Effects of Display Layout on Spatial Memory for Immersive Environments

JIAZHOU LIU, Monash University, Australia
ARNAUD PROUZEAU, Inria, France
BARRETT ENS, Monash University, Australia
TIM DWYER, Monash University, Australia

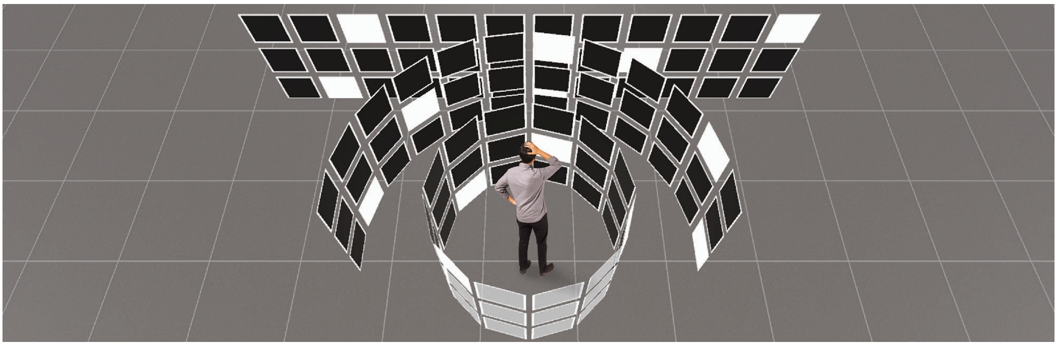

Fig. 1. The extensive 3D space surrounding users in immersive environments allows for various display layouts. In this paper we consider layouts which wrap around the user to varying degrees: Flat, Semi-Circle, and Full-Circle. We hypothesise that the degree of curvature affects users' spatial memory during navigation. We explore this hypothesis through a sequence of studies.

In immersive environments, positioning data visualisations around the user in a wraparound layout has been advocated as advantageous over flat arrangements more typical of traditional screens. However, other than limiting the distance users must walk, there is no clear design rationale behind this common practice, and little research on the impact of wraparound layouts on visualisation tasks. The ability to remember the spatial location of elements of visualisations within the display space is crucial to support visual analytical tasks, especially those that require users to shift their focus or perform comparisons. This ability is influenced by the user's spatial memory but how spatial memory is affected by different display layouts remains unclear. In this paper, we perform two user studies to evaluate the effects of three layouts with varying degrees of curvature around the user (flat-wall, semicircular-wraparound, and circular-wraparound) on a visuo-spatial memory task in a virtual environment. The results show that participants are able to recall spatial patterns with greater accuracy and report more positive subjective ratings using flat than circular-wraparound layouts. While we didn't find any significant performance differences between the flat and semicircular-wraparound

Authors' addresses: Jiazhou Liu, Monash University, Wellington Road, Melbourne, Australia, jiazhou.Liu@monash.edu; Arnaud Prouzeau, Inria, Bordeaux, France, arnaud.prouzeau@inria.fr; Barrett Ens, Monash University, Wellington Road, Melbourne, Australia, barrett.ens@monash.edu; Tim Dwyer, Monash University, Wellington Road, Melbourne, Australia, Tim.Dwyer@monash.edu.

2573-0142/2022/12-ART576 $15.00
https://doi.org/10.1145/3567729

layouts, participants overwhelmingly preferred the semicircular-wraparound layout suggesting it is a good compromise between the two extremes of display curvature.

CCS Concepts: • **Human-centered computing** → **Virtual reality**; **User interface design**; *Empirical studies in HCI*.

Additional Key Words and Phrases: spatial memory, spatial layout, immersive analytics, virtual reality, interface design

ACM Reference Format:
Jiazhou Liu, Arnaud Prouzeau, Barrett Ens, and Tim Dwyer. 2022. Effects of Display Layout on Spatial Memory for Immersive Environments. *Proc. ACM Hum.-Comput. Interact.* 6, ISS, Article 576 (December 2022), 21 pages. https://doi.org/10.1145/3567729

1 INTRODUCTION

In immersive systems for data visualisation, it is common to position data displays in a *wraparound* configuration, such that the data displays surround the position of the user [3, 27, 29, 46]. This design choice may descend from early room-size projection Virtual Reality (VR) systems (e.g., the CAVE [9]), capable of projecting stereo imagery in every direction the user may face. Arranging the visuals to completely surround the user, as in modern CAVE2 environments [13], is then a fairly logical way to utilise the display capability. A display that fully surrounds the user may also have the advantage that the user can simply turn on the spot to navigate. Such rotational navigation may be less physically demanding than walking to change one's viewpoint in a traditional flat layout. Direct manipulation with immersive interface elements is also well-suited to a wraparound configuration, as arranging objects around a common centre point reduces the effort needed when reaching from object to object [12, 34]. Another source for this design choice to 'wrap' the visuals around the user may be VR headsets that are tethered (connected by a physical cable to a PC). Headset tethers physically limit the distance a user can comfortably walk and introduce a fear of tripping over the cable and can therefore be a disincentive for users to walk any significant distance.

Recent results from multiple studies that allow users of immersive visualisation systems to freely create visualisations, have reported that people tend to position visuals in a circle at arms' length around them [3, 27, 29, 46]. Some of these papers have therefore concluded that circular arrangements of data visualisations around users should be supported as a key aspect of immersive visualisation systems. But once again, an alternative explanation for this observed user behaviour may be an unwillingness by users to walk in the virtual environment due to a tether or unfamiliarity with physical navigation in VR. It should also be noted that the immersive studies above typically involved free-form data exploration, rather than controlled data understanding tasks – so the observed user preference for wraparound displays in these contexts does not necessarily indicate support for analysis tasks.

While there is some evidence from existing work that display curvature affects visualisation tasks, these findings are overall rather inconclusive. To give a couple of examples, a study by Shupp et al. has shown advantages of curved 2D displays in path-following tasks on maps [49], but was inconclusive for comparison tasks. While, a recent study by Liu et al. [31] with small multiples in VR found that participants were able to perform visual comparison and search tasks better with flat displays than wraparound displays. As a result, there are no clear guidelines for creators of immersive visualisation systems on this rather fundamental question of how to layout displays in immersive visualisation spaces. In their concluding remarks, Liu et al. hypothesise that effects of layout curvature on spatial memory may help to explain their observed differences, but to our knowledge, no existing studies have tested whether curvature of information displays influence spatial memory.

In this paper, we study lower-level spatial memory tasks to see how spatial memory is supported by display layout. The hope is that clear findings in this regard can lead to more concrete design guidelines for immersive data visualisation and potentially other sensemaking activities in immersive environments. Spatial memory seems particularly relevant to recall and comparison tasks in visualisation, because these tasks typically involve the comparison between multiple data encodings to identify patterns or anomalies. Since immersive environments allow data visualisations to be spread out over a large region, navigation between multiple objects for comparison requires users to temporarily remember their physical locations. If we can minimise the effort required for such context switches, users will switch more often to reduce demand on their visual working memory [42]. For instance, participants in Liu's study with small multiples seemed to find switching easier with a flat than with a curved layout [31]. Beyond data visualisation tasks, users' ability to remember locations of objects in immersive environments has implications for many other applications, from group work to gaming [17, 38, 40].

In Sections 3 and 4, we present two user studies to test the effect of different display layouts (*Flat* vs *Full Circle*, *Flat* vs *Semicircle*) by investigating user's ability to recall locations of items within the layout for a straight-forward visuo-spatial memory task. As detailed in Section 3.6, results of the first study clearly show that participants are able to recall room-scale patterns of cards more accurately with a *Flat* than a *Full Circle* display layout. Subjective feedback also reports better performance and less mental effort and frustration with the *Flat* than the *Full Circle* layouts. In Study 1 we also introduced conditions with visual modifiers to try to isolate the effect of physical navigation from other differences between *Flat* and *Full Circle*, such as the possibility of getting an overview by stepping back from *Flat* or the obvious landmarks provided by the edges of the *Flat* layout. Neither of these visual modifiers made a significant difference, meaning that the primary benefit of *Flat* over *Full Circle* is likely the physical navigation, i.e. walking rather than rotation. The results from the second study show non-significant differences between the *Flat* and the *Semicircle*. However, our participants prefer the *Semicircle* layout suggesting it is a good compromise between the *Flat* and *Full Circle* layouts. Overall the findings from our two studies suggest that, when the tasks depend on the users' spatial memory of the layout, layouts of information displays in immersive environments that completely surround the user should be avoided.

Apart from the empirical findings as mentioned above, our paper also presents a methodological contribution. This is the first rigorous study, inspired by the design of spatial memory tests from psychology, to test the effects of layout of displays in immersive environments on spatial memory. We have also explored the effects of landmarks and the ability to have an overview as subordinate factors. We hope this paper will establish a connection between spatial memory and analytic tasks in information visualisation, and be used as a foundation for future work expanding this fundamental topic.

2 RELATED WORK

In this section we first present previous work regarding the use of different layouts to present information in immersive environments. We then discuss the relationship between spatial memory and the design of interfaces.

2.1 Layouts in Immersive Environments

The CAVE2 was one of the first VR setups to be used for immersive analytics [13], with an array of high-resolution displays that provide visualisations wrapping around the user. Since then, both with CAVEs and HMDs, researchers have tended to imitate this layout [5, 7, 26]. Several recent

VR visualisation studies that allowed participants to arrange their own display space have found that participants have a tendency to arrange the displays in a wraparound configuration. Batch et al. [3] found this effect in a longitudinal study of economists creating free-form visuals across multiple sessions. Satriadi et al. [46] found a similar effect for participants arranging multiview maps. Other recent research [23, 29, 36, 37] also found similar wraparound effects during document arrangement tasks.

