The Effect of Egocentric Body Movements on Users' Navigation Performance and Spatial Memory in Zoomable User Interfaces

Roman Rädle¹, Hans-Christian Jetter², Simon Butscher¹, Harald Reiterer¹

¹ Human-Computer Interaction Group, University of Konstanz, Germany
{Roman.Raedle,Simon.Butscher,Harald.Reiterer}@uni-konstanz.de

² Intel ICRI Cities, University College London, London, United Kingdom
h.jetter@ucl.ac.uk

ABSTRACT

We present two experiments examining the impact of navigation techniques on users' navigation performance and spatial memory in a zoomable user interface (ZUI). The first experiment with 24 participants compared the effect of egocentric body movements with traditional multi-touch navigation. The results indicate a 47% decrease in path lengths and a 34% decrease in task time in favor of egocentric navigation, but no significant effect on users' spatial memory immediately after a navigation task. However, an additional second experiment with 8 participants revealed such a significant increase in performance of long-term spatial memory: The results of a recall task administered after a 15-minute distractor task indicate a significant advantage of 27% for egocentric body movements in spatial memory. Furthermore, a questionnaire about the subjects' workload revealed that the physical demand of the egocentric navigation was significantly higher but there was less mental demand.

Author Keywords

Spatial memory; navigation performance; egocentric body movements; zoomable user interface

ACM Classification Keywords

H.5.3. Group and Organization Interfaces: Collaborative computing.

General Terms

Experimentation; Human Factors; Measurement; Theory.

INTRODUCTION

Space is an important cognitive resource. During epistemic activities, people lay out information in space and scale to externalize and spatially reflect their current thoughts [18], or to provide a shared space for in-group discussion [12]. This behavior is documented for many activities, for example for sensemaking and analysis of data [1], ideation in creative design [9], or general practices of knowledge

work [17]. For this reason, humans' spatial memory and navigation performance are crucial for overall user performance in a variety of application domains. A good spatial memory allows users to keep track of the location and identity of frequently used objects and can also increase the speed of navigation.

Figure 1. Egocentric navigation to navigate in large information spaces projected on high-resolution displays.

Since the advent of interactive systems with large screens, navigation is no longer constrained to moving in virtual space alone. The benefit of large interactive displays is that they blend the advantages of the digital world (e.g., saving and restoring content, dynamic arrangements, and automatic layouts; rendering multimedia content) with that of familiar physical navigation [1,3,4,24]. For example, they enable users to navigate the information space from an egocentric perspective by glancing around or walking in front of the large display so that view management becomes an entirely familiar physical activity that often happens beneath the conscious awareness of the user.

However, the users' visual acuity is limited, restricting the number of details that they can perceive on a large display depending on their current position, viewing angle [6], and the display's resolution. One solution to this problem is to provide users with an additional personal mobile display, e.g., a tablet. Typically, a tablet is held at an optimal distance from one's eyes so that it can provide users with a view of a section of the content of the large display and it enables natural manipulation and annotation of items with the tablet's touch or stylus input. Furthermore, such mobile displays can be used as a high-resolution viewport for a zoomable user interface that enables users to navigate the

large display's content in space and scale. This enables natural ways of accessing detailed information or extended functionality that is only visible when zooming in, for example with semantic zooming [5]. A potential disadvantage, however, is the need to reintroduce navigation techniques for view management (i.e., zooming and panning) that potentially limit the positive effect of the physical navigation in front of a large display.

In this paper, we were interested in determining the effect the navigation technique for view management on the tablet has on the users' navigation performance and spatial memory. To address this, we implemented two navigation techniques that facilitate navigation in a ZUI. The first navigation technique enables panning and zooming based on egocentric body movements that utilize proxemic measures [10]. This navigation technique was compared with the well-established "drag-to-pan" and "pinch-to-zoom" multi-touch gestures for ZUIs in two controlled experiments.

We describe the two experiments that explored the effect on users' navigation performance and spatial memory when navigating in virtual information spaces. Each experiment is explained in detail and the results are reported and then discussed. The paper concludes with an interpretation of the results, and a discussion of its potential contributions to existing application domains and future work.

RELATED WORK

Our work here draws upon the existing studies that measured the influence of interaction and navigation techniques or visualizations on users' spatial memory and navigation performance. The second part of the related work summarizes those projects that make use of spatial interaction techniques but which do not focus on research in spatial memory.

Studies of Spatial Memory in HCI

An early and well-known study of spatial memory in HCI is Data Mountain that concentrated on the effect that visualizations had on users' spatial memory performance in a document management task [21]. Robertson et al. compared the management of browser bookmarks with a traditional browser (Microsoft IE4) and the Data Mountain interface that allowed users to arrange thumbnails of bookmarks in a virtual 2.5D space. The results showed that, on average, users recalling bookmarks quicker and more reliable when using the spatial interface and hinted at the importance of designing for better spatial memory in HCI.

