

A High-End Virtual Reality Setup for the Study of Mental Rotations

Alexandre Lehmann, Manuel Vidal, Heinrich H Bü

▶ **To cite this version:**

Alexandre Lehmann, Manuel Vidal, Heinrich H Bü. A High-End Virtual Reality Setup for the Study of Mental Rotations. Presence: Teleoperators and Virtual Environments, 2008, 17, pp.365 - 375. 10.1162/pres.17.4.365 . hal-01519785

HAL Id: hal-01519785

https://amu.hal.science/hal-01519785

Submitted on 9 May 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Alexandre Lehmann†
LPPA-CNRS
Collège de France
Paris, France

Manuel Vidal*,†
Max Planck Institute for Biological Cybernetics
Tübingen, Germany
and
LPPA-CNRS
Collège de France
11 place Marcelin Berthelot
75005 Paris, France

Heinrich H. Bülthoff
Max Planck Institute for Biological Cybernetics
Tübingen, Germany

A High-End Virtual Reality Setup for the Study of Mental Rotations

Abstract

Mental rotation is the capacity to predict the orientation of an object or the layout of a scene after a change in viewpoint. Previous studies have shown that the cognitive cost of mental rotations is reduced when the viewpoint change results from the observer's motion rather than the object or spatial layout's rotation. The classical interpretation for these findings involves the use of automatic updating mechanisms triggered during self-motion. Nevertheless, little is known about how this process is triggered and particularly how sensory cues combine in order to facilitate mental rotations. The previously existing setups, either real or virtual, did not allow disentangling the different sensory contributions, which motivated the development of a new high-end virtual reality platform overcoming these technical limitations.

In the present paper we will start by a didactic review of the literature on mental rotations and expose the current technical limitations. Then we will fully describe the experimental platform that was developed at the Max Planck Institute for Biological Cybernetics in Tübingen. The setup consisted of a cabin mounted on the top of a six degree-of-freedom Stewart platform inside of which was an adjustable seat, a physical table with a screen embedded, and a large projection screen. A 5-PC cluster running Virtools was used to drive the platform and render the two passive stereovision scenes that were displayed on the table and background screens. Finally, we will present the experiment using this setup that allowed replicating the classical advantage found for a moving observer, which validates our setup. We will conclude by discussing the experimental validation and the advantages of such a setup.

1 Introduction

As humans navigate through unfamiliar terrain, they have to efficiently extract and encode spatial information about their body and the environment. Several mechanisms driving human behavior and actions within the natural habitat, such as the construction of spatial representations, rely on updating processes that are continuously at work. One of these mechanisms, acting during changes in viewpoint, is mental rotation. This capacity generally involves redundant sensory information about the world and our motion that can come either from external cues such as optic flow or internal cues (idiothethic) such

Presence, *Vol. 17, No. 4, August 2008,* 365–375
© 2008 by the Massachusetts Institute of Technology

*Correspondence to manuel.vidal@college-de-france.fr.
†The first two authors contributed equally to this research.

as the vestibular or proprioceptive signals. We developed an innovative experimental platform in order to study mental rotations. Before presenting this platform, we will first review the literature on mental rotations, starting from the viewpoint dependency of spatial memory to the general finding that moving observers perform better. In the last section, we will describe current issues in the mental rotation community, some being controversial. In our view, the ongoing debated issues can be explained by the lack of appropriate means to investigate them. Throughout this section, the limitations of previous setups will be pointed out. The hybrid system presented here, halfway between physical and virtual setups, was specifically designed to overcome many of these existing limitations. It will allow future unique experiments addressing current issues in mental rotation research. The purpose of the present paper is to describe this new platform and present the results of its experimental validation through the replication of classical mental rotation findings.

1.1 Viewpoint Dependency of Spatial Memory

It has been shown that passive recognition of objects depends on the observer's viewpoint. Shepard and Metzler (1971) used 3D objects with intrinsic spatial structure to test people's ability to recognize novel objects across multiple viewpoints. They used a task in which subjects were simultaneously shown two pictures depicting an object and they had to decide whether those objects were the same or different. The two pictures could arise from different viewpoints varying from 0° to 135°. Subjects performed faster when tested with the learned view than with a novel viewpoint. Moreover, Shepard and Metzler found that participants' reaction time (RT) was a linear function of the angle between learned and presented view. This led them to suggest that people have to perform a mental rotation in order to align the test stimulus with the learned stimulus in the observer's reference frame. They hypothesized that this process was done at constant angular speed, thus explaining the linear RT relationship. Since then, these findings have been extended to the recognition of arrays of objects, either visually (Diwadkar & McNamara, 1997) or haptically learned (Newell, Woods, Mernagh, & Bülthoff, 2005), and more generally to large layouts (Roskos-Ewoldsen, McNamara, Shelton, & Carr, 1998; Shelton & McNamara, 1997). In these studies, spatial relationships within a configuration of objects showed a dependency on the viewpoint adopted while learning the spatial layout.