Wraparound layouts allow users to have elements within arm's reach, and reduce visual distortion of far-away elements [43]. A flat layout would in contrast require more physical movements. Using a wall display, Shupp et al. [49] compared semi-circular with flat layouts in three different map tasks (search, route tracing and image comparison), and showed that the semi-circular one led to improved performance for search and route tracing but was inconclusive for comparison. In the Personal Cockpit, Ens et al. [12] arranged a set of virtual displays equidistant from the user's shoulder to support direct input. They showed the importance of having this curved layout fixed in the world instead of moving with the user's body as it could provoke incessant small movements. For Menu selection in AR, Lubos et al. [34] showed that such a curved layout is more efficient if centred on the wrist, thus at the border of the kinesphere, rather than on the head. In recent work, Reipschlaeger proposed an immersive AR system to complement a physical flat display display wall. They explore a design space which includes a virtual extension to the display wall which can wrap around the user. They hypothesised that this can reduce perspective foreshortening effects and bring the visuals closer to the user, but the prototype is not evaluated with a study [44].

While proposed many times in such design explorations it seems the effectiveness of such a wraparound visualisation has until recently been rarely studied. However, recently Liu et al. [31] did compare a flat, semi-circle and full-circle layout for small-multiples visualisation comparison tasks in VR. Their results suggest that while the flat layout was faster with a small number of elements, there was no statistically significant difference between the three layouts with a large number of multiples. However, in the latter case, participants complained that the flat layout led to too much physical effort and that the full circle was disorienting. Overall, their findings indicated that semi-circle layout was a good compromise for the comparison tasks tested.

To summarise, circular layouts have been proposed in various visualisation contexts and applications. Recent studies suggest that curvature of information displays in immersive environments may affect task performance, but there is no clear explanation for these effects. We hypothesise that spatial memory is a significant component in visual sensemaking activity, particularly the kind of comparison tasks explored by Liu et al. [31], since moving between individual visualisations to compare them requires navigating back and forth between them. Therefore, in this paper, we focus on the question of whether spatial memory is significantly affected by layout curvature. Such a finding would help to explain past observed differences in performance on more general visualisation tasks and would provide clear guidance for how to use curved displays for sense-making tasks more generally.

2.2 Spatial Memory

Spatial memory is responsible for collecting and storing information about people's surrounding environment as well as facilitating navigation. It plays a part in both working memory (to remember position and orientation of objects around a person) and long-term memory (to build a mental map of specific locations). In this paper, we focus on spatial memory representation in working memory, as working memory is used to store and process information about the current environment. One well-known model of working memory is the multi-component model defined by Baddeley and Hitch [2]. Their *Visuospatial Sketchpad* model proposes a component of working memory that manages visual and spatial information about the current environment. While there is an interdependence between visual and spatial information, the visuospatial sketchpad treats these visual and spatial components as distinct [32]. There are two commonly-applied tasks to measure visuospatial capabilities in the working memory: (1) the visual pattern span, which focuses on the visual aspect of this memory [10], and (2) the Corsi block tapping task [8], which focuses mostly on the spatial aspect.

Researchers in VR and HCI have since explored how spatial memory in both working and long-term memory can improve interfaces. With CommandMaps, Scarr et al. [48] showed that a spatial grid of commands is more efficient than hierarchical menus for expert users. Li et al. [28] used the same technique in VR with the command visualised in an egocentric layout. This technique can be improved by the use of visual landmarks on desktops [51], and in VR [19]. Virtual landmarks have also been shown to improve remote collaboration in AR as it provides common ground [39]. Perrault et al. [41] used the method of Loci to associate commands to specific physical locations. This is extended to virtual environments by Fruchard et al. [15]. Using a similar method called memory palace, Krokos et al. [24] showed that spatial memory could provide more benefits in VR than on a desktop. Yang et al. [54] also found that using a VR-based memory palace variant increased the effectiveness and performance of retrieving and retaining knowledge. With Data Mountain, Robertson et al. [45] used the 3D location on an inclined plane to classify documents and showed that it improved the search for a specific document. Cockburn et al. [6] showed that the same technique in 2D is more efficient as less cluttered. And Jansen et al. [22] showed that physical navigation and the availability of an overview also improved performance. Zagermann et al. [55] also used a similar task to investigate the effect of different input modalities and display sizes on spatial memory but did not report clear findings. In recent work, Friedrich et al. [14] found that there is a modest benefit to retrieval performance when users had to use locomotion to place windows.

2.3 Studying Spatial Memory

This section details the historical context and rationale for studying spatial memory for immersive visualisation and for our chosen task. While past work in data visualisation explored the effects of layout in the context of high-level visual data analytics tasks, either qualitatively [3] or quantitatively [31], their conclusions regarding spatial memory were inconclusive due to the confounds introduced by the complex nature of such tasks. Thus, we choose an abstract task, which is adapted from two common tasks developed by psychologists to assess visuo-spatial memory: the visual memory span task [33, 53], and the Corsi block tapping task [8].

In the visual memory span task as used by Logie et al. [33] and Wilson et al.[53], participants were presented with a grid pattern of squares (half filled with black and half with white) for a short duration. They were then asked to reproduce this pattern on an empty grid. The difficulty was increased over time by adding two squares incrementally. The Corsi block tapping task is similar to visual memory span task but requires participants to recall a specific sequence. In this task, participants are asked to tap a set of blocks (among 9) in a specific order. The first task begins with two blocks, and the number of blocks is incremented by one every time participants successfully recall the pattern. The experiment ends when the participants fail twice.

In summary, spatial memory has been shown to influence interface performance both on desktops and in VR. We can also see that the interface itself has an impact in this spatial memory (2D vs. 3D, impact of navigation, etc.). However, little is known about how the arrangement of visual elements in the space around the user in VR environments affects their working memory, for example, whether certain layouts reinforce or detract from spatial memory. In this paper, we take a further step in this direction and assess the impact of the layout of visual elements on spatial memory in VR. This requires participants to move within the tracked volume, which allows us to

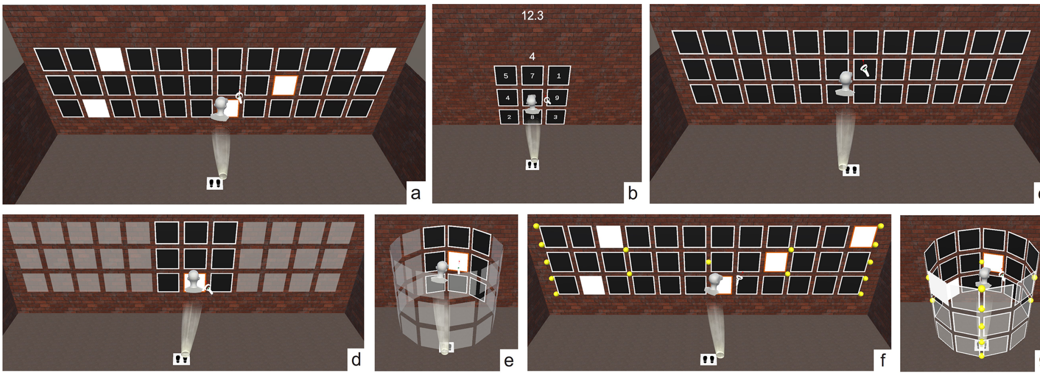

Fig. 2. (top) In the study, participants had to remember 5 cards highlighted in a grid in a learning phase (a), then they performed a distractor task to decay short term memory (b), and finally in a recall phase, they had to select the memorised cards in an empty grid (c). (bottom) The between-subjects Visual Modifier factor has three levels: *Regular* (a), *Restricted FoV* (d,e), and *Landmark* (f and g) for both layouts.

easily vary conditions across factors such as the degree to which the layout of grids wraps around the participant.

3 STUDY 1: FLAT VERSUS FULL CIRCLE

Our initial study focuses on examining the effectiveness and efficiency (see 'measures' subsection below) on a task that requires spatial memory with a comparison of *Flat* and *Full Circle* layouts. While we are also interested in studying partial-wraparound layouts, early pilot testing with a visuo-spatial memory task led us toward a minimal study design with 2 conditions per participants in order to limit their exposure to VR. Thus we first investigate our primary research questions with these two 'extreme' cases before including *Semicircle* layouts in a follow-up study.

To limit the time spent in VR and mitigate possible effects of simulator sickness, we keep the study duration relatively short (to maximum 1 hour). Due to high inter-participant variability, we limited our design to a single factor and controlled the remaining variables. Thus, we use a single task difficulty with a fixed size grid (3 rows x 12 columns) and fixed number of items that need to be recalled (5 items), as described below.

3.1 Task

We use a grid of 36 virtual cards arranged into 12 columns and 3 rows, as shown in Figure 2. This arrangement follows the 'large scale comparison' study design of Liu et al. [31], which used 36 small multiples in a 3x12 grid. In each task, participants must learn and recall the locations of 5 cards, which is slightly higher than the known capacity of visuo-spatial working memory (4 objects) [35]. The number of cards to recall was chosen to control the level of difficulty after pilot tests (using 3–6 cards) and power analysis. As described in Section 2.3 the task chosen for our study is inspired by this past work.

Each card is 0.4×0.4 m, with a horizontal and vertical offsets of 0.1 m between each pair (Fig. 1). The *Flat* layout has a width of 5.9 m and a height of 1.4 m (see Figure 2-a). The circumference of the *Full Circle* layout is equal to the width of the *Flat* layout, resulting in an approximate radius of 0.95 m. With this radius, the cards are within arm reach and thus, we speculate that walking will be minimised (see Figure 2-e and g). Similar dimensions of wraparound layout (i.e. an approximately arms-length radius) are also observed to emerge naturally from user placement of objects in exploratory studies from [3, 16, 46], as discussed in Section 2. The height of the *Full Circle* layouts is same as the *Flat* layouts. The participant starting position in the *Flat* layout has a 0.95 meter distance to the center of the layout grid while the starting position of the *Full Circle* layout is the center point of the circle.

To reduce variability in the study data, we create a fixed sequence of patterns for the grids by generating them in a constrained-random manner, then validate them manually. There are 3 constraints for the generation: (1) no 2 adjacent cards can be included in the same pattern; (2) at least 1 card is included on each row to balance the pattern vertically; and (3) at least 2 cards are included on each side (left or right) to balance the pattern horizontally.