According to embodied views of cognition, memory performance can be influenced by body movements. For example, a study by Tan et al. [25] reported a 19% better spatial memory performance in favor of touch input compared to mouse input when recalling objects that have been positioned by the users in a preceding task. The authors reason that the increase of precision in the spatial recall task is attributable to "kinesthetic cues."

Inspired by these findings, Jetter et al. conducted two consecutive experiments in 2012 that investigated the effect of touch vs. mouse input on navigation performance and users' spatial memory for panning user interfaces (panning-only UI) and zoomable user interfaces (ZUI). In their first experiment, they found that multi-touch interaction for a panning-only UI on a tabletop leads to better navigation performance as well as better spatial memory compared to traditional interaction with a mouse. Conversely, the second experiment does not show such significant differences for a ZUI. The navigation performance is even worse than with a mouse. For our experiment, we adapted Jetter et al.'s study design and their approach of measuring navigation performance in a ZUI as "[...] a measure of length and not of time or motor costs since they would introduce noise that is not related to memory (e.g., user-specific experience with using multi-touch, individual preferences for mouse velocity)." [13]

Later, Kaufmann and Ahlström [16] conducted a similar study on navigation performance and spatial memory on projector phones with physical peephole interaction. They conducted a user study on the effect of peephole navigation vs. multi-touch interaction on spatial memory. Unlike Jetter et al., they also introduced an observer as bystander next to the individual carrying out the tasks. They found that observers were more accurate with the peephole interaction when they were asked to recall previously visited locations on a map. However, Kaufmann and Ahlström, did not study the effect on humans' spatial cognition when using a ZUI.

Spatial Interaction Techniques & Bodily Movement

Fitzmaurice demonstrated spatial interaction with a 3D virtual environment in his Chameleon prototype [8]. Although he did not study the effect on users' spatial memory when using spatial interaction, the Chameleon prototype enabled an early experimental evaluation of the advantages of spatial interaction and he observed a positive user experience.

Later, Peephole [27] interaction used a small mobile device and no additional overview projection to navigate in a large virtual canvas. The researchers implemented different usage scenarios for interaction with 2D and 3D environments. The focus of their work is on the study of different interaction techniques, including selection tasks, navigation, and manipulation such as drawing. Although they mention that spatially-aware displays enable the use of spatial memory for navigation, they have not conducted an experiment that supports this statement.

Spindler et al. combined a spatially-aware display with a magic lens [7]. Their tangible PaperLens [23] enables physical navigation in volumetric, layered, or zoomable information spaces. In consecutive work, they also studied the minimum thickness of layers, users' holding accuracy, and the physical boundaries of the interaction volume [22]. However, they have not studied the effect of tangible magic lenses on users' spatial memory.

Ball et al. [3] conducted a user study that measured the effect of physical vs. virtual navigation in a 2D virtual space when interacting in front of a large, high-resolution display. Participants were allowed to freely move in front of the display and all interaction with it was performed with a Gyration GyroMouse. Ball et al. found that physical navigation outperforms virtual navigation for tasks such as navigation, search, or pattern matching, but the effects on users' spatial memory were not tested.

In contrast to the afore-mentioned related work, our research specifically investigates the effect of egocentric body movements vs. multi-touch on users' navigation performance and spatial memory when interacting with a zoomable user interface.

RESEARCH BACKGROUND

Recent work in HCI indicates potential advantages of physical navigation over virtual navigation in large information spaces both for single users [1,3,4] and for multi-user collaborative work [11]. However, as discussed above, even a large high-resolution display can only provide a limited amount of information, for example because of the users' visual acuity, and it also restricts the ways in which users can interact and manipulate the information, e.g., using touch or stylus input.

ZUIs (or "multi-scale UIs") on a secondary personal device can provide the user with a second, mobile window to the primary large display. The large display's entire content becomes a virtual "canvas" that can be navigated through the mobile display using zooming and panning techniques and manipulated by touch or stylus input. This is particularly useful when navigating or collaborating in large virtual workspaces with very detailed data, e.g., high-resolution maps, satellite imagery, large documents, image collections, or large virtual whiteboards. In collaborative scenarios, ZUIs with personal devices can also enable different coupling styles of collaboration, e.g., parallel work vs. tightly-coupled collaboration [26].

Unfortunately, these advantages come at a cost: ZUIs introduce a greater cognitive load due to the frequent switches between application level (e.g., reading, annotating, search) and system level tasks (e.g., view management with panning and zooming) [1]. To reduce this load, Bederson formulates design principles for ZUIs [5]: It must be possible for users to navigate within information spaces without any sort of training and it must be difficult or impossible for them to get lost in Desert Fog [15]. He also states that spatial layout and movement by panning and zooming alone are not sufficient cues for memorizing content and locations. Therefore, we want to test whether an egocentric navigation can increase the efficiency and reduce the cognitive load of ZUI navigation and if its greater kinesthetic and proprioceptive feedback offer additional cues that facilitate the encoding of the location and identity of objects in the users' memory.