It follows that mental rotations can be defined as the capacity to predict the new spatial relationships of a layout after its rotation in the observer's reference frame. For a stationary observer, as is the case in almost all object recognition studies, this definition coincides with the process involved in the studies mentioned above. It is interesting to note that a vast majority of object recognition studies has focused solely on a passive observer receiving different visual inputs. Even though this paradigm provides experimental convenience and allowed researchers to build recognition models that were successfully applied to computer vision, it fails to explain the growing body of literature on active observer motion, perhaps because it lacks some ecological validity. Indeed, in real life, we rarely observe arrays of independent objects changing their orientation in synchrony. In fact, changes in orientation usually occur when the observer moves around the object or scene.

1.2 Better Performance for Moving Observers

A given change in viewpoint of a spatial layout can occur either as a result of the rotation of the layout in front of a stationary observer (e.g., on a rotating table) or the rotation of the observer around a stationary layout. The resulting change in retinal projection being exactly the same, traditional models of object recognition would predict a similar behavioral outcome. Simons and Wang (1998) designed an experiment in which people sat in front of an array of five objects arranged on a table. Subjects had to learn the layout of the objects for 3 s. The objects were then occluded for 7 s during which one of the objects was moved and the observer's viewpoint was changed either by rotating the table or by having the subject walk to another viewing position. Subjects then had to tell which of the five objects had been displaced. Two control conditions (same

retinal projection before and after retention interval) were used: one in which neither the observer nor the table moved and one in which both the observer and the table moved in the same direction. They found that, for a static observer, performance was less accurate when the retinal projection changed than when the view remained identical. In contrast, they found no significant difference with or without change in the viewpoint when observers moved. Their results were completely unexpected in light of classical object recognition theories and suggested that different mechanisms are at stake for static and moving observers.

This advantage for a moving observer has been reported in the case of imagined rotations. Early work by Amorim and Stucchi (1997) has shown that mental explorations, namely rotations, of an imagined clock were not equivalent in object-centered and viewer-centered task. They found an extra processing cost in the object-centered condition relative to the viewer-centered condition, also suggesting different underlying mechanisms. Wraga, Creem, and Proffitt (2000) found a viewer advantage for subjects imagining object/viewer rotations of arrays that were on a table or around them in the room. This effect also holds true for a single object but seems to decrease with object familiarity.

Experiments based on virtual reality setups also replicated this effect. By using a purely visual replication of Simons and Wang's task, Christou and Bülthoff (1999) did not observe the viewer advantage in subjects' accuracy, although there was a significant cost reduction in the reaction times. This might arise from the fact that the virtual setup they used was poorly immersive: subjects viewed the virtual room on a computer monitor, subtending a narrow field of view, providing subjects with a weaker optic flow cue than in real-world situations. The fact that they used a two choice discrimination task (change/no change) could also partly account for the lack of significant effect on performance. In another study that used motion tracking to update a virtual scene, subjects were asked to point to learned locations of alcoves after body/display rotations (Wraga, Creem-Regehr, & Proffitt, 2004). Performance was significantly improved (higher accuracy and lower response latencies) in the viewer task as compared to the display task, thus in favor of the viewer advantage. In light of the literature mentioned above, it appears that retinal information alone is not enough for building a theory of object recognition. In fact, other sources of information such as proprioceptive and vestibular cues, but also optic flow cues, have to be taken into account.

1.3 Spatial Updating: Explaining the Viewer Advantage

Spatial updating refers to people's ability to update spatial relationships between themselves and the environment during self-motion. Typical tasks used to test the spatial updating capacity are the pointing to locations of remembered objects (Wang & Spelke, 2000; Loomis, Klatzky, Golledge, & Philbeck, 1999), triangular completion tasks (Klatzky, Loomis, Beall, Chance, & Golledge, 1998; Amorim, Glasauer, Corpinot, & Berthoz, 1997; Riecke, van Veen, & Bülthoff, 2002) and foraging tasks (Ruddle & Lessels, 2006). To date, the most widely accepted interpretation is that viewer mode situations benefit from a spatial updating process. This egocentric updating capacity, using an unknown combination of sensory cues, would allow for greater performance in natural situations where changes in viewpoint arise from observer's motion. Indeed, such capacity is not available in object mode where the observer's viewpoint remains the same.

