile screen influences the nature or the fidelity of the resulting representation. Although past research shows that people become accustomed to simulated movement by touch on mobile devices even before the age of 2 (Rideout & Saphir, 2013), it is not yet known whether such simulated movement has disadvantages for the perception and memory of depicted environments compared to natural rotations of the self.

Previous research on spatial updating has documented the importance of physical body movement for monitoring changes in egocentric relations following learning (e.g., Loomis, Lippa, Golledge, & Klatzky, 2002; Presson & Montello, 1994; Rieser, 1989; Rieser, Guth, & Hill, 1986). Rieser (1989) argued that the idiothetic information (i.e., proprioceptive information, vestibular signals, and efference copy) that is available during physical movement allows for the effortless updating of egocentric locations concurrently with movement. In contrast, when idiothetic information is lacking, as in the case of imaginal or simulated movement, localizing objects from novel points of observation entails, according to Rieser (1989), effortful computational processing at the end of the movement to determine how egocentric spatial relations have changed.

Overall, findings from studies on spatial updating suggest that physical movement is critical for maintaining our orientation within our surroundings following learning. However, being oriented within an environment must be also important during encoding; when we start to explore an environment, being oriented to our surroundings may allow us to integrate information across successive views in order to construct an accurate mental representation for the space around us. Based on this assumption, a hypothesis worth exploring is that constructing a spatial representation is more difficult with simulated than physical movement, due to the lack of idiothetic information to support the quick and accurate orientation during exploration. Moreover, in addition to easier learning, it is possible that studying a spatial scene from multiple perspectives adopted through physical movement, stores in memory sensorimotor information that could serve as cues for subsequent retrieval, leading to better overall spatial performance than when studying the same spatial scene through simulated movement. Finally, another possibility is that experiencing the scene from multiple physical perspectives gives rise to orientation-independent performance. Although such a prediction is not supported by studies with immediate scenes (e.g., Kelly et al., 2007; but see Shelton &

McNamara, 1997 for evidence that more than one experienced perspectives may exhibit a performance benefit), the tethering of physical movement to distal scenes during learning, may yield different results. This possibility is examined in the current study.

In summary, in the present study we investigated 1) whether spatial memories of distal environments experienced as 360° panoramas on mobile screens are stored in a preferred orientation, 2) whether actual self-rotations during the encoding of the scene, compared to simulated movement via swiping, leads to a) faster encoding, and b) superior spatial memory overall and/or orientation-independent performance.

To examine these hypotheses, we carried out an experiment in which we asked participants to study and memorize the locations of objects of a simple spatial scene depicted on a tablet. All participants started exploring the scene from the same initial orientation. In the *move condition*, participants changed their viewpoint in the scene by physically rotating their body while holding the tablet whereas in the *no-move condition* they remained stationary and rotated their viewpoint in the scene by swiping the screen with their finger. In order to assess whether physical movement leads to faster encoding than simulated movement, we compared the time participants took to learn the scenes in the two conditions. We then asked participants to carry out a computer-based task that required them to point to the locations of memorized objects in the spatial scene from imagined perspectives. Pointing error and response latency were recorded and analyzed to examine whether physical movement leads to more accurate and/or faster overall performance than simulated movement. Moreover, based on past research showing that, in the absence of other cues, the first orientation experienced determines the preferred orientation of spatial memory (Avraamides & Kelly, 2005; Hatzipanayioti, Galati, & Avraamides, 2015), we compared error and latency for perspectives that were aligned vs. misaligned to participants’ initial orientation during encoding. Specifically, our main interest was to assess whether an alignment effect (i.e. the difference between aligned and misaligned perspectives) would be present and whether its magnitude would differ across the move and no-move conditions.

Method

Participants

Twenty-four young adult volunteers (24-35 years old, M= 29.08; 12 female) with normal or corrected to normal vision participated in the experiment.

