Title
Factors That Contribute to Individual and Sex Differences in Perspective Taking Performance

Author
Gunalp, Peri Nicole

Publication Date
2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Santa Barbara

Factors That Contribute to Individual and Sex Differences in Perspective Taking Performance

A Thesis submitted in partial satisfaction of the

requirements for the degree Master of Arts

in Psychological and Brain Sciences

by

Peri Nicole Gunalp

Committee in charge:

Professor Mary Hegarty, Chair

Professor Tamsin German

Professor Richard Mayer

Professor Dan Montello

January 2018

The thesis of Peri Nicole Gunalp is approved.

Tamsin German

Richard Mayer

Dan Montello

Mary Hegarty, Committee Chair

July 2017

Factors That Contribute to Individual and Sex Differences in Perspective Taking Performance

ACKNOWLEDGEMENTS

I would like to thank the members of my committee, Drs. Tamsin German, Richard Mayer, and Dan Montello for their thoughtful feedback on earlier versions of this manuscript. I would also like to thank my advisor Dr. Mary Hegarty for her thorough comments and guidance throughout the execution of this research and the production of this manuscript.

ABSTRACT

Factors That Contribute to Individual and Sex Differences in Perspective Taking Performance

by

Peri Nicole Gunalp

Previous research has demonstrated a sex difference favoring males in perspective taking ability (e.g. Eisenberg & Lennon, 1983; Fields & Shelton, 2006; Meneghetti, Pazzaglia, & De Beni, 2012). Factors that influence sex differences in perspective taking ability in favor of males have been under-explored, and no unified explanation of the extant sex differences in this ability exists currently. Task components (for example including a directional cue in an Spatial Orientation Test [SOT] array), social nature of the task (presence of a human figure in an array), or embodied nature of the task (ease of imaging self in task) may each shape perspective taking ability. Experiment 1 examined how perspective taking ability was influenced by both a social directional cue, and an abstract, non-human, directional cue. The social condition included a human avatar in the SOT array. The spatial condition included an arrow. Results indicated that females and males performed best in the social condition and no better in the spatial condition than the control condition, indicating that social task components are influential in this ability. Experiment 2 compared a replicated social condition to a different non-human directional cue (chair). Results showed that there was no significant difference between the avatar and the chair conditions for males and females.

This suggests that the “social” effect found in Experiment 1 is nuanced, and that perspective taking ability may rather be influenced by ease of embodiment of the focal/central task object. This may indicate that prior evidence of sex differences in this ability have reflected task components rather than inherent ability.

Keywords: perspective taking, sex differences, social influence, embodiment, spatial ability

TABLE OF CONTENTS

Factors That Contribute to Individual and Sex Differences in Perspective Taking Performance

Perspective taking can be categorized as either social perspective taking or visuospatial perspective taking. Galinsky, Ku, and Wang (2005) define perspective taking as, "…the process of imagining the world from another's vantage point or imagining oneself in another's shoes," (p. 110). Visuospatial perspective taking can be thought of as taking a specific physical perspective in space and manipulating or observing something about that perspective. For example, the Money Road Map task requires participants to imagine walking on a path through a city and asks participants at each turn of the path whether they would be turning right or left (Money, Alexander, & Walker, 1965). Social perspective taking, on the other hand, can be more closely related to intuitive definitions of empathy, wherein one imagines taking the psychological perspective of another. This might be more closely related to the processes that occur when giving someone directions; one must imagine what the other person sees and knows about the environment to be successful. Clearly, these two types of perspective taking are intertwined/not entirely distinct from each other. There are further distinctions that can be made within this broad categorization—for example dissociating between visual perspective taking and spatial perspective taking.

The present paper focuses primarily on individual differences in visuospatial perspective taking performance and how they are affected by social and other cues. Often we measure performance but term it ability. Thus in this paper when we refer to performance on measures, we suspect that it is indicative of, but not necessarily a direct measure of, inherent ability. Much research has demonstrated individual differences in spatial perspective taking (e.g. Kozhevnikov & Hegarty, 2001; Hegarty & Waller, 2004), as well as robust sex differences favoring males (e.g. Eisenberg & Lennon, 1983). The purpose of the present research is to examine factors that may contribute to these individual and sex differences, specifically focusing on the role of social and directional cues in perspective taking ability.

While some consider perspective taking to be an effortful and intentional task (e.g. Shelton & McNamara, 2004), others consider perspective taking an extension of an innate or implicit ability (e.g. Tversky & Hard, 2009). Still others have related perspective taking to embodied cognition (e.g. Kessler & Wang, 2012). Despite being characterized as visuospatial, there are several theories regarding the processes behind this type of perspective taking that rely on social mechanisms. For example Tversky and Hard (2009) interpret perspective taking ability as being influenced by social-sense making.

In their study Tversky and Hard (2009) examined the effect an actor in a scene might have on a participant's descriptions of the spatial relationships between two objects. In addition to a control condition, two conditions included an actor in the scene. In the control condition, participants most frequently used the self or egocentric perspective, describing the spatial relationship between two objects from their own perspective. However, in both conditions that included an actor in the scene, participants took the actors' perspective (non-egocentric) significantly more frequently than in the control condition. Tversky and Hard suggested that together these experiments showed that including an actor in the scene, and drawing attention to the actor's action or agency as part of the scene, tapped into an implicit social sense-making ability that underlies perspective taking.

When a person is present in a scene being viewed we may implicitly try to make sense of their presence by determining the role they play in the social scope of the scene and in the environment. This attempt at sense making may even precede our more explicit goals to describe the scene or part of the scene, if that is our task (Tversky & Hard, 2009). In this way, the social underpinnings of this type of sense making are clear; in order to resolve possible discrepancies between another's perspective and our own, an implicit mechanism may trigger us to instinctively take a perspective that is non-egocentric. What gives the actor enough agency to trigger this sense making; his being human or his action in the scene? If an object were added to the scene that was not perceived to have agency (perhaps a non-human object such as a chair) would this social sense making still be elicited?

The present study aims to address these questions. Related research has also examined how object agency influences perspective taking ability. Shelton, Clements-Stephens, Lam, Pak and Murray (2011) examined how the agency of the target in a perspective taking task might influence accuracy and response time on a modified version of a three-mountain task (Piaget & Inhelder, 1967). In this task, participants were shown an image of three buildings and had to indicate which target around a matching display of three buildings would have that view. Each target had what Shelton et al. call a different degree of agency: one target was a human figure artist's model (high agency), one was a camera (moderate agency), and another was a triangle stacked on top of a block (low agency). These levels of agency seem to rather reflect a degree of interactivity, in that they might help a person imagine the perspective because they could be interacted with. The present experiments also consider this issue, and discuss it in terms of interactivity rather than agency. For each condition, participants saw displays with only one type of target around the buildings. It was hypothesized that participants in the high agency target conditions would have the highest accuracy and shortest response time. Responses on the Autism Quotient

(AQ) were also analyzed, as the AQ is considered a self-report measure of social skills (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) to test the idea that if a social object with agency is central to the task, participants' social skills might be triggered and subsequently influence performance on the task.

Shelton et al. found that contrary to their hypothesis there was no significant difference in accuracy or response time across agency (target type) conditions. Participants were equally accurate and fast in responding to trials with all three targets (artist model, camera, and triangle block). Critically, however, there was a significant correlation between the Autism Quotient responses, particularly sections regarding social skills and communication, and accuracy in the artist model condition. These results further promote the idea of a social mechanism behind perspective taking, suggesting that there is a connection between an individual's self reported social skills and the accuracy with which an individual can take a targets' perspective, particularly when that target has a high level of agency (Shelton et al., 2011). Furthermore, these results relate to the social sense-making ability described by Tversky and Hard (2009), as self-reported social ability could very well be correlated with individual differences in social sense-making ability.

Research by Sulpizio, Committeri, Metta, Lambrey, Berthoz, and Galati (2015) also found evidence for automatic perspective taking focused on an agent. Sulpizio et al. used a virtual environment that depicted a room with eight different camera viewpoints (a camera was at each 45 degree interval around a 360 degree circle). Participants were shown a view of the room at study, a view of the room at test, and were asked to determine if there was movement of a plant (the target object) between the two views. In addition to the control condition, which was simply a view of the room that always matched from study to test, in the primed condition there was an avatar standing in the scene. On some trials this avatar was positioned to prepare participants for a large perspective shift between the study and test images. In the unprimed condition, there was no avatar in the scene, but in contrast to the control condition, there was occasionally a change between views at study and test. It was hypothesized that if a large change in perspective on the scene from study to test was anticipated by the avatar (how the avatar would have viewed the plant), then participants would be more accurate and faster in judging if the study and test images were the same than in unprimed condition.

The hypothesis was supported, suggesting that affording participants the opportunity to encode the specific perspective of the avatar, and in a sense prioritize the avatar's perspective over their own, allowed participants to resolve greater angular differences between study and test images more quickly and accurately. Even when participants were asked to encode the position of another object in the scene (the plant), which could have been a non-social prime by acting as a focal object around which participants could have noticed change, the shifts between study and test images were more difficult to judge (Sulpizio et al., 2015). Thus, again presence of an agent seems to contribute strongly to perspective taking ability.