This hypothesis is given much credit in light of brain imaging studies such as the one by Zacks, Vettel, and Michelon (2003). Using an fMRI paradigm where people had to imagine an array of objects rotating or imagine themselves rotating around the array, the authors found a dissociation between object-based spatial transformations and egocentric perspective transformations. Imagined object rotations led to selective increases in the right parietal cortex and decreases in the left parietal cortex, whereas viewer rotation led to selective increases in the left temporal cortex. Their results argue against the view that mental image transformations are performed by a unitary neural processing system, and fits well with behavioral data from mental rotation studies. The viewer advantage thus seems to arise because a partly different neural system is selectively engaged for egocentric spatial transformation (i.e., spatial updating capacity).

1.4 Debated Issues and Related Technical Limitations

Current directions of mental rotation research point toward the relative contributions of many different cues to this egocentric updating capacity. Therefore, it is of critical importance to be able to fully disentangle the contribution of each sensory modality in the viewer advantage for mental rotations. Except for the study of Christou and Bülthoff (1999), with the previously existing setups one could not trigger a pure visual rotation of the observer. The platform presented here not only allows this visual manipulation but also allows researchers to combine it with other sensory cues. It will thus be possible to design future experiments addressing the following unresolved issues.

1.4.1 Role of Visual Cues. Many studies have indeed tried to control for various other interpretations of this effect. Simons and Wang (1998) investigated whether background visual cues could explain this difference by using phosphorescent objects in a dark room but they still found a viewer advantage. In a later study, the authors conclude that the difference between viewpoint and orientation changes cannot be fully explained by visual cues (Simons, Wang, & Roddenberry, 2002). They go even further, claiming that visual cues do not contribute at all to this advantage. Such a claim is difficult to hold, knowing that the visual information provided in this study was very poor compared to real-life situations. The background they used was uniform, subjects had a limited field of view, and in particular no dynamic visual information of their rotation. Indeed, only static snapshots before and after the change in viewpoint were provided. Christou, Tjan, and Bülthoff (2003) used a paper clip–type stimulus, a multipart stimulus designed to have intrinsic spatial relationships equivalent to those found in arrays of objects. Subjects had to recognize such objects learned from different viewpoints and they could be provided with background cues (fixed frame of reference) or with new viewpoint cues. They found that extrinsic visual context facilitated the accuracy of subjects and that such facilitation could not be attributed to spatial updating as pointed out by Simons and Wang. Along this line, another study by Burgess, Spiers, and Paleologou (2004) also shows contradictory results. They used a similar paradigm to Simons and Wang's in which they introduced a phosphorescent landmark external to the array. This landmark could move congruently with the egocentric or with the allocentric frame of reference. They conclude that part of the effect attributed to egocentric updating by Simons and Wang can be explained by the use of allocentric knowledge provided by the room. Further research needs yet to be done to fully address this issue. The experimental setup described in this paper provides an ideal tool for those studies.

1.4.2 Role of Active Control and Rotation Magnitude Availability. In order to investigate whether active control of the change in viewpoint could explain the viewer advantage, Wang and Simons (1999) first tested whether adding active control of the table rotation for static observers could improve the mental rotation. The table rotation was produced by a handle that could be manipulated either by the experimenter or the subject. They found no significant improvement when subjects controlled the table rotation. Secondly, they tested whether removing active control in a moving observer would impair performance in this condition. They passively rolled subjects in wheelchairs and found no significant difference to active movement. These two results suggest that active control is not central to the process underlying the viewer advantage. By adding the view of a passive handle they claim to have also ruled out differences arising from the availability of the continuous rotation magnitude information. One could still argue that this cue did not provide sufficient object-related online-rotation information. In contrast, Wraga et al. (2000) found that adding passive haptic information—subjects could feel an object rotating in their hands during the rotation phases—decreased the cost of object mental rotation to the level of a moving observer. Wexler, Kosslyn, and Berthoz (1998) reported a consistent facilitation effect in similar conditions. To cope with these conflicting results, we plan to investigate this issue further in one experiment by providing participants with a strong and continuous cue about the rotation magnitude—leaving the table visible during its

rotation but hiding the objects. Such a manipulation can easily be achieved using our experimental platform.

2 Materials and Methods

2.1 Participants

Twelve naïve subjects (five females and seven males, aged from 21 to 33) participated in this experiment. They received a financial compensation for their time of 12 euros. All subjects were right-handed and had a normal or corrected-to-normal vision. They all gave prior written consent. The whole experiment lasted approximately 75 min and was approved by the local ethics committee.

2.2 Apparatus

The experiment was conducted using an interactive virtual reality setup that allowed immersing participants in a partially virtual environment. The setup consisted of a closed cabin mounted on the top of a 6 DOF Stewart platform (see Figure 1). Inside of this cabin was an adjustable seat where participants rested, with a physical table placed in between the seat and a large projection screen. The seat position was adjusted in height and longitudinally in order to have a constant specific viewing position across participants. The viewpoint was set to 50 cm away from the table vertical axis, 38.5 cm above the table-screen surface, and 138 cm away from the front projection screen (subtending 61° of horizontal FOV). A distributed application was developed with the Virtools VR Pack in order to drive a cluster of five PCs connected via a gigabit network. This real-time application was controlling synchronously: body motion, visual background, table layout with response interface, and monitoring of the experimental progression.



