Human Interaction and Emerging Technologies (IHIET 2024), Vol. 157, 2024, 324–333
https://doi.org/10.54941/ahfe1005492

The Construction of Egocentric and Allocentric Spatial Representations in Visual-Spatial Working Memory in Highly Immersive Virtual Reality (CAVE)

Olga A. Saveleva[1,2], Boris B. Velichkovsky[2,3], Galina Ya Menshikova[2], and Grigory S. Bugriy[2]

[1]I.M. Sechenov First Moscow State Medical University, Moscow, Russia
[2]Moscow State University (Lomonosov MSU), Moscow, Russia
[3]Moscow State Linguistic University, Moscow, Russia

ABSTRACT

In solving spatial tasks, neurocognitive egocentric and allocentric spatial representations storing in the visual-spatial working memory. Egocentric representations encode the visual scenes in self-centered coordinates and allocentric representations – in world coordinates regardless of the observer's position. Previously studies showed a good consistency in spatial processing about real environments compared to virtual reality environments. A presentation method was developed for memorizing and reconstructing 3D scenes using the highest immersive CAVE virtual reality system. A space for task, library of objects and virtual scenes were designed, each containing seven virtual objects located in different 3D positions. Three viewpoints were given for reproduction: ≪the front≫ viewpoint (to reproduce the memorized scene from the imaginary egocentric position), ≪the left≫ and ≪the above≫ viewpoints (to reproduce the scene from the left or above imaginary allocentric positions, respectively). The participant had to reconstructed memorized scene in a natural way by choosing objects from the library and placed it in virtual space in accordance with the given imagine viewpoint. The score of object localizations was estimated separately by three parameters — topology, metrics, and depth. The results showed, that for both types of spatial representations schematic topological properties were preserved better in visual-spatial working memory than the exact metric information (especially for the egocentric representations). Overall, the egocentric representations were more effective in the reconstruction of 3D scenes than allocentric representations. It was also found that when using an allocentric representations, the need to add a height axis (vertical rotation) diminishes the effectiveness of the scene reconstruction from visual-spatial working memory, compared to rotations in the horizontal plane. The results suggest that both egocentric representations and allocentric representations can be formed in visual-spatial working memory, but that egocentric representations are more basic in the solution of spatial tasks using visual-spatial working memory. These results not only have theoretical significance in cognitive psychology, but also have the potential for wide practical application in healthcare, education, developmental and sports psychology, human factor research and related interdisciplinary fields.

Keywords: Egocentric spatial representations, Allocentric spatial representations, Spatial memory, Visual-spatial working memory, Topology, Metric, Depth, Immersive virtual reality

INTRODUCTION

Spatial Representations (SRs)

Modern approaches for studying spatial cognition using the concept of spatial representations (SRs), usually understood as cognitive units of spatial cognition (Menshikova et al., 2020). Also, very often the concept of spatial representations using synonymously with the concept of processing strategies, cognitive systems, frames, etc. Two types of SRs are identified in the literature: egocentric representations (ESRs) and allocentric representations (ASRs) (Colombo et al., 2017; Derbie, 2021; Klatzky, 1998; Serino et al., 2014). The dissociation of cognitive systems that ensure the formation of egocentric and allocentric spatial representations has been proven in many studies. ESRs encode object locations relative to the observer and ASRs encode object locations by specifying the relative positions between objects. The two types of spatial representations are related in a complex fashion, ASRs being considered more abstract than ESRs (Klatzky, 1998). Both ESRs and ASRs are involved in the solution of many spatial tasks including wayfinding and spatial memory. An important topic in ESR/ASR research is the study of their cognitive and neurocognitive mechanisms. This is the more important as deficits in ESRs and, especially, ASRs have been shown in a variety of conditions including old age and Alzheimer's Disease (Colombo et al., 2017; Serino et al., 2014; Tuena, 2021), autism (Presley, 2021), and Williams syndrome (Broadbent, 2014).

