Memory in Motion: The Influence of Gesture- and Touch-Based Input Modalities on Spatial Memory

Johannes Zagermann, Ulrike Pfeil, Daniel Fink, Philipp von Bauer, and Harald Reiterer
HCI Group, University of Konstanz,
{johannes.zagermann,ulrike.pfeil,daniel.immanuel.fink,philipp.bauer,harald.reiterer}@uni-konstanz.de

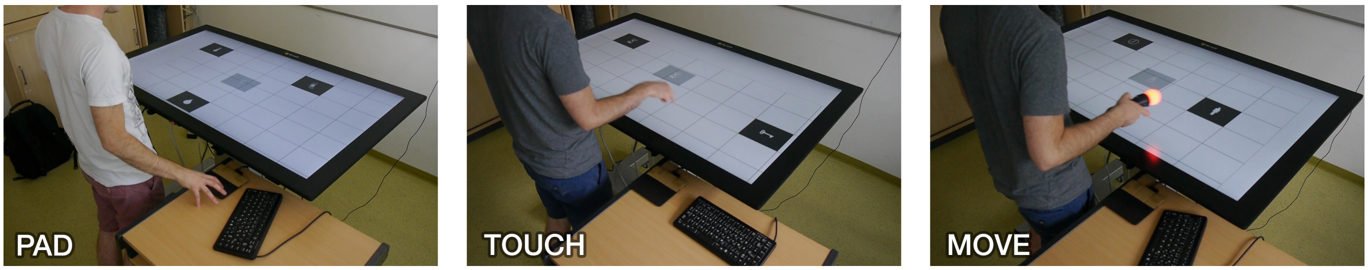

Figure 1. Our three different input modalities PAD, TOUCH, and MOVE used on LARGE

ABSTRACT
People's ability to remember and recall spatial information can be harnessed to improve navigation and search performances in interactive systems. In this paper, we investigate how display size and input modality influence spatial memory, especially in relation to efficiency and user satisfaction. Based on an experiment with 28 participants, we analyze the effect of three input modalities (trackpad, direct touch, and gesture-based motion controller) and two display sizes (10.6″ and 55″) on people's ability to navigate to spatially spread items and recall their positions. Our findings show that the impact of input modality and display size on spatial memory is not straightforward, but characterized by trade-offs between spatial memory, efficiency, and user satisfaction.

ACM Classification Keywords
H.5.2. User Interfaces: Input devices and strategies (e.g., mouse, touchscreen). Evaluation/methodology

Author Keywords
spatial memory; input modality; output device

INTRODUCTION
Spatial memory is a crucial aspect for successful interaction design as it takes advantage of users' ability to memorize object locations to ease navigation and search processes. Interfaces that harness users' spatial memory can profoundly reduce cognitive and physical load, support fluid interaction, and free resources to allow for better task performance [24]. Spatial memory is a fundamental component of HCI design across a wide span of everyday interactive systems. For example, it can be utilized in desktop applications to support menu navigation and icon search [21, 24]. In particular, spatial memory is an important aspect for interfaces consisting of a large digital landscape that extends the visible display space (e.g. pan and zoom interfaces in public displays, overview + detail interfaces in navigation systems). In regular use, spatial memory eventually reduces visual search leading to automatic navigation and freeing cognitive resources for the task [24].

Previous research has shown that the display size and the input modality impact on users' spatial memory and navigation performance [11, 20, 23, 24, 27, 28]. Yet, these studies investigated in- and output modalities separately, making it difficult to apply the results to real world scenarios. We believe that we need to consider the interplay of different in- and output modalities to provide valuable insights on spatial memory for everyday interactive systems (e.g. desktop applications, navigation systems, or public displays). In our work, we thus focus on the combination of in- and output modality that map current off-the-shelf interactive systems (e.g. tablets, laptops, and tabletops supporting input via trackpad, direct touch as well as gesture). Also, previous work suggests that different aspects of embodiment (e.g. touch vs. mouse [11, 28], body vs. touch panning [13], body movement [22], direct vs. indirect touch [20]) positively affect spatial memory performance. However, existing studies mostly compare input modalities that differ profoundly in their interaction techniques, not explaining how the underlying aspects of embodiment influence spatial memory. Applying the findings to settings beyond the specific conditions used in these studies is difficult. Understanding how aspects of embodiment (e.g. haptic feedback, kinesthetic cues, or direct interaction) influence spatial memory can help system designers to apply the findings to the in- and output modalities of their specific setting. Furthermore, previous research has focused on navigation path length and object location when investigating spatial memory. However, for designers to be useful, it is important to understand how spatial memory relates to aspects of usability. Looking at the definition of usability, which is divided into the aspects effectiveness, efficiency, and user satisfaction [7], we argue

CHI 2017, May 06–11, 2017, Denver, CO, USA
©2017 ACM. ISBN 978-1-4503-4655-9/17/05...$15.00
DOI: http://dx.doi.org/10.1145/3025453.3026001

that current methods can be mapped to effectiveness (e.g. how well users perform a task). However, the relations of spatial memory to efficiency and user satisfaction are unclear. In real-life scenarios, spatial memory is not detached from usability, but designers have to gauge these aspects depending on the context and the task at hand (e.g. prefer efficiency in desktop application or user satisfaction in leisure activities). In these cases, relating spatial memory to aspects of usability facilitates the incorporation of spatial memory into system design.

Our work contributes a detailed understanding of how the combination of different in- and output modalities impact on spatial memory. We also investigate the role of embodiment (e.g. kinesthetic cues, haptic feedback, direct interaction) in these effects. Analysing the influence of these aspects of embodiment not only on spatial memory but also on efficiency and user satisfaction, we tackle the question to what extent these measurements lead to coherent results. Our multi-faceted analysis accounts for complex combinations of in- and output modalities. Thus, results can be mapped onto real-world scenarios and implications can explain user behavior and support navigation and search processes in interactive systems.

RELATED WORK

In this section, we first define our understanding of spatial memory, especially in the context of HCI. Further, we summarize related studies regarding the impact of in- and output modalities on spatial memory. We explain how existing studies informed our work and led us to our research questions.

Spatial Memory

Spatial memory is described as the aspect of human cognition that enables us to acquire spatial information, for example the location of objects and places and the way to navigate to them. Once this spatial information is successfully learned and stored, spatial memory can be harnessed in order to perform navigation and search tasks with low cognitive effort [1, 24].

The prospects of a more fluid interaction when spatial memory is harnessed can also be applied to HCI. Here, spatial memory can support a user's interaction with a system by easing the cognitive effort required for navigating and manipulating an information space. Giving an extensive review of spatial memory in HCI research, Scarr et al. [24] distinguish between the components of navigation and object location memory. Interfaces that support spatial memory make use of people's ability to memorize object locations and navigation paths spatially which allows them to increase their interaction performance.