Figure 1. *A schematized model of the experimental setup: the moving platform with a front projection screen, a table-embedded screen with a touch-screen, and a noise-cancellation headphone.*

2.2.1 Body Motion. In trials where participants' viewpoint was to be changed, a smooth off-axis yaw rotation of the body was performed around the table vertical axis. The rotation amplitude was 50° to the right and the duration was 5 s using a raised-cosine velocity profile. The repositioning to the starting position was performed using a trapezoidal velocity profile with a maximum instant velocity of 10 °/s, and onset duration of 500 ms.

2.2.2 Visual Background Stimulation. The virtual environment consisted in a detailed model of a rectangular room (2 m wide × 3 m long). The furniture in the room was designed with Autodesk 3ds Max modeling software and rendered in real-time with Virtools' engine. The table displaying the test objects was physically and virtually located in the center of this room. In trials where participants were changing viewpoint, a smooth visual rotation corresponding to the physical rotation around the table was simulated in the front projection screen.

The scene was displayed using a passive stereoscopic vision technique based on anaglyphs, with an interpupilar distance of 6.5 cm separating the left and right

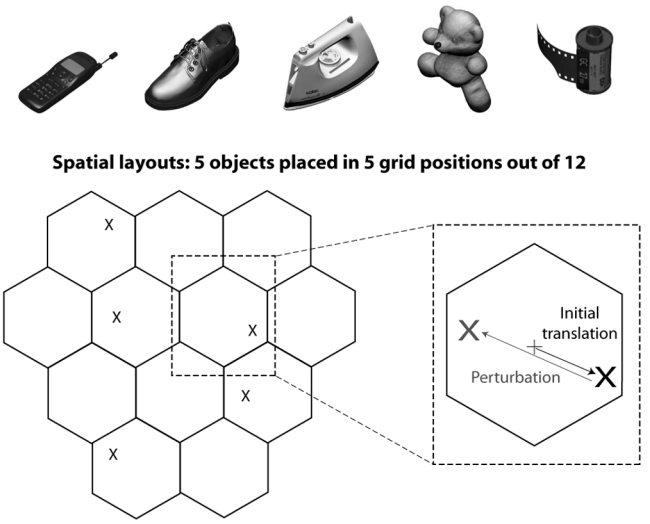


Figure 2. *(Top) The five objects used in the mental rotation task: a mobile phone, a shoe, an iron, a teddy bear, and a cannister of film. (Bottom) An example of object layout generation: the five objects are placed at the center of five distinct cells of an alveolus grid, and then translated randomly so as to remain within the cell (initial translation, black vector). In order to maximize the number of possible movable objects (avoiding their collapse) for the detection task, the perturbation vector of a given object followed the opposite direction of its initial translation in the cell (gray vector).*

camera in the virtual scene. One image for each eye (left and right) was rendered and superimposed using appropriate color filters (red and cyan) before being displayed. The left-eye image corresponded to the red channel of the rendering of the left-camera and the right-eye image corresponded to the sum of the blue and green channel of the rendering of the right-camera. This color technique was selected in order to preserve the perception of colors of the environment, and was computed in real-time with a DirectX® 9.0 HLSL pixel shader. During the entire experiment, participants wore a pair of red/cyan spectacles in order to correctly filter each image that was to be displayed for each eye.

2.2.3 The Table Displaying the Objects and the Response Interface. The same five objects were always used for the mental rotation task (mobile phone, shoe, iron, teddy bear, and film; see Figure 2, top). The size and familiarity of these objects were matched. The size was adjusted so as to have equivalent projected surfaces on the horizontal plane and limited discrepancy in height across objects. As for the familiarity, we chose objects from everyday life that could potentially be found on a table. The table placed in front of the participants consisted of a black cylindrical frame resting on a single central leg, inside of which a 21-in. TFT screen was embedded in order to display the object layouts (see Figure 1). The frame was hiding all but a central circular portion sustaining a diameter of 29 cm of the screen surface. The same stereoscopic rendering technique as for the front screen was used, and the asymmetrical frustum was adjusted according to the participant's viewing position. The geometrical configuration of the objects was generated automatically according to specific rules. The object layouts used for each trial were based on a 12-alveolus beehive-like grid (see Figure 2, bottom), with a distance of 6 cm between two cell-centers. Five cells were picked randomly in which the five objects were placed at a random distance from the center of the cell (initial translation vector in Figure 2). The mental rotation task required participants to select one of the objects resting over the table. In order to make the answering as natural and intuitive as possible, we decided that participants should be able to simply touch the object with their dominant-hand index, as illustrated in Figure 1. Therefore, the table screen was equipped with a 3M ClearTek™ II touch screen. A detection algorithm was implemented to compute which object was the closest to the contact point within a maximum range of 1.5 cm. The name of the selected object was displayed on the table-screen at the end of the response phase, in order to provide a possible control for errors in the automatic detection process and to provide feedback with the response interface in general.