Indeed, it may be that empathetic ability is a key component of the implicit sense-making mechanism as suggested by Tversky and Hard (2009). Activation of this mechanism may predispose us to take a non-egocentric perspective in social situations, and in turn support our desire to resolve differences between another's perspective and our own. If this is true, would populations who perform poorly on perspective taking tasks benefit from explicit activation of empathy in some way? Specifically, because women tend to perform more poorly than men on perspective taking tasks (e.g. Baron-Cohen & Wheelwright, 2004), would framing a perspective taking task with empathy improve female performance? If so, one might even conjecture that extant sex differences favoring males in perspective taking abilities could be reflecting an exclusion factors that activate social/empathetic mechanisms in task format, rather than task difficulty or inherent individual ability. Importantly, the involvement of empathy in perspective taking may be tied to *what* individuals are taking the perspective of. If taking the perspective of an *object*, perhaps empathy is not a component of perspective taking ability overall, but if taking the perspective of another *person*, empathy may be a key component of this processing. This issue is examined in the present study.

Research conducted by Tarampi, Heydari, and Hegarty (2016) began to explore this connection by including an empathy component in their perspective taking task. In their experiments, Tarampi et al. used the Spatial Orientation Test (Kozhevnikov & Hegarty, 2001) which requires participants to estimate an angle between three objects in an array. In the task participants are asked to imagine standing at one object in the array, facing a second, and point to a third (see Figure 1a). Participants give their responses by drawing a line on an answer circle. The answer circle contains a line that is labeled with both the object at which participants were asked to imagine standing and the object toward which they would be facing. The participant's task is then to draw the line indicating where they would "point" to the third target object given their imagined location and facing direction in the array, as shown in Figure 1a. Angular error scores were calculated for each trial then averaged to yield a participant's average error score (Tarampi, et al., 2016).

In their first study Tarampi et al. incorporated stereotype threat in the task instructions, with either a male positive or female positive bias. In the male positive condition the instructions indicated that the task was a measure of spatial ability, and that men tended to have higher spatial ability than women. In the female positive condition the instructions said that the task was a measure of empathetic ability and that women tended to have higher empathetic ability than men. Additionally, in the female positive condition the instructions were always paired with arrays that included a human figure that replaced one of the objects in each trial. The human figure was shown facing in the direction of the object the participants were told to imagine facing in each trial (see Figure 1b). Participants were then asked to imagine taking the perspective of the person in the array, as opposed to imagining being at an object in the array. In the male positive condition on the other hand male bias instructions were always paired with arrays that had no human figure (Tarampi et al., 2016).

Women performed better in the female positive condition than the male positive condition. Critically, there was a sex difference in the male positive condition but this was eliminated in the female positive condition. Men performed equally well in both the female and male positive conditions (Tarampi et al., 2016). Subsequent experiments were thus conducted in which only the instructions were manipulated and paired with the original (no-human) array, or only the array was manipulated and had non-biased instructions. While the instructions alone were not sufficient to significantly affect performance on the SOT it was shown that the presence of a human figure in the array was what differentially bolstered female performance (Tarampi et al., 2016). Tarampi et al. suggested that including a human figure in the array could be engaging embodied or social processes that influence perspective taking ability.

What is it about including a human figure in the array that improved the accuracy with which women completed the task? Could it be that the human figure in the array is providing additional directional cues that subsequently make the task easier? Or, like Tarampi et al. suggest, could the human figure be activating some social mechanism that makes taking the perspective of the human figure more intuitive/easier? As the human figure included in the array used by Tarampi et al. was facing in the direction of the object the participant was to imagine facing, the human figure inherently provided an additional directional cue that was not present in the original array. This cue could have made the task easier in general and could have been a cue that women selectively noticed and used to improve their performance on the task. It is also possible that the human-ness of the human figure included in the array could have been the influential factor, somehow tapping into a social mechanism reliant on empathetic ability.

Thus, there are two theories to explain the effects of the human figure: it contributed an additional directional cue that made the task easier, or tapped into a social mechanism that females use more readily as a task solving strategy. Experiment 1 employs two conditions to directly test these theories, an arrow condition and a social condition. In the arrow condition, participants are asked to imagine standing at an arrow (a non-human directional cue), and in the social condition participants are asked to imagine standing at a person (human directional cue). In Experiment 2, we contrast a human with a chair to examine intermediate levels of agency.

Additionally, the experiments reported here utilize a new virtual reality format to see if the results from Tarampi et al. generalize from the original paper SOT to a more naturalistic and immersive SOT. The abstract nature of the original paper SOT was unrealistic and could have made the task more difficult for women, which could in turn be affecting the extant reports of sex differences in perspective taking ability. We also include the original paper SOT as a posttest to examine the correlation between the VR and original paper and pencil versions of the task.

Experiment 1

In Experiment 1 participants completed the SOT in the immersive VR environment, either in the control array, with an arrow in the array, or with a person in the array. If the social mechanism described above does support perspective taking ability and the human figure taps into that social mechanism, I hypothesized that females would perform better in the social condition than either the arrow or control conditions. However, if the person in the array simply provides an additional directional cue that improves female performance, I hypothesized that men and women would do better in both the social and arrow conditions than the control condition (as both the social and arrow conditions provide an additional directional cue). Lastly, I aimed to compare main effects (of gender and condition) of the VR SOT to the paper SOT to see if both the paper and VR versions of this task were roughly equivalently difficult, and if participants performed equally on both the paper and VR versions. There are differences between the original SOT used by Tarampi et al. and the present VR SOT, but if the paper and VR SOTs measure the same ability we might see a sex difference in performance such that men might perform better than women, as Tarampi et al. found. Regardless of which hypothesis is supported, these experiments shed light on the social components of perspective taking tasks that aid female performance

Method

Participants and Design

One hundred thirty-five undergraduate students from the University of California, Santa Barbara participated in this study for course credit. One participant was excluded because they declined to report their gender. Therefore, our final sample was 134 participants, ages 17-33 (M = 19.30, SD = 1.74).

This study was a single-factor between subjects design. The independent variables were VR condition, which had three levels: social, spatial, and control gender (female and male). There were 64 males (21 social, 22 arrow, 21 control) and 70 females (26 social, 22 arrow, 22 control) in the sample. The dependent variable measured was angular error, and each participant's angular error across trials was averaged for their total score.

Materials

Participants used an Oculus Rift DK2 (60 Hz refresh rate) headset during the Virtual Reality Spatial Orientation Test (VR SOT) trials, which viewed a virtual environment that was programmed using Vizard software. The single work station consisted of two monitors used for the VR SOT, which were Dell P24124 (60 Hz refresh rate) with Nvidia GeForce GTX (660). An Xbox controller was used in the VR SOT trials for participants to submit their answers. The virtual environment arrays differed by condition, but always contained seven objects. The environment was made to look naturalistic and resemble a park scene. The ground of the environment was green/grassy, and the objects included in the array could have all been found in a park (see Figure 2). All trials in the virtual environment were the same format as the trials in the original paper SOT, such that all trials read, "Imagine you are standing at <u>X</u> facing <u>Y</u>, point to <u>Z</u>," where X, Y, and Z are all objects in the array. The control virtual environment had no extra cue in the array (see Figure 2a). A trial in the control condition read, “Imagine you are standing at the mailbox facing the picnic table, point to the crate.” On each trial of the control condition, participants were asked to imagine standing at a different object in the array, facing another object, and to point to a third. This condition has a higher working memory load than in either the arrow or social conditions, as in the latter conditions participants imagine standing at the same object on every trial, and that object faces in the direction participants are to imagine facing.

The arrow environment had an arrow that acted as a non-human directional cue and replaced one in the objects in the array for each of the trials (see Figure 2b). A trial in the arrow condition read, “Imagine you are standing at the *arrow* facing the picnic table, point to the crate.” For every trial participants imagined standing at the arrow—the arrow moved accordingly to a new location on each trial, and would face the way participants were to imagine facing. The social environment had a human avatar that acted as a human directional cue that would replace one object in the array for each of the trials (see Figure 2c). A trial in the social condition read, “Imagine you are standing at the *person* facing the picnic table, point to the crate.” In all conditions the 12 trials were presented in a randomized order for each participant. Again here, the participants imagined standing at the person on each trial, and the person moved in the array as stated in each trial.

On each trial an answer circle was provided above the array that had an arrow pre-drawn in it (see Figure 2f). The arrow was labeled with both the object at which participants were asked to imagine standing and the object toward which they would be facing. The participant’s task was then to move an arrow in the answer circle using the XBox controller indicating where they would point to the third target object given their imagined location in the array and facing direction. Scoring for this task was based on the absolute value of the difference between the angle of the arrow on the answer circle submitted by the participant and the correct angle between the three objects in the array. It is important to note that the greatest amount of error a participant could have on any given trial is 180 degrees—the farthest off of the correct pointing direction is the exact opposite direction. As such, scoring for error always measured the shortest arc between the participant's estimated angle and the target angle, maintaining this maximum error of 180 degrees. Each participant was given an angular error score for each trial. Average angular error was then calculated to account for number of trials completed by the participant.

If participants did not complete all the trials of the SOT, the trials that were left incomplete received an error score of 90 degrees, which would be chance performance, because the absolute error of a participants' response can range from 0-180. The number of trials that were not completed by the participant was then taken into account in addition to the error scores of their attempted trials. For example, if a participant did not complete three of the 12 trials, their overall average error would be the following: (total of angular error scores on trials 1-9 + 90x3)/12.