Influence of Input Modality

We identified previous studies investigating the impact of input modality on users' spatial memory. Tan et al. [28] focused on the influence of touch vs. mouse interaction on users' spatial memory. They asked 28 participants to perform a memorization task on a 18.1″ screen with an invisible grid and concluded with a 19% increased memorization accuracy for touch input. Building on this work, Jetter et al. [11] investigated the effect of multi-touch vs. mouse input on participants' navigation performance and object location memory in *panning* and *panning & zooming* interfaces. Results show that multi-touch results in better spatial memory and navigation performance for *panning* interfaces but not for *panning & zooming* interfaces. Palleis and Hussmann [20] distinguished between the influence of kinesthetic cues and direct interaction on spatial memory and navigation performance. Results show that neither of these aspects influences spatial memory, but a smaller touch input surface in the indirect touch condition resulted in a better navigation performance.

Focusing on whole body movements, Klinkhammer et al. [13] compared users' spatial memory for panning using single touch vs. body panning on a horizontal user interface. Although body panning did not result in an increased navigation performance, participants performed better with regard to object location memory. Similarly, Rädle et al. [22] investigated the effect of egocentric movements in peephole navigation of zoomable peephole interfaces, concluding with a positive effect of egocentric body movements on spatial memory.

The above mentioned studies indicate that users' spatial memory can be positively influenced by a suitable input modality, however, results on the characteristics of the optimal input modality are contradicting. Conditions that used touch performed better than mouse [11, 28]. However, when replacing the mouse with an indirect touch modality, findings show no difference in spatial memory and an improvement for navigation performance for the indirect interaction on a small input surface [20]. Conditions including whole body movement (e.g. body panning, egocentric body movements) performed better than those with less movement [13]. Our goal is to build on this existing research to further understand the influence of different aspects of embodiment (kinesthetic cues, haptic feedback, direct interaction) on spatial memory including not only touch-based but also gesture-based input modalities.

Influence of Output Modality

Previous research investigated how different visualization techniques influence spatial memory. For example, fisheye lenses [25], peephole navigation [12, 16, 23] and systems that allow for an overview [8] or focus and context view [3, 19] have been analysed to further understand how these techniques impact on navigation and object relocation. Visualizations like grids and structural alignments of items on a landscape have been found to positively effect spatial memory [15, 24].

Analyzing the influence of screen properties on spatial memory has shown that large displays lead to a greater sense of presence and a higher level of immersion which positively influences users' spatial memory performance [26, 29]. For example, Tan et al. [26, 27] showed that physically large displays increase performance on spatial tasks such as 3D navigation, mental map formation and memory tasks. Further, they found that the benefit of using a large display strongly depends on the kind of tasks performed on it: The benefit of large screens was found to be restricted to tasks in which users have to take on an egocentric strategy [26]. However, this finding could not be replicated by Tyndiuk et al. [29] who found that large displays are beneficial for both, ego- and exocentric tasks.

Tan et al. [27] also investigated users' spatial memory performance when navigating a virtual environment. Results show that users perform better in mental map formation and memory tasks on a large display because they feel more immersed into the environment compared to a small screen. Tyndiuk et al. [29] investigated the influence of display size on spatial memory and showed that not all users benefit from the use of large displays as only users with high visual selective attention abilities performed better on large displays. Rädle et al. [23] focused on peephole navigation and analyzed the effect of peephole size on users' map navigation behavior, navigation performance, and task load. They confirmed a sweet spot of a tablet-sized peephole where learning speed, navigation speed, and task load were improved.

We believe that the contradicting findings concerning the influence of display size on spatial memory performance might be due to the different input modalities that participants of previous studies worked with. Thus, we analyze the influence of display size in relation to the input modality to further understand their combined influence on spatial memory.

RESEARCH QUESTIONS

The goal of our study was to better understand how aspects of embodiment influence spatial memory. In addition, we were interested in the relation of these measures to the efficiency of task completion and user satisfaction. To this end, we defined input modality and display size as independent variables. In order to address our dependent variables, we focused on spatial memory, efficiency, and user satisfaction resulting in the following three research questions:

RQ1: How do different levels of embodiment influence spatial memory? (measured in the navigation performance and object location memory performance)

RQ2: How do different levels of embodiment influence the efficiency of task completion? (measured in navigation speed and task completion time)

RQ3: How do different levels of embodiment influence the user satisfaction? (measured in perceived task load and user experience)

Concerning the input modality we defined the three conditions (1) **PAD** (using a trackpad), (2) **TOUCH** (using direct single touch), and (3) **MOVE** (using a gesture-based motion controller). We chose these three conditions as they constitute three types of state-of-the-art input modalities: Trackpads are common features of laptops and can replace the use of a mouse especially to provide mobility. Touch technology can be seen as the main input modality for mobile devices, tabletops, and public displays. As the number of consumer devices that allow for gesture-based interaction either using additional sensing hardware like controllers (e.g. Sony PlayStation 3 Move motion controller or Nintendo's Wii Remote) or based on computational image processing (e.g. Microsoft Kinect or laptops with built-in Intel RealSense cameras) increases, we chose gesture-based interaction as a third input modality. In addition to off-the-shelf devices that allow for the individual use of the named input modalities, there is also a growing market of devices like e.g. laptops that enable all three input modalities to be used in sequence or even in combination, bringing up the question when to use which of them. We tested several arrangements and carefully considered comparable conditions to make sure that participants can handle the conditions equally well to ensure comparability and prevent problems with target acquisition. Regarding **MOVE**, we decided for an additional sensing hardware as the controller showed a higher accuracy and robustness for panning operations compared to the computational detection of a hand's position. This was also the case when dealing with arbitrary and task-dependent movements like panning or clutching. The activation of the panning-mode was more robust using the trigger of the controller compared to the detection of an additional finger gesture.

	PAD	TOUCH	MOVE
Kinesthetic Cues	–	×	×
Direct Interaction	–	×	–
Haptic Feedback	×	×	–
Activation	Touch	Touch	Trigger

Table 1. Comparison of the three different input modalities

Table 1 shows an overview of the characteristics of the three different input modalities with respect to embodiment. Our choice of input modalities allows to compare them on the basis of four aspects: (1) *kinesthetic cues* – an awareness of body's (or parts of it) position with respect to itself or to the environment [28]. For our study, we describe *kinesthetic cues* as the extent to which the user's arm movement corresponds to the actual path length, (2) *direct interaction* – here we distinguish between direct panning manipulations (TOUCH) and indirect panning manipulations using a mediator (PAD and MOVE), (3) *haptic feedback* – we consider this as direct feedback to panning a specific surface, and (4) *activation* – switching the mode between arbitrary and task-oriented movements.