We isolated two main tasks that are relevant when interacting with large high-resolution displays and which often occur in epistemic processes. First, in the navigation task, users search for specific objects on a large spatial canvas, and second, a recall task in which users need to precisely recall the locations of objects based on their spatial memory. In consequence of the open research about the advantages of egocentric body movements, we formulated our research question as follows:

Do egocentric body movements better support the navigation performance and increase users' spatial memory capabilities in a zoomable user interface?

Based on the literature review and on our research question, we generated the following four hypotheses:

Hypothesis 1 (H1): Navigation Performance

Users perform better in navigation tasks within a zoomable user interface when using egocentric navigation compared to multi-touch navigation.

Hypothesis 2 (H2): Spatial Memory

The object location and identity recall capabilities of users increase with egocentric body movement compared to multi-touch navigation.

Hypothesis 3 (H3): Subjective Workload

The workload assessment of users significantly differs between egocentric body movement and multi-touch navigation.

Hypothesis 4 (H4): Long-term Memory

The interaction with egocentric body movement has a positive effect on users' long-term memory compared to multi-touch navigation.

EXPERIMENT E1

The first experiment compares two navigation techniques and allows us to confirm or reject Hypotheses H1-H3. Both techniques allow panning and zooming of a large virtual canvas. However, the baseline technique uses traditional touch input and the second technique uses egocentric body movements. According to our H1 and H2, we assume that egocentric body movements will result in better *spatial memory performance* and *navigation performance*. Our experiment compares the impact of multi-touch vs. egocentric navigation on a zoomable UI. For the experiment, we decided to use more abstract and generic spatial tasks instead of focusing on a specific real-world application, thus increasing the internal validity. Still, it could resemble an epistemic session where design artifacts were created on a tablet and then collected and arranged in space on a wall-sized display in order to project and share one's internal thoughts.

In the experiment, participants had to perform two tasks: (1) a navigation task and (2) a recall task. In the navigation task, they had to search for eight presented objects similar to a

memory game¹ (see Figure 2). This task is analogous to many real-world applications in which users have to search for a particular information object in a large information space. The objects were repeated, which solidifies them in users' spatial memory. The second task then exemplified a task in which users had to recall the spatial location and identity of previously searched objects.

Figure 2. Both sides of a memory card. The memory card hid the symbol (left) until a certain zooming level was reached and then uncovered the underlying symbol (right).

Figure 1 provides an overview of the environmental setup. Both the high-resolution display and the mobile device are connected and share a large virtual canvas. Objects (memory cards) are spatially distributed on the large canvas. The mobile device provides a dynamic view of the canvas whereas the view on the large display visualizes the whole canvas statically. Two navigation techniques were provided to manipulate the dynamic view on the tablet. They are explained in the following paragraphs. In addition, to raise the awareness of users, a purple rectangle was presented on the large display that represents the view currently visible on the tablet (see Figure 5). With this, users have the chance to orient themselves spatially and keep their global orientation without getting lost in Desert Fog.

Figure 3. A participant solving the navigation task in the multi-touch interaction condition.

Multi-touch condition (Figure 3). The participants are provided with a PC desk and an office chair. Both were arranged by the participants at the beginning of the multi-touch condition so that they could comfortably sit and see all of the objects on the large canvas. The tablet is laid out on the desk and users are able to navigate the large canvas on the tablet using conventional multi-touch *drag-to-pan* and

pinch-to-zoom gestures. While there are alternative interaction techniques for pan and zoom, we chose this technique because it reflects the dominant design in ZUIs for multi-touch devices. To not penalize the multi-touch condition, we let participants decide whether they would like to hold the tablet in an upright position or leave it lying on the desk in front of them. All participants chose to place the tablet in the horizontal position on the desk.

Egocentric condition (Figure 4). In this condition, the users pan and zoom in the large information canvas employing proxemic measures. The technique uses the position and the distance of the tablet in relation to the large display for navigation. This technique is similar to PaperLens [23]. However, in this condition, we used a vertical display instead of a horizontal display and enlarged the interaction space to force participants to move their entire body instead of relying only on arm movements. Lateral and perpendicular movements of the device resulted in panning and orthogonal movements in zooming (move towards large display = zoom in, move away from large display = zoom out). By doing so, we could observe the effect of body movements on the users' navigation performance and spatial memory and compare it with multi-touch navigation without body movement in the other condition.

Figure 4. A participant solving the navigation task in the egocentric interaction condition.

Method

We conducted a controlled lab experiment and recruited 24 subjects. The navigation technique (m = multi-touch, e = egocentric) and the object pools (A = PoolA, B = PoolB) were independent variables. The navigation performance (cost factor), the recall error rate, the navigation task completion time, and the subjective reports were dependent variables. The study was conducted as a within-subjects design with repeated measures. It was systematically counterbalanced and each participant was randomly assigned to one of the four groups ($mAeB$, $mBeA$, $eAmB$, $eBmA$).