Paper materials included the Spatial Orientation Test (SOT) (Kozhevnikov & Hegarty, 2001) and the MRM (Money, Alexander, & Walker, 1965). The paper SOT always showed the same array of seven objects from an aerial view (180° above the array) and was always the control condition of the SOT, such that it showed only the original array (no human figure or arrow). Below the array there was an answer circle in the same format as the VR SOT on which participants could draw their answer line (see Figure 1a). Scoring for this task was the same as the VR SOT, such that each trial completed by each participant was given an angular error score.

The Money Road Map Test (MRM) shows an aerial view of a path (represented by a dashed line) through a city (see Figure 3). The participants were asked to imagine walking along the path as it changes direction and label each turn with “R” or “L” indicating the required direction of the turn. A timer was used for both the paper SOT and the MRM to maintain the task-appropriate time limits (five minutes for the SOT, 30 seconds for the MRM). The Qualtrics online survey platform was used for the questionnaire that was distributed at the end of the study (see Appendix A for a complete list of questions). Questions included demographics, the Santa Barbara Sense of Direction Scale (Hegarty, Richardson, Montello, Lovelace, & Subbiah, 2002), and mental strategy used in both the VR and paper SOTs (see Appendix A). The strategies measured in this questionnaire were categorized as either abstract or embodied. The abstract strategies indicated that the participant imagined manipulating the array in some way, for example superimposing a clock on the array, to solve each trial. In contrast the embodied strategies indicated that the participants imagined themselves in the array somehow, for example imaging standing in the array to solve each trial.

Procedure

Upon arrival participants were shown to the experimental running room and given an informed consent sheet. After signing the consent sheet, participants began the VR SOT. To begin this task participants completed two controller practice trials. In the virtual environment, a text box with an answer circle and instructions appeared above the ground of the environment, which told participants how to use the controller to move the arrow in the answer circle. The instructions indicated that the participants could use the left analog joystick and both the left and right trigger buttons to move the arrow around in the answer circle. The experimenter read these instructions aloud to the participant, and pointed to the joystick and trigger buttons on the controller. To submit a response, participants were instructed to press the “A” button on the controller. For the practice trials the participant’s goal was to align the arrow in the answer circle with a red triangle on the outside of the circle. When the arrow and the triangle aligned, the triangle turned green and participants were instructed to press “A” to submit their answer. These practice trials were merely to acclimate participants to the joystick controller used for this task. Participants completed two of these controller practice trials without the Oculus headset on.

After completing the two controller practice trials participants were instructed by the experimenter to put on the Oculus headset. Any adjustments in band size or lens magnification were made at this point if necessary. After the headset was adjusted to fit comfortably, participants were handed the controller again and began the task practice trials. To complete the task practice trials, participants were first asked to learn the names of the four objects in a practice array and inform the experimenter when they were ready to continue. When the participants were ready, the experimenter advanced the screen and read the first practice trial aloud. The participants completed four practice trials before beginning the test trials.

The experimenter then advanced the program to the first test array screen. Before beginning the test trials, participants were asked to memorize the names of items in the test array. The experimenter then advanced the program through each of the seven objects and made sure each object was correctly identified by the participant. The participants were then told that they would have about five minutes to complete the trials, and the experimenter left the experimental running room. The participants were also instructed to tell the experimenter when they had completed the trials.

After completing the VR SOT trials, the experimenter then gave the participants the paper SOT packet and after reading the instructions, reiterated that the participant would have five minutes to complete the 12 trials. The experimenter then flipped to the first trial in the packet, started the timer, and left the experimental running room. The trials of this task were the same format as the VR SOT such that the instructions for each trial read, "Imagine you are standing at X facing Y, point to Z," where X, Y, and Z are all objects in the array. The participants were allowed to move on to the next task if they completed the trials in fewer than five minutes and did not want to check their work. After five minutes had passed, the experimenter stopped the participant's progress regardless of how many trials had been completed. Experimenters took the paper SOT packet and gave the participant the MRM packet. Participants were given the standard MRM instructions, which said that they would be shown a map of a city with a path drawn through it, and were supposed to imagine walking along that path and label each of the turns on the path as either a right or left turn. Participants were allowed 30 seconds to complete the task.

The experimenter then opened the survey on the computer and told the participant that they had as much time as needed to answer the questions. Participants were instructed to tell the experimenter when they had completed the survey, after which the participants were debriefed and shown out of the lab.

Results

We predicted that females would perform as well as males in the social condition, and if the social mechanism is activated by the avatar, should perform significantly better in the social condition than either the arrow or control conditions. If the additional directional cue offered by either the arrow or the person boosted performance, we hypothesized that females and males would perform better in both the arrow and social conditions than the control condition. Lastly, we hypothesized that the paper SOT would be strongly correlated with each of the VR SOT conditions, as the two tests should largely measure the same ability.

Virtual Reality Spatial Orientation Test

As our data were non-normal, a logarithm transformation was completed on each individual trial raw error score, and then a new mean error score was calculated for each participant. However, after the log transformation the data were still bimodal, with one mode containing most of the subjects and the other mode containing the few subjects that performed at chance levels of error (90 degrees). There were 15 subjects that performed at or worse than chance, 3 (2 female, 1 male) from the control condition, 4 from the arrow condition (2 female, 2 male), and 8 from the social condition (6 female, 2, male). The analyses were conducted with and without these subjects. With all participants included, (except for one who declined to report gender), for the VR SOT, a between subjects 2 (gender: male, female) x 3 (condition: control, arrow, social) ANOVA found a significant main effect of gender, $F(1,128) = 4.4$, $p = .034$, $\eta_p^2 = .035$, such that men ($M = 2.55$, $SD = 0.79$) performed better than women ($M = 2.84$, $SD = 0.83$) over all, with an effect size of $d =$

.36. There was no significant main effect of condition, p = .249, and no significant interaction, p = .896 (see Figure 4).

With participants who completed the task at or above chance levels of error excluded from the analysis, a 2 (gender: male, female) x 3 (condition: control, arrow, social) ANOVA on the VR SOT angular error means found a significant main effect of gender, $F(1,113) = 6.27$, $p = .014$, $\eta_p^2 = .053$, such that men (M = 2.37, SD = 0.54) performed better than women (M = 2.62, SD = 0.6) over all, d = .44. There was also a significant main effect of condition, $F(2,113) = 7.93$, $p = .001$, $\eta_p^2 = .123$, such that participants performed better in the social condition (M = 2.22, SD = 0.5) than either the arrow (M = 2.64, SD = 0.64) or control (M = 2.62, SD = 0.5) conditions. For a complete list of means by gender and condition, see Table 1. There was no significant interaction between gender and condition, p = .515.

These results support the social mechanism hypothesis, in that all participants performed better in the social condition than either the arrow or control conditions. Additionally, LSD post-hoc pairwise comparisons indicated that there was no significant difference between the arrow and control conditions, p = .846. The current results seem to more closely support the social mechanism theory, in that participants performed best in the social condition, and performance in the arrow and control conditions did not differ significantly. As such, it seems that the social aspect of including a human figure in the array may be what is differentially improving performance on this task. Additionally, our hypothesis predicted an interaction between gender and condition, which was not found to be significant.

Paper Spatial Orientation Test

For the paper SOT, our analyses revealed the same findings whether all participants were included or not. As such the following analysis includes all participants. A 2 (gender) x 3 (condition of VR SOT) ANOVA found a significant main effect of gender, $F(1,128) = 12.24$, $p = .001$, $\eta_p^2 = .087$, such that males ($M = 2.53$, $SD = 0.55$) performed better than females ($M = 2.90$, $SD = 0.67$), $d = .60$ (see Figure 5), replicating previous research on this task (Tarampi et al., 2016). There was no main effect of condition, $F(2,128) = .27$, $p = .763$, which is to be expected as all participants completed the original version of the paper SOT, so an effect of condition would reflect a difference in paper SOT scores only if it was influenced by which VR condition the participant was in. There was no significant interaction between gender and condition, $F(2,128) = .09$, $p = .916$.

A paired-samples t-test also found that there was no significant difference in angular error between the control condition of the VR SOT and the paper SOT, $t(39) = -.28$, $p = .783$. Additionally the mean angular error across VR conditions ($M = 23.97$, $SD = 14.44$) was not largely different from the mean angular error across paper conditions ($M = 26.42$, $SD = 16.65$). This suggests that although the angular error was slightly higher in the VR SOT conditions than in the paper SOT, this difference was not statistically significant.

Survey

Additional analysis revealed notable findings from the survey portion of the experiment (see Appendix A for questions and scoring). One section of the questionnaires regarded mental strategies used to solve trials on each of the two (paper and VR) versions of the SOT. For analysis, we categorized participant responses as consistent with either an embodied or abstract mental strategy. Two coders categorized these responses, and inter-rater reliability was k = 1.00, p = .014. It was found that almost all participants reported using an embodied strategy to complete both the paper SOT (110 embodied, 9 abstract) and VR SOT (104 embodied, 15 abstract). A chi-square analysis indicated that there was no significant difference between conditions for the embodiment strategy in the VR SOT $\chi^2(2) = 1.958$, $p = .376$, thus the vast majority of participants employed this strategy in VR regardless of condition.