The independent variable *input modality* is a within-subjects variable allowing for inner-subjects ratings of comparisons and preferences.

Concerning the output modality, we define the conditions (1) **SMALL** (display size: 10.6″, a typical size for a personal device) and (2) **LARGE** (display size: 55″, a typical size for a tabletop/public display) to account for two different scales to exhaust the respective level of embodiment. This independent variable is a between-subject variable. We expect that especially the positive influence of kinesthetic cues will result in greater spatial memory performance on the LARGE condition compared to the SMALL condition.

EXPERIMENT

This section provides an overview of the employed methods, including the apparatus, the task, participants and the procedure that we followed in our experiment.

Apparatus

Figure 1 shows the setting of our study with LARGE and the three input modalities PAD, TOUCH, and MOVE. We chose two differently sized output devices: a 10.6″ Microsoft Surface 2 Pro and a 55″ Microsoft Perceptive Pixel. Both devices had the same display resolution of 1920×1080 pixels to guarantee for a higher internal validity of the study. In the condition with the 10.6″ display, the device was placed on top and in the center of the 55″ display as our pre-test showed that it eased interaction with MOVE and TOUCH (people stepped back when placing the tablet closer). Also, this placement ensured that participants' postures were comparable for both display sizes. Next to the display, we placed a small mobile desk, which could be easily reconfigured by participants depending on their handedness. This desk was used to fill out questionnaires and to place the input devices on.

In the **TOUCH** condition, participants used the output device specific touch capabilities. For our study, we asked participants to use single touch only to ensure comparability between the three conditions. We used a Logitech T650 trackpad for the **PAD** condition, adjusting it in three ways: from a hardware perspective, we limited the area where participants could interact with it to a 16:9 ratio. This allowed us to have the same ratio of participants' movements in all conditions as the displays have the same ratio. We did this by placing an overlay on top of the trackpad itself. From a software perspective, we changed the control-display-ratio. Panning from the outer-most left position of the trackpad to its outer-most right position was equivalent to the same panning operation using TOUCH resulting in a 1:1 mapping with the output device. We adjusted the trackpad to work without additional pressing to have a comparable haptic feedback with TOUCH. These three design decisions regarding the trackpad led to a better comparability between the three conditions.

For the **MOVE** condition, we used a Sony PlayStation 3 Move motion controller (weight: 277g). We attached a PlayStation Eye Camera in bird's-eye-view directly above the display and adjusted the distance to the display to allow for an interaction space corresponding to the size of the display, which allowed us to have a comparable movement range to the TOUCH condition. The camera detected the position of the visual marker of the controller, and participants were able to pan the canvas by pulling and holding a trigger at the backside of it.

Task

We used a well-established and validated task [11, 13, 15, 16, 17, 22] consisting of a navigation phase and a subsequent object location memory phase (see the video in supplemental materials). Using one of the input modalities, a canvas could be navigated using panning operations. The canvas itself contained a 24×18 grid with a spatial configuration of 18 items in total. In the default position, a 8×6 section of the grid was visible – containing no items.

In the navigation task, the system started at the default position in the center of the canvas where no items were located. In the middle of this default position, a semi-transparent overlay of the item to be searched was shown. Participants had to find this item on the canvas and pan to its position using the condition-specific panning possibilities. Items to be searched were auto-selected as participants panned to their position with a tolerance of 100px to their center. After the item was found, the task continued at the default position with the next item. For each input modality, participants had to find 8 out of 18 items. The same sequence was repeated 6 times (6 blocks), resulting in 48 search trials. In total, each participant had three of these 48 search trials (one per input modality), resulting in 144 trials per participant and 4032 trials over all sessions.

To avoid possible learning effects, we used several item sets: (1) a demo item set containing 12 letters of the alphabet, (2) an exercise item set containing 18 symbols, and (3) three task item sets containing 18 symbols each. All item sets were exclusive without overlaps. Additionally, the items were divided equally between the different item sets with respect to their complexity and theme (devices, animals, vehicles, etc.). Also, the position of the different items and thus the total optimal navigation path length was taken into consideration with a similar length in all item sets (16729px, 16125px, and 15956px) with a randomization factor of +-3% to prevent possible influencing factors like symmetry or memorization of specific areas.

In the second phase – the object location phase – participants had to locate the presented items on their specific position on the canvas. In this phase, the entire canvas was shown, with no need for panning interaction. To locate the items, participants used the arrow keys of a keyboard to avoid influences of the specific input modality (cf. [11]). After placing an item to its position, the specific item was invisible to avoid possible mental constructs. We provide the application with all item sets and a supplemental video online.[1]

Participants

28 participants (12 female) were recruited for the experiment. As a precondition, we only invited participants that had German as their mother tongue. By this, we avoided possible influences on the memorization of the items due to different linguistic representations of foreign languages similar to [11, 15]. The mean age was 24.57 years (SD = 3.62, aged 18-32). All participants were right-handed and had normal or corrected to normal eye-sight, consequently they had no problems with the employed size of items. 21 participants were undergraduate students, 4 PhD students and 3 university employees. Participants had a mixed background ranging from computer science, to psychology, biology, educational science, and law. All participants were familiar with touch interfaces and 18 participants reported to already have used gesture-based input modalities, for example Nintendo's Wii Remote. To recruit participants, we used postings and flyers looking for people who are interested in playing games like Memory/Pairs[2], which helped us to find participants who were motivated by our chosen task. We assigned 14 participants to each display size.

Procedure

At the beginning of the study, each participant was asked to fill out a demographic questionnaire including questions about tech-savviness and eye-sight. Then, each participant received an introduction into the navigation and object relocation phases with their first input modality and a demo item set. Participants were given time to explore the functionalities of the system and input modality with an exercise data set until they felt comfortable using it. After that, participants were asked to start the navigation phase with the first input modality with the remark to navigate to the required items as fast as possible. After that, participants were asked to fill in a NASA TLX [6] and User Experience Questionnaire (UEQ) [14]. Then, participants had to relocate the position of all items on the canvas as precisely as possible. This sequence was the same for all three input modalities. In order to avoid

[1] http://hci.uni-konstanz.de/trr161

[2] https://www.ravensburger.de/start/memory/index.html

learning effects, we counterbalanced the order of input devices and the assignment of item sets by using a random selection of 14 out of 36 possible combinations for each output modality. Finally, each participant filled out a post-questionnaire concerning subjective preferences. Each session lasted about one hour in total, and participants were compensated for their time.