2.2.4 Monitoring of the Experiment. An infrared camera allowed monitoring of the participants on the motion platform in the control area of the experimental room. Data from each ongoing trial were displayed on a monitor in order to check the correct progression of the experiment, and allow intervening in case of touch-screen dysfunction (incapacity to touch an object). Subjects were warned about this possible dysfunction and were instructed to simply say their answer

aloud if it happened. Participants wore noise-cancellation headphones in order to suppress most of the environmental noise (mainly platform leg motions). The experimenter could speak to participants through these headphones via a microphone/loudspeaker system in the control area.

2.3 Procedure

2.3.1 Time-Course of a Trial. On each trial, participants viewed a new layout of the five objects on the table for 3 s (learning phase). Then the objects and the table disappeared for 7 s (rotation phase). During this period, participants might rotate or not in the environment and one of the five objects was systematically translated 4 cm (perturbation vector in Figure 2). Object translations followed two constraints: first, to ensure that at least 3 cm separated the moved object from each of the remaining objects in the tested layout, and second, not to move closer than 1.5 cm from the edge of the visible circular portion of the screen. The objects were then shown again, and participants were asked to pick the object they thought had moved by simply touching the object on the surface of the screen (test phase).

2.3.2 Conditions. Each participant experienced four different kinds of trials according to the tested condition (see Figure 3). Conditions were defined as a combination of two factors: the viewing position (unchanged or changed) and the retinal projection (same or different). For half of the trials, observers remained at the same viewpoint in the (virtual) room for both the learning and test phase (unchanged viewing position). For the other half of the trials, participants learned the object layout from the first viewpoint, and during the hidden phase they were passively rotated around the table by 50° to the second viewpoint (changed viewing position). In trials where the retinal projection was the same, participants were tested with the same view of the table as in the learning phase. When there was a change in viewing position, the table was rotated so as to compensate for the observer rotation. In trials where the retinal projection was different, they were tested with the same 50° view change, resulting either from the rotation of the table or the rotation of the observer around the table. The two conditions with a change in retinal projection correspond to the situation where a mental rotation is required, whereas the others are control conditions that will be used to evaluate the cost of the mental rotation according to a change or not in the viewing position.

2.3.3 Experimental Design. Each experimental condition was tested 20 times for a total of 80 trials. In order to avoid on the one hand the difficulty in switching from one condition to another, and on the other hand, order effects when presenting each condition in separate blocks, trials were partially blocked. The order was counterbalanced both within and across subjects using a nested Latin-Square design as in Wang and Simons (1999). Trials were arranged into blocks of five trials from the same condition, and these blocks of five were arranged into blocks of 20 trials with five trials from each condition. Four different orders of the conditions within a block of 20 were created using a Latin-Square design, and each subject experienced all four of these blocks of 20. The order of the blocks of 20 was counterbalanced across subjects, also using a Latin-Square design. At the beginning of each block, participants were informed about the test condition with a text message displayed on the front screen. This defined whether and how the view of the layout would change ("The table will rotate" or "You will rotate around the table") or remain the same ("Nothing will rotate" or "You and the table will rotate").

2.3.4 Data Analysis. For each trial, the accuracy in detecting the moved object and the reaction time (RT) were recorded. A two (viewing position) × two (retinal projection) repeated-measures ANOVA design was used to analyze accuracy and RT. Gender was introduced as a categorical factor and post hoc analyses were performed with Scheffe's test. On a few occasions (11 times out of 960), subjects could not answer by touching the object because of a malfunctioning touch screen (see Section 2.2.4, Monitoring of the Experiment). For further analysis, response times for these few data points were replaced by the average response time in the same condition.