Apparatus

The system consisted of two displays: a large high-resolution and non-interactive wall-sized display and a tablet providing

¹ Memory, also known as Concentration, Pelmanism, or Pairs.

multi-touch input. The large display consisted of 2 eyevisCube EC-67-SXT+ (67" each) and a total resolution of 3.840 x 1.080 pixels. The physical dimensions of the horizontally concatenated cubes were 292 x 82 cm. The tablet was an Acer Iconia Tab W500 (10.1" TFT LCD Display LED Backlight) with a resolution of 1.280 x 800 pixels. It provided a capacitive multi-touch input and weighed approximately 1.6 kg. A rigid body marker was attached to the tablet in order to track the egocentric body movements. The rigid body was present in both conditions to avoid any bias. For the egocentric navigation, the total interaction volume in front of the wall-sized display was approximately 292 x 160 x 82 cm. In the multi-touch navigation, participants were asked to position the PC desk at a distance so that they could comfortably see the entire large canvas. The mean distance of the PC desk to the wall-sized display was approximately 160 cm.

The software was implemented in Microsoft .NET 4.0/WPF with C#. The ZUI, semantic zoom, and view synchronization were implemented with the ZOIL API² [14]. An OptiTrack motion capture tracking system from NaturalPoint with 18 cameras (10 default lens, 8 wide-angle lens) allowed for precise tracking of the tablet for the egocentric body movement condition. It tracks motion within a tolerance less than 0.5 mm. We implemented an OptiTrackInputModule for the ProximityToolkit [19] to measure proxemic dimension, such as the position, distance, and orientation of the tablet in relation to the large display. The mobile device was calibrated in the interaction volume so that the minimum scale factor of .75 on the mobile device was reached at a distance of 190 cm and the maximum scale factor of 8.0 was reached at a distance of 30 cm in front of the wall-sized display.

Participants

In all, 24 subjects (12 female) participated in the experiment. The age of participants ranged from 19 to 57 years ($M = 25.42$, $SD = 8.87$). Age was non-normally distributed, with skewness of 2.97 ($SE = 0.47$) and kurtosis of 8.51 ($SE = 0.92$). Two participants were left-handed and one participant had red-green color blindness. However, this did not bias our study since all tasks could be accomplished simply by perceiving shapes and monochromatic colors. All participants were members of the university but none of them had a background in computer science or were affiliated with the computer science department (18 students, 2 PhD students, 2 apprentices, 2 administrative staff).

Task Design

We used a different object pool with numbers to introduce the tasks and two training pools with letters that were different from the object pools that were used in the actual tasks.

During the actual tasks, 22 objects were equally distributed across the virtual canvas in each pool. Each object had a size of 50 x 50 pixels. Since a within-subjects design was applied, we created two different object pools with different symbols and varying object locations to prevent learning effects across conditions (see Figure 2). The object pools were also systematically counterbalanced throughout the two navigation conditions. Both object pools were tested in a pre-test to guarantee that the pools would be easily distinguished from each other (to avoid confusion with a previous pool) and their objects (e.g., avoid having an airplane and a helicopter in the same pool since both are aircrafts).

During the *navigation* task, participants had to search for 8 objects in 8 blocks (= 8 objects x 8 blocks = 64 trials). The other 14 objects served as distractors to increase the difficulty of the task. In total there were 3072 trials (24 participants x 2 conditions x 8 objects x 8 blocks).

Task 1: Navigation Task

The first task focused on the navigation performance in a ZUI. It compared navigation cost and navigation time of the two navigation techniques: multi-touch and egocentric. In the navigation task, the participants had to pan and zoom to the center of the overview display, which is referred to as 'home position' in the following. The home position was visualized as a red rectangle. Figure 3 and Figure 4 illustrate the rectangle in the picture-in-picture at the top right.

Figure 5. A user navigating in the large canvas using the egocentric navigation. The purple rectangle on the overview display indicates the current position of the viewport on the large canvas.

After the tablet's view matched with the home position, it disappeared and a translucent object (see Figure 5) was presented in the center of the tablet's screen. Then, the participant had to find the corresponding on the large virtual canvas. Figure 5, for instance, illustrates the view on the tablet displaying the uncovered *flag* object on the top left and the requested and translucent object in center of the tablet's screen. A match was accepted as a sufficient if it matched

² ZOIL Framework – <http://zoil.codeplex.com/>

with the presented translucent object on the tablet within an offset of 10 pixels in position and 30 pixels in size. The tablet revealed memory cards content only if the scale factor of the view is > 4.0 or above while the objects on the large display permanently show the back of the memory card. In addition, the distance between the objects assured that only one uncovered symbol was visible on the tablet at a time, to avoid building a spatial memory based on relative relationships (e.g., object X is to the right of object Y). After matching the requested object, the home position appeared again and the participants had to match it again in order to continue with the next object. Each match was signaled by a compound audio feedback for both the homing action and the correctly identified object.