One question on the survey asked participants in the social VR condition (n = 39) to rate how much they identified with the avatar on a five-point scale, from 1-not at all to 5-all of the time. The mean identification scores for males (M = 3.05, SD = 1.10) and females (M = 2.32, SD =1.20) were marginally different $t(37) = -1.99$, $p = .054$, though this difference was not significant. One might suspect that because the avatar used was male, though gender was not explicitly stated but rather apparent, male subjects might report identifying with the avatar more than females.

Correlations

Our analysis (with participants excluded who had at or above chance levels of error examining log transformed means) also revealed several significant correlations (see Table 3). Overall, the VR and paper SOTs were significantly positively correlated ($r = .377$, $p <$ .001). In the control condition, performance on the VR and paper versions of the SOT was moderately positively correlated ($r = .47$, $p = .002$). In the arrow condition, the VR and paper SOT were moderately correlated ($r = .43$, $p = .006$). However, in the social condition the VR and paper SOT were not significantly correlated ($r = .26$, $p = .11$), which could be due to small sample size. Additionally, the VR SOT was significantly correlated with both the MRM ($r = -.36$, $p < .001$), and the SBSOD ($r = .20$, $p = .028$). This suggests that the VR

SOT is a good measure of perspective taking performance as it correlates with another established measure of this ability and self-reported sense of direction.

Discussion

Findings from Experiment 1 indicate that performance was best in the social condition, which was consistent with our hypothesis. Although it was expected that the social condition would improve only female performance, both males and females performed best in this condition. This suggests that the presence of the human figure eased participants' ability to imagine taking a perspective different from their own. Additionally, analyses demonstrated a larger main effect of sex in the paper SOT ($\eta_p^2 = .087$) than the VR SOT ($\eta_p^2 = .053$). This may suggest that the more abstract birds-eye view in the paper SOT makes the paper format of task more difficult for women than the VR version, which utilizes an approximately 125° viewpoint. Perhaps the original view of the paper task was a contributing factor to sex differences found in prior research.

Correlational analyses revealed that the control VR condition and paper SOT were moderately correlated, suggesting that both tasks measure the same ability, which logically follows as neither version introduced a new figure into the array (no arrow or human figure). Interestingly the social VR condition was not significantly correlated with the paper SOT. Also, for both the VR and paper SOTs, a majority of participants reported using embodied strategies to complete the trials. This suggests that the use of an embodied strategy was not necessarily reliant on the cue type included in the array (what VR condition participants saw) as we might have suspected. Together with the results from the correlational analyses, it seems that the social VR condition is both distinct from (non-significant correlation between social VR SOT and paper SOT) and similar (a majority of participants reported using embodied strategies in all conditions not just the social condition) to the arrow and control conditions.

However, Experiment 1 is not without limitations. The arrow condition could have been problematic for participants, in that when the trial referenced an arrow, it could have been referencing either the arrow in the array or the arrow in the answer circle. This ambiguity could have resulted in poorer performance in the arrow condition for both males and females. Experiment 2 resolves this ambiguity by utilizing a chair as the non-human directional cue. Additionally, the original SOT includes one sample item that has the correct answer shown. The VR SOT conditions in Experiment 1 do not give such a sample item. Considering that approximately 11% of our sample was excluded in some analyses due to error at or above chance, it may be important to include a completed sample item in the VR to check for understanding of the task. This was added in Experiment 2. Lastly, the wording used by Tarampi et al. was slightly different from the present experiment. Where the present experiment says, "Imagine you are standing at the person, facing…" Tarampi et al. said, "Take the perspective of the person, facing…" Though the main effect of gender was significant in this experiment, this difference in wording may be a possible explanation of why this gender difference in the social condition was not eliminated (as in Tarampi et al., 2016). The latter wording is more active than the former, and could have either drawn more attention to the human figure included in the array or helped participants realize that the human figure could be used to help solve each trial. This wording is also more social, in that it suggests the participants actually take the perspective of another person, rather than just imagine standing at another persons' location. Experiment 2 addresses this as well as the other limitations mentioned above to more precisely examine the influences on perspective taking ability.

Experiment 2

Experiment 2 addresses the limitations of Experiment 1 by adding feedback to a practice trial, utilizing a chair as the non-human directional cue, and using the exact wording in the social condition as Tarampi et al. (2016). We hypothesize that the main effects from Experiment 1 will replicate—specifically, we hypothesize that participants will perform better in the social condition than either the control or chair conditions. As in Experiment 1, Experiment 2 correlations were examined among the VR SOT, paper SOT, MRM, and the SBSOD to ensure replication and validity of the VR measurement. As such, we predict that the VR and paper SOTs will be correlated as in Experiment 1. For comparability sake, though the results are the same with everyone included, we have taken out the participants who performed at or above chance.

Method

Participants and Design

One hundred and ninety-one students from the University of California Santa Barbara participated in this experiment for course credit, ages 17-30 (M = 19.10, SD = 1.70).

This study was a single-factor between subjects design. The independent variables were VR condition, which had three levels: social, spatial, and control and gender (female and male), and there were 98 females (34 social, 31 chair, 33 control) and 93 males (31 social, 31 chair, 31 control) in the sample. The dependent variable measured was angular error, and each participant's angular error across trials was averaged for their total score.

Materials

The materials for this experiment were the same as Experiment 1, except for a few changes: in the directional cue condition the arrow in the task array was replaced with a chair (see Figure 2d for a screenshot), and the wording of the instructions for the social condition were changed from “Imagine standing at the person, facing X, point to Y,” to read, “Take the perspective of the person, facing X, point to Y.” The survey questions regarding strategy used during the task were altered to reflect this change. Additionally, “feedback” was given to participants on the first practice trial, in the form of a line displaying the correct answer for the trial (see Figure 2g).

Procedure

The procedure of this experiment was the same as Experiment 1, with one change. Before the practice trials, participants were told that on the next screen they would see an example trial with the correct answer drawn in. After advancing to the first practice trial, participants were instructed to move their arrow to where they thought the answer was, but to wait to submit their response until instructed by the experimenter. After the participants had moved their arrow, the experimenter reiterated that the line shown on the circle was the correct answer. Regardless of where the participant’s arrow was, the experimenter said, “Can you see that [if you were standing at the crate (control condition) /standing at the chair (chair condition)/ if you took the perspective of the person (social condition)], facing the duck, you would point to the wheel barrow in the direction indicated by the line?” This was reiterated until the participant affirmed that it made sense.

Results

As in Experiment 1, we predicted that females would perform as well as males in the social condition, and if the avatar activates the social mechanism, participants should perform significantly better in the social condition than either the chair or control conditions. If the additional directional cue offered by either the chair or the person boosted performance, we hypothesized that females and males would perform better in both the chair and social conditions than the control condition.

Virtual Reality Spatial Orientation Test

It should be noted again that the data were not normally distributed; rather they were negatively skewed with most participants performing the task with low-to-moderate angular error. For subsequent analyses, the data were log transformed in that each measure of error on each trial of both the paper and VR SOTs was log transformed. An average error per participant was thus calculated from the log transformed trial error scores, and was used in the following analyses. These analyses were conducted on the whole sample and with participants who scored at or above chance levels of error excluded. Both analyses revealed the same findings, and to maintain comparability between the two experiments the findings reported in this section exclude participants who performed at or above chance levels of error, which leaves 182 participants.

For the VR SOT, a between subjects 2 (gender: male, female) x 3 (condition: control, chair, social) ANOVA found a significant main effect of gender, $F(1,176) = 13.10$, $p < .001$, $\eta_p^2 = .069$, such that males had significantly lower error on average ($M = 2.46$, $SD = .55$) than females ($M = 2.77$, $SD = .65$) across conditions, $d = -.52$. There was also a significant main effect of condition, $F(2,176) = 4.99$, $p = .008$, $\eta_p^2 = .054$, such that participants performed with significantly lower angular error in the chair (*M* = 2.47, *SD* = .67) and social (*M* = 2.57, *SD* = .57) conditions than the control (*M* = 2.80, *SD* = .58) condition (see Figure 6). For a complete list of means by gender and condition, see Table 2. There was no significant interaction of condition and gender $p = .54$. Additionally, LSD post-hoc pairwise comparisons revealed that while there was a significant difference between the control and chair ($p = .002$) and the control and social conditions ($p = .035$), there was no significant difference between the chair and social conditions ($p = .334$).

These results support the directional cue hypothesis, in that men and women performed better in the chair and social conditions than the control condition. Importantly there was no significant difference between the chair and social conditions. In some ways these results more closely align with the directional-cue explanation of the findings from Tarampi et al. (2016), suggesting that the directional cue offered by the chair allowed females to complete this task with lower error that was not significantly different from males, and with less error than the control condition. While one may argue that the improvement in the social condition supports the social mechanism theory, the important consistency between the chair and the avatar as cues is that they both provide a *directional* cue. If one factor is contributing to performance, these results seem to suggest that the directional nature of the cue is what is important; otherwise, the social condition would have had lower error than both the control and chair conditions.

Paper Spatial Orientation Test

In the paper SOT, a between subjects 2 (gender: male, female) x 3 (condition: control, chair, social) ANOVA found a significant main effect of gender, $F(1,176) = 19.23$, $p < .001$, $\eta_p^2 = .099$, such that males completed the task with significantly lower angular error on average (M = 2.44, SD = .57) than females (M = 2.86, SD = .7) across conditions, d = -.66 (see Figure 7). As expected, there was no main effect of condition p = .815, and no interaction of gender and condition, p = 859. Again, it should be noted that there were not different conditions of the paper task; rather the "conditions" here reflect which VR SOT the participants saw before completing the paper SOT. This replicates the findings Experiment 1.