Current research has recently suggested the involvement of hippocampus and surrounding areas in the formation of SRs (Danjo, 2020; Fidalgo, Martin, 2016). This may be well related to the findings that hippocampus is heavily involved in spatial processing in both animals and humans (Bird, 2008). ESR/ASR formation has also been shown to be associated with the parietal lobes which are related to attention management and multimodal sensory integration (Iachini et al., 2009). This indicates that a complex network of anatomical structures and cognitive processes supports the establishment of ESR/ASR while solving complex spatial tasks (Danjo, 2020). These may include (spatial) memory processes as indicated by the involvement of the hippocampus. In this research, we would like to extend the research on the neurocognitive underpinnings of ESR/ASR by the study of the role of working memory in their formation.

This study is interdisciplinary. The results are very important in healthcare (especially in children and involutionary age), education, medical rehabilitation of cognitive and neurobiological disorders, sports psychology, as well as the study of human factors and the development of artificial intelligence systems used in cognitive ergonomics. Well known systematic review point out in conclusion that the study of cognitive deficits based on cognitive concepts of ESR/ASR might be crucial to make accurate diagnosis and rehabilitation (Tuena, 2021).

Spatial Representations (SRs) of Virtual Reality (VR) Environments

In the modern digital world, it is especially important to study the psychological processes of processing spatial information not only in the real environment, but also in a virtual reality. Highly immersive technologies allow to conduct cognitive experiments in unique conditions that are difficult to create using standard laboratory methods, for example, to study the construction of ESR/ASR at the earliest stages of their formation.

Among the methods for studying spatial representations, the method using virtual reality (VR) technologies is increasingly being used. Despite subtle differences, real-world and virtual versions showed good overlap for the assessment of spatial memory even in clinical subjects, not just healthy human subjects (Tuena, 2021).

On the other hand, in the digitalizing modern world it is very important to study the processing of spatial representations in virtual reality. We used the CAVE virtual reality system, which provides the highest immersive conditions for conducting a cognitive experiment in working memory methodology.

Visual-Spatial Working Memory (VSWM) and SRs

Working memory (WM) is a hypothetical cognitive system involved in the temporary storage and processing of information (Velichkovsky, 2017; Baddeley, 2012). Within WM, a visual-spatial component (VSWM) has been identified, among others. VSWM is used for processing of spatial and (independently) of visual information and is heavily involved in solving spatial tasks (McAfoose, Baune, 2009). It can be assumed that VSWM should be involved in processing of SRs. Indeed, recent research has shown a correlation between individual differences in VSWM and the effectiveness in SRs encoding (Ishikawa, 2023; Wen et al., 2013). Below, we report on a study about how information about 3D scenes is preserved and processed in VSWM over short-periods of time when a 3D scene reconstruction task is performed involving an ESR (no need for a perspective change, reconstructing the scene "from the front") or an ASR (with a need for perspective change, reconstruction the scene "from the left" or "from above").

MATERIALS AND METHODS

Participants

Thirty nine volunteers (21F, 18M, age range 22±3 years) took part in the experiment. All subjects had normal or corrected to normal vision and had no vestibular dysfunctions or injuries.

Equipment

We used the CAVE system Barco Ispace 4.0, which consisted of four large flat screens (each 2.5×2.5 m in size) combined to form a cube. Active eyewear CrystalEyes 3 Stereographics was used to create stereo effects. The manipulations of virtual objects were carried out using a Flystick 2. The A.R.T. DTrack 2 tracking system was used to record the participant's positions with update rate of 8 Hz (see Figure 1). The virtual scenes were developed with the use of VirTools 4.0 software.

Figure 1: CAVE virtual reality system and the typical scene view (in the upper right corner).

Stimuli

Six virtual scenes were designed, each containing 7 virtual objects located in different 3D positions (see Figure 1). All objects were organized in a virtual volume of (20°×20°×7°), the angular size of each object was about 2°×2°. The average brightness of the objects varied slightly within 15 – 20 Cd/m2. The participants were recommended to keep constant position (2.3 m) relative to the virtual scene to control angular sizes of the virtual objects. The objects were presented against a background which represented a dark 3D space, in which a lot of small white balls (0.5°×0.5°×0.5°) were uniformly distributed. Their density in the space was about 6 cpd.