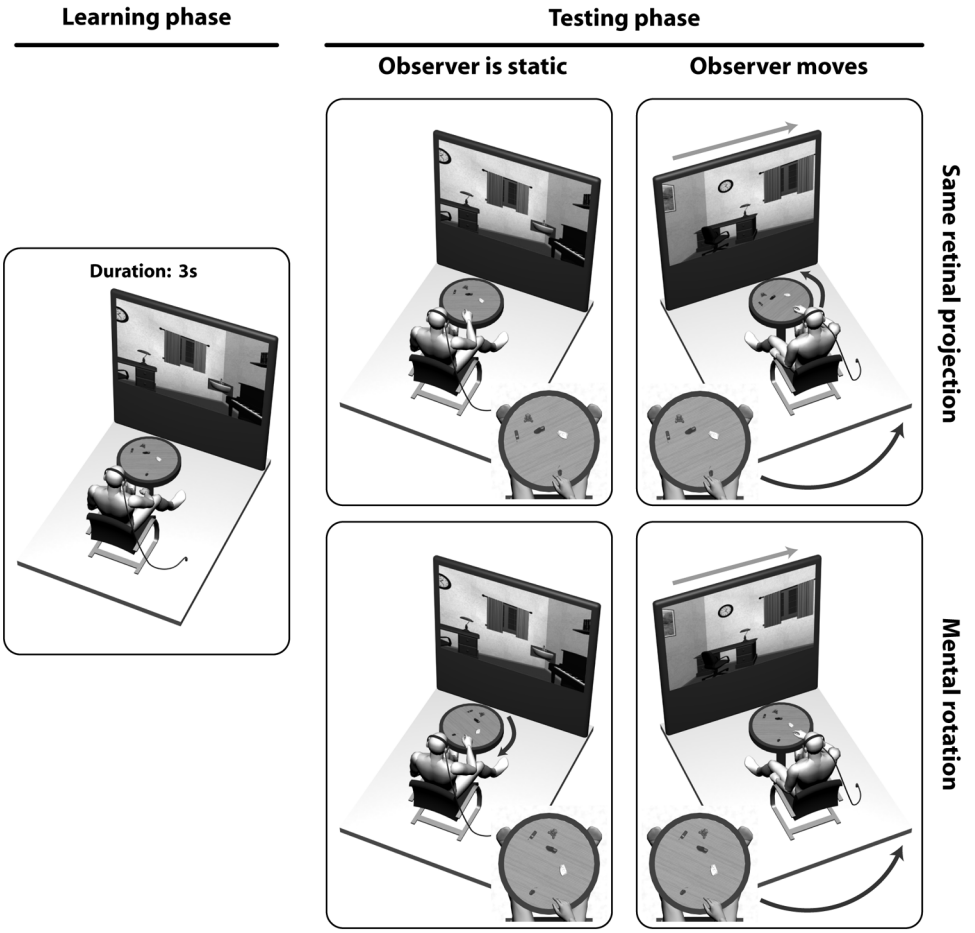


Figure 3. *Illustration of the four conditions, defined as a combination of retinal projection (same or different) and viewing position (unchanged or changed) factors. The inset shows the table as seen from the subjects' viewpoint.*

3 Results

First of all, there was no significant effect of participants' gender on the mental rotation performance, neither for the accuracy (70.1% and 70.4% for females and males, respectively) nor for the reaction times (4.58 s and 4.84 s, respectively). Data for both genders will be pooled together in the following analyses.

When subjects had to execute a mental rotation (following a change in retinal projection), performance was significantly disrupted (see Figure 4). Accuracy in detection dropped by 23.8% ($F(1,10) = 55.48$; $p < .0001$) and the measured reaction times increased by 1.78 s ($F(1,10) = 42.11$; $p < .0001$). Post hoc tests on accuracy revealed that this difference was significant both for a table rotating in front of a static observer ($p < 0.0001$) and a rotation of the observer around a static table ($p < 0.05$).

For each of the viewing position factors (unchanged or changed), the cost of the mental rotation was computed as the difference between the average performance when there was no change of the view of the table (same retinal projection) and when there was a rotation of the table in the observer's reference frame (different retinal projection). The gray arrows in Figure 1 illustrate these costs; note that the sign convention for the reaction times was reversed so as to deal with positive values only. This cost analysis is mathematically