ANALYSIS

We chose a 2×3 factor split-plot counterbalanced study design consisting of a between-subjects part with the display size as independent variable with two conditions (SMALL and LARGE) and a within-subjects part with the input modality as independent variable with three conditions (PAD, TOUCH, and MOVE)

Data collection

We employed several data collection methods in order to investigate our research questions:

RQ1 - Spatial Memory: We focused on the task-dependent aspect of spatial memory: Navigation performance and object location memory performance. For the navigation performance, we calculated the ratio between users' actual navigation path to the shortest possible path (excluding the first search block as results for this are influenced by the randomness of initial navigation trials, cf. [11]). For the object location memory, participants were asked to position items on an empty canvas at the correct position. Here, we logged the position and calculated the Euclidean distance to its actual position in pixels.

RQ2 - Efficiency: In order to analyze the efficiency of task completion, we logged the duration participants needed to find the specific items (e.g. the task completion time in seconds for each item) and the navigation speed (pixels per second) that participants operated at.

RQ3 - Satisfaction: To analyze the user satisfaction for the investigated conditions, we used three questionnaires in total: a NASA TLX [6] was used to measure the perceived task load. The UEQ [14] provided information on the experience using the system in the different conditions. Additionally, we asked participants at the end of the experiments to rank the three input modalities and justify their preferences.

RESULTS

In this section, we report our findings in relation to our three research questions. In particular, we focus on the differences across the three different input modalities (PAD, TOUCH, and MOVE) and output devices (SMALL and LARGE) with respect to participants' spatial memory (RQ1), their efficiency (RQ2), and their subjective satisfaction (RQ3). As the data violated the assumption of normal distribution, we decided to use a non-parametric approach to analyse the data statistically. For the overall tests on statistical significance concerning the three different input modalities, we used a Friedman test. If this test showed statistical significant differences, we employed a Wilcoxon signed-rank test as a post-hoc analysis including Bonferroni correction. Additionally, a Mann-Withney-U-test was used to test for statistical significant differences between the two different display sizes.

RQ1: Spatial Memory

For spatial memory, we differentiate between navigation path and object location memory.

Navigation Path Length

Mean values for *Navigation Path Length* are shown in Table 2. Results show a statistically significant difference in the navigation path length depending on the input modality, $\chi^2(2) = 6.929$, $p < .05$, irrespectively of the display size. However, a Wilcoxon signed-rank test showed no statistical significant effect between pairs after applying Bonferroni correction. A Friedman test showed no statistically significant differences on the levels of SMALL and LARGE for the three input modalities.

A Mann-Whitney-U-Test showed a statistical significant difference for the display size when comparing navigation path length for TOUCH ($U = 55.0$, $p < .05$) and MOVE ($U = 52.0$, $p < .05$).

Focusing on the navigation path across blocks, Figure 2 shows that participants performance increased over time with all input modalities. Whereas with PAD and MOVE, the navigation path seems to level off in block 5, this was found to occur earlier in TOUCH as here, the curve already bends in block 4. In addition, the lines indicating the decrease in mean navigation path length per block differ for TOUCH and MOVE when comparing SMALL and LARGE.

	PAD	TOUCH	MOVE
M_{TOTAL}	2.65 (1.10)	2.11 (0.85)	2.43 (0.97)
M_{SMALL}	2.74 (1.07)	2.43* (0.93)	2.69* (0.75)
M_{LARGE}	2.56 (1.17)	1.78* (0.64)	2.18* (1.13)

Table 2. Mean navigation path length (1.0 = optimal path length). Standard deviation (SD) is shown in brackets, statistically significant differences are indicated via asterisks (for between).

Object Location Memory

Mean values for *Object Location Memory* are shown in Table 3. Results show no statistically significant difference in the object location memory depending on the input modality ($\chi^2(2) = 3.429$, $p > .05$), irrespectively of the display size. However, there was a statistically significant difference in the object location memory for LARGE, ($\chi^2(2) = 7.429$, $p < .05$) – but no difference for SMALL. A Wilcoxon signed-rank test showed a statistically significant difference for a pairwise comparison of PAD and MOVE for LARGE ($Z = -2.919$, $p < .016$).

A Mann-Whitney-U-Test showed no statistically significant difference for the display size when comparing object location memory.

	PAD	TOUCH	MOVE
M_{TOTAL}	1175.12 (457.97)	1075.80 (352.63)	1059.50 (534.19)
M_{SMALL}	1107.45 (521.41)	1156.96 (399.52)	1250.64 (638.24)
M_{LARGE}	1242.79^g (392.26)	994.64 (290.61)	868.36^g (326.53)

Table 3. Mean Euclidian distance in pixels. SD is shown in brackets, statistically significant differences are indicated via raised letters (for within).

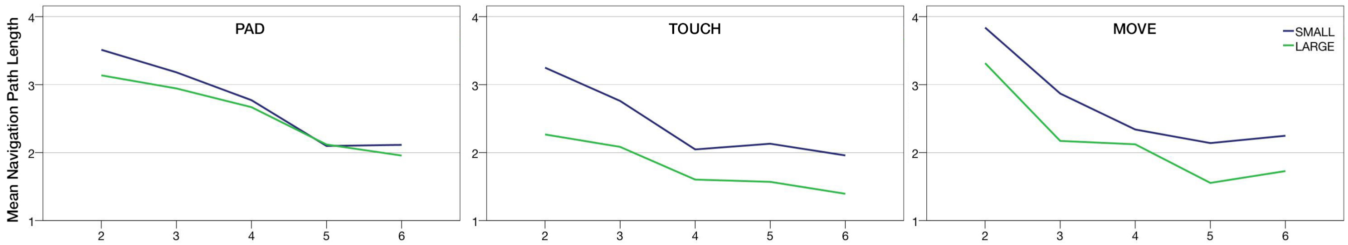

Figure 2. Mean Navigation Path Length for PAD, TOUCH, and MOVE per block. Optimal navigation path length: 1.0

RQ2: Efficiency

To provide a nuanced description of efficiency, we differentiate between navigation speed (in pixels per second) and average task completion time for each item search (in seconds).

Navigation Speed

Mean values for *Navigation Speed* are shown in Table 4. There was a statistically significant difference in the navigation speed depending on which type of input device was used, $\chi^2(2) = 30.071$, $p < .05$, irrespectively of the display size. A Wilcoxon signed-rank test showed that the navigation speed among all three input modalities was statistically significantly different: PAD-TOUCH: ($Z = -3.575$, $p < .016$), PAD-MOVE: ($Z = -4.486$, $p < .016$), TOUCH-MOVE: ($Z = -4.144$, $p < .016$).