Procedure

Each of six virtual scenes was presented 3 times for 25 seconds. The order of scenes presentation was quasi-random. The participants were asked to remember the objects and their location and then to reproduce the memorized scene. The algorithm of reproduction was as follows. Immediately after the scene presentation an arrow was shown for 3 seconds. Its orientation showed from which point of view the participant had to imagine and reproduce a 3D scene that he had just seen. Three viewpoints were given for reproduction: ≪the front≫ viewpoint (to reproduce the memorized scene from the imaginary egocentric position), ≪the left≫ and ≪the above≫ viewpoints (to reproduce the scene from the left or above imaginary allocentric positions, respectively). Then, the library consisting of 21 familiar and unfamiliar objects was shown. The participant had to choose objects from it and placed them in virtual space in accordance with the given point of view. The object locations in the virtual space were recorded. Each scene was reproduced using each of three points of view. A series of training exercises had been carried out before the experiment to allow participants to get acquainted with manipulations of virtual objects. The average time required to reproduce a scene was about 20–30 sec. The training exercises and the main experiment took 20–25 minutes.

Data Processing

The score of scenes reproduction was estimated as the sample-average of the number of correctly reproduced virtual objects, separately for each mental viewpoint position. The score of object localizations was estimated separately by three parameters—topology, metrics, and depth.

RESULTS

The accuracy of spatial memory was assessed by computing the probability of objects being correctly identified during 3D scene reconstruction (averaged over probes, subjects, and experimental conditions, see Table 1). The scores were very high meaning that, generally, around 6–7 visual objects were successfully held in subjects' VSWM. This is well in line with the magical number of 7 plus/minus 2 as identified by Miller (1965). However, this is distinctively over the 4 elements limit of VSWM identified by (Luck & Vogel, 1997) which may be related by an absence of an addition attention-demanding task as employed in the Luck & Vogel study. Overall, we think the using 7 objects in our experiment didn't exceed typical limitations of VSWM and we obtained results pertaining to visual scene processing in VSWM without the need for recruiting additional storage mechanisms as was demonstrated for verbal WM (Oberauer, 2002). Importantly, the accuracy of object identification didn't differ significantly between experimental conditions (all ps>0.1).

Table 1. Theaccuracy of SRs from all imaginary viewpoints.

Variable	Front Viewpoint (ESRs)	Left Viewpoint (ASRs)	Above Viewpoint (ASRs)	F-test	F-value	p-value
The accuracy of objects identification	0,94(0,09)	0,90(0,11)	0,93(0,09)	0,12	2,14	0,58
The accuracy of topology	0,88(0,10)	0,81(0,15)	0,67(0,24)	0,01	--	--
The accuracy of metric	0,48(0,23)	0,41(0,20)	0,39(0,23)	0,19	1,71	0,36
The accuracy of depth	0,54(0,15)	0,39(0,18)	0,31(0,14)	0,01	--	--

The analysis of how the topological properties of the remembered visual scenes were reconstructed revealed some differences between experimental conditions. Topological properties were not preserved if the relative location in any pairs of objects was wrongly reproduced. The number of wrongly reproduced relative locations was averaged over probes, subjects, and experimental conditions. These averages were submitted to pairwise t-tests. The accuracy of topology (probabilities of preserving the topology of the remembered visual scene) are presented in Table 1. It can be seen that topology is well preserved for ESR (about 90% correct in the front condition), but not in both ASR conditions (81% correct in the left condition and 67% correct in the above condition). T-test revealed significant differences between all experimental conditions (≪front≫–≪left≫, t(38) = 3,88, p < 0,01; ≪left≫–≪above≫, t(38) = 3,15, p < 0,01, ≪front≫–≪above≫, t(38) = 5,33; p < 0,01). These results suggest a gradual decrease in the accuracy of topological information in ASRs as they require more mental rotation.