Materials and Equipment

Two spatial scenes were created using professional 3D modeling software (3Ds Max, Autodesk, San Rafael, CA) and exported to 360° spherical images. Each scene included 5 different objects occupying positions around the central viewpoint (Figure 1). The objects appeared on columns with a height of 1.1m that were placed within an 8m x 8m virtual room (Figure 2). The distance of each column from the center of the virtual room was fixed at 3m. Participants viewed the scenes on a 7-inch tablet (Google Nexus 7). A script in the Unity3D game engine (Unity Technologies, San Francisco, CA) was used to control the presentation of the 360° spherical panoramas and to record the encoding time during the learning phase. In the move condition the tablet’s built-in gyroscope was used to track participants’ orientation and update the graphics accordingly. In the no-move condition the tablet’s touch screen was used to update the graphics according to participant’s finger swiping. A Python script in the Vizard VR Toolkit (Wolrdviz, Santa Barbara, CA) was used to display stimuli and control the experimental task during the testing phase. The task was presented on a desktop computer and participants responded with a joystick placed at a comfortable position in front of them.

Design

The experiment followed a 2 (sensorimotor condition: move vs no-move) x 2 (imagined perspective: aligned vs misaligned) within-participants design. The order in which the two sensorimotor conditions were presented and the assignment of layout to each condition were counterbalanced across participants.

Procedure

Participants signed informed consents prior to the experiment and were thoroughly debriefed afterwards.

At the beginning of the experiment participants stood in the center of a dim-lit room holding the tablet. They were then asked to study a layout of 5 objects displayed on the tablet and memorize the location of each object. All participants started studying the layout facing towards the same orientation, indicated as 0° in Figure 1. Once participants were given the instructions about the experiment and indicated their readiness to start the task, the experimenter pressed an on-screen button in the upper left corner of the display to start the timer.

In the move condition, participants physically rotated their body while holding the tablet to view the scene. In the no-move condition, they remained stationary while holding the tablet and rotated the scene by swiping their finger to the right or left on the screen. They were instructed to spend as much time as necessary to memorize object locations and then press the on-screen button to stop the timer. Once they did so, participants were asked to re-adopt the initial facing orientation (0°) in the scene (also the initial physical orientation in the move condition) and hand the tablet back to the experimenter. A short test was then carried out to ensure that participants could indeed remember all object locations. While occupying the initial orientation, they were asked to point by extending their arm towards the location of each of the 5 objects as they were announced by the experimenter in a random order. The experimenter assessed visually participants' accuracy in order to restart the learning procedure if necessary. However, this was not necessary as all participants could remember correctly the 5 object locations after a single learning experience.

Participants were then guided to a nearby room to carry out the memory trials. This testing phase involved a perspective-taking task in which participants had to point to an object after imagining standing in the center of the memorized spatial scene facing another object. As shown in Figure 3, each trial was presented in 3 steps: The Orientation, the Response Computation, and the Response Execution. In the *Orientation* step, the picture of an object appeared and participants were asked to pull the trigger of the joystick when they had mentally adopted the perspective that corresponded to that object. When they did so, in the *Response Computation* step, they were presented with the target object and pulled the trigger button again as soon as they knew the location of the target relative to their imagined perspective. After doing so, in the *Response Execution* step, an on-screen pointer appeared which participants manipulated with the joystick to provide their response. Measuring pointing latency in 3 steps allowed us to assess separately the potential effect of movement type on the process of adopting an imagined perspective and that of computing a response vector from it[1]. Pointing error, measured as the unsigned angular deviation from the correct response, and response latency in each of the three steps of the trial were recorded and analyzed offline. Once the sequence of learning and testing was completed, it was repeated after a mandatory 5-minute break for the other sensorimotor condition using the other layout of objects.

Each block in the testing script contained 20 trials that involved all possible pairs of

---

[1] No differences in response execution latency should be present if participants complied fully with the instruction to compute the response before they proceeded to the response execution step.

objects presented in a random order for each participant. Participants carried out 3 blocks in each sensorimotor condition.