Each trial is divided into 4 phases: *preparation*, *learning*, *distraction*, and *recall*. In the *preparation* phase, participants are required to stand at the starting position facing forward, as indicated by a pair of footprints. Once in position, participants trigger the *learning* phase (Figure 2-a) by pressing a button on the controller. In this phase, a pattern of 5 white cards is revealed in the grid. Participants are given 15 seconds to tap each white card with a controller held in their dominant hand. Changes to the card boundary colour provide feedback to track participants' progress in this learning phase. Specifically, a green border means the card has been tapped which helps participants ensure they touch each white card. Also, an orange border means the card has appeared within the participant's point of view; although this is primarily for experimenters as participants will only see cards when they already have an orange border.

Since short term memory decays within 15–30 seconds [1], we include a *distraction* phase lasting at least 15 seconds between the learning and recall phases. In this phase, a distractor task requires participants to tap a new set of randomly numbered cards in a given sequence (Figure 2-b). Participants will see a countdown timer on top of the task board. During the distractor task, if participants have tapped the wrong cards or idled for 3 seconds, they will be penalised by adding 3 seconds to their current timer.

Finally, in the *recall* phase, participants are asked to recreate the pattern shown in learning phase by tapping on an empty layout. We do not set a time limit for this phase (Figure 2-c). Participants need to confirm their answer by pressing a button on the controller. The number of correctly selected cards is then shown to them.

3.2 Design

Layout. The primary motivation of this study is to examine the effect of layout on participants' ability to recall locations of items within the layout. As described above, we design two layouts with extreme curvatures: *Flat* and *Full Circle* (see Figure 1).

Visual Modifier. In a visual and spatial memory task, as the one described above, we expect the *Flat* layout provide some advantages over the *Full Circle* one. First, participants can take a step back to get an overview of the *Flat* wall, which is not possible in the *Full Circle* layout. The *Flat* layout also provides implicit landmarks at its edges and corners. While prior work suggest benefits of explicit landmarks [18, 19, 51], we speculate that, to some extent, this result is also generalised to implicit landmarks. To evaluate the impact of these advantages on participants' performance, we include a Visual Modifier factor that will either (1) limit the capabilities of an overview of the layout (*Restricted FoV*), (2) add landmarks to the grid (*Landmark*), or (3) leave the two layouts unchanged (*Regular*). The *Restricted FoV* condition reduced the benefit of *Flat* layout over *Full Circle* by only showing cards on three nearest columns from the participant (see Figure 2-d and e). The other columns show only semi-transparent backgrounds in place of the cards. The *Landmark* condition adds a set of yellow spheres in fixed positions on the grid as landmarks (see Figure 2-f and g). This design was inspired by a similar study by Gao et al. [19]. To help orient users, we put different number of landmarks in different positions: middle (*Flat*) or front (*Full Circle*) with a single landmark, left/right edges (*Flat*) or back (*Full Circle*) with 5 landmarks, and 2 landmarks in-between (see Figure 2-f). These visual modifiers are enabled in both learning and recall phases.

Our study design is a 2 (Layout) x 3 (Visual Modifier) mixed-factors study. Layout is treated as a within-subjects factor with 2 conditions (*Flat* and *Full Circle*). Visual Modifier is a between-subjects factor with 3 conditions (*Regular*, *Restricted FoV*, and *Landmark*). Each participant performs both Layout but only one Visual Modifier. The reason for our mixed-factor study is that a fully within-subject study would result in an excessive amount of VR exposure for participants, especially while doing a cognitively intensive task.

In total, we collect data from 648 completed trials (12 participants × 9 repetitions × 2 Layouts × 3 Visual Modifiers). We treat Layout and Visual Modifier as independent variables. Dependent variables include completion time, cards accuracy, Manhattan distance of the errors, walking distance, head positions and rotation angles, and subjective ranking.

3.3 Participants and Apparatus

In total, we recruit 36 participants (18 males and 18 females; mean age = 25.11, SD = 3.89). All participants are students from our university. Each Visual Modifier group consists of 12 participants with the same male and female ratio (6 males and 6 females). Of the participants, 22 have at least some experience with VR and 3 of them rate themselves as VR experts. Participants sign up voluntarily and are rewarded a gift card (A$20) and small gifts (candies and chocolates) as a sign of appreciation.

We run our study using a tethered HTC Vive Pro in a 10 m x 5 m empty room. We develop our study software using the Unity development environment (2019.3.0f3). We leverage VRTK [4] for interactive components. The prototype runs on a Windows 10 PC with an Intel I7 7800X (3.5GHz) processor and an NVIDIA GeForce GTX 1080 (32GB RAM) graphics card. The source code is publicly available and may be downloaded via GitHub: [30].

3.4 Procedure

After completing a consent form and demographic questionnaire, participants are given a verbal explanation of the trial workflow. Next, participants put on the VR headset and performed a series of training scenes such as interactions and the trial workflow. After that, participants complete the trials with the two Layout conditions in an alternating order. The order of the starting layout is also counter-balanced. Participants first complete 2 practice trials (one with each Layout condition), followed by 3 blocks of 6 trials each. Participants are asked to remove the VR headset to take a short break between blocks. Following the completion of all 20 trials, participants complete a short questionnaire with (1) NASA-TLX [20], (2) rankings on Layout preference, and (3) the general strategy they use to complete the tasks. The total study duration is about 45 minutes, including roughly 30 minutes in VR. All participants complete the full set of trials successfully.

The experiment environment includes surrounding walls, carpet floor, a starting position sign, and the experimental grid, as shown in Figure 2. The starting position and surrounding walls are visible at all times during the experiment.

The vertical position of the grid is adjusted using a standard calibration for every participant before the experimental trials. It is used to normalise the individual height differences and make sure that every participant has the same ability with the controls for selection. These configurations are adapted from Liu's work [27] and validated through our pilot testing of different variations.

3.5 Measures

For each trial, we record the number of correctly chosen cards, along with their positions in the grid. We also record the *Recall Time* taken to select the 5 cards in the *recall* phase. The *Recall*

Time is calculated from the start of each *recall* phase immediately after the timer of distrator task ends, to the time the participant presses the button on the controller to indicate task completion and see the results. In our analysis, we use two methods to measure participants' recall accuracy: *Cards Incorrect* and *Manhattan Distance Error*. *Cards Incorrect* measures the average number of cards selected incorrectly in each trial (also expressed as an open unit interval). To reveal deeper granularity in the responses, we further include the Manhattan Distance Error measure, which measures the sum of Manhattan distances from incorrectly selected cards to the correct cards. The Manhattan distance is a common distance matrix for discrete space and has been used more often to vectors that describe objects on a uniform grid [19, 31, 47], than other distance measurements, such as a Euclidean distance. Because the selection is non-sequential, there are many possible solutions to this measure, so we take the solution with the minimum distance as calculated using the Hungarian Algorithm [25].

Participants' head pose is tracked throughout each trial and is used to calculate the *Walking Distance* travelled by participants, *Head Rotation*, and *Head Pointer Intersections* with the plane of the wall display.

3.6 Results of Study 1

We report on the following measures: First, we include three performance measures, (1) *Cards Incorrect*, or proportion of cards chosen incorrectly in each trial; (2) (optimal) *Manhattan Distance Error* between the participants' answers and the solutions; and (3) *Recall Time*, total time of the recall phase. Then we include three physical movement measures collected during the learning phase: (4) *Walking Distance* travelled by participants; (5) *Head Rotation* performed by participants calculated by cumulative angular distances from their tracked head movements; and (6) *Relative Head Position* compared with the projection of the participant head pointer on the plane of the wall display. Lastly, (7) subjective measures include overall *Preference*, as well as the *Mental* and *Physical* workload of participants.

For the *Cards Incorrect*, *Manhattan Distance Error*, *Time*, *Travelled Distance*, and *Head Rotation*, we first assess the normality of the data using the Shapiro-Wilk normality test. Then we assess the homogeneity of variances by Levene's test and sphericity. When the criteria are met, we use a mixed ANOVA to test the differences between the Visual Modifier independent groups whilst subjecting participants to repeated measures within the two Layout conditions. A t-test is used to compare the means of Layout pairs within each Visual Modifier condition.

When the data is not normally distributed, we use a Kruskal-Wallis test to compare the data against the Visual Modifier factor. We also use a Friedman test to compare the data against the Layout factor. For the *subjective measures*, we use a Kendall's W test for each Visual Modifier condition to see the significant effects between the Layout variables.

Cards Incorrect. The data for the incorrect cards are normally distributed and meet the assumption of homogeneity of variances. The mixed ANOVA test shows no significant effect of Visual Modifier ($p = 0.48$, effect size $\eta^2_G = 0.04$) among *Regular* (*mean* = 0.45, *SD* = 0.16), *Restricted FoV* (*mean* = 0.49, *SD* = 0.15), and *Landmark* (*mean* = 0.44, *SD* = 0.17). However, it shows a significant effect of Layout factor ($p = 2.28 \times 10^{-10}$, effect size $\eta^2_G = 0.32$), with a lower overall error for *Flat* (*mean* = 0.37, *SD* = 0.13) vs *Full Circle* (*mean* = 0.55, *SD* = 0.13). A closer inspection in Figure 3-a shows the results of paired T-tests of the Layout factor for each Visual Modifier group.

Participants on average have a *Cards Incorrect* rate of 0.36 ($SD = 0.12$) in the *Regular-Flat* condition, and a *Cards Incorrect* rate of 0.55 ($SD = 0.13$) in the *Regular-Full Circle* condition. A paired T-test reports a significant difference between the two conditions ($p = 1.02 \times 10^{-3}$, effect size $|d| = 1.58$).