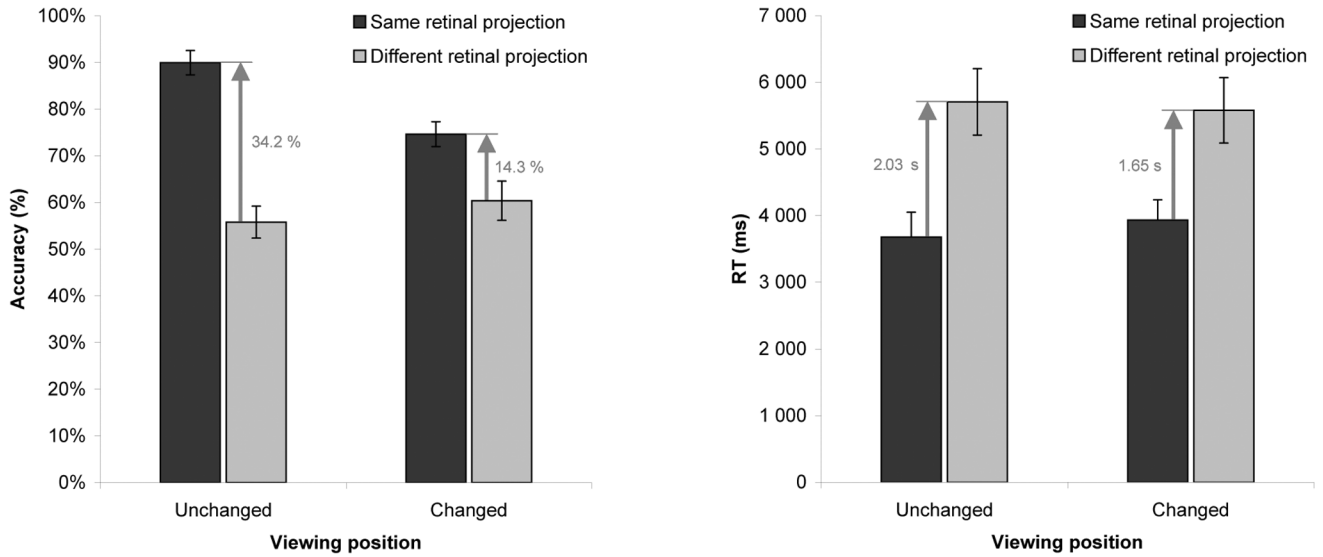


Figure 4. *The mental rotation task performance: the average accuracy (left) and reaction time (right) plotted as a function of the change in viewing position and retinal projection. The error bars correspond to the inter-individual standard error. The costs of the mental rotation are shown with gray arrows for both the accuracy and reaction time.*

equivalent to the analysis of the interaction between viewing position (unchanged or changed) and retinal projection (same or different) factors, as done in the previous studies (Wang & Simons, 1999). There was a significant cost reduction for the accuracy when observers rotated around the table (14.3%) as compared to static observers (34.2%), showing that the mental rotation performance is enhanced when there is a change in viewpoint ($F(1,10) = 10.99$; $p < .008$). Consistently, a small decrease in the average RT cost was also measured (1.65 s for rotating observers against 2.03 s for static observers), which was not statistically significant ($F(1,10) = 3.02$; $p = .11$).

4 Discussion

4.1 Empirical Validation

In line with classical results (Diwadkar & McNamara, 1997), we found a viewpoint dependency effect in static object recognition. Indeed, in the situation where the viewing position remained unchanged, subjects' performance was disrupted when they had to accomplish the task after the rotation of the object layout. This static mental rotation cost is reflected by the significant decrease in accuracy and the increasing trend of the reaction time.

On the other hand, for moving observers, the cost of dynamic mental rotations was significantly smaller than for static mental rotations. This means that our results yield a significant cost reduction in the case of a moving observer, thus replicating the findings of Simons and Wang (1998). The fact that the mental rotation cost for a moving observer was not null as it was sometimes reported using physical setups can be explained in terms of cue richness. Indeed, our multimodal VR setup still lacked some information compared to real-world situations: the available visual cues were restricted to a 61° FOV and subjects had less complex vestibular and somatosensory information. These findings are consistent with the results of Simons and Wang's second experience: partly removing visual cues increased the dynamic mental rotation cost, but still showed a viewer advantage. In conclusion, we have replicated the classical effects that had been previously reported, hence showing the validity of our virtual reality setup for further studies of mental rotations.

4.2 Technical and Methodological Advances

The fact that our platform uses mixed elements of a physical and virtual setup allows for better control of the various stimulations, full experimental monitoring, innovative conditions, faster experimental sessions, and ecological immersion.

Using virtual 3D objects offers many advantages. Objects' stimuli can be controlled with great accuracy in terms of contrast, saliency, and subtending visual angle. Furthermore, the virtual setup allows systematic rules for the definition of the object positions in the layout. This precise positioning removes the possibility of subjects relying on a grid in order to memorize the layout (e.g., using verbal strategies), and a large number of spatial layouts can be generated according to a set of constraints. The translation vector of the target object can be controlled with great precision. This opens up new research possibilities in the future such as the systematic study of distance thresholds using, for instance, staircase procedures. Testing different types of array perturbations, such as object swapping or rotating objects around their center, will allow further questioning the nature of the information encoded by subjects. The same applies to environmental cues that can be tailor-made to fit the scientific question. If we add to this the fact that trials can be fully automated, many new experimental conditions can be introduced and properly counterbalanced using Latin-Square designs. The duration of trials is strongly shortened (the experimenter doesn't have to rush to place the objects when preparing configurations or when translating an object during the 7 s of rotation) and potential sources of variability are avoided. It becomes possible to have more trials in shorter experimental time. Furthermore, such an experimental platform offers a total monitoring of the ongoing experiment. The experimenter has online information about the current trial and receives a video as well as audio feedback from the subject. Errors and bugs in trial management as well as subjects' troubles (headphones not worn, glasses falling, subject being distracted) can thus be instantly noticed and taken care of in order to avoid collecting useless data. Selecting an object with a simple touch is fast and very intuitive. This allows for measuring reaction times, and processing the responses automatically. This is quite a breakthrough since, to date, real-setup studies not only had to manually account for verbal responses but they had no access to response latencies. On the other hand, previous studies using virtual mental rotations only had access to reaction times using a simplified task consisting of a two alternative forced choice (i.e., among two of the objects or in a change detection task). Furthermore, knowing subjects' performance after each trial allows for the use of psychophysical adaptive methods to generate the subsequent trials.