Results

To examine potential differences in learning, we compared the learning time (i.e., the time participants needed to encode the object locations) across the two sensorimotor conditions with a paired-samples t-test. Pointing error and response latency for the 3 steps of a trial (i.e., Orientation, Response Computation, and Response Execution) were analyzed with separate Repeated Measures Analyses of Variance (ANOVA) with terms for sensorimotor condition (move vs. no-move) and imagined perspective (aligned vs. misaligned). Aligned trials were those in which participants imagined adopting the initial orientation (0° in Figure 1) while misaligned trials were those involving the remaining 4 perspectives in each layout. Responses from misaligned perspectives were averaged to a single value.

Learning time.

The paired-samples t-test for learning time indicated that participants were faster to memorize the spatial layout in the move condition (M=114.16s, SD=62.38s) than in the no-move condition (M=145.64s, SD=71.10), t(23)=2.77, p=.011.

Pointing Error.

The ANOVA on pointing error revealed a significant effect for imagined perspective, F(1, 23)=40.53, p<.001, η^2=.64. As shown in Figure 4, participants were more accurate in localizing objects from aligned (M=12.40, SE=.94) than from misaligned imagined perspectives (M=23.35, SE=2.10). Neither the main effect for sensorimotor condition nor the interaction of imagined perspective and sensorimotor condition were significant, F(1,23)=2.67, p=.12, η^2=.10 and F(1,23)=.13, p=.73, η^2=.01 respectively.

Pointing Latency.

As shown in Figure 5, all three latency measures yielded a pattern of results that replicated the one obtained for pointing error. That is, participants were overall faster to adopt the aligned than the misaligned perspective, and also to compute and execute the response from it. This finding was corroborated with significant main effects for imagined perspective in all three measures, F(1, 23)= 47.89, $p<.001$, $\eta^2=.68$ for Orientation, F(1, 23)= 44.05, $p<.001$, $\eta^2=.66$ for Response Computation, and F(1, 23)= 10.73, $p=.003$, $\eta^2=.32$ for Response Computation. Neither a main effect for sensorimotor condition nor an interaction of imagined perspective with sensorimotor condition were found in any of the three latency measures.

Discussion

The primary aim of the study was to investigate whether spatial memories acquired from experiencing distal environments on mobile devices are orientation dependent. Indeed, in line with previous studies (e.g., Kelly, Avraamides, & Loomis, 2007) and theories on spatial memory (e.g., McNamara, 2003), our results indicated that spatial information acquired from mobile devices is stored in memory from a preferred orientation, in this case the initial orientation participants had when they first entered the spatial scene.

Results showed that participants were faster to adopt an imagined perspective that was aligned than misaligned with the learning orientation but also to compute and execute a response from it. That an effect was found in response execution suggests that not all participants conformed to our instruction to move the joystick only after they were sure about the response they would make. But, as the pattern of latency in this step mimics that of the Response Computation and Response Execution step, this does not compromise the overall finding that localizing objects from imagined perspectives is easier from the aligned than the misaligned perspective. This finding replicates past findings with immersive environments (Kelly et al., 2007) and described scenes (Hatzipanayioti et al., 2015) and documents that spatial environments viewed on mobile devices are also orientation dependent.

Notably, although strong alignment effects were found in pointing error and latencies for all the 3 processing steps of a trial, the size of these effects was equal between the two sensorimotor conditions. In addition, overall error and latency did not differ between the two conditions. These findings go against the prediction that sensorimotor information elicited by physical movement could provide additional cues for retrieval that would benefit either overall performance or performance from misaligned perspectives in particular.

However, results indicated that forming a spatial representation by viewing a scene on a tablet was faster when participants physically rotated their body towards each object than when they simulated the movement by turning their viewpoint via swiping the touchscreen. To rule out the possibility that this is a general result of physical rotation being faster to execute than simulated rotation via finger swiping, we carried out a follow-up experiment. In this experiment, we asked 24 new participants to study the same scenes and simply memorize the names of the objects, without us making any reference in the instructions about locations. Results from this supplementary experiment did not replicate the advantage for the move condition. Instead, participants were numerically faster in the no-move than in the move condition[2]. This finding suggests that the advantage for the move condition in the main experiment is not a general effect but was more likely related to the process of encoding locations in memory.