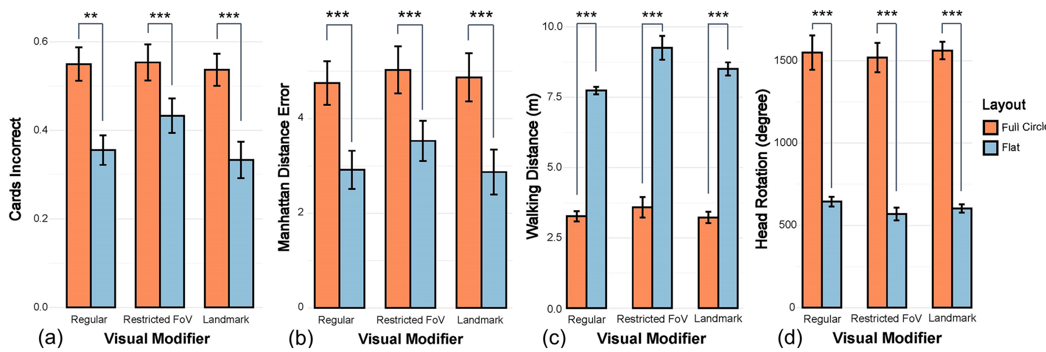

Fig. 3. (a) Cards Incorrect (lower is better), (b) Manhattan Distance Error (lower is better), (c) Walking Distance, and (d) Head Rotation between the VISUAL MODIFIER factor within the *Flat* layout and *Full Circle* layout . Error bars denote standard error. Asterisks in this figure and following figures represent the level of significance: * means $p < 0.05$, ** means $p < 0.01$, and *** means $p < 0.001$.

Participants on average have a *Cards Incorrect* rate of 0.43 ($SD = 0.13$) in the *Restricted FoV-Flat* condition, and a Cards Incorrect cards rate of 0.55 ($SD = 0.14$) in the *Restricted FoV-Full Circle* condition. A paired T-test reports a significant difference between the two conditions ($p = 1.68 \times 10^{-4}$, effect size $|d| = 0.87$).

Participants on average have a *Cards Incorrect* rate of 0.33 ($SD = 0.14$) in the *Landmark-Flat* condition, and a *Cards Incorrect* rate of 0.54 ($SD = 0.13$) in the *Landmark-Full Circle* condition. A paired T-test reports a significant difference between the two conditions ($p = 3.67 \times 10^{-5}$, effect size $|d| = 1.52$).

Manhattan Distance Error. The Manhattan Distance provides a secondary evaluation of the error measure. Whereas *Cards Incorrect* provides an absolute measure, Manhattan Distance rewards card selections by summing the distance from the selected cards to the correct ones.

The data for the *Manhattan Distance Error* is normally distributed and meet the assumption of homogeneity of variances. The mixed ANOVA test shows no significant effect of VISUAL MODIFIER factor ($p = 0.73$, effect size $\eta^2_G = 0.02$) among *Regular* ($mean = 3.83$, $SD = 1.74$), *Restricted FoV* ($mean = 4.28$, $SD = 1.75$), and *Landmark* ($mean = 3.87$, $SD = 1.96$). However, the mixed ANOVA test shows significant effect of LAYOUT factor ($p = 1.33 \times 10^{-11}$, effect size $\eta^2_G = 0.25$), with a lower overall error for *Flat* ($mean = 3.10$, $SD = 1.50$) vs *Full Circle* ($mean = 4.88$, $SD = 1.65$). Figure 3-b shows the paired T-test of the LAYOUT factor for each VISUAL MODIFIER group.

The *Manhattan Distance Error* is on average 2.92 ($SD = 1.40$) in the *Regular-Flat* condition, and 4.75 ($SD = 1.59$) in the *Regular-Full Circle* condition. A paired T-test reports a significant difference between the two conditions ($p = 8.29 \times 10^{-4}$, effect size $|d| = 1.22$).

The *Manhattan Distance Error* is on average 3.53 ($SD = 1.48$) in the *Restricted FoV-Flat* condition, and 5.03 ($SD = 1.73$) in the *Restricted FoV-Full Circle* condition. A paired T-test reports a significant difference between the two conditions ($p = 1.90 \times 10^{-4}$, effect size $|d| = 0.93$).

The *Manhattan Distance Error* is on average 2.87 ($SD = 1.64$) in the *Landmark-Flat* condition, and 4.87 ($SD = 1.77$) in the *Landmark-Full Circle* condition. A paired T-test reports a significant difference between the two conditions ($p = 1.06 \times 10^{-6}$, effect size $|d| = 1.17$).

Recall Time. The recall time data is not normally distributed. The Kruskal-Wallis test shows no significant effect of VISUAL MODIFIER factor ($p = 1$, effect size $\eta^2 = 0.03$) among *Regular* ($mean = 21.77$, $SD = 12.31$), *Restricted FoV* ($mean = 28.04$, $SD = 17.33$), and *Landmark* ($mean = 24.92$,

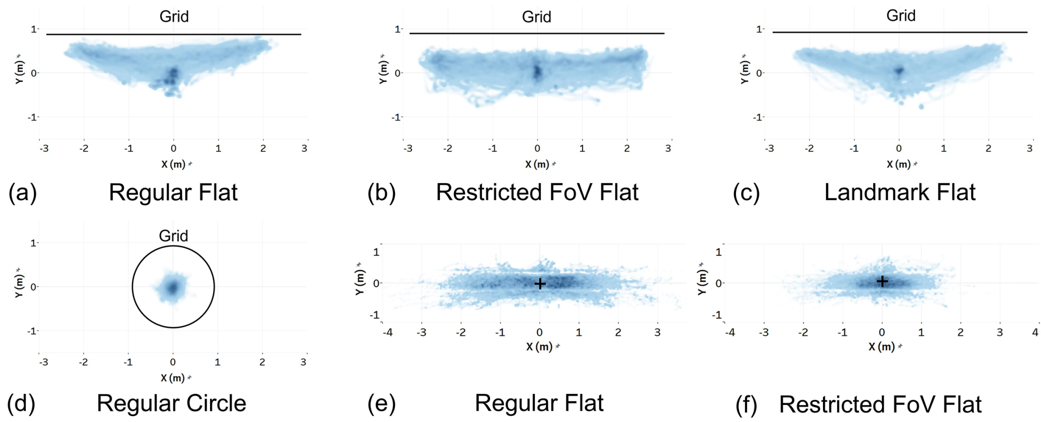

Fig. 4. Density maps showing participants' positions during the learning phase in (a) *Regular* condition, (b) *Restricted FoV* condition, (c) *Landmark* condition for *Flat* layouts, and (d) *Regular* condition for *Full Circle* layouts. And density maps show where the participant head pointer (ray cast from from headset's forward direction) intersects the plane of the wall display, relative to the user's head position in (e) *Regular* condition and (f) *Restricted FoV* condition during the learning phase for *Flat* layouts.

$SD = 9.87$). The Friedman test shows no significant effect of Layout factor ($p = 0.05$, effect size $W = 0.11$) between *Flat* (*mean* = 24.55, $SD = 12.83$) and *Full Circle* (*mean* = 25.27, $SD = 14.47$).

Walking Distance. Overall, we use a non-parametric test to analyse the walking distance data because the data in each condition is not normally distributed. The Kruskal-Wallis test shows no significant effect of Visual Modifier factor ($p = 0.52$, effect size $\eta^2 = -0.01$) among *Regular* (*mean* = 5.50, $SD = 2.35$), *Restricted FoV* (*mean* = 6.43, $SD = 3.19$), and *Landmark* (*mean* = 5.87, $SD = 2.80$). The Friedman test shows significant effect of Layout factor ($p = 1.97 \times 10^{-9}$, effect size $W = 1$) between *Flat* (*mean* = 8.50, $SD = 1.16$) and *Full Circle* (*mean* = 3.36, $SD = 0.89$). Figure 3-c shows the paired T-test of the Layout factor for each Visual Modifier group.

Participants travel on average 7.74 ($SD = 0.47$) meters in the *Regular-Flat* condition, and 3.27 ($SD = 0.63$) meters in the *Regular-Full Circle* condition. The data are normally distributed in this condition. A paired T-test reports a significant difference between the two conditions ($p = 3.05 \times 10^{-10}$, effect size $|d| = 8.06$).

Participants travel on average 9.26 ($SD = 1.47$) meters in the *Restricted FoV-Flat* condition, and 3.59 ($SD = 1.26$) meters in the *Restricted FoV-Full Circle* condition. The data are not normally distributed in this condition. The Friedman test shows a significant difference between the two conditions ($p = 5.32 \times 10^{-4}$, effect size $W = 1$).

Participants travel on average 8.51 ($SD = 0.81$) meters in the *Landmark-Flat* condition, and 3.23 ($SD = 0.70$) meters in the *Landmark-Full Circle* condition. The data are normally distributed in this condition. A paired T-test reports a significant difference between the two conditions ($p = 1.04 \times 10^{-11}$, effect size $|d| = 6.99$).

Figure 4-a, b, and c show density maps of all user positions (top view) collected during the learning phase for the *Flat* layout. The origin represents the start position. We can see that, in general, participants step back further than the starting position in all conditions, especially in the *Regular* condition. This is indicated by the second dark spot below the origin. Moreover, we can also see that the participants tend to move left and right in the *Restricted FoV* condition, while move directly to the card patterns in the other two conditions. This might explain why they form different density patterns on the graphs. On the other hand, user positions (top view) collected

during the learning phase for the *Full Circle* layout don't differ much among the Visual Modifier conditions. Figure 4-d shows an example of the density plot of users' positions in the *Regular* condition (density maps for the other full circle conditions appear similar and are omitted).

Head Rotation. The data for the Head Rotation during the learning phase is normally distributed. The mixed ANOVA test shows no significant effect in the Visual Modifier factor ($p = 0.75$, effect size $\eta_G^2 = 0.01$) among *Regular* (*mean* = 1097, *SD* = 531), *Restricted FoV* (*mean* = 1044, *SD* = 539), and *Landmark* (*mean* = 1082, *SD* = 510), but a significant effect in the Layout factor ($p = 3.25 \times 10^{-20}$, effect size $\eta_G^2 = 0.83$) between *Flat* (*mean* = 605, *SD* = 111) and *Full Circle* (*mean* = 1544, *SD* = 287). Figure 3-d shows the paired T-test of the Layout factor for each Visual Modifier group.