A paired-samples t-test indicated that there was a significant difference in the levels of error between the control condition of the VR SOT and the paper SOT $t(59) = -2.40$, p =.02. This suggests that the angular error was significantly higher in the VR SOT conditions than in the paper SOT, which does not replicate the findings from Experiment 1. The mean angular error across VR conditions (M = 27.04, SD = 18.38) was not largely different from the mean angular error across paper conditions (M = 28.67, SD = 18.75). However, the overall patterns of the data between experiments is the same, which suggests that the significantly higher error in the VR SOT compared to the paper SOT in this experiment may be due to differences in sample sizes between the two experiments.

Survey

Additional analysis revealed notable findings from the survey portion of the experiment (see Appendix A for questions and scoring). Due to experimenter error, survey data were not collected for 14 subjects. As such, the following analyses were conducted with a sample of n = 177. One section of the questions participants answered regarded mental strategies used to solve trials on each of the two (paper and VR) versions of the SOT. For analysis, we categorized participant responses in the same way as in Experiment 1, as consistent with either an embodied or abstract mental strategy. Two people categorized these responses, and inter-rater reliability was k = 1.00, $p = .008$. It was found that a majority of participants reported using an embodied strategy to complete both the paper SOT (159 embodied, 18 abstract) and VR SOT (159 embodied, 18 abstract). A chi-square analysis indicated that there was no significant difference between conditions for the embodiment strategy in the VR SOT, $\chi^2(2) = 3.131$, $p = .209$; thus, the vast majority of participants employed this strategy regardless in VR of condition. In the paper SOT, a chi-square analysis indicated that there was no significant difference between conditions $\chi^2(2) = 1.663$, $p = .435$.

One question on the survey asked participants in the social VR condition (n = 63) to rate how much they identified with the avatar on a five-point scale, from 1-not at all to 5-all of the time. As in Experiment 1, there was no difference in identification with the avatar between males (2.5, SD) and females (2.4, SD), $t(60) = -.232$, $p = .817$.

Correlations

Correlational analyses of log-transformed data revealed that the paper and VR SOTs were significantly positively correlated across conditions ($r = .54$, $p < .001$) (see Table 4). Each individual condition of the VR was highly positively correlated with the paper: control ($r = .58$, $p < .001$), chair ($r = .59$, $p < .001$), and social ($r = .47$ $p < .001$). Additionally, the VR SOT across conditions was highly negatively correlated to the MRM ($r = -.551$, $p < .001$) indicating that participants who completed more of the MRM had lower average angular error scores on the VR SOT. Interestingly, the VR SOT was not significantly correlated with the SBSOD ($r = .035$, $p = .63$). This may suggest that the VR SOT is not robustly correlated with self-reported sense of direction, as the correlation seen in

Experiment 1 was not found here. This may reflect a general weakness of correlation between self-reported sense of direction and perspective taking ability.

Discussion

The results of Experiment 2 support the directional cue hypothesis, in that participants performed better in the chair and social conditions than the control condition. Importantly the chair and social conditions were not significantly different from each other. This suggests that the directional cue offered by both the chair and the avatar improved performance on this task relative to the control condition. Additionally, as mentioned above, the "embodied" strategy used when a person is included in the array may be reliant on empathy, or the activation of an automatic social mechanism. Empathy for a chair is not possible or is at least uncommon, and as such the "embodied" strategy used here would be more reliant on possible embodiment or interaction with the object.

An interesting aspect of this experiment could be that the chair afforded participants the opportunity to imagine themselves *sitting in* the chair in the array. This could make the task easier for participants in the chair condition to complete by prompting use of an interactive embodied strategy (imagining themselves sitting in the array) rather than a non-interactive embodied (imagining being a different person) or abstract strategy. To control for this potential strategy suggestion, future experiments should explore other directional non-human cues, such as the stacked block figure used by Shelton et al. (2011) described above. To further examine the influence of interactivity, future research should include other cues that allow for interaction, such as a camera on a tripod.

General Discussion

Taken together, these experiments both demonstrate robust sex differences in performance on the Spatial Orientation Test (SOT), both in original paper format and the new VR format. The two versions of the task were also moderately correlated in both experiments, and the VR SOT tended to be more difficult than paper, though this effect only reached significance in Experiment 2. Although there was a significant difference in error between VR and paper SOT in Experiment 2, across the two experiments the two task formats were highly correlated, which suggests that they measure the same ability. Additionally, in both experiments the social condition was completed with significantly less error than the control condition.

It is also important to consider the inconsistencies between Experiments 1 and 2. In Experiment 1 participants performed better in the social condition than either the arrow or control conditions, while in Experiment 2 participants performed better in the chair and social conditions than the control condition. Thus, one inconsistency between these two experiments was the differing effects of the non-human directional cues (arrow and chair) on performance. Though the arrow did not significantly improve performance in Experiment 1, the chair (a different non-human directional cue) used in Experiment 2 did significantly improve performance, which could be due to the chair's interactive nature.

It could be that the arrow used in Experiment 1 was too abstract a cue of direction for people to use to make the task easier. Previous research has found a male-favored sex difference in abstract spatial ability (e.g. Voyer, Voyer, & Bryden, 1995), specifically for tasks like the Rod and Frame Test (Asch & Witkin, 1948), which can be considered a measure of spatial perception. If this abstract ability plays a role in perspective taking, and this ability is greater in males, it is possible that males were more readily able to utilize the arrow cue in Experiment 1 than females. This could also explain improvement in the chair condition relative to the control condition of Experiment 2, in that the chair may have been a more concrete directional cue that both males and females were equally disposed and able to use.

Overall the results of the present experiments provide support for both the directional cue and social mechanism hypotheses. More research is necessary to examine the multitude of factors that could influence performance (and specifically female performance) on this task. Future research should examine other non-human directional cues, particularly those that are and are not interactive, such as weather veins, cameras, and abstract block figures. The interactive-ness of a directional cue, as discussed above, may be elemental to the ease with which participants take the perspective of an inanimate object. Future studies should also consider ambiguously or incorrectly oriented human figures, to further isolate the social components that may be influential to performance. For example, the array could include a human figure that always faces forward, therefore not facing any of the objects and not changing its facing direction between trials.

Future research should also examine avatars varied in race, sex, age, and other qualities to fully explore the link between avatar-participant identification and performance (e.g. Guadagno, Blascovich, Bailenson, & McCall, 2007; Swinth & Blascovich, 2001). That only a male avatar was used in the social condition of the VR SOT could be considered a limitation of the present study. On a larger scale, future studies should utilize the immersive capabilities of VR and create a life-size immersive array within which participants can stand or walk before completing trials. Testing perspective taking ability at the environmental level would be an important component of developing training methods for perspective taking abilities that are applicable to environmental space in the real world.

The present experiments serve as a starting point from which to begin understanding the cognitive processes that underlie perspective taking ability. Future research should build upon these results by incorporating the ideas mentioned above, ultimately guiding the development of training strategies for perspective taking and other spatial abilities. These results help to identify the trainable components of perspective taking, which may in the future influence training of spatial abilities in general.

References

1. Asch, S. E., & Witkin, H. A. (1948). Studies in space orientation: I. Perception of the upright with displaced visual fields. *Journal of Experimental Psychology*, *38*(6), 762-782.
2. Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"?. *Cognition*, *21*(1), 37-46.
3. Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *34*(2), 163-175.
4. Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5-17.
5. Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin*, *94*(1), 100.
6. Fields, A. W., & Shelton, A. L. (2006). Individual skill differences and large-scale environmental learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(3), 506-515.
7. Galinsky, A. D., Ku, G., & Wang, C. S. (2005). Perspective-taking and self-other overlap: Fostering social bonds and facilitating social coordination. *Group Processes & Intergroup Relations*, *8*(2), 109-124.
8. Guadagno, R. E., Blascovich, J., Bailenson, J. N., & Mccall, C. (2007). Virtual humans and persuasion: The effects of agency and behavioral realism. *Media Psychology*, *10*(1), 1-22.
9. Hegarty, M., Richardson, A. E., Montello, D. R., Lovelace, K., & Subbiah, I. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence*, *30*(5), 425-447.
10. Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, *32*(2), 175-191.
11. Kessler, K., & Wang, H. (2012). Spatial perspective taking is an embodied process, but not for everyone in the same way: differences predicted by sex and social skills score. *Spatial Cognition & Computation*, *12*(2-3), 133-158.
12. Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & Cognition*, *29*(5), 745-756.
13. Money, J., Alexander, D., & Walker, H. T. (1965). *A standardized road map test of direction sense manual*. Baltimore: Johns Hopkins Press.
14. Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of personality*, *40*(4), 525-543.
15. Piaget, J. & Inhelder, B. (1967) The coordination of perspectives, (209-246). *In The child's conception of space*. New York: Norton & Co.
16. Schneider, W., Eschman, A., and Zuccolotto, A. (2012). *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools, Inc.
17. Shelton, A. L., Clements-Stephens, A. M., Lam, W. Y., Pak, D. M., & Murray, A. J. (2012). Should social savvy equal good spatial skills? The interaction of social skills

with spatial perspective taking. *Journal of Experimental Psychology: General*, *141*(2), 199.