Our conjecture is that the advantage of the move condition over the no-move condition for learning is due to idiothetic information that is available with physical movement, and possibly to allothetic (i.e., visual) information as well, that allowed participants to keep track of their orientation while rotating in the spatial scene. As documented by the literature on spatial updating (e.g. Rieser, 1989), proprioceptive and vestibular signals, as well as stored

---

[2] M=17.554, SD=7.90 for the move condition, M=15.18, SD=6.22 for the no-move condition, t(23)=2.36, p=.16.

copies of efferent commands, allow the moving observer to effortlessly update self-to-object information during movement. In our case, such information could have allowed participants to quickly turn to inspect the objects and automatically update egocentric relations. Moreover, visual information from the room could have provided optic flow or other visual cues that allowed participants to easily monitor the extent of their movement. Effortless updating of one's viewpoint in the virtual scene could have in turn allowed participants to integrate more easily into the developing spatial representation the locations of objects observed at successive views.

Interestingly, the benefit of physical movement for encoding did not transfer to retrieval. Instead, it seems that once constructed, either by physical or simulated movement, the spatial representation could support the execution of the perspective taking task in the same way. This is in fact in line with the functional equivalence hypothesis, which posits that although spatial representations may be more difficult to construct from certain inputs (e.g., audition than vision) or induce modality-specific biases, once constructed they can support a spatial task in the same way regardless of the input (Loomis, Lippa, Klatzky, & Golledge, 2002).

In summary, the present study contributes 3 key findings to the literature of spatial cognition. First, it documents that spatial representations of remote scenes viewed on mobile devices are orientation dependent. Second, it shows that, compared to purely simulated movement, concurrent physical movement confers an advantage for the encoding of spatial relations in scenes experienced on a mobile device. Finally, it indicates that despite this encoding advantage for physical movement, the spatial representation resulting from simulated movement is just as efficient in supporting memory-based perspective taking.

In addition to informing the field of spatial cognition, our findings may have important implications for the design of modern technologies, such as mobile applications and games. In cases where the encoding of a spatial configuration is important – as when memorizing the directions of possible escape routes in an action game or the layout of a building in an AR app for Architecture or Interior Design – relying on the gyroscope of the device rather than on finger swiping may lead to faster encoding in memory. Notably though, once encoded in memory, spatial information could be retrieved and used in much the same way regardless of the mode of encoding. This suggests that, despite the greater encoding time, navigating simple interior environments on mobile devices by interacting with the touchscreen is an efficient means of committing spatial information to memory for later use. Future research may examine whether this is also the case with larger and more complex spatial environments than the ones used in the current study.

Compliance with Ethical Standards

Ethical approval: All procedures performed in the reported study involving human participants were in accordance with the ethical standards of the Cyprus National Biothetics Committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

Conflict of Interest: Author S.A declares that he has no conflict of interest. Author A.H declares that she has no conflict of interest. Author M.N.A declares that he has no conflict of interest.

References

Avraamides, M. N., & Kelly, J. W. (2008). Multiple systems of spatial memory and action. *Cognitive Processing, 9,* 93-106. Doi: 10.1007/s10339-007-0188-5

Avraamides, M.N., & Kelly, J. W. (2005). Imagined perspective-changing within and across novel environments. Spatial Cognition 2004-Lecture Notes in Artificial Intelligence, (pp.245-258), Springer-Verlag, Berlin-Heidelberg. Doi: 10.1007/978-3-540-32255-9_15

Hatzipanayioti, A., Galati, A., & Avraamides, M.N. (2015). The Protagonist's First Perspective Influences the Encoding of Spatial Information in Narratives. *Quarterly Journal of Experimental Psychology, 69*, 505-520. Doi: 10.1080/17470218.2015.1056194.