Participants perform on average 644 (*SD* = 103) degrees of head rotation in the *Regular-Flat* condition, and 1549 (*SD* = 362) degree in the *Regular-Full Circle* condition during the learning phase. A paired T-test shows a significant difference between the two conditions ($p = 4.82 \times 10^{-7}$, effect size $|d| = 3.41$).

Participants perform on average 569 (*SD* = 134) degrees of head rotation in the *Restricted FoV-Flat* condition, and 1519 (*SD* = 310) degree in the *Restricted FoV-Full Circle* condition during the learning phase. A paired T-test shows a significant difference between the two conditions ($p = 4.82 \times 10^{-7}$, effect size $|d| = 3.98$).

Participants perform on average 602 (*SD* = 87) degrees of head rotation in the *Landmark-Flat* condition, and 1562 (*SD* = 182) degree in the *Landmark-Full Circle* condition during the learning phase. A paired T-test shows a significant difference between the two conditions ($p = 7.35 \times 10^{-11}$, effect size $|d| = 6.72$).

Relative View Position. Figure 4-e,f show density plots of the Relative View Position, calculated by projecting the head pointer (ray cast from headset's forward direction) onto the *Flat* grid, with the head position as the origin point. This helps us to investigate how frequently participants look at far away or nearby objects. We use a density map to observe the density of the relative positions for *Flat* layouts. The *Landmark* condition (density map omitted) has a similar plot as the *Regular* condition, while there are differences between the *Restricted FoV* condition and the *Regular* condition.

Results show that the range of head rotation was more smaller with the *Restricted FoV* than in the *Regular* condition, which means that participants look more often at the objects near them than the object faraway in the *Restricted FoV* condition compared with the other Visual Modifier conditions.

Subjective Rating. Overall, for the learning phases, 28 out of all 36 participants prefer the *Flat* layout , while for the recall phases, 31 out of all 36 participants prefer the *Flat* layout (see Figure 5-d,e, and f). Figure 5-a, b, and c show the Kendall's W test for each Visual Modifier condition.

For the NASA-TLX scores in the *Regular* condition, the Kendall's W test shows significant effects for mental ($p = 7.0 \times 10^{-3}$, effect size $r = 0.61$), performance ($p = 7.0 \times 10^{-3}$, effect size $r = 0.61$), effort ($p = 0.03$, effect size $r = 0.38$), frustration ($p = 8.0 \times 10^{-3}$, effect size $r = 0.58$), and overall mean ($p = 4.0 \times 10^{-3}$, effect size $r = 0.69$). The results show that the participants perform the best in the Flat layout with the least mental effort and frustration.

For the NASA-TLX scores in the *Restricted FoV* condition, the Kendall's W test shows significant effects for mental effort only ($p = 0.02$, effect size $r = 0.45$). The results show that the participants perceived that the combination of *Flat* layout and *Restricted FoV* as the least mentally demanding compared to other conditions.

For the NASA-TLX scores in the *Landmark* condition, the Kendall's W test shows significant effects for mental ($p = 3.0 \times 10^{-3}$, effect size $r = 0.75$), performance ($p = 3.0 \times 10^{-3}$, effect size $r = 0.75$), effort ($p = 5.0 \times 10^{-3}$, effect size $r = 0.67$), frustration ($p = 8.0 \times 10^{-3}$, effect size $r = 0.58$),

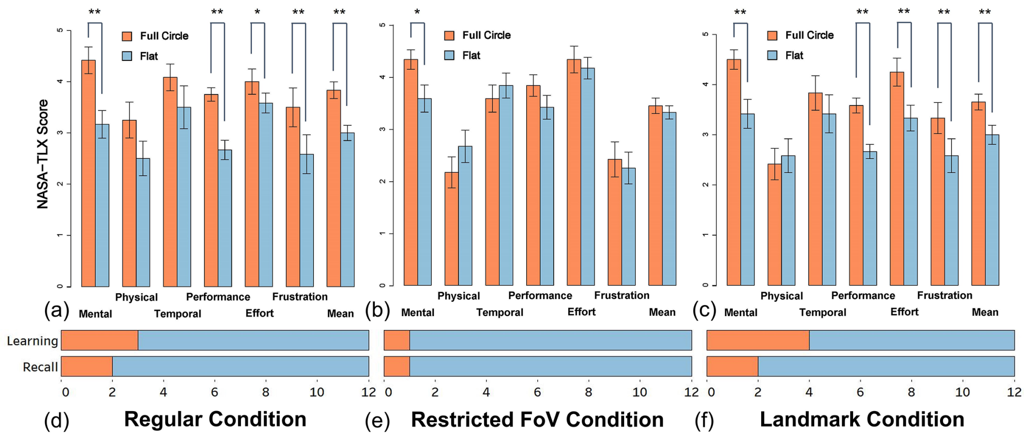

Fig. 5. Participant responses on the NASA-TLX for each layout in (a) *Regular* condition, (b) *Restricted FoV* condition, and (c) *Landmark* condition. Participant preference for each layout in (d) *Regular* condition, (e) *Restricted FoV* condition, and (f) *Landmark* condition. In the NASA-TLX, performance was rated in reverse order (lower is better). Error bars denote standard error.

and overall mean ($p = 4.0 \times 10^{-3}$, effect size $r = 0.69$). The results show that the participants perform the best in the Flat layout with the least mental effort and frustration.

3.7 Discussion

Our accuracy results suggest that the use of *Flat* layout leads to better user performance than *Full Circle* layout. In the *Flat* layout condition, participants made less errors, as measured by *Card Incorrect* and *Manhattan Distance Error*, than the *Full Circle* layout condition. We found no significant difference in recall time between the two layout conditions. These results support our initial conjecture that spatial memory is an important factor in the findings of Liu et al. [31] in their study showing that *Flat* better supported complex visualisation tasks than *Full Circle*.

Looking at the density maps in Figure 4, we notice that participants seemed to take a step back during the task to get an overview of the workspace in the *Flat* layout for both the *Regular* and *Landmark* conditions (this was not possible in the *Restricted FoV*). This could suggest that the positive results of the *Flat* layout may be due to the ability for participants to get an overview of the workspace. To understand the contribution of this overview effect in our observed results, we included the *Restricted FoV* condition in our study. In this condition, our results indicate that the *Flat* layout still leads to better performance compared to the *Full Circle* one meaning that the overview provided by the *Flat* is not the main reason for the difference between the two layouts.

Comments from participants mentioned another potential explanation: with its corners, the *Flat* layout provides natural landmarks that can support participants' spatial memory (the corners were also visible in the *Restricted FoV* condition). This is confirmed by Uddin et al. [51] and Gao et al. [19], where they found that artificial landmarks play an important role in assisting spatial memorisation and retrieval of items in a grid of interface components. Therefore, in the *Landmark* condition, artificial landmarks were added in both layouts to limit any such inherent advantage of the *Flat* one. Our results in this condition still indicate a significantly better performance with the *Flat* layout; therefore eliminating landmarks as a main factor for the difference between the two layouts.

In summary, our results suggest that the main impact of the type of layout on spatial memory is not due to either overview or landmarks. Rather, the analysis of the tracking data in our study shows that the *Full Circle* layout condition requires more head rotation than the *Flat* layout condition. Participants confirmed this by mentioning that the *Full Circle* was more disorienting (P2, P8, P20, and P25), and made locating elements more difficult (P8, P24, P33, and P36). This effect is also supported by the fact that the rotation participants performed during and after the learning phase facilitated errors as a result of an inaccurate update of spatial mental representation [52]. Given all of the apparent advantages of the *Flat* layout, it is no surprise that it was preferred by most participants. Therefore, our study points to the conclusion that disorientation due to the rotational movement required to explore the workspace using the *Full Circle* layout is the main factor that impacts spatial memory.

4 STUDY 2: FLAT VERSUS SEMICIRCLE

Study 1 shows that *Flat* layouts outperform *Full Circle* layouts and are preferred by participants in all 3 Visual Modifier conditions. However, in recent research, Liu et al. [31] report that participants also prefer the *Semicircle* layouts. It is reasonable to hypothesize that such a *Semicircle* arrangement [5, 21, 49] is a good compromise between *Full Circle* and *Flat*. We therefore run another study to investigate the effects on visio-spatial memory between the *Flat* layout and *Semicircle* layout.

4.1 Task

The *Flat* condition has a width of 5.9 m and a height of 1.4 m, which is same as study 1. The *Semicircle* condition has an approximate radius of 1.9m, which results in an arc length the same length as the width of the flat grid. The starting position in the *Flat* layout has a 1.9 meter distance to the grid while the starting position of the *Semicircle* layout is the centre point of the semicircle. The task and the procedure are the same as in Study 1.

4.2 Design

We conduct a within-subjects design study with the Layout (*Flat* and *Semicircle*) as the main independent variable. All participants complete the full set of trials successfully. In total, we collect data from 216 trials (12 participants × 9 repetitions × 2 Layouts). Dependent variables are same as the study 1 including completion time, cards accuracy, Manhattan distance of the errors, walking distance, head positions and rotation angles, and subjective ranking.

4.3 Participants and Apparatus

In total, we recruit 12 participants (6 males and 6 females; mean age = 24.42, *SD* = 2.50) from our university. All participants were students and did not participate in our study 1. 6 participants have at least some experience with VR and 1 of them rate himself as a VR expert. Participants sign up voluntarily and are rewarded a gift card (A$20) and small gifts (candies and chocolates) as a sign of appreciation. We use the same apparatus as in study 1.

4.4 Measures

Measures for study 2 are similar to study 1. For each trial, we record the number of incorrectly chosen cards, along with their positions in the grid. We also record the *Recall Time* taken to select the 5 cards in the answer phase. In our analysis, we use two methods to measure participants' recall accuracy: *Cards Correct* and *Manhattan Distance*.