The accuracy of metric properties in VSWM was also analyzed for different types of SR. To this end, we computed the averaged probabilities of there being no metric errors in the visual scene reconstruction. A metric error was coded as the deviation of an objects position to more than 20% from its original position. The averages for the probability of metrically correct reconstructions are given in Table 1 and were also submitted to a series of pairwise t-test. It can be seen from Table 1 that metric accuracy in VSWM is much lower than topological accuracy and that the results don't differ much between SRs (if anything, there is a tendency for the ESR to be more slightly effective in this respect than both ASR). T-tests support this conclusion and reveal no significant differences between experimental conditions (all ps> 0.1). However, the standard deviations reported in Table 1 for metric accuracies are relatively high (around 0.2, which mean the individual accuracies vary on average as much as from 20% to 60%) which suggest relatively high variability of metric accuracy across probes and subjects. The last point suggests a possible influence of individual differences in the spatial abilities.

Finally, we assessed ESR/ASR differences in preserving depth relations. To this end, we computed the relative number of depth errors. Depth errors were identified as false ordering of the reconstructed objects along the depth axis (Z-axis in the front condition, X-axis in the left condition, and Y-axis in the above condition). The average probabilities of depth reconstruction accuracy are presented in Table 1. It can be seen that the depth-related accuracy is generally relatively low (around 50%-40%). It is highest in the ESR front condition (54%) with a gradual decrease of the both ASR conditions (left – 39%, above – 32%). Pairwise t-test comparisons revealed that all experimental conditions differ significantly with respect to depth-related accuracy ("front"–"left", t(38)=3,587, p = 0,01; "left"–"above", t(38)=2,32, p = 0,03; "front"–"above", t (38)=6,02, p = 0,01). As maintaining depth relationships are related to the maintenance of metric information, these data suggest that VSWM is not very good in keeping exact spatial information about the perceived visual scenes. They also suggest that the depth information is best reconstructed based on an ESRs.

Differences in spatial abilities are consistently reported for men and women (Chen et al., 2020; Conrad, Hull, 1964; Hedges, Nowell, 1995), but see, for example, (Bartlett, Camba, 2023; Self et al., 1992) for a critical review of gender differences in spatial abilities. Given that we had a similar number of man and women in our study (21F, 18M), we also performed a brief check of possible gender differences in our data. No gender differences were found for the recognition accuracy. This conclusion is supported by a t-test comparison (p>0.1). Also, no differences were found neither for topological accuracy, nor for metric accuracy, nor for depth-related accuracy (all ps>0.1). These results strongly suggest that the formation of SRs in VSWM is a basic cognitive process not dependent on the specifics of the male or female brain (e.g., hemispheric asymmetries). This is not surprising given that the ecological and evolutionary requirement in representing basic relationships within the spatial environment are most likely shared for man and women. Our data shows that there are no gender differences in the basic processes of ESRs and ASRs build-up in VSWM.

CONCLUSION

Considering our data, several general conclusions can be drawn. First, it seems that the characteristics of ESRs and ASRs created in VSWM are very similar to that typically reported for SRs in long-term memory (LTM). In fact, we have seen that while topological relationships within the retained visual scene are very well preserved, the more exact metric and depth information is lost to a large extent. That is, our data show that SRs in VSWM tend to be schematic – a trend often observed for spatial LTM but also for verbal LTM (e.g., Bartlett's schema theory, see Wagoner, 2013). This raises the question of how VSWM and spatial LTM interact when building ESRs and ASRs. In accordance with classic cognitive memory models (Atkinson, Shiffrin, 1968), we would advocate the position that SRs (at least, ESRs) are initially built within the VSWM and are later transferred to spatial LTM. Clarifying this relationship would surely somewhat change the tradition of preferably studying ESRs/ASRs as representations within LTM. Of course, the relationship between VSWM and spatial LTM may be very complex with respect to ASRs as ASRs may be abstracted from many instances of ESRs sored in spatial LTM (Röhrich et al. 2014, for a recent model on the ESR/ASR-related interplay of VSWM and spatial LTM).