Task 2: Recall Task

The second and consecutive recall task requested that participants recall the positions of the objects in the previous navigation task. Therefore, the participants were provided with an empty canvas on the large display. Then, the first object of the sequence from the previous navigation task was presented in the center of the large display. Participants had to place the object at their recalled position based on their spatial memory. The object could be moved with help of the arrow keys on a keyboard. The keyboard was used to avoid any effects originating from the use of muscle memory. If they were satisfied with the current position, they confirmed it with the Enter key and the object disappeared. Thereafter, the next object appeared in the center of the screen and they had to continue the task until all eight objects were positioned. Each time a participant pressed an arrow key, the object moves by 25 pixels in the given direction. This helped to speed up the reproduction process instead of positioning objects pixel-wise. The original positions of the objects, however, can be matched exactly.

Figure 6. Calculation of the recall error between the recalled location and the original location. The objects are 50x50 pixels in size and the grid cells are 25x25 pixels.

Each x- and y-offset to the original object was considered an error. Figure 6 illustrates the grid and the calculation of the overall error that was measured as the Euclidian distance between the original location and the participants' recalled location.

Procedure

At the beginning, the participants were welcomed by the experimenter and asked to fill out a pre-test questionnaire. The questionnaire included questions regarding demographic data, their experiences of touch technology, and their handedness, which could be important for later video analysis. Then, all participants were introduced to the first navigation technique according to their group. They were allowed to practice before the actual tasks began. Through this, we wanted to avoid any learning effects based on the interaction technique and enable participants to familiarize themselves with the input method. For the practice training, each participant had to complete 4 objects with 4 repetitions ($= 4 \text{ objects} \times 4 \text{ blocks} = 16 \text{ trials}$). Afterwards, participants were asked if they felt comfortable with the navigation technique; if they concurred, the study began.

Results

During each trial, we logged the navigation time and path including the x- and y-position as well as the zoom factor s . Although navigation time was measured, we did not apply a Fitts' law [2] test because our research did not focus on the input device and its index of performance for navigation.

Both navigation techniques had different sampling rates. Therefore, we had to reduce the input data of both interaction techniques to a 10Hz sampling rate. On the sampled data, we applied a ZUI cost metric (see Figure 7) that is similar to the metric proposed by Jetter et al. [13].

$$c = H \cdot \Delta x + W \cdot \Delta y - \Delta x \Delta y + W \cdot H \cdot |\log \Delta s|$$

$$c = H \cdot \Delta x + W \cdot \Delta y + W \cdot H \cdot |\log \Delta s|$$

Equation 1. The original ZUI cost metric. [13] (top) Modified ZUI cost metric. (bottom)

Figure 7. Illustration of modified ZUI cost metric, which approximates the continuously updated pixels when the participants shifted the view.

The original ZUI cost metric (Equation 1, top) calculates the updated pixels between a view t_{n-1} and the view t_n discretely. Since horizontal and vertical panning happens in parallel (Figure 7, red rectangle), Jetter et al. included the term $-\Delta x \Delta y$. We removed the term $-\Delta x \Delta y$ in our modified cost metric (Equation 1, bottom) to achieve a better approximation of the continuous movement and shifting during user interaction (Figure 7, green triangle).

Results for Navigation Performance

The resulting navigation performance per trial using the ZUI cost metric was 5,509,852.07 pixels² (SD = 1,513,327.43) for the multi-touch condition and 2,910,584.94 pixels² (SD = 671,967.12) for the egocentric condition. Block 1 was removed in both conditions, since participants were navigating the objects for the first time and therefore their chosen navigation paths were entirely random. A 2x7x8 (navigation technique x block x object) analysis of variance (ANOVA) with repeated measures was performed on the travelled navigation paths. The analysis revealed a statistically better navigation performance for the egocentric navigation, $F_{1,23} = 117.03$, $p < .001$, $partial \eta^2 = .84$. The results show that the navigation with egocentric navigation is 47% more efficient than with multi-touch navigation; therefore, we can confirm our Hypothesis H1.

Figure 8. The navigation cost per trial based on the ZUI cost factor for Blocks 2 to 8.

Tan et al. [25] reported in their study that females may benefit more from kinesthetic cues provided by touch screen devices than males. Our results, however, did not show a significant difference for the navigation technique between the genders ($F_{1,22} = .342$, $p = .565$, $partial \eta^2 = .015$).

The design of the experiment can be regarded as successful since the analysis did not show a significant interaction between navigation technique and group order ($F_{3,20} = .683$, $p = .573$, $partial \eta^2 = .093$). The mean navigation performance for Blocks 2 to 8 are plotted in Figure 8. As mentioned above, Block 1 was omitted in both conditions since participants were unfamiliar with the virtual canvas and object locations at the beginning. It illustrates a learning effect for the virtual canvas in both conditions. However, the navigation paths for egocentric navigation are shorter in an earlier stage compared to the multi-touch navigation.

Results for Navigation Time

The navigation time was recorded starting when the participants matched the red rectangle and ending when they matched the target object (see Figure 1). The mean navigation time per trial for the multi-touch condition was 10,460.82 ms (SD = 4,102.97) and the mean for the egocentric condition was 6,868.94 ms (SD = 2,260.19). A 2x7x8 ANOVA (navigation technique x block x object) with repeated measures was performed on the collected data. The analysis revealed that the technique had a statistically significant effect on the navigation time, $F_{1,23} = 13.96$, $p < .05$, $partial \eta^2 = .38$. Again, there was no interaction between the navigation technique and gender ($F_{1,22} = .102$, $p = .752$, $partial \eta^2 = .005$). Also, the analysis of variance did not show a significant interaction between navigation technique and group order ($F_{3,20} = .1853$, $p = .170$, $partial \eta^2 = .218$).