Participants' head pose are also tracked during the duration of each trial, which we use to calculate the *Walking Distance* travelled by participants and *Head Rotation* during trials. In the post-study questionnaire, the subjective task load is measured using questions derived from the

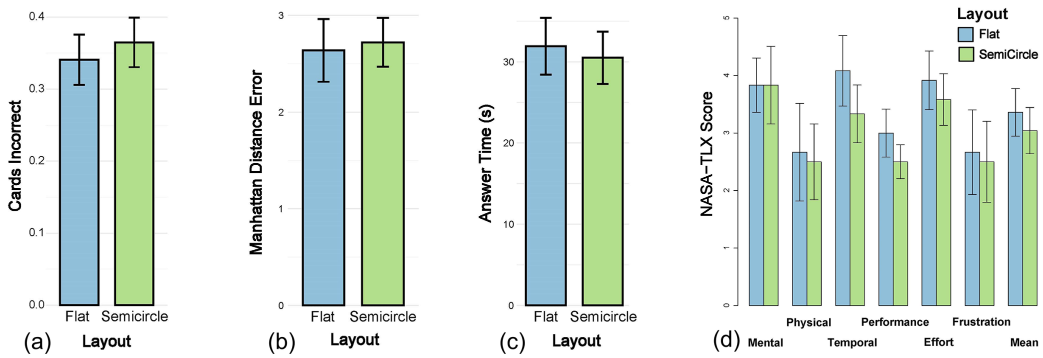

Fig. 6. (a) Cards Incorrect (lower is better), (b) Manhattan Distance (lower is better), (c) Recall time during the recall phase, and (d) Each participant responses on the NASA-TLX in *Flat* layout and *Semicircle* layout. In the NASA-TLX questionnaire, performance was asked in a reverse order (lower is better).

NASA-TLX [20]. Participants also indicate their preferred layout and were asked about their strategy to solve the task.

4.5 Results of Study 2

For quantitative measures, we use the same statistical test method as the first study. Measures are broken down into three categories for reporting, as follows.

Performance: (1) *Cards Incorrect*, or proportion of cards chosen incorrectly in each trial; (2) (optimal) *Manhattan Distance Error* between the participants' answers and the solutions; and (3) *Recall Time*, total time of the recall phase.
Learning Phase: (4) *Walking Distance* travelled by participants; (5) *Head Rotation* performed by participants calculated by cumulative angular distances; and (6) *Relative Head Position* compared with the projection of the participant head pointer on the plane of the wall display.
Subjective: We collected *Preference*, as well as the *Mental* and *Physical* workload of participants.

Cards Incorrect. The data for the incorrect cards are normally distributed. Participants on average have an incorrect cards error rate of 0.34 ($SD = 0.12$) with the *Flat* condition and an incorrect cards error rate of 0.36 ($SD = 0.12$) with the *Semicircle* condition (see Figure 6-a). A paired T-test doesn't find a significant difference between the two conditions ($p = 0.95$, effect size $|d| = 0.20$).
Manhattan Distance Error. The data for the Manhattan Distance Error between the participant answer and the solution is normally distributed. This distance is on average 2.64 ($SD = 1.13$) with the *Flat* layout, and 2.72 ($SD = 0.88$) with the *Semicircle* layout (see Figure 6-b). A paired T-test doesn't find a significant difference between the two conditions ($p = 0.78$, effect size $|d| = 0.08$).
Recall Time. The data for time taken to select the cards in the recall phase is normally distributed. On average, participants took 31.92 ($SD = 12.11$) seconds to select the cards the *Flat* condition, and 30.50 ($SD = 11.18$) seconds in the *Semicircle* condition (See Figure 6-c). A paired T-test doesn't find a significant difference between the two conditions ($p = 0.41$, effect size $|d| = 0.12$).
Walking Distance. The data for the Walking Distance during the learning phase are normally distributed. During the learning phase, participants travel on average 9.38 ($SD = 0.70$) meters in the *Flat* condition, and 6.74 ($SD = 0.61$) meters in the *Semicircle* condition (See Figure 7-a). A

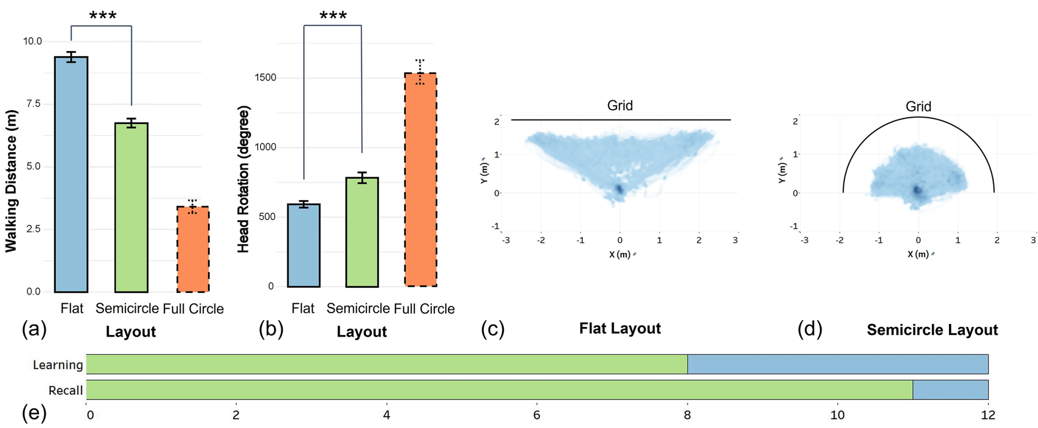

Fig. 7. (a) Walking distance and (b) Head rotation in *Flat* layout and *Semicircle* layout during the learning phase (dashed bars represent mean walking distance and head rotation in study 1 for *Full Circle* layout). Density plots for participants positions during learning phase in (c) *Flat* layout, (d) *Semicircle* layout. And (e) Subjective preference in two layouts. Error bars denote standard error.

paired T-test shows a significant difference between the two conditions ($p = 4.62 \times 10^{-9}$, effect size $|d| = 4.02$).

Figures 7-c and d show density maps of all user positions (top view) collected during the trials. The origin represents the start position. We can see that, in general, participants spend similar time on the original position and on the path to touch the cards between the two layouts.

Head Rotation. The data for the Head Rotation during the learning phase is normally distributed. During the learning phase, participants perform on average 592 ($SD = 81$) degree of head rotation in the *Flat* condition, and 782 ($SD = 133$) degree in the *Semicircle* condition (See Figure 7-b). A paired T-test shows a significant difference between the two conditions ($p = 4.47 \times 10^{-6}$, effect size $|d| = 1.73$).

Subjective Rating. Overall, for the learning phases, 8 out of 12 participants prefer the *Semicircle* layout, while for the recall phases, 11 out of 12 participants prefer the *Semicircle* layout (See Figure 7-e). Figure 6-d shows the NASA-TLX score assessed by all participants for each Layout condition. The Kendall's W test showed no significant effects among these criteria. From the figure, we can see that *Semicircle* layout is rated higher than *Flat* layout in all criteria, which supports our finding on the preference for the *Semicircle* layout.

4.6 Discussion

Contrary to what has been shown by Shupp et al. [49] on their spatial memory study in wall-displays environments, we did not find any difference between the *Flat* and *Semicircle* layout. A notable number of participants prefer the *Semicircle* layout and mention that *Semicircle* layout is easier to see and memorise than *Flat* layout (P3, P5, P8, P10, and P11). They feel it is easier to rotate their heads than walking around to browse (P8). However, some participants report that *Flat* layout has the benefit of providing a complete overview (P4, P6, and P7) but requires more time to touch and memorise than *Semicircle* layout (P1, P2, and P6). The mixed feedback indicates a balance of pros and cons for these two layout types, which may be the reason for their similar performance.

Overall, the *Semicircle* layout provides less walking distance but more head rotation than *Flat* layout. *Semicircle* layout is a compromise between the *Flat* layout and *Full Circle* layout and our follow up study does not show a significant negative impact on spatial memory due to the limited rotation required by *Semicircle* layout.

5 LIMITATIONS AND FUTURE WORK

One limitation of our study design is that our *Full Circle* wraparound condition only tests views with a close distance. While the diameter of around 2 metres is chosen to keep the same grid size as in the *Flat* layout, enabling arms reach distance between the participants and the grid, and to minimise the need for walking (and therefore physical effort), we don't know if the effect on spatial memory remains the same in other wraparound layout such as CAVE2 environments. Also, to limit the experiment duration and participant fatigue we only test one study setup with 36 (3x12) cards and one task difficulty level: memorising a pattern of 5 cards. As discussed in Section 3.1, the study setup follows Liu et al. [31], where 36 small multiples were used in a 3x12 grid, while the task difficulty level was chosen after extensive piloting to find a level difficult enough but doable. Another limitation of our study design is that we choose to focus on one type of selection techniques: direct tapping. As shown on both small [50] and large 2D displays [11], the type of input to select an element (either up-close or from a distance) has an influence on spatial memory. Future work should explore the impact of remote selection techniques such as hand pointer or gaze on spatial memory compared to the direct tapping method we use.

There are, of course, several more factors that should be taken into account in future studies. Firstly, we can summarise from the two studies that among the three layout types, *Flat* layout requires the most translational movement while *Full Circle* requires most head rotation. While our result suggests that full body rotation has a negative effect on spatial memory, further studies could help to isolate the limits of rotation more precisely or whether walking could have a beneficial kinaesthetic memory effect. Secondly, the current experiment relies purely on physical navigation of the participants to interact with the virtual world. However, the effects on spatial memory of physical navigation versus virtual navigation (such as zooming and panning) for this particular task is unknown. Finally, we feel it is important to more deeply explore the connection between spatial memory and analytic tasks in information visualisation. While immersive environments provide users with more 'space to think' [29] than traditional displays, designers must be able to understand the implications of such environments on spatial understanding, when creating tools to facilitate user understanding.