Second, generally we have seen a higher effectiveness of ESRs in the reconstruction of spatial information in VSWM. This could be expected as ESRs are by definition more sensory-related and seem to provide a more exact replica of the perceived visual scene. This finding is in line with the idea that VSWM storage is more in terms of sensory features than LTM, where storage is semantically organized (Conrad, Hull, 1964). It is interesting that we also see the presence of much more abstract ASRs in VSWM which means VSWM storage may be multimodal and abstract (Baddeley, 2012). Generally, our data on the higher efficiency of ESRs over ASRs in VSWM under the condition of limited stimuli presentation and response suggest that (1) ASRs are built-up on the basis of ESRs in the VSWM and the (2) the build-up of ASRs in VSWM is a time-consuming process. Thus, manipulating the time intervals within spatial VSWM task would be a strong experimental manipulation for the study of ASRs abstraction in VSWM in future research.

Third, within the ASRs we have consistently found that spatial information is better preserved in the "left" condition than in the "above" condition. This is a controversial result since we actually had expected that the "above" condition would be most effective. This could be expected as (modern) humans seems to be very well adjusted to form 2D "maps" for the bird's perspective out of 3D environmental information. However, exact the opposite was found. Speaking in ecological terms, abstracting ASRs staying "on the ground" is easier than adding a height direction and abstracting an ASRs "from above" (as humans naturally are not used to flying). This distinction can be cast in terms differences within 3D mental rotation as mental rotation within a horizontal plane may be easier than mental rotation within a vertical plane (although, generally, 3D mental rotation is more effective than 2D mental rotation (Bartlett, Camba, 2023; Neubauer et al., 2010; Paraskeva et al., 2010)). We think these ideas should be elaborated on more thoroughly in future research.

Fourth and last, we can envisage a differential psychology perspective in the ESR/ASR research. First, a gender perspective may be advanced. Our data shows that there are no gender differences in the basic processes of ESRs and ASRs build-up in VSWM. Still, we had a relatively small sample, so this null result may be due to the low power of our research. On the whole, we think the lack of gender differences can be expected as the processes we study are of fundamental importance for effective spatial interactions for all humans regardless of gender. However, we would hypothesize that if some gender differences will be found in this area of spatial cognition, these would be located more within the ASR domain. This is because ESRs are so basic and immediately related to sensory experience that it would hard to expect fundamental differences between genders within these basic spatial mechanisms. Also, while there were no gender differences, we found indications of substantial individual differences in the reconstruction of exact metric information. This can be a basic and very specific inter-individual variation. Future research should show how it is related to the formation of ESRs and ASRs in VSWM and how it is related individual differences in spatial abilities.

This study is, to the best of our knowledge, the first study of ESR/ASR formation in VSWM using ecologically valid 3D visual scenes presentation in an immersive virtual reality. We found that the formation of ESRs and ASRs is possible in VSWM as a prerequisite for storing a SR in the spatial LTM. We also found that schematic topological information is better preserved than exact metric information already at the level of ESRs, and that there the ASRs built from horizontal rotations are more exact than ASRs built from vertical rotations. We also found evidence that there may be individual differences in the ability to form ESRs/ASRs in VSWM. Overall, our results suggest that VSWM may be involved in the formation of ESRs and ASRs as an important aspect of human spatial cognition.

ACKNOWLEDGMENT

The authors would like to acknowledge Russian Scientific Fund. This study was supported by RSF, grant #22-18-00689.

REFERENCES

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In Psychology of learning and motivation (Vol. 2, pp. 89–195). Academic press. 89–195.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. Annual review of psychology, 63, 1–29.

Bartlett, K. A., &Camba, J. D. (2023). Gender differences in spatial ability: A critical review. Educational Psychology Review, 35(1), 8.

Bird, C. M., & Burgess, N. (2008). The hippocampus and memory: Insights from spatial processing. Nature reviews neuroscience, 9(3), 182–194.

Broadbent, H. J., Farran, E. K., & Tolmie, A. (2014). Egocentric and allocentric navigation strategies in Williams syndrome and typical development. Developmental science, 17(6), 920–934.