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. Doi: 10.1037/0278-7393.33.6.1092

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, & Cognition, 28,* 335–345. Doi: 10.1037/e501882009-156

McNamara, T. P. (2003). How are the locations of objects in the environment represented in memory? In. C. Freksa, W. Brauer, C. Habel, K. F. Wender (Eds.), Spatial Cognition III: Routes and navigation, human memory and learning, spatial representation and spatial reasoning, LNAI 2685 (pp. 174–191). Berlin: Springer-Verlag. Doi:10.1007/3-540-45004-1_11

Mou, W., Biocca, F., Owen, C. B., Tang, A., Xiao, F., & Lim, L. (2004). Frames of reference in mobile augmented reality displays. *Journal of Experimental Psychology: Applied*, *10*(4), 238-244. Doi: 10.1037/1076-898X.10.4.238

Mou, W., McNamara, T. P., Valiquette, C. M., & Rump, B. (2004). Allocentric and egocentric updating of spatial memories. Journal of Experimental Psychology: *Learning, Memory & Cognition, 30*, 142– 157. Doi:10.1037/0278-7393.30.1.142

Presson, C. C, & Montello, D. R. (1994). Updating after rotational and translational body movements: Coordinate structure of perspective space. *Perception, 23*, 1447-1455. Doi: 10.1068/p231447

Rideout V. & Saphir, M. (2013). *Zero to Eight: Children's Media Use in America 2013*. San Francisco, CA: Common Sense Media.

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 15,* 1157-1165. Doi: 10.1037//0278-7393.15.6.1157

Rieser, J. J, Guth, D. A, & Hill, E. W. (1986). Sensitivity to perspective structure while walking without vision. *Perception, 15,* 173-188. Doi: 10.1068/p150173

Shelton, A. L., & McNamara, T. P. (1997). Multiple views of spatial memory. Psychonomic Bulletin & Review, 4, 102-106. Doi: 10.3758/bf03210780

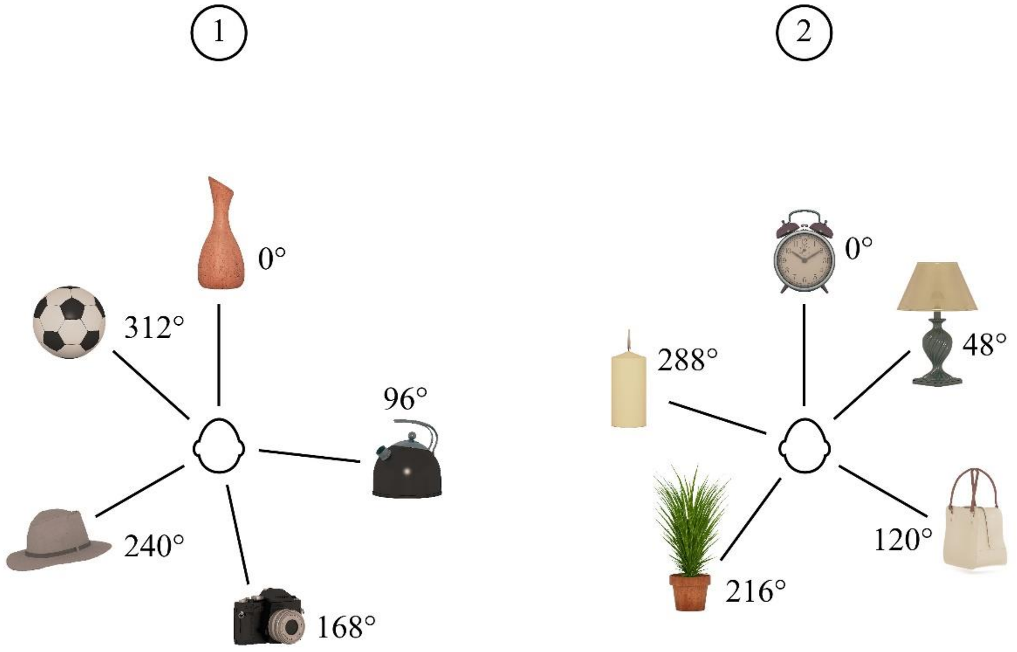


Figure 1. Schematic depiction of the two spatial layouts used in the study.