18. Shelton, A. L., & McNamara, T. P. (2004). Spatial memory and perspective taking. *Memory & cognition*, *32*(3), 416-426.
19. Sulpizio, V., Committeri, G., Metta, E., Lambrey, S., Berthoz, A., & Galati, G. (2015). Visuospatial transformations and personality: evidence of a relationship between visuospatial perspective taking and self-reported emotional empathy. *Experimental Brain Research*, *233*(7), 2091-2102.
20. Tarampi, M. R., Heydari, N., & Hegarty, M. (2016). A Tale of Two Types of Perspective Taking Sex Differences in Spatial Ability. *Psychological Science*, *27*(11), 1507-1516.
21. Tversky, B., & Hard, B. M. (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition*, *110*(1), 124-129.
22. Uttal, D. H., & Cohen, C. A. (2012). Spatial thinking and STEM education: When, why and how. *Psychology of Learning and Motivation*, *57*(2), 147-181.
23. Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*(2), 250-270.

Figures

A. B.

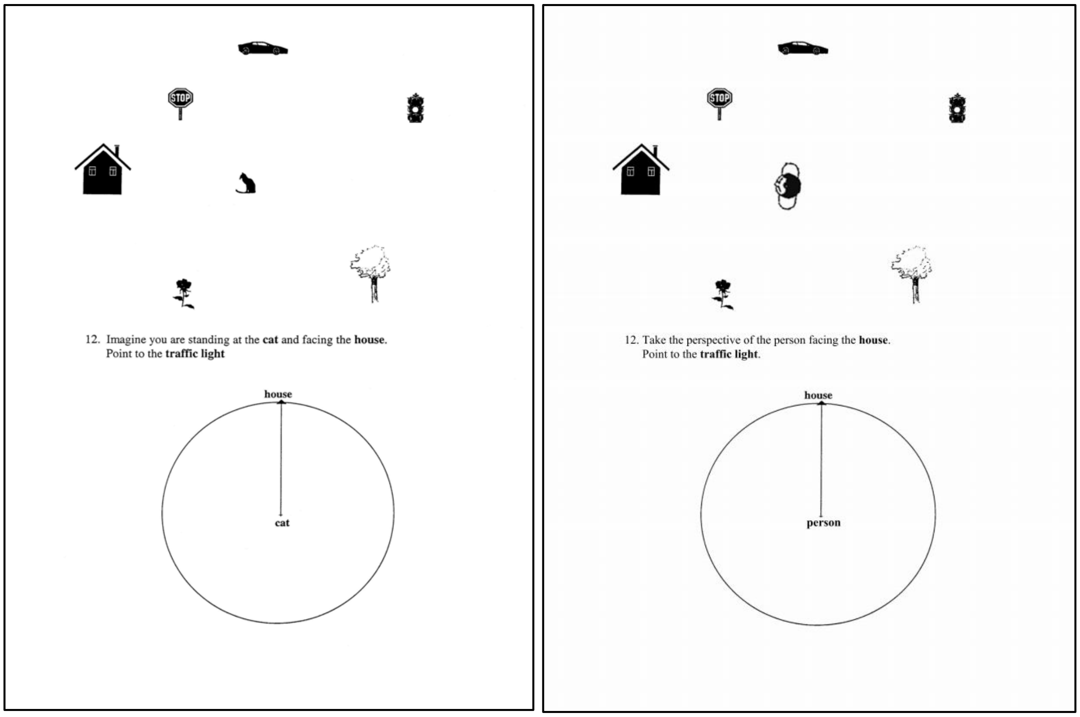

Figure 1. a) Shows a sample trial from the control condition of the Spatial Orientation Test (Kozhevnikov & Hegarty, 2001), b) shows the same trial from the social condition (with a human figure in the array) used by Tarampi et al. 2016. Only the original array (A) was used in the present experiments.

Figure 2. Shows screenshots of the VR Spatial Orientation Test from the control (A), arrow (B), social (exp. 1) (C), chair (D), and social (exp. 2) (E) conditions, an up close look at the answer circle (F), and the "feedback" given in all conditions of Experiment 2 during practice trials (G).

Figure 2A. Control condition Experiment 1.

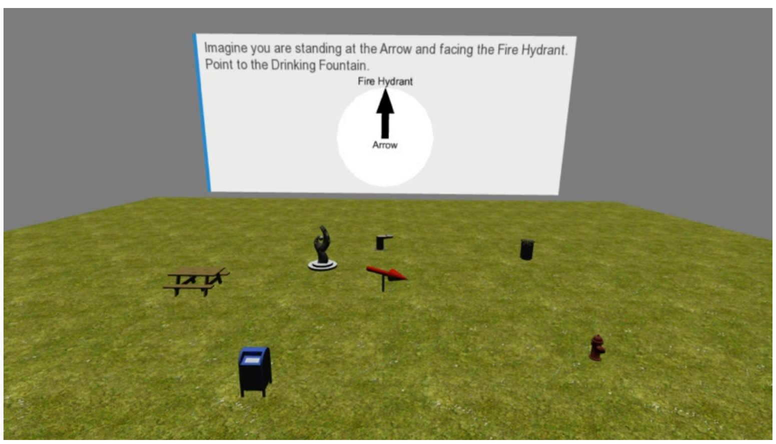

Figure 2B. Arrow condition Experiment 1.

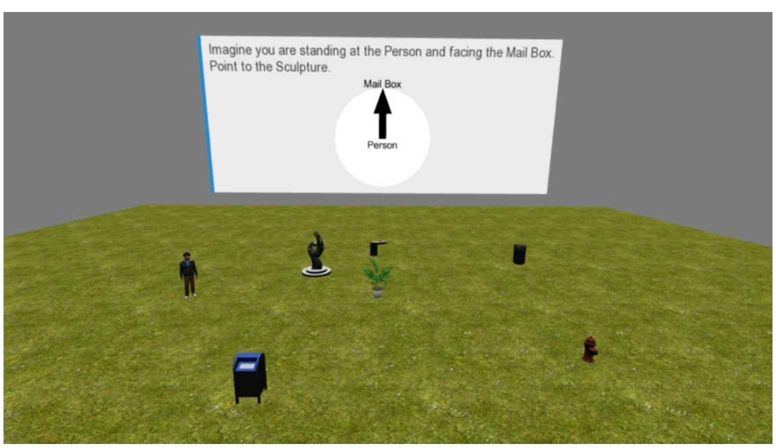

Figure 2C. Social condition Experiment 1.

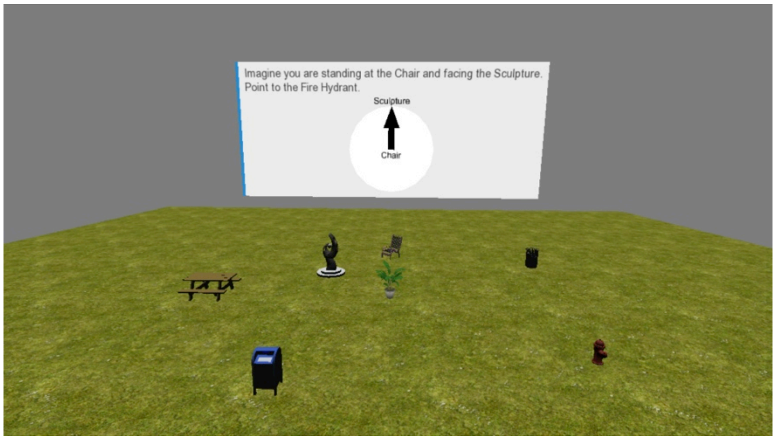

Figure 2D. Chair condition Experiment 2.

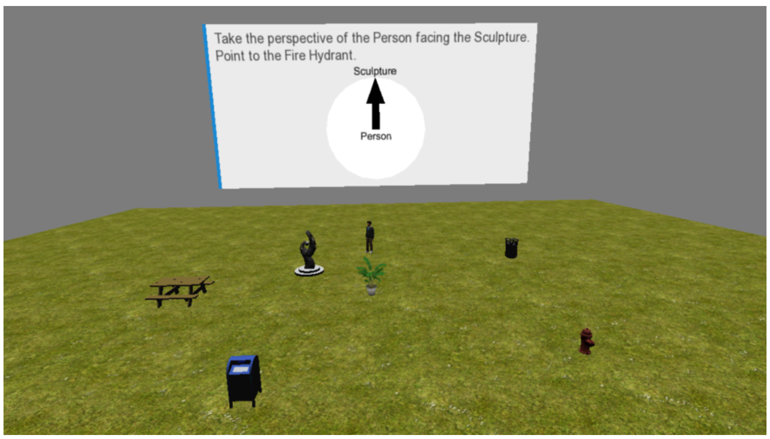

Figure 2E. Social condition Experiment 2.

Figure 2F. Close-up screenshot of the answer circle.