Figure 9. The mean navigation time per trial for Block 2 to Block 8.

As shown in Figure 9, the mean navigation time in the multi-touch condition improved quickly with every block. To ensure that the significant effect of navigation time was also present in the final block, we also performed a 2x8 ANOVA on Block 8 only. This still revealed that the navigation technique had a statistically significant effect on the navigation time, $F_{1,23} = 5.72$, $p < .05$, $partial \eta^2 = .199$.

The mean navigation time of both conditions is illustrated in Figure 9. Again, Block 1 was omitted in both conditions since participants were unfamiliar with the virtual canvas. Remarkably, all participants learned object locations and identities as the task progressed. The results show that egocentric navigation is 34% faster than in the multi-touch condition.

Results for Spatial Memory / Recall

The comparison of the results of the recall task with multi-touch and egocentric navigation did not show significant differences ($F_{1,23} = .120$, $p = .732$, $partial \eta^2 = .005$).

Participants performed equally well in multi-touch and egocentric conditions when they had to recall the locations and identities of object they had to search for in a previous task. The mean error in grid units was 5.69 (SD = 2.63) for the multi-touch condition and 5.39 (SD = 2.97) for the egocentric condition. In our analysis, we also tested the serial position effect [20] (primacy and recency effect) and found no significant differences between the first and the last objects and middle objects 2 to 7.

NASA TLX

Participants rated their subjective workload at the end of each condition. For this, we used the standardized NASA TLX. We did not find any significant difference in the overall subjective workload when comparing the navigation techniques. However, if we look closely into the subscales and compare them pairwise, the two subscales *mental demand* and *physical demand* denote significant differences. All calculated statistical significance values were based on a two-tailed paired-samples *t*-test.

Figure 10. The subjective workload assessment was measured using the NASA Task Load Index. Pairwise comparison showed significant differences for mental demand and physical demand.

On average, participants reported a significantly greater mental demand for multi-touch navigation ($M = 53.75$, $SE = 4.67$) than for egocentric navigation ($M = 40.83$, $SE = 4.17$), $t(23) = 2.262$, $p < .05$, $r = .43$. Probably, this is reasonable because users needed to transform recalled spatial information into 2D touch gestures for the multi-touch navigation whereas in the egocentric navigation users applied knowledge used in everyday life.

Unsurprisingly, on average, participants reported significantly less physical demand for multi-touch navigation ($M = 36.25$, $SE = 5.267$) than for egocentric navigation ($M = 55.42$, $SE = 4.94$), $t(23) = -3.960$, $p < .05$, $r = .64$. Of course, participants had to move around and hold

the tablet in laborious positions in the egocentric navigation, while in the multi-touch condition, the participants were seated and could rest their arms on the table.

Thus, even though the pairwise comparison of subscales did reveal significant differences, we can reject Hypothesis H3 that the overall subjective workload of users does not significantly differ between egocentric body movement and multi-touch.

Results of Subjective Preference

Nineteen out of 24 participants preferred the egocentric navigation over the multi-touch navigation. The participants that preferred multi-touch reported very frequent usage of touch devices in the pre-test questionnaire (e.g., smartphones). In addition, two participants who do not have any multi-touch devices reported that egocentric navigation “feels more natural” to them.

Discussion

The key finding of our experiment was that egocentric navigation performs significantly better in terms of travelled path length and time. Apparently, in the egocentric condition, participants naturally applied a navigation style with combined zooming and panning operations that resulted in more efficient navigation paths. We observed how users navigated between two targets by moving the device back or forth and simultaneously moving laterally in front of the large display. This resulted in navigation paths that are close to the “optimal” paths in a ZUI with exponential savings [13]. In the multi-touch condition, however, users made use of this combined movement less often and more often alternated between pure zooming or panning operations. This resonates with the findings of Jetter et al. [13] that observed the same for the multi-touch condition in their study of navigation performance in a ZUI on a tabletop.

Since the results showed significantly better navigation performance with egocentric navigation, but no significant differences in spatial memory performance, we cannot attribute the improved navigation to better spatial memory. This still could mean that there was an additional memory source applied during the spatial interaction tasks, e.g., motor memory. However, this assumption has not been verified with our experiment and remains an open research question.

In addition, a later review of the recorded video exposed an interesting observation: Participants, especially those with less experience with multi-touch devices, approached incorrect targets repeatedly within the same trial in the multi-touch condition. This happened even though participants were provided with the current view on the large display (see Figure 1, purple rectangle). We assume that this was the case because users in the multi-touch condition tended to lose their global orientation, i.e., the awareness of where their current view on the tablet is located within the canvas. Users in the egocentric condition, however, used their physical position in front of the large display as a spatial cue to maintain this global orientation.