Results show statistically significant differences in the navigation speed across the input modalities for SMALL ($\chi^2(2) = 9.000$, $p < .05$) and for LARGE ($\chi^2(2) = 24.571$, $p < .05$). A Wilcoxon signed-rank test showed that the navigation speed using SMALL was statistically significant different for pairwise comparisons of PAD and MOVE ($Z = -2.919$, $p < .016$) and for TOUCH and MOVE ($Z = -2.982$, $p < .016$). A Wilcoxon signed-rank test showed that the navigation speed on LARGE was statistically significant different for all three pairwise comparisons of input modalities: PAD-TOUCH: ($Z = -3.296$, $p < .016$), PAD-MOVE: ($Z = -3.296$, $p < .016$), TOUCH-MOVE: ($Z = -2.982$, $p < .016$).

A Mann-Whitney-U-test showed a statistical significant difference for the display size when comparing navigation speed for TOUCH ($U = 2.0$, $p < .05$) and MOVE ($U = 1.0$, $p < .05$).

	PAD	TOUCH	MOVE
M_{TOTAL}	$1421.96^{a,c}$ (456.62)	$1159.53^{b,c}$ (394.74)	$967.79^{a,b}$ (327.92)
M_{SMALL}	1499.03^{d} (581.62)	$1474.18^{e,*}$ (299.29)	$1224.72^{d,e,*}$ (266.30)
M_{LARGE}	$1344.89^{g,h}$ (286.30)	$844.89^{g,i,*}$ (144.30)	$710.87^{h,i,*}$ (101.20)

Table 4. Mean navigation speed (in pixels per second). SD is shown in brackets, statistically significant differences are indicated via asterisks (for between) or raised letters (for within).

Task Completion Time

Mean values for *Task Completion Time* for single items are shown in Table 5. Results show a statistically significant difference in the task completion time depending on the input modality, $\chi^2(2) = 22.071$, $p < .05$, irrespectively of the display size. A Wilcoxon signed-rank test showed that task completion time using MOVE was statistically significantly longer than using PAD ($Z = -4.099$, $p < .016$) and TOUCH ($Z = -2.755$, $p < .016$). There was no statistically significant difference between PAD and TOUCH.

Results also show a statistically significant difference in the task completion time depending on the input modality – for SMALL ($\chi^2(2) = 15.857$, $p < .05$) and for LARGE ($\chi^2(2) = 9.000$, $p < .05$). A Wilcoxon signed-rank test showed that the navigation time on SMALL was statistically significantly different for pairwise comparisons of PAD and MOVE ($Z = -2.480$, $p < .016$) and for TOUCH and MOVE ($Z = -2.480$, $p < .016$). A Wilcoxon signed-rank test showed that only the pairwise comparison for the task completion time using LARGE for PAD and MOVE are statistically significantly different ($Z = -3.233$, $p < .016$).

A Mann-Whitney-U-Test showed a statistically significant difference for the display size when comparing navigation time for TOUCH ($U = 55.0$, $p < .05$).

	PAD	TOUCH	MOVE
M_{TOTAL}	3.71^{a} (1.21)	3.90^{b} (1.54)	$5.31^{a,b}$ (2.17)
M_{SMALL}	3.56^{d} (0.89)	$3.37^{e,*}$ (1.14)	$4.48^{d,e}$ (1.02)
M_{LARGE}	3.87^{g} (1.48)	4.43^{*} (1.74)	6.13^{g} (2.70)

Table 5. Mean task completion time (in seconds). SD is shown in brackets, statistically significant differences are indicated via asterisks (for between) or raised letters (for within).

RQ3: User Satisfaction

For user satisfaction, we differentiate between participants' task load, their user experience and the subjective preferences of input modalities.

Task Load

Mean values for *Task Load* are shown in Table 6. Results show a statistically significant difference in the task load measured by the NASA TLX questionnaire depending on the input modality, $\chi^2(2) = 14.721$, $p < .05$, irrespectively of the display size. A Wilcoxon signed-rank test showed that the task load using MOVE was statistically significantly higher than for PAD ($Z = -3.952$, $p < .016$) and for TOUCH ($Z = -3.063$, $p < .016$). There was no statistically significant difference between PAD and TOUCH ($Z = -0.445$, $p > .016$).

Results also show a statistically significant difference when comparing the overall task load for the three different input modalities on the level of SMALL, $\chi^2(2) = 10.145$, $p < .05$. Here, a Wilcoxon signed-rank test revealed statistically significant differences for a pairwise comparison of PAD and MOVE ($Z = -3.077$, $p < .016$). However, no statistically significant difference was found on the level of LARGE.

A Mann-Whitney U-test revealed no statistically significant differences between the two display sizes.

	PAD	TOUCH	MOVE
M_{TOTAL}	31.15[a] (15.36)	32.13[b] (12.91)	44.96[a, b] (19.46)
M_{SMALL}	32.44[d] (15.76)	33.37 (9.35)	46.45[d] (19.88)
M_{LARGE}	29.87 (15.43)	30.88 (15.99)	43.46 (19.65)

Table 6. Mean user task load. SD is shown in brackets, statistically significant differences are indicated via raised letters (for within).

Analysis of mean values in the TLX subscales showed statistically significant differences for *physical demand* ($\chi^2(2) = 22.691$, $p < .05$), *effort* ($\chi^2(2) = 16.074$, $p < .05$), and *frustration* ($\chi^2(2) = 15.340$, $p < .05$). Table 7 shows the mean values for these subscales.

A Wilcoxon signed-rank test showed that the *physical demand* was statistically significantly higher for MOVE when compared to PAD ($Z = -4.251$, $p < .016$) and TOUCH ($Z = -3.354$, $p < .016$). The *effort* was also statistically significantly higher for MOVE when compared to PAD ($Z = -3.321$, $p < .016$) and TOUCH ($Z = -3.153$, $p < .016$). The level of *frustration* was statistically significantly higher for MOVE when compared to PAD ($Z = -2.909$, $p < .016$) and TOUCH ($Z = -2.791$, $p < .016$). Subscales that indicated statistically significantly different mean values are shown Table 7.

A Mann-Whitney-U-Test showed a statistically significant difference for display size when comparing the subscale *performance* ($U = 54.0$, $p < .05$). For MOVE, participants rated their performance better on LARGE with $M_{LARGE} = 33.57$, ($SD = 28.04$) and $M_{SMALL} = 49.86$, ($SD = 20.29$).

	PAD	TOUCH	MOVE
Physical Demand	19.89[a] (15.76)	31.57[b] (23.56)	53.50[a, b] (24.58)
Effort	33.54[a] (21.55)	35.07[b] (19.75)	50.54[a, b] (19.94)
Frustration	23.54[a] (20.88)	22.82[b] (17.38)	39.21[a, b] (25.41)

Table 7. TLX subscales. SD is shown in brackets, statistically significant differences are indicated via raised letters (for within).

User Experience

User Experience was evaluated based on the subscales of the UEQ [14]: *attractiveness*, *perspicuity*, *efficiency*, *dependability*, *stimulation*, and *novelty*.