Chen, W., Liu, B., Li, X., Wang, P., & Wang, B. (2020). Sex differences in spatial memory. Neuroscience, 443, 140–147.

Colombo, D., Serino, S., Tuena, C., Pedroli, E., Dakanalis, A., Cipresso, P., & Riva, G. (2017). Egocentric and allocentric spatial reference frames in aging: A systematic review. Neuroscience &Biobehavioral Reviews, 80, 605–621.

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. British journal of psychology, 55(4), 429–432.

Danjo, T. (2020). Allocentric representations of space in the hippocampus. Neuroscience research, 153, 1–7.

Derbie, A. Y., Chau, B. K., Wong, C. H., Chen, L. D., Ting, K. H., Lam, B. Y.,... & Smith, Y. (2021). Common and distinct neural trends of allocentric and egocentric spatial coding: An ALE meta-analysis. European Journal of Neuroscience, 53(11), 3672–3687.

Fidalgo, C., & Martin, C. B. (2016). The hippocampus contributes to allocentric spatial memory through coherent scene representations. Journal of Neuroscience, 36(9), 2555–2557.

Hedges, L. V., &Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. Science, 269(5220), 41–45.

Iachini, T., Ruggiero, G., Conson, M., &Trojano, L. (2009). Lateralization of egocentric and allocentric spatial processing after parietal brain lesions. Brain and cognition, 69(3), 514–520.

Ishikawa, T. (2023). Individual differences and skill training in cognitive mapping: How and why people differ. Topics in Cognitive Science, 15(1), 163–186.

Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In Spatial cognition: An interdisciplinary approach to representing and processing spatial knowledge (pp. 1–17). Berlin, Heidelberg: Springer Berlin Heidelberg.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. Nature, 390(6657), 279–281.

McAfoose, J., &Baune, B. T. (2009). Exploring visual–spatial working memory: A critical review of concepts and models. Neuropsychology review, 19, 130–142.

Menshikova. G. Ya., Saveleva. O. A., Velichkovskiy. B. B. & Bugriy. G. S. (2020). Formirovaniyeegotsentricheskikhiallotsentricheskikhprostranstvennykhreprezentatsiy v rabocheypamyati. Voprosypsikhologii. (6). 131–140.

Neubauer, A. C., Bergner, S., & Schatz, M. (2010). Two-vs.three-dimensional presentation of mental rotation tasks: Sex differences and effects of training on performance and brain activation. Intelligence, 38(5), 529–539.

Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. Journal of Experimental Psychology: Learning, Memory, and Cognition, 28(3), 411.

Paraskeva, F., Mysirlaki, S., &Papagianni, A. (2010). Multiplayer online games as educational tools: Facing new challenges in learning. Computers & Education, 54(2), 498–505.

Presley, A. (2021). The extent to which autistic traits are predictive of impairments in allocentric spatial navigation.

Röhrich, W. G., Hardiess, G., &Mallot, H. A. (2014). View-based organization and interplay of spatial working and long-term memories. PloS one, 9(11), e112793.

Self, C. M., Gopal, S., Golledge, R. G., &Fenstermaker, S. (1992). Gender-related differences in spatial abilities. Progress in Human Geography, 16(3), 315–342.

Serino, S., Cipresso, P., Morganti, F., & Riva, G. (2014). The role of egocentric and allocentric abilities in Alzheimer's disease: a systematic review. Ageing research reviews, 16, 32–44.

Tuena, C., Mancuso, V., Stramba-Badiale, C., Pedroli, E., Stramba-Badiale, M., Riva, G., & Repetto, C. (2021). Egocentric and allocentric spatial memory in mild cognitive impairment with real-world and virtual navigation tasks: A systematic review. Journal of Alzheimer's Disease, 79(1), 95–116.

Wagoner, B. (2013). Bartlett's concept of schema in reconstruction. Theory & Psychology, 23(5), 553–575.

Wen, W., Ishikawa, T., & Sato, T. (2013). Individual differences in the encoding processes of egocentric and allocentric survey knowledge. Cognitive science, 37(1), 176–192.