This resonates with the results of the NASA TLX questionnaire, in which egocentric navigation was rated as significantly less mentally demanding than multi-touch; thus, it leaves more cognitive resources (e.g., resources that were used for system-level tasks such as zooming and panning) for application-level tasks. This benefit also shows in the willingness of the users to accept the significantly greater physical demand (53%), since 19 out of 24 participants preferred the egocentric navigation technique over multi-touch for the given task.

EXPERIMENT E2

In a second experiment E2, we wanted to study if the greater proprioceptive and kinesthetic feedback of body interaction would lead to stronger encoding as well as fixation of object location and identity in users' spatial memory as formulated in our Hypothesis H4.

The experiment measures the effect of the two navigation techniques egocentric vs. multi-touch on users' long-term memory. We randomly selected eight participants from the first experiment E1 (two participants from each of the four groups: mAeB, mBeA, eAmB, and eBmA). At the end of Experiment E1, these participants were asked to take part in an additional short experiment. The eight participants (3 female) were aged between 19 to 30 years ($M = 22.75$, $SD = 3.20$). Participants were distracted for 15 minutes showing them a game on a tabletop. This was to ensure that participants did not think or talk about the previous experiment. Participants were then asked to recall the objects from the endmost navigation technique after the 15 minutes of distraction. The recall task was the same as in Experiment E1. Again, the error was calculated based on the original location of the object and the recalled location (Figure 6).

Results

The mean error for the multi-touch condition was 6.26 ($SD = 1.09$, $SE = .54$) and the mean error for the egocentric condition was 4.56 ($SD = .32$, $SE = .16$). A one-way ANOVA revealed significant differences for the navigation technique, $F_{1,6} = 8.979$, $p < 0.05$, $\omega = .71$, indicating that egocentric body movements increase long-term spatial memory (see [26]). The spatial memory indicates an improvement of 27% in favor of the egocentric navigation.

Discussion

Although the results may show a significant improvement for long-term spatial memory in egocentric navigation, it is too early to generalize this for the entire user population because of the small sample size ($n = 8$). However, this second experiment is a good starting point for further investigation of the effect of egocentric body movements on users' long-term spatial memory and to examine whether the underlying mechanisms differ from those for spatial memory from E1.

CONCLUSION

In this paper, we presented two experiments that research the effect of egocentric body movements on users' navigation performance, spatial memory, and long-term spatial memory.

The results of Experiment E1 showed a significantly better navigation performance in a ZUI for egocentric navigation in terms of path length (47%) and task time (34%). We provided possible explanations based on a more frequent use of combined zooming and panning operations that resulted in more efficient navigation paths. Participants also reported a significantly lower mental demand which resonates with our observation that in the egocentric condition, they used their physical position in front of the large display as a spatial cue to maintain their global orientation.

Despite the fact that the results show a significant difference in navigation performance, we could not observe any statistically significant differences for the spatial memory. However, Experiment E2 hinted at a potential effect of the greater proprioceptive and kinesthetic feedback on long-term spatial memory, which we will investigate in a follow-up study. In future work, we want to investigate whether the implicitly better spatial memory during the navigation task for the egocentric navigation and not significantly different spatial memory in the explicit recall task is due to muscle memory. Moreover, we want to investigate how egocentric navigation performs in collaborative environments, for example, multi-display environments for mixed-focus collaboration [26]. In doing so, we want to learn about the advantages or disadvantages of egocentric and proxemic interactions in such collaborative environments, such as the problems that occur when multiple users concurrently manipulate and navigate a shared space.

Admittedly, our study has some limitations that could be addressed in future studies. In E1, all participants chose a horizontal orientation of the mobile device, which probably could have some effect on their mental demand because of the mental adjustments of the different orientations (vertical wall-sized display and horizontal mobile device). This is something that could be explored in future work.

Although the experiments were designed to study more abstract and generic spatial tasks instead of focusing on a specific real-world application, we believe that these results can be applied to various application domains. For example, we recommend using egocentric navigation in those contexts that require the user to invest great cognitive resources for application-level tasks, e.g., during time critical decision-making or epistemic processes. However, this must be balanced with the greater physical demand which prohibits the use of egocentric navigation when navigation operations are executed very frequently over a longer period of time and thus result in physical strain and fatigue.

ACKNOWLEDGMENTS

This work was partially supported by DFG Research Training Group GK-1042 and by the Ministry for Science, Research, and Art Baden-Württemberg under the project Blended Library. We would like to thank Stephan Huber for the implementation and support throughout the study. Further, we would like to thank Nicolai Marquardt for providing us with the source code of the Proximity Toolkit.