Irrespectively of the display size, results show no statistically significant difference in *attractiveness* depending on the input modality, $\chi^2(2) = 3.448$, $p > .05$. However, we found statistically significant differences in *perspicuity* ($\chi^2(2) = 14.673$, $p < .05$), *efficiency* ($\chi^2(2) = 12.250$, $p < .05$.), *dependability* ($\chi^2(2) = 18.780$, $p < .05$.), *stimulation* ($\chi^2(2) = 6.928$, $p < .05$.), and *novelty* ($\chi^2(2) = 21.189$, $p < .05$.). Mean values for these subscales are shown in Table 8.

A Wilcoxon signed-rank test showed that the *perspicuity* was ranked better for PAD compared to MOVE ($Z = -2.916$, $p < .016$), and for TOUCH compared to MOVE ($Z = -3.078$, $p < .016$). Also, the *efficiency* was evaluated likewise with $Z = -3.623$, $p < .016$ for the comparison of PAD with MOVE and $Z = -3.180$, $p < .016$ for the comparison of TOUCH and MOVE. The *dependability* was ranked higher for PAD in comparison to MOVE ($Z = -3.746$, $p < .016$) and for TOUCH in comparison to MOVE ($Z = -3.993$, $p < .016$). The post-hoc analysis showed no statistically significant differences for the scale of *stimulation* after Bonferroni correction. The *novelty* was ranked highest for MOVE and lowest for PAD. Pairwise comparisons showed statistically significant differences between PAD and TOUCH ($Z = -2.502$, $p < .016$), PAD and MOVE ($Z = -3.695$, $p < .016$), and TOUCH and MOVE ($Z = -3.182$, $p < .016$).

Additionally, a Mann-Whitney U-test showed a statistically significant difference for display size when comparing *novelty* for PAD ($U = 48.5$, $p < .05$, with $M_{LARGE} = 0.57$, ($SD = 0.89$) and $M_{SMALL} = -0.30$, ($SD = 0.84$)), and TOUCH ($U = 47.5$, $p < .05$, with $M_{LARGE} = 1.00$, ($SD = 0.73$) and $M_{SMALL} = 0.09$, ($SD = 1.02$)). In these two conditions, LARGE led to a more novel impression of PAD and TOUCH.

	PAD	TOUCH	MOVE
Perspicuity	1.87[a] (1.16)	2.12[b] (0.78)	1.21[a, b] (1.21)
Efficiency	1.23[a] (0.92)	1.16[b] (0.88)	0.21[a, b] (1.04)
Dependability	1.65[a] (0.76)	1.81[b] (0.56)	0.69[a, b] (1.15)
Novelty	0.13[a, c] (1.27)	0.55[b, c] (0.99)	1.29[a, b] (1.12)

Table 8. Subscales of the UEQ. SD is shown in brackets, statistically significant differences are indicated via raised letters (for within). Semantic Differential: Min: −3, Max: 3.

Subjective Preferences

Concerning the subjective preferences of input modalities, we asked participants to rank them. For analysis, we gave input modalities scores based on their ranked position (e.g. 1 for the input device ranked best, 2 for the one ranked second, and 3 for the least preferred input modality). Table 9 shows the mean scores based on the given ranks.

	PAD	TOUCH	MOVE
M_{TOTAL}	1.71 (0.71)	1.68 (0.77)	2.61 (0.63)
M_{SMALL}	1.79 (0.70)	1.50 (0.76)	2.71 (0.47)
M_{LARGE}	1.64 (0.74)	1.86 (0.77)	2.50 (0.76)

Table 9. Subjective Preferences. SD is shown in brackets.

Results show that participants preferred PAD and TOUCH compared to the MOVE condition. PAD was rated best by 12 participants. Concerning the PAD, participants specifically preferred the "quick, efficient and easy interaction" (P14) and the possibility for "fast navigation" (P19). In addition, participants liked the fact that it required "less movement" (P10) than the other conditions. Fourteen participants rated TOUCH as their favourite input modality. Concerning TOUCH, participants specifically liked the "intuitive interaction" (P20) and the "direct manipulation" (P4) that allowed for a "good overview over the canvas" (P28) and a "direct way of navigating towards the items" (P11). Out of 28 participants, 19 ranked MOVE to be the least preferred input device. Reasons for that were mostly the "cumbersome" (P32) interaction as participants often felt that navigation was "jerky" (P9). In particular, the use of the trigger was perceived as "disruptive" (P15).

DISCUSSION

In our discussion, we address the three research questions (RQ1 - RQ3) regarding the aspects of spatial memory, efficiency, and user satisfaction.

Spatial Memory

Regarding the performance of spatial memory, we measured the navigation performance (ratio of actual navigation path and shortest possible path) and participants' object location memory (distances of the placed items to the correct location in the reconstruction task). Our results concerning the two measurements match, however they contradict the results regarding the efficiency of task completion.

Our results show that the display size has a positive impact on the navigation performance in the conditions TOUCH and MOVE. With these two input modalities, participants performed statistically significantly better in the LARGE condition. These findings are similar to previous research that a larger display positively influences spatial memory [27], however our findings show that this is only the case for interaction modalities that provide kinesthetic cues (TOUCH and MOVE), suggesting that this effect only occurs when arm movements are sufficiently large. We could also observe that TOUCH resulted in a steeper learning curve for navigation path length across the blocks, levelling off earlier than those for PAD and MOVE (see Figure 2). Our findings are in line with previous research that kinesthetic cues and direct interaction improve navigation performance in panning interfaces [11, 28].

Results regarding the object location memory go even further, showing more accuracy for MOVE. This effect was statistically significantly different between PAD and MOVE for LARGE. However, differences between TOUCH and PAD were not statistically significant. This finding resonates with [20] as spatial memory was not found to be influenced by direct interaction in touch-based interaction. Nonetheless, the statistically significant improvement for MOVE in LARGE suggests that kinesthetic cues in gesture-based interaction indeed positively impact on spatial memory.

In summary, we conclude that embodiment impacts on the performance of spatial memory, as kinesthetic cues in the LARGE condition lead to better navigation performance when working with TOUCH and MOVE. In addition, object location memory performance is increased when interacting with MOVE suggesting that – in gesture-based interaction – the involvement of kinesthetic cues does not only lead to a better navigation performance but also positively influences object location memory. This finding could not be replicated for touch-based interaction.

Efficiency

Regarding the efficiency, we were particularly interested in how the different aspects of embodiment impact on participants' navigation speed and task completion time. Although, we assumed that these two measurements should be correlated, our results showed unexpected differences regarding these two measurements between the different input modalities and display sizes.