Our study was conducted in Virtual Reality such that the participants could not see any background of their real-world environment. It would be interesting to see how well our results carry over to Augmented Reality headsets, where additional landmarks from the environment may be visible behind the virtual imagery [36, 37]. We would also expect that the more limited field of view provided by existing AR headsets compared to the VR headsets, would place greater demands on spatial memory, though the technology continues to improve.

6 CONCLUSION

We contribute two user studies, each of which evaluates the effect of display layout on spatial memory in immersive environments. We use an abstract task in both studies using layouts of 2D cards based on visuo-spatial memory studies from psychology to determine which layout leads to the best retention of card patterns. In the first study, we examine the effects of the *Flat* layout and the *Full Circle* layout on spatial memory with a mixed-factors study. This study focuses on the two extremes of the layout curvature spectrum as the within-subjects factor. We also investigate the subordinate factors in this study such as the overview advantage of the *Flat* layout and the natural edges and corners of the *Flat* layout as a between-subjects factor. The result shows that the *Flat* layout outperformed the *Full Circle* layout with better accuracy. Moreover, the subjective result shows that the participants perform the best in the *Flat* layout with least mental effort and frustration. This general result held regardless of the subordinate factors (visual modifiers) that we introduced, implying that the main factor that influence the performance is the type of physical navigation. In other words, that walking in front of a *Flat* display is less detrimental to spatial memory than rotation.

In the second study we test whether the *Semicircle* layout provides a good compromise. The second study reveals no significant difference between *Flat* and *Semicircle* while participants prefer the *Semicircle* layout. Past work [31] has suggested that *Semicircle* layout provides advantages over *Flat* in terms of reduced total walking distance and perspective distortion without requiring the full rotation navigation of *Full Circle*. Our finding suggests that these advantages of *Semicircle* can be achieved with no significant negative impact on spatial memory.

In summary, the clear take away from our studies is that full wrap-around displays and the rotation required to navigate them is disorienting to users and should be used with caution for immersive information presentation. This finding is contrary to use patterns we see emerging in literature, so we hope this paper will influence future system implementers to be aware of this limitation, and further research to find nuance.

ACKNOWLEDGEMENT

We thank our anonymous reviewers for their feedback. This research was supported under the Australian Research Council's Discovery Projects funding scheme (project number DP180100755) and by an Australian Government Research Training Program (RTP) Scholarship.

REFERENCES

[1] R.C. Atkinson and R.M. Shiffrin. 1968. Human Memory: A Proposed System and its Control Processes. In *Psychology of Learning and Motivation*, Kenneth W. Spence and Janet Taylor Spence (Eds.). Psychology of Learning and Motivation, Vol. 2. Academic Press, http://www.sciencedirect.com/science/article/pii/S0079742108604223, 89 – 195. https://doi.org/10.1016/S0079-7421(08)60422-3

[2] Alan D. Baddeley and Graham Hitch. 1974. Working Memory. In *Psychology of Learning and Motivation*, Gordon H. Bower (Ed.). Psychology of Learning and Motivation, Vol. 8. Academic Press, http://www.sciencedirect.com/science/article/pii/S0079742108604521, 47 – 89. https://doi.org/10.1016/S0079-7421(08)60452-1

[3] Andrea Batch, Andrew Cunningham, Maxime Cordeil, Niklas Elmqvist, Tim Dwyer, Bruce H Thomas, and Kim Marriott. 2019. There is no spoon: Evaluating performance, space use, and presence with expert domain users in immersive analytics. *IEEE Transactions on Visualization and Computer Graphics* 26, 1 (2019), 536–546. https://doi.org/10.1109/TVCG.2019.2934803

[4] Christopher-Marcel Boddecker. 2018. ExtendRealityLtd/VRTK. https://github.com/ExtendRealityLtd/VRTK. Accessed: 2019-07-08.

[5] M. Cavallo, M. Dholakia, M. Havlena, K. Ocheltree, and M. Podlaseck. 2019. Dataspace: A Reconfigurable Hybrid Reality Environment for Collaborative Information Analysis. , 145-153 pages. https://doi.org/10.48550/arXiv.1903.03700

[6] Andy Cockburn and Bruce McKenzie. 2002. Evaluating the Effectiveness of Spatial Memory in 2D and 3D Physical and Virtual Environments. In *Proc. SIGCHI on Human Factors in Computing Systems* (Minneapolis, Minnesota, USA) *(CHI '02)*. ACM, New York, NY, USA, 203–210. https://doi.org/10.1145/503376.503413

[7] Maxime Cordeil, Tim Dwyer, Karsten Klein, Bireswar Laha, Kim Marriott, and Bruce H Thomas. 2016. Immersive collaborative analysis of network connectivity: Cave-style or head-mounted display? *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2016), 441–450. https://doi.org/10.1109/TVCG.2016.2599107

[8] Philip M Corsi. 1973. *Human memory and the medial temporal region of the brain.* Ph.D. Dissertation. ProQuest Information & Learning.

[9] Carolina Cruz-Neira, Daniel J. Sandin, Thomas A. DeFanti, Robert V. Kenyon, and John C. Hart. 1992. The CAVE: Audio Visual Experience Automatic Virtual Environment. *Commun. ACM* 35, 6 (June 1992), 64–72. https://doi.org/10.1145/129888.129892

[10] S Della Sala, C Gray, A Baddeley, and L Wilson. 1997. *The Visual Patterns Test: A new test of short-term visual recall.* Thames Valley Test Company, Thames Valley Test Company.

[11] Achim Ebert, Matthias Deller, Daniel Steffen, and Matthias Heintz. 2009. "Where Did I Put That?" – Effectiveness of Kinesthetic Memory in Immersive Virtual Environments, Vol. 5616. Springer, 179–188. https://doi.org/10.1007/978-3-642-02713-0_19

[12] Barrett M. Ens, Rory Finnegan, and Pourang P. Irani. 2014. The Personal Cockpit: A Spatial Interface for Effective Task Switching on Head-Worn Displays. In *Proc. SIGCHI on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. ACM, New York, NY, USA, 3171–3180. https://doi.org/10.1145/2556288.2557058

[13] Alessandro Febretti, Arthur Nishimoto, Terrance Thigpen, Jonas Talandis, Lance Long, J. Pirtle, Tom Peterka, Alan Verlo, Maxine Brown, Dana Plepys, Daniel Sandin, Luc Renambot, Andrew Johnson, and Jason Leigh. 2013. CAVE2: A Hybrid Reality Environment for Immersive Simulation and Information Analysis. *Proc. SPIE - The International Society for Optical Engineering* 8649 (03 2013), 03–. https://doi.org/10.1117/12.2005484

[14] Thibault Friedrich, Arnaud Prouzeau, and Michael McGuffin. 2021. The Effect of Increased Body Motion in Virtual Reality on a Placement-Retrieval Task. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology* (Osaka, Japan) *(VRST '21)*. Association for Computing Machinery, New York, NY, USA, Article 14, 5 pages. https://doi.org/10.1145/3489849.3489888

[15] Bruno Fruchard, Eric Lecolinet, and Olivier Chapuis. 2018. Impact of Semantic Aids on Command Memorization for On-Body Interaction and Directional Gestures. In *Proceedings of the International Conference on Advanced Visual Interfaces* (Castiglione della Pescaia, Grosseto, Italy) *(AVI '18)*. ACM, New York, NY, USA, Article 14, 9 pages. https://doi.org/10.1145/3206505.3206524

[16] Alexia Galati, Riley Schoppa, and Aidong Lu. 2021. Exploring the SenseMaking Process through Interactions and fNIRS in Immersive Visualization. *IEEE Transactions on Visualization and Computer Graphics* 27, 5 (2021), 2714–2724. https://doi.org/10.1109/TVCG.2021.3067693

[17] Ronak Dipakkumar Gandhi and Dipam S Patel. 2018. Virtual reality–opportunities and challenges. *Virtual Reality* 5, 01 (2018), 2714–2724.

[18] BoYu Gao, Zhiguo Chen, Xian Chen, Huawei Tu, and Feiran Huang. 2021. The effects of audiovisual landmarks on spatial learning and recalling for image browsing interface in virtual environments. *Journal of Systems Architecture* 117 (2021), 102096. https://doi.org/10.1016/j.sysarc.2021.102096

[19] B. Gao, H. Kim, B. Kim, and J. Kim. 2018. Artificial Landmarks to Facilitate Spatial Learning and Recalling for Curved Visual Wall Layout in Virtual Reality. In *IEEE International Conference on Big Data and Smart Computing (BigComp)*. IEEE, IEEE, 475–482. https://doi.org/10.1109/BigComp.2018.00076

[20] Sandra G. Hart. 2006. NASA-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (2006), 904–908. https://doi.org/10.1177/154193120605000909 arXiv:https://doi.org/10.1177/154193120605000909

[21] Devamardeep Hayatpur, Haijun Xia, and Daniel Wigdor. 2020. DataHop: Spatial Data Exploration in Virtual Reality. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '20)*. Association for Computing Machinery, New York, NY, USA, 818–828. https://doi.org/10.1145/3379337.3415878

[22] Yvonne Jansen, Jonas Schjerlund, and Kasper Hornbæk. 2019. Effects of Locomotion and Visual Overview on Spatial Memory When Interacting with Wall Displays. In *Proc. CHI on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. ACM, New York, NY, USA, Article 291, 12 pages. https://doi.org/10.1145/3290605.3300521

[23] Dylan Kobayashi, Nurit Kirshenbaum, Roderick S Tabalba, Ryan Theriot, and Jason Leigh. 2021. Translating The Benefits Of Wide-Band Display Environments Into An XR Space. In *Symposium on Spatial User Interaction* (Virtual Event, USA) *(SUI '21)*. Association for Computing Machinery, New York, NY, USA, Article 9, 11 pages. https://doi.org/10.1145/3485279.3485294