REFERENCES

1. Andrews, C., Endert, A., and North, C. Space to Think: Large, High-Resolution Displays for Sensemaking. *In Proc. of CHI '10*, ACM Press (2010), 55–64.
2. Balakrishnan, R. “Beating” Fitts’ law: virtual enhancements for pointing facilitation. *International Journal of Human-Computer Studies* 61, 6 (2004), 857–874.
3. Ball, R., North, C., and Bowman, D.A. Move to Improve: Promoting Physical Navigation to Increase User Performance with Large Displays. *In Proc. of CHI '07*, ACM Press (2007), 191–200.
4. Ball, R. and North, C. The effects of peripheral vision and physical navigation on large scale visualization. *In Proc. of GI '08*, Canadian Information Processing Society (2008), 9–16.
5. Bederson, B.B. The promise of zoomable user interfaces. *Behaviour & Information Technology* 30, 6 (2011), 853–866.
6. Bezerianos, A. and Isenberg, P. Perception of Visual Variables on Tiled Wall-Sized Displays for Information Visualization Applications. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (2012), 2516–2525.
7. Bier, E.A., Stone, M.C., Pier, K., Buxton, W., and DeRose, T.D. Toolglass and magic lenses: the see-through interface. *In Proc. of SIGGRAPH '93*, ACM (1993), 73–80.
8. Fitzmaurice, G.W. Situated information spaces and spatially aware palmtop computers. *Communications of the ACM* 36, 7 (1993), 39–49.
9. Geyer, F., Pfeil, U., Höchtel, A., Budzinski, J., and Reiterer, H. Designing reality-based interfaces for creative group work. *In Proc. of C&C '11*, ACM Press (2011), 165–174.
10. Greenberg, S., Marquardt, N., Ballendat, T., Diaz-Marino, R., and Wang, M. Proxemic Interactions: The New Ubicomp? *interactions* 18, January (2011), 42–50.
11. Jakobsen, M. and Hornbæk, K. Proximity and physical navigation in collaborative work with a multi-touch wall-display. *In Proc. of CHI EA '12*, (2012), 2519–2524.
12. Jetter, H.-C., Gerken, J., Zöllner, M., Reiterer, H., and Milic-Frayling, N. Materializing the Query with Facet-Streams – A Hybrid Surface for Collaborative Search on Tabletops. *In Proc. of CHI '11*, ACM (2011), 3013–3022.
13. Jetter, H.-C., Leifert, S., Gerken, J., Schubert, S., and Reiterer, H. Does (multi-)touch aid users’ spatial memory and navigation in ‘panning’ and in ‘zooming & panning’ UIs? *In Proc. of AVI '12*, ACM (2012), 83.
14. Jetter, H.-C., Zöllner, M., Gerken, J., and Reiterer, H. Design and Implementation of Post-WIMP Distributed User Interfaces with ZOIL. *International Journal of Human-Computer Interaction IJHCI (Special Issue on Distributed User Interfaces)*, (2012).
15. Jul, S. and Furnas, G.W. Critical zones in desert fog: aids to multiscale navigation. *In Proc. of UIST '98*, (1998), 97–106.
16. Kaufmann, B. and Ahlström, D. Studying spatial memory and map navigation performance on projector phones with peephole interaction. *In Proc. of CHI '13*, ACM Press (2013), 3173–3176.
17. Kidd, A. The marks are on the knowledge worker. *In Proc. of CHI '94*, ACM Press (1994), 186–191.
18. Kirsh, D. Thinking with external representations. *AI & Society* 25, 4 (2010), 441–454.
19. Marquardt, N., Diaz-Marino, R., Boring, S., and Greenberg, S. The proximity toolkit: prototyping proxemic interactions in ubiquitous computing ecologies. *In Proc. of UIST '11*, ACM (2011), 315–326.
20. Murdock Jr., B.B. The serial position effect of free recall. *Journal of Experimental Psychology* 64, 5 (1962), 482–488.
21. Robertson, G., Czerwinski, M., Larson, K., Robbins, D.C., Thiel, D., and van Dantzich, M. Data Mountain: Using Spatial Memory for Document Management. *In Proc. of UIST '98*, ACM Press (1998), 153–162.
22. Spindler, M., Martsch, M., and Dachsel, R. Going Beyond the Surface: Studying Multi-Layer Interaction Above the Tabletop. *In Proc. CHI '12*, ACM Press (2012), 1277–1286.
23. Spindler, M., Stellmach, S., and Dachsel, R. PaperLens: Advanced Magic Lens Interaction Above the Tabletop. *In Proc. of ITS '09*, ACM Press (2009), 69–76.
24. Tan, D.S., Gergle, D., Scupelli, P., and Pausch, R. Physically large displays improve performance on spatial tasks. *ACM Transactions on Computer-Human Interaction* 13, 1 (2006), 71–99.
25. Tan, D.S., Pausch, R., Stefanucci, J.K., and Proffitt, D.R. Kinesthetic cues aid spatial memory. *In Proc. of CHI EA '02*, ACM Press (2002), 806–807.
26. Tang, A., Tory, M., Po, B., Neumann, P., and Carpendale, S. Collaborative coupling over tabletop displays. *In Proc. of CHI '06*, ACM Press (2006), 1181–1190.
27. Yee, K. Peephole displays: pen interaction on spatially aware handheld computers. *In Proc. of CHI '03*, ACM Press (2003), 1–8.