As expected, PAD resulted in statistically significantly higher navigation speed compared to TOUCH and MOVE. We believe that this result is due to the 1:1 mapping of PAD to the display size. Thus, kinesthetic cues can hamper efficiency by forcing the user to engage in larger arm movements. Interestingly, there was also a statistically significant difference in navigation speed between TOUCH and MOVE, showing that TOUCH was statistically significantly faster. This finding is interesting as the difference in motor activity between these conditions is marginal. However, in addition to the arm movement, participants had to handle MOVE (e.g. pull the trigger) which might have caused them to move at a slower speed. In the TOUCH condition, there was no such device interference. In addition, MOVE required participants to lift their arms, which might have caused fatigue as another possible reason for slower navigation speed.

Our results show that display size positively influences spatial memory when using TOUCH and MOVE, as a smaller display increased the navigation speed for these modalities. This finding was expected as SMALL requires less motor activity. This was not the case for PAD which allowed for equally fast interaction across SMALL and LARGE.

The increase of navigation speed was only partially transferred into task completion time. Although interacting with PAD still resulted in faster task completion time, TOUCH reached similar results and the statistically significant difference found between these input modalities concerning navigation speed diminished when measuring task completion time. This is in line with previous research who did not find statistically significant differences in task completion time between direct and indirect touch input [20]. In our study, PAD and TOUCH resulted in statistically significant faster task completion times than MOVE. These results suggest that the haptic feedback in TOUCH allowed for similar task completion times by compensating for the slower navigation speed. We suspect that it might be a more accurate way of navigating to the item that allowed participants using TOUCH to reach similar task completion times than participants using the PAD. Alternatively, another interpretation might be that PAD allowed for navigation at such a fast speed that users were not able to keep up with. This resulted in the fact that the headstart in navigation speed could not be held concerning the task completion time.

MOVE resulted in statistically significant longer task completion times than both, PAD and TOUCH. This shows that participants were not able to apply a strategy that compensated for the slower navigation speed resulting in correlating results of slower navigation speed and slower task completion time for MOVE. We suspect that they had to focus on the interaction with the device when using MOVE, which negatively affected both, the navigation speed and task completion time.

In summary, we conclude that embodiment also affects the efficiency of task completion, as larger motor activities result in slower navigation speed. However, we also showed that parts of the slower navigation speed can be compensated for in TOUCH but not in MOVE. This suggests that haptic feedback allows participants to compensate for slower navigation speed resulting in faster task completion times.

User Satisfaction

Regarding the user satisfaction, we were particularly interested in users' subjective rating of their task load (measured with the NASA TLX [6]) and their user experience (measured with the UEQ [14]). Results of these two questionnaires show that irrespectively of the display size, participants rated MOVE differently compared to PAD and TOUCH.

Results of the NASA TLX show that there are statistically significant differences regarding the task load across the three input modalities. Post-hoc tests revealed that MOVE resulted in a statistically significantly higher task load. Looking at the subscales, MOVE was rated with a statistically significantly higher task load onto the user with respect to *physical demand*, *effort*, and *frustration*. Physically handling the device in addition to the kinesthetic movement might have caused the physical demand. We believe that the novel way of interaction required users to put more effort into using the device and problems during usage might have caused frustration. When looking at the results of the UEQ, we can observe similar patterns: Here, TOUCH and PAD scored statistically significant better in the subcategories *perspicuity*, *efficiency*, and *dependability*, reflecting the benefits of the approved and long-learned interaction modalities. As MOVE was also rated to be statistically significantly more novel than the other two modalities, we believe that it might be due to this novelty, that it was scored lower on the other scales of user experience. As expected, the display size did not profoundly impact on the user satisfaction, as the only difference was that PAD and TOUCH were rated statistically significant more novel in the LARGE condition.

In summary, we conclude that both TOUCH and PAD are rated equally well, both regarding the subjective task load, and the user experience. However, MOVE was perceived to acquire a higher task load and resulted in a lower user experience. Although this is in line with our findings regarding the efficiency of interaction, it contradicts our findings regarding spatial memory. Higher efficiency and user satisfaction for PAD and TOUCH was not reflected in the spatial memory performance.

LIMITATIONS

The accuracy and the general behavior of the different input modalities might have influenced participants when solving the task: The 1:1 mapping of PAD and thus the resulting possible navigation speed might have been unfamiliar to participants as most of them are used to classic control-display ratios of off-the-shelf laptops. Also, the novelty and unfamiliar experience using MOVE might have influenced participants' interaction behavior. The visual representation of items, which differed in their size regarding the two display sizes, can be seen as another possible influence. The perceived visual representation of items on LARGE was larger in contrast to SMALL, which might have biased the memorization of items.

IMPLICATIONS FOR DESIGN AND RESEARCH

Summarizing, our results show that embodiment indeed impacts on spatial memory, however not in a straightforward way. Our measurements regarding efficiency showed that larger kinesthetic cues (arm movements) resulted in slower navigation speed which also negatively influenced the overall task completion time. However, TOUCH (with its direct interaction and haptic feedback) mediated this effect by guiding towards a more accurate way of interaction. We name this phenomena the *Efficiency vs. Spatial Memory* trade-off. The longer it took participants to use the modality and complete the task, the more they engaged with the items and remembered them resulting in better spatial memory performance. In addition, our results also showed that embodiment as implemented in MOVE (gesture-interaction with additional device) was rated quite negatively with regard to user experience and task workload. However, it also had a positive impact on the performance of spatial memory (maybe due to the inconvenience of the interaction). We name this phenomena the *User Satisfaction vs. Spatial Memory* trade-off.

In the following sections, we discuss how the two trade-offs can be exploited to influence users' interaction patterns with regard to spatial memory. We discuss the implications of our results for practitioners and show how they can be used to inform design decisions.

Efficiency vs. Spatial Memory

This trade-off is especially reflected in the strategy employed when working with PAD. For PAD, participants made use of the increased interaction speed, resulting in fast interaction and short task completion times. However, this high speed came at a cost regarding spatial memory performance, as users seemed to be less accurate when navigating towards an item and performed less accurate in object relocation. Regarding navigation performance, Figure 3 (participants had to navigate to the item at the left) exemplifies that participants using PAD tended to overshoot and diverted from the optimal path towards the end (when they could already see the item). This suggests that with PAD, participants were less careful to navigate accurately. Contrary to that, participants working with TOUCH and MOVE interacted in a slower pace, however being more careful in their navigation precision (compare Figure 3 in the middle and on the right). Working with TOUCH, this precision led to surprisingly fast task completion times. Working with MOVE, it resulted in high object location memory.

Regarding the display size, participants working with TOUCH and MOVE benefit from larger kinesthetic movements and thus a larger display as it improves their navigation performance, however no statistically significant differences between the two display sizes could be found regarding the object location memory. Using PAD, no benefits could be gained when working on a larger display.