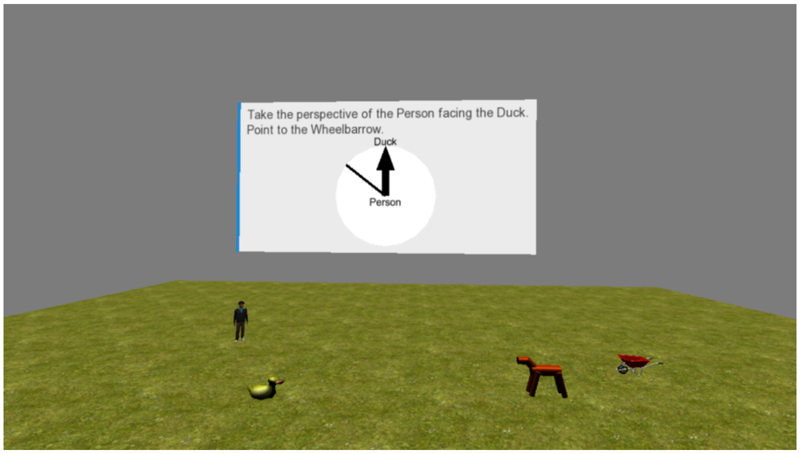

Figure 2G. "Feedback" given in Experiment 2 for first practice trial.

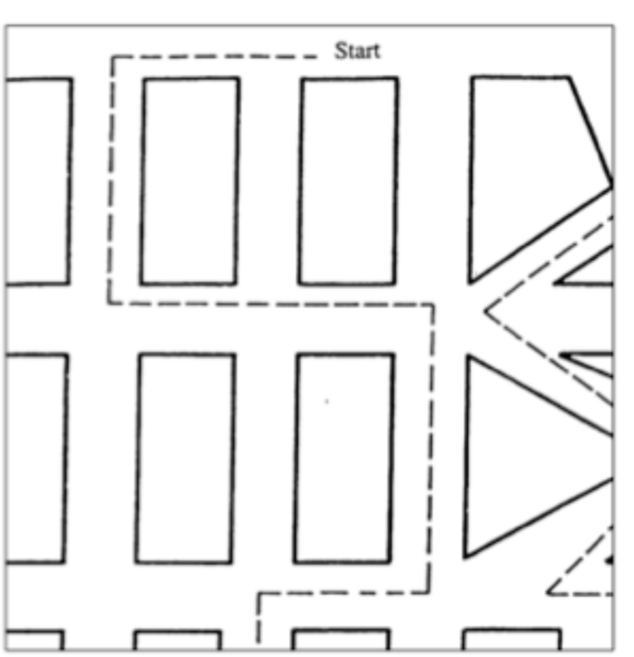

Figure 3. Sample section of the MRM.

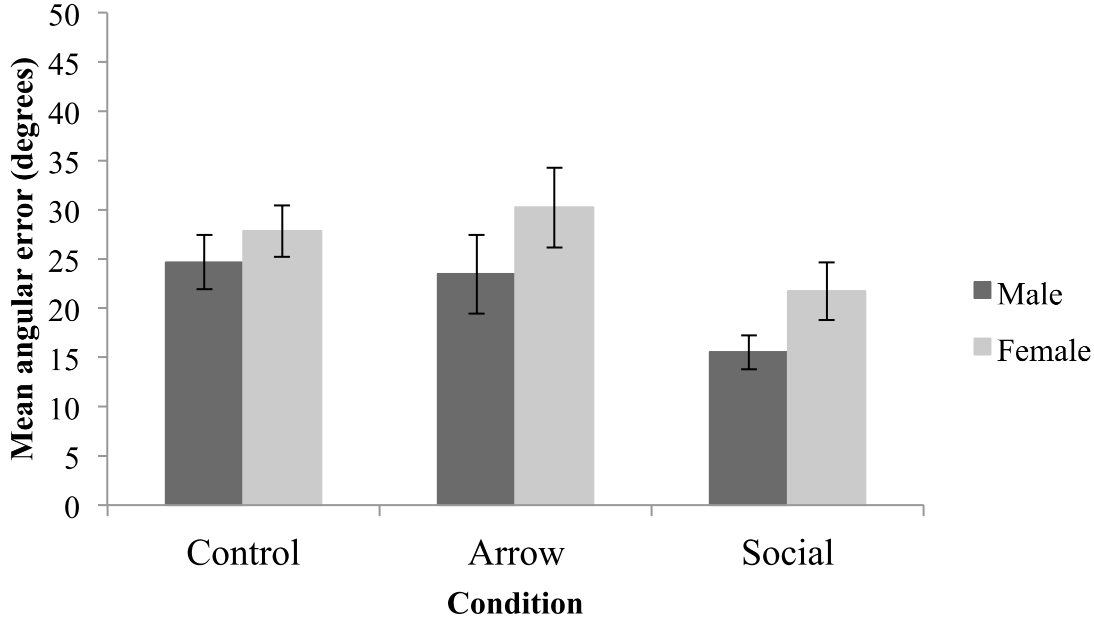

Figure 4. Results from Experiment 1 showing mean angular error of males and females in the control, arrow, and social VR conditions, with standard error bars. Note: shown here are the raw means, however analyses were conducted on the log-transformed error scores, as the data were non-normal. Fifteen subjects were excluded due to error above chance, and one subject was excluded as they declined to report their gender (n = 119).

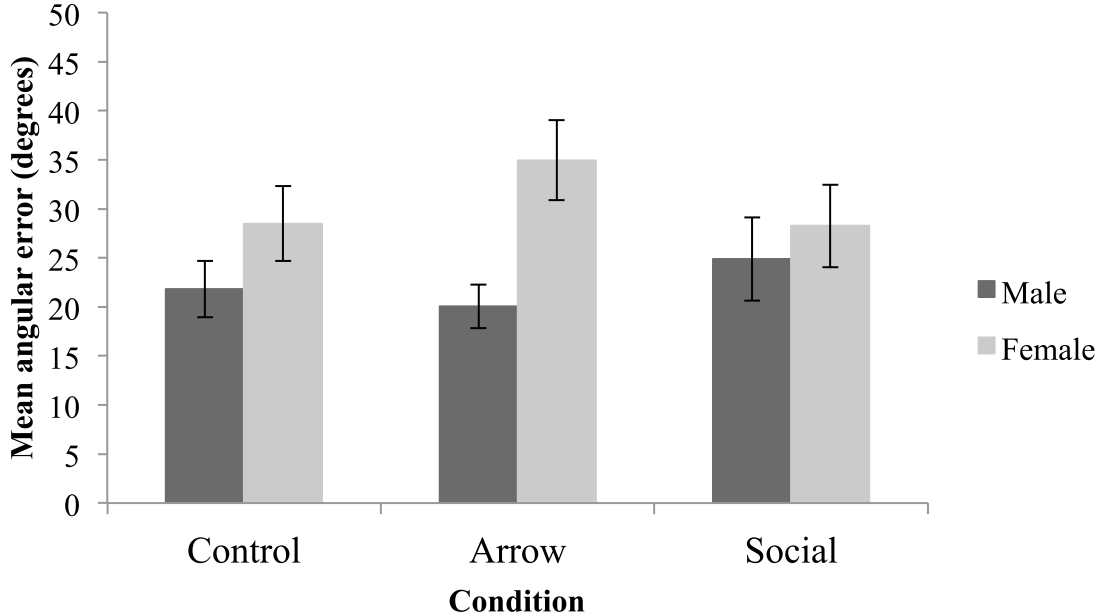

Figure 5. Results from Experiment 1 showing mean angular error of males and females in the paper SOT with standard error bars, as grouped by previous VR condition. Note: shown here are the raw means, however analyses were conducted on the log-transformed error scores, as the data were non-normal. One subject was excluded as they declined to report their gender (n =134).

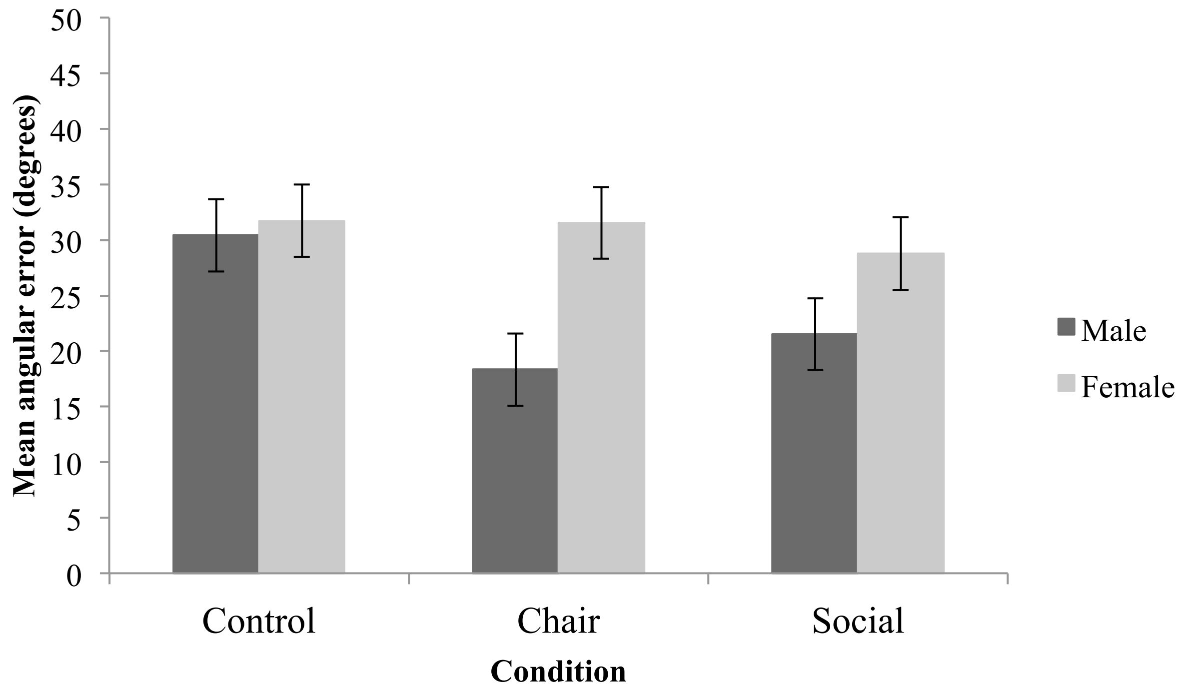

Figure 6. Results from Experiment 2 showing mean angular error of males and females in the VR SOT with standard error bars. Note: shown here are the raw means, however analyses were conducted on the log-transformed error scores, as the data were non-normal. Participants at or above chance levels of error were excluded (n =182).