Figure 2. Example view from a spatial scene as presented on the tablet.

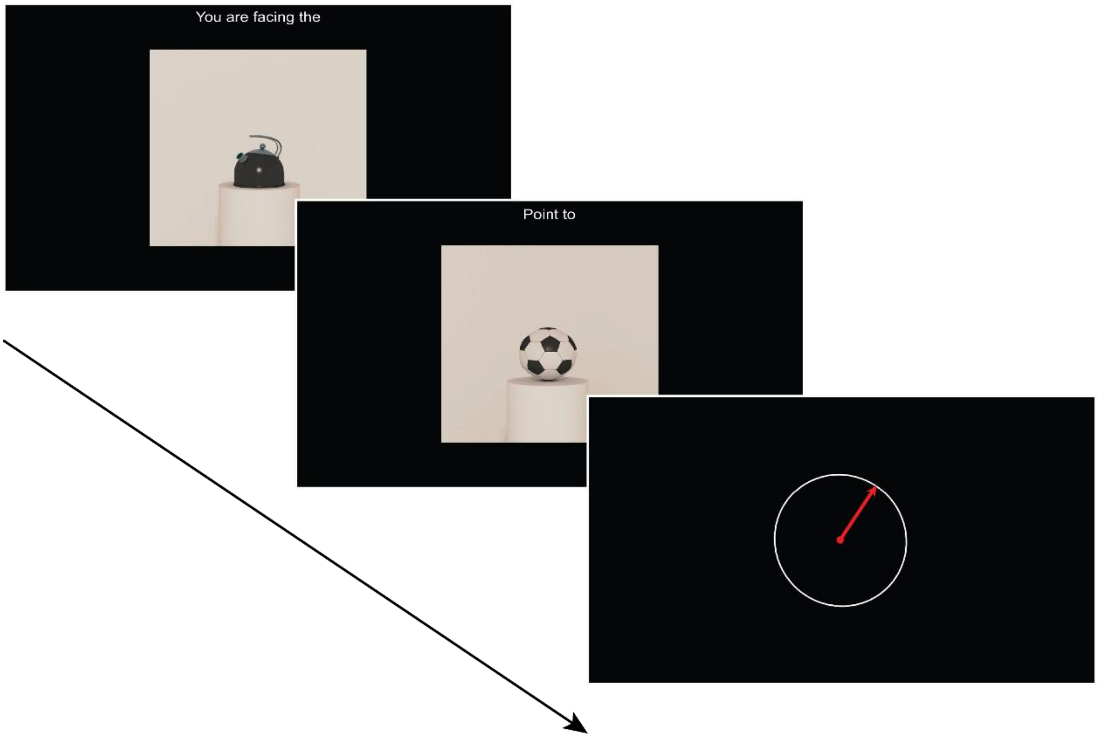


Figure 3: Example trial in the testing phase

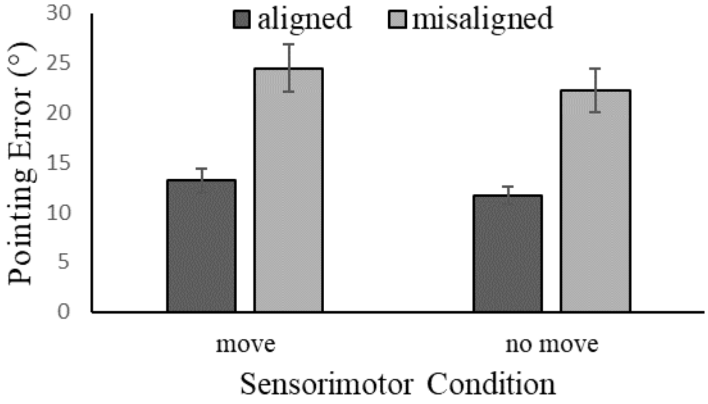
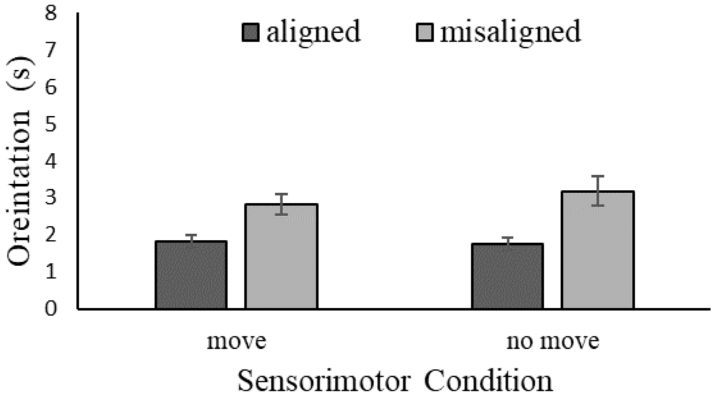
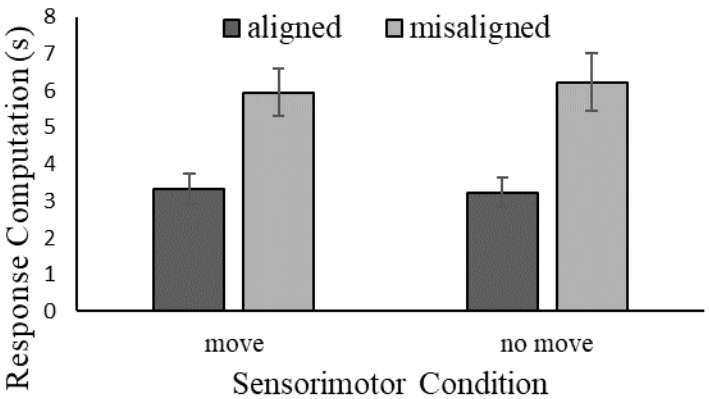


Figure 4: Pointing Error as a function of sensorimotor condition and imagined perspective.

a.


aligned
misaligned
Oreintation (s)
8
7
6
5
4
3
2
1
0
move
no move
Sensorimotor Condition

b.


aligned
misaligned