[24] Eric Krokos, Catherine Plaisant, and Amitabh Varshney. 2019. Virtual Memory Palaces: Immersion Aids Recall. *Virtual Real.* 23, 1 (March 2019), 1–15. https://doi.org/10.1007/s10055-018-0346-3

[25] Harold W. Kuhn. 2010. *The Hungarian Method for the Assignment Problem.* Springer Berlin, Heidelberg, Berlin, Heidelberg, 29–47. https://doi.org/10.1007/978-3-540-68279-0_2

[26] Oh-Hyun Kwon, Chris Muelder, Kyungwon Lee, and Kwan-Liu Ma. 2015. Spherical layout and rendering methods for immersive graph visualization. In *IEEE Pacific Visualization Symposium (PacificVis)*. IEEE, IEEE, 63–67. https://doi.org/10.1109/PACIFICVIS.2015.7156357

[27] Benjamin Lee, Xiaoyun Hu, Maxime Cordeil, Arnaud Prouzeau, Bernhard Jenny, and Tim Dwyer. 2021. Shared Surfaces and Spaces: Collaborative Data Visualisation in a Co-located Immersive Environment. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (Feb 2021), 1171–1181. https://doi.org/10.1109/tvcg.2020.3030450

[28] Frank Chun Yat Li, David Dearman, and Khai N. Truong. 2009. Virtual Shelves: Interactions with Orientation Aware Devices. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology* (Victoria, BC, Canada) *(UIST '09)*. ACM, New York, NY, USA, 125–128. https://doi.org/10.1145/1622176.1622200

[29] Lee Lisle, Xiaoyu Chen, J.K. Edward Gitre, Chris North, and Doug A. Bowman. 2020. Evaluating the Benefits of the Immersive Space to Think. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*

(VRW). IEEE, New York, NY, USA, 331–337. https://doi.org/10.1109/VRW50115.2020.00073

[30] Jiazhou Liu. 2022. Spatial Memory Test. https://github.com/JiazhouLiu/SpatialMemoryTest.

[31] Jiazhou Liu, Arnaud Prouzeau, Barrett Ens, and Tim Dwyer. 2020. Design and Evaluation of Interactive Small Multiples Data Visualisation in Immersive Spaces. In *IEEE Conference on Virtual Reality and 3D User Interfaces (VR) (VR '20)*. IEEE, Atlanta, Georgia, USA, 588–597. https://doi.org/10.1109/VR46266.2020.00081

[32] Robert H Logie and RH Logie. 1995. *Visuo-spatial working memory*. Psychology Press, Psychology Press.

[33] Robert H. Logie, Gesualdo M. Zucco, and Alan D. Baddeley. 1990. Interference with visual short-term memory. *Acta Psychologica* 75, 1 (1990), 55 – 74. https://doi.org/10.1016/0001-6918(90)90066-O

[34] Paul Lubos, Gerd Bruder, Oscar Ariza, and Frank Steinicke. 2016. Touching the Sphere: Leveraging Joint-Centered Kinespheres for Spatial User Interaction. In *Proceedings of the Symposium on Spatial User Interaction* (Tokyo, Japan) *(SUI '16)*. ACM, New York, NY, USA, 13–22. https://doi.org/10.1145/2983310.2985753

[35] Steven J Luck and Edward K Vogel. 1997. The capacity of visual working memory for features and conjunctions. *Nature* 390, 6657 (1997), 279–281.

[36] Weizhou Luo, Anke Lehmann, Hjalmar Widengren, and Raimund Dachselt. 2022. Where Should We Put It? Layout and Placement Strategies of Documents in Augmented Reality for Collaborative Sensemaking. In *CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 627, 16 pages. https://doi.org/10.1145/3491102.3501946

[37] Weizhou Luo, Anke Lehmann, Yushan Yang, and Raimund Dachselt. 2021. Investigating Document Layout and Placement Strategies for Collaborative Sensemaking in Augmented Reality. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, Article 456, 7 pages. https://doi.org/10.1145/3411763.3451588

[38] Muhanna A. Muhanna. 2015. Virtual reality and the CAVE: Taxonomy, interaction challenges and research directions. *Journal of King Saud University - Computer and Information Sciences* 27, 3 (2015), 344–361. https://doi.org/10.1016/j.jksuci.2014.03.023

[39] Jens Müller, Roman Rädle, and Harald Reiterer. 2016. Virtual Objects as Spatial Cues in Collaborative Mixed Reality Environments: How They Shape Communication Behavior and User Task Load. In *Proc. CHI on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. ACM, New York, NY, USA, 1245–1249. https://doi.org/10.1145/2858036.2858043

[40] Huansheng Ning, Hang Wang, Yujia Lin, Wenxi Wang, Sahraoui Dhelim, Fadi Farha, Jianguo Ding, and Mahmoud Daneshmand. 2021. A Survey on Metaverse: the State-of-the-art, Technologies, Applications, and Challenges. *CoRR* abs/2111.09673 (2021), 2714–2724. arXiv:2111.09673 https://arxiv.org/abs/2111.09673

[41] Simon T. Perrault, Eric Lecolinet, Yoann Pascal Bourse, Shengdong Zhao, and Yves Guiard. 2015. Physical Loci: Leveraging Spatial, Object and Semantic Memory for Command Selection. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. ACM, New York, NY, USA, 299–308. https://doi.org/10.1145/2702123.2702126

[42] Matthew D. Plumlee and Colin Ware. 2006. Zooming Versus Multiple Window Interfaces: Cognitive Costs of Visual Comparisons. *ACM Trans. Comput.-Hum. Interact.* 13, 2 (June 2006), 179–209. https://doi.org/10.1145/1165734.1165736

[43] Arnaud Prouzeau, Anastasia Bezerianos, and Olivier Chapuis. 2016. Visual Immersion in the Context of Wall Displays. In *Proceedings of the ACM Companion on Interactive Surfaces and Spaces* (Niagara Falls, Ontario, Canada) *(ISS Companion '16)*. ACM, New York, NY, USA, 33–39. https://doi.org/10.1145/3009939.3009945

[44] Patrick Reipschläger, Tamara Flemisch, and Raimund Dachselt. 2021. Personal Augmented Reality for Information Visualization on Large Interactive Displays. *IEEE Transactions on Visualization and Computer Graphics* 27 (2 2021), 11. Issue 2. https://doi.org/10.1109/TVCG.2020.3030460

[45] George Robertson, Mary Czerwinski, Kevin Larson, Daniel C. Robbins, David Thiel, and Maarten van Dantzich. 1998. Data Mountain: Using Spatial Memory for Document Management. In *Proceedings of the 11th Annual ACM Symposium on User Interface Software and Technology* (San Francisco, California, USA) *(UIST '98)*. ACM, New York, NY, USA, 153–162. https://doi.org/10.1145/288392.288596

[46] Kadek Ananta Satriadi, Barrett Ens, Maxime Cordeil, Tobias Czauderna, and Bernhard Jenny. 2020. Maps Around Me: 3D Multiview Layouts in Immersive Spaces. *Proc. ACM Hum.-Comput. Interact.* 4, ISS, Article 201 (nov 2020), 20 pages. https://doi.org/10.1145/3427329

[47] Joey Scarr, Andy Cockburn, and Carl Gutwin. 2013. Supporting and Exploiting Spatial Memory in User Interfaces. *Found. Trends Hum.-Comput. Interact.* 6, 1 (dec 2013), 1–84. https://doi.org/10.1561/1100000046

[48] Joey Scarr, Andy Cockburn, Carl Gutwin, and Andrea Bunt. 2012. Improving Command Selection with CommandMaps. In *Proc. SIGCHI on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. ACM, New York, NY, USA, 257–266. https://doi.org/10.1145/2207676.2207713

[49] Lauren Shupp, Robert Ball, Beth Yost, John Booker, and Chris North. 2006. Evaluation of Viewport Size and Curvature of Large, High-Resolution Displays. In *Proceedings of Graphics Interface* (Quebec, Canada) *(GI '06)*. Canadian Information

Processing Society, CAN, 123–130. https://doi.org/10.5555/1143079.1143100

[50] Desney S. Tan, Randy Pausch, Jeanine K. Stefanucci, and Dennis R. Proffitt. 2002. Kinesthetic Cues Aid Spatial Memory. In *CHI Extended Abstracts on Human Factors in Computing Systems* (Minneapolis, Minnesota, USA) *(CHI EA '02)*. ACM, New York, NY, USA, 806–807. https://doi.org/10.1145/506443.506607

[51] Md. Sami Uddin, Carl Gutwin, and Andy Cockburn. 2017. The Effects of Artificial Landmarks on Learning and Performance in Spatial-Memory Interfaces. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 3843–3855. https://doi.org/10.1145/3025453.3025497

[52] David Waller, Daniel Montello, Anthony Richardson, and Mary Hegarty. 2002. Orientation Specificity and Spatial Updating of Memories for Layouts. *Journal of experimental psychology. Learning, memory, and cognition* 28 (12 2002), 1051–63. https://doi.org/10.1037//0278-7393.28.6.1051

[53] J. T. Lindsay Wilson, Jane H. Scott, and Kevin G. Power. 1987. Developmental differences in the span of visual memory for pattern. *British Journal of Developmental Psychology* 5, 3 (1987), 249–255. https://doi.org/10.1111/j.2044-835X.1987.tb01060.x arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.2044-835X.1987.tb01060.x

[54] Fumeng Yang, Jing Qian, Johannes Novotny, David Badre, Cullen D. Jackson, and David H. Laidlaw. 2021. A Virtual Reality Memory Palace Variant Aids Knowledge Retrieval from Scholarly Articles. *IEEE Transactions on Visualization and Computer Graphics* 27, 12 (2021), 4359–4373. https://doi.org/10.1109/TVCG.2020.3009003

[55] Johannes Zagermann, Ulrike Pfeil, Daniel Fink, Philipp von Bauer, and Harald Reiterer. 2017. Memory in Motion: The Influence of Gesture- and Touch-Based Input Modalities on Spatial Memory. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 1899–1910. https://doi.org/10.1145/3025453.3026001