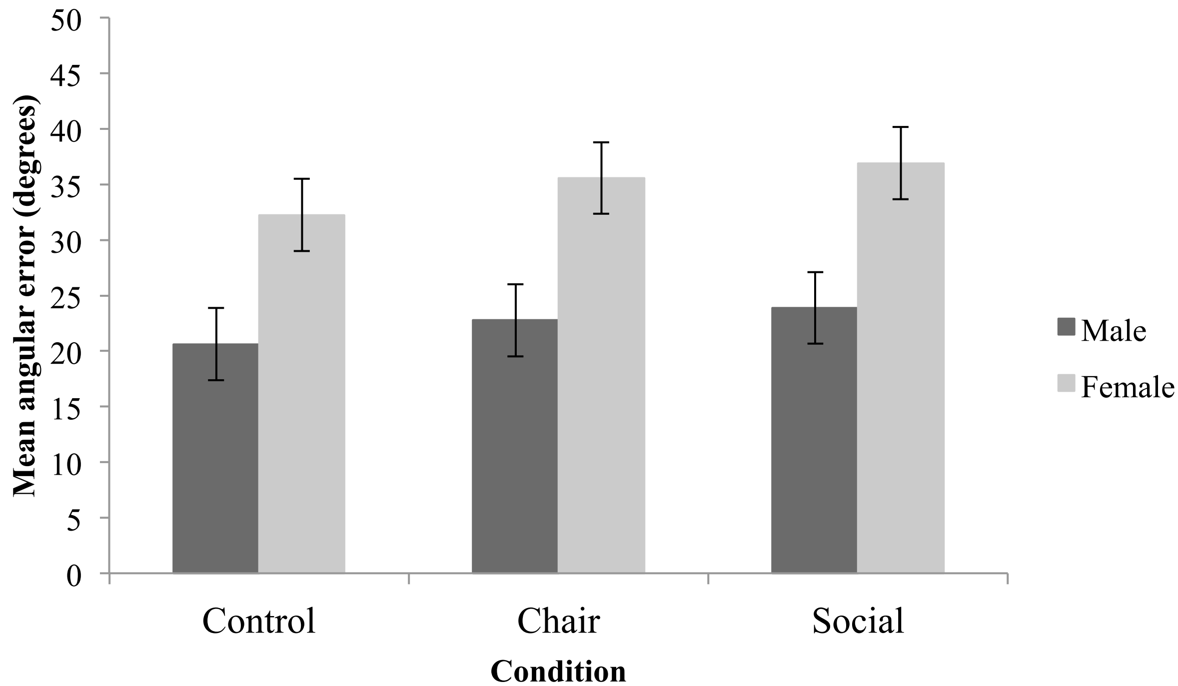

Figure 7. Results from Experiment 2 showing mean angular error of males and females in the paper SOT with standard error bars, as grouped by previous VR condition. Note: shown here are the raw means, however analyses were conducted on the log-transformed error scores, as the data were non-normal. Participants at or above chance levels of error were excluded (n =182).

Table 1. Results from Experiment 1 showing mean VR angular error of males and females in the control, arrow, and social VR conditions. Note: shown here are the raw means and log-transformed means. Fifteen subjects from Experiment 1 were excluded due to error above chance, and one subject was excluded as they declined to report their gender (n = 119).

	Raw Means		Log Means	
	Males *M (SD)*	Females *M (SD)*	Males *M (SD)*	Females *M (SD)*
Control	24.66 (12.37)	27.81 (11.59)	2.57 (0.52)	2.68 (0.49)
Arrow	23.45 (17.88)	30.22 (18.14)	2.45 (0.56)	2.84 (0.67)
Social	15.51 (7.53)	21.7 (13.09)	2.09 (0.42)	2.34 (0.53)

Table 2. Results from Experiment 2 showing mean VR angular error of males and females in the control, arrow, and social VR conditions. Note: shown here are the raw means and log-transformed means. All participants from Experiment 2 were included (n = 191).

	Raw Means		Log Means	
	Males *M (SD)*	Females *M (SD)*	Males *M (SD)*	Females *M (SD)*
Control	32.67 (22.91)	40.46 (32.6)	2.71 (0.61)	3.07 (0.76)
Chair	20.33 (15.18)	31.52 (25.29)	2.34 (0.76)	2.65 (0.75)
Social	21.51 (11.44)	37.76 (30.98)	2.37 (0.43)	2.94 (0.77)

Table 3. Correlations for Experiment 1 between log transformed mean angular errors on the VR SOT, log transformed mean angular errors on the paper SOT, the Money Road Map test (MRM), and the Santa Barbara Sense of Direction Scale (SBSOD), with all subjects excluded who had at or above chance levels of error.

	Paper SOT	VR SOT	MRM	SBSOD
VR SOT	.38**	--	--	--
VR SOT Control	.47**	--	-.49**	.44**
VR SOT Arrow	.43**	--	-.54***	.11
VR SOT Social	.26	--	-.49**	.08
MRM	-.50***	-.36***	--	--
SBSOD	.23*	.20*	-.30**	--

Note. Lower scores of angular error indicate better performance on the SOTs. **p* < .05; ***p* < .01, ****p* < .001.

Table 4. Correlations for Experiment 2 between log transformed mean angular errors on the VR SOT, log transformed mean angular errors on the paper SOT, the Money Road Map test (MRM), and the Santa Barbara Sense of Direction Scale (SBSOD), with 14 subjects not included due to experimenter error (n = 177).

	Paper SOT	VR SOT	MRM	SBSOD
VR SOT	.54***	--	--	--
VR SOT Control	.58***	--	-.54***	.09
VR SOT Chair	.59***	--	-.57***	-.24
VR SOT Social	.49***	--	-.49***	.20
MRM	-.55***	-.55***	--	--
SBSOD	.08	.04	-.02	--

Note. Lower scores of angular error indicate better performance on the SOTs. **p* < .05; ***p* < .01, ****p* < .001.

Appendix

Survey Questions

1. Other students have reported a range of strategies that they used to do these tasks. Please indicate which strategy is closest to the one that you used when doing each of the tasks. For this task what strategy did you use the most: [VR array] [VR answer circle].
    a. I mentally superimposed the whole array of objects on the answer circle to formulate my answer.
    b. I superimposed the answer circle on the array of objects to formulate my answer.
    c. I imagined drawing the angle between the object I was facing, my imagined location and the target object. I then moved and rotated that angle and superimposed on the answer circle to formulate my answer.
    d. I imagined myself being in the array, at the location I was told to imagine, turning my body to the imagined facing direction and figured out where the target object would be in relation to my body.
    e. I did not use any of these strategies. I used the following strategy:
2. For this task what strategy did you use the most [paper array] [paper answer circle]:
    a. I mentally superimposed the whole array of objects on the answer circle to formulate my answer.
    b. I superimposed the answer circle on the array of objects to formulate my answer.

c. I imagined drawing the angle between the object I was facing, my imagined location and the target object. I then moved and rotated that angle and superimposed on the answer circle to formulate my answer.
d. I imagined myself being in the array, at the location I was told to imagine, turning my body to the imagined facing direction and figured out where the target object would be in relation to my body.
e. I did not use any of these strategies. I used the following strategy:

3. Santa Barbara Sense of Direction Scale (SBSOD)
4. What is your age?
5. What is your gender?

   Male Female Other Decline to answer

6. What is your major?
7. What is the highest level math course you have taken?
8. What is your ethnicity? (Select all that apply)

   White, African American or Black, Hispanic, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Other

9. (Social condition only) How much did you identify with the avatar?

   Not at all Not much Moderately Most of the time All of the time

Note: The strategy questions were scored as follows: strategy choices a and b were coded as abstract, while choices c and d were coded as embodied. Any text entries were coded by two independent raters as either abstract or embodied. The SBSOD scale was scored as per author recommendation. See the following, "The recommended scoring procedure for the scale is to first reverse score the positively phrased items. This ensures that all items are coded such that a high number indicates more ability and a low number indicates less ability. The items that should be reverse scored are items 1, 3, 4, 5, 7, 9, and 14. After reverse scoring, then sum the scores for all of the items together, and then divide the total by the number of items (15) to compute the overall score for the scale (average score across items). Using this technique, the score will be a number between 1 and 7 where the higher the score, the better the perceived sense of direction," (Hegarty et al., 2002).