For designers, knowing about the trade-off between efficiency and spatial memory can bear important implications: Depending on the task and context, designers need to account for the specific input modality. For example, desktop applications (especially for laptops) are often designed for the interaction with a trackpad. Knowing that this modality scores high on efficient interaction, designers are encouraged to compensate for the flipside of the coin: a lower performance in spatial memory. As the indirect mapping of this input leads to more speed, long navigation distances can be easily compensated for. However, this comes at the cost of a less accurate navigation and a lower accuracy in object relocation. Thus, great care should be taken to support spatial memory from an application perspective, for example by easing target acquisition (e.g. with iceberg targets [31, p. 76], accounting for low accuracy) or

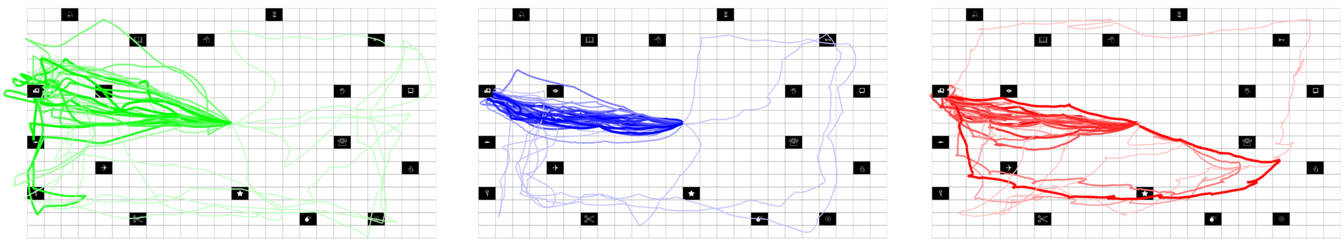

Figure 3. Visualizations of navigation paths for the subset of participants navigating to the same item on LARGE using PAD, TOUCH and MOVE. The thickness of the paths visualize the sequence of trials. Early trials are thinner.

by providing additional visual cues like e.g. landmarks [30], Halos [2], or Wedges [5], accounting for lower object location memory performance.

On the other hand, when designing systems for large interactive surfaces in public settings (e.g. in a museum or in a tourist information store), touch interaction is often chosen to allow for direct manipulation. Our findings show that such a direct mapping of input and output leads to slower navigation speed, especially for large displays. On the other hand, we could also show that the haptic feedback in TOUCH supports a high precision in navigation performance. We therefore encourage designers to compensate for the slower navigation speed. We recommend to allow for interaction on different levels of detail to increase the efficiency of task completion (e.g. by using interactive overview + detail visualizations [4]).

User Satisfaction vs. Spatial Memory

This trade-off is best reflected in the interaction patterns with MOVE. Interacting with MOVE not only resulted in a low efficiency but also led to frustration. The novelty of the device hampered user interaction, and caused the user to focus more on the device than on the task. This led to slower task completion times and lower user satisfaction. Although using MOVE on LARGE led to even lower efficiency (slower navigation speed and longer task completion time), it had a positive impact on spatial memory (navigation path length and object location memory). Especially when working on LARGE, the combination of kinesthetic cues and additional effort to interact with the device led to a shorter navigation path and better object location memory. This discrepancy in our findings between user satisfaction and spatial memory performance is in line with the "performance-preference dissociation" [9, 10, 18] that states that "people are not necessarily performing best with the interfaces they prefer" [10]. Although spatial memory performance was best when interacting with MOVE, participants clearly rated this input modality lowest.

We strongly recommend designers to carefully take into account the trade-off between user satisfaction and spatial memory in the design of interactive systems. Especially in use cases that work with gesture input (e.g. in-car settings or in sterile environments like the operating theatre), care has to be taken to make use of the high spatial memory performance in order to alleviate the low subjective rating concerning task load and user experience. We therefore suggest to allow for the customization of the interface by the user as this increases satisfaction [24] but also nurtures spatial memory due to conscious and explicit individual configuration activities.

Our findings show that especially with gestural input, participants can well remember the navigation path to get to an item as well as the absolute position of that item on a digital landscape. However, the complexity of the input procedure itself was considered to be cumbersome and frustrating. The realization of a gesture should therefore place as little effort as possible to the user [24] to allow for easy and smooth interaction. Future research could also focus on the role of performing gestures without an additional device.

CONCLUSION

In this paper, we studied the effects of different levels of embodiment on participants' spatial memory, from different perspectives that formed our research questions: spatial memory performance, efficiency, and user satisfaction. We conducted an experiment using a well-established task and item set and studied individuals working on three different input modalities and two output devices. In detail, we focused on aspects like navigation paths and object location memory (spatial memory), navigation speed and task completion time (efficiency), and subjective preferences, user experience and task load (user satisfaction).

We were able to identify effects of different combinations of off-the-shelf in- and output modalities, which helped us to come up with a nuanced description of the influences of different aspects of embodiment. Relating traditional measurements for spatial memory to additional, usability-based measurements provides a first step towards interactive systems that rely on both: Exploitation of a user's spatial memory while providing efficient and satisfying interaction.

Rather than rating interaction modalities and display sizes with respect to their impact on spatial memory, our findings show that participants engage in different interaction strategies when working in different conditions. Depending on the interaction modality and the display size, we identified two trade-offs: The *Efficiency vs. Spatial Memory* trade-off shows that navigation speed negatively correlates with the navigation and spatial memory performance, and the *User Satisfaction vs. Spatial Memory* trade-off shows that a subjective user satisfaction score negatively correlates with the performance of spatial memory. Thus, the interrelation of user satisfaction, efficiency, and spatial memory leads to different usage patterns that participants take on to fulfil the task at hand.

ACKNOWLEDGMENTS

We thank the German Research Foundation (DFG) for financial support within project C01 of SFB/Transregio 161.

REFERENCES

1. Gary L Allen. 2004. *Human spatial memory: Remembering where*. Psychology Press.
2. Patrick Baudisch and Ruth Rosenholtz. 2003. Halo: A Technique for Visualizing Off-screen Objects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, USA, 481–488. DOI: http://dx.doi.org/10.1145/642611.642695
3. Simon Butscher and Harald Reiterer. 2016. Applying Guidelines for the Design of Distortions on Focus+Context Interfaces. In *Proceedings of the International Working Conference on Advanced Visual Interfaces (AVI '16)*. ACM, New York, NY, USA, 244–247. DOI: http://dx.doi.org/10.1145/2909132.2909284
4. Andy Cockburn, Amy Karlson, and Benjamin B. Bederson. 2009. A Review of Overview+Detail, Zooming, and Focus+Context Interfaces. *ACM Comput. Surv.* 