Response Computation (s)
8
7
6
5
4
3
2
1
0
move
no move
Sensorimotor Condition

c.

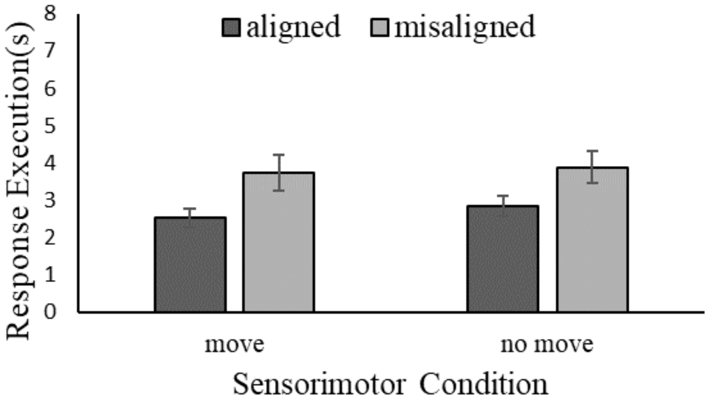

aligned
misaligned
Response Execution(s)
8
7
6
5
4
3
2
1
0
move
no move
Sensorimotor Condition

Figure 5: Orientation Latency (a), Response Computation Latency (b), and Response Execution Latency (c), as a function of sensorimotor condition and imagined perspective.

Data Deposition Information INCLUDING WEBLINK

https://osf.io/8uj7k/?view_only=fab28577190e4114b14dfdd59901f174