Our setup seems to overcome many other methodological approaches in terms of practical use and scientific possibilities, and it allows researchers to replicate previous results. One of the main motivations behind the development of this experimental platform was that visual, vestibular, and acoustic information can easily be suppressed or manipulated independently. Previous experimental platforms offered poor (real setups) to null (imagined) assessment of the modality contributions.

Acknowledgments

This work was supported by a post-doctoral research scholarship paid to Manuel Vidal from the Max Planck Society and by a doctoral research scholarship paid to Alexandre Lehmann from the Centre Nationale pour la Recherche Scientifique. We are grateful to the workshop of the Max Planck Institute for the construction of the table setup.

References

Amorim, M. A., Glasauer, S., Corpinot, K., & Berthoz, A. (1997). Updating an object's orientation and location during nonvisual navigation: A comparison between two processing modes. *Perception Psychophysics, 59,* 404–418.

Amorim, M. A., & Stucchi, N. (1997). Viewer- and object-centered mental explorations of an imagined environment are not equivalent. *Cognitive Brain Research, 5,* 229–239.

Burgess, N., Spiers, H. J., & Paleologou, E. (2004). Orientational manoeuvres in the dark: Dissociating allocentric and egocentric influences on spatial memory. *Cognition, 94,* 149–166.

Christou, C., & Bülthoff, H. H. (1999). *The perception of spatial layout in a virtual world* (Tech. Rep. No. 75). Tübingen, Germany: Max Planck Institute.

Christou, C. G., Tjan, B. S., & Bülthoff, H. H. (2003). Extrinsic cues aid shape recognition from novel viewpoints. *Journal of Visualization, 3,* 183–198.

Diwadkar, V. A., & McNamara, T. P. (1997). Viewpoint dependence in scene recognition. *Psychological Science, 8,* 302–307.

Klatzky, R. L., Loomis, J. M., Beall, A. C., Chance, S. S., & Golledge, R. G. (1998). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science, 9,* 293–298.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., & Philbeck, J. W. (1999). Human navigation by path integration. In R. G. Golledge (Ed.), *Wayfinding behavior* (pp. 125–151). Baltimore: The John Hopkins University Press.

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

Riecke, B. E., van Veen, H. A. H. C., & Bülthoff, H. H. (2002). Visual homing is possible without landmarks: A path integration study in virtual reality. *Presence: Teleoperators and Virtual Environments, 11,* 443–473.

Roskos-Ewoldsen, B., McNamara, T. P., Shelton, A. L., & Carr, W. (1998). Mental representations of large and small spatial layouts are orientation dependent. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24,* 215–226.

Ruddle, R. A., & Lessels, S. (2006). For efficient navigational search, humans require full physical movement, but not a rich visual scene. *Psychological Science, 17,* 460–465.

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

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171,* 701–703.

Simons, D. J., & Wang, R. F. (1998). Perceiving real-world viewpoint changes. *Psychological Science, 9,* 315–320.

Simons, D. J., Wang, R. X. F., & Roddenberry, D. (2002). Object recognition is mediated by extraretinal information. *Perception & Psychophysics, 64,* 521–530.

Wang, R. F., & Spelke, E. S. (2000). Updating egocentric representations in human navigation. *Cognition, 77,* 215–250.

Wang, R. X. F., & Simons, D. J. (1999). Active and passive scene recognition across views. *Cognition, 70,* 191–210.

Wexler, M., Kosslyn, S. M., & Berthoz, A. (1998). Motor processes in mental rotation. *Cognition, 68,* 77–94.

Wraga, M., Creem, S. H., & Proffitt, D. R. (2000). Updating displays after imagined object and viewer rotations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 151–168.

Wraga, M., Creem-Regehr, S. H., & Proffitt, D. R. (2004). Spatial updating of virtual displays during self- and display rotation. *Memory and Cognition, 32,* 399–415.

Zacks, J. M., Vettel, J. M., & Michelon, P. (2003). Imagined viewer and object rotations dissociated with event-related fMRI. *Journal of Cognitive Neuroscience, 15,* 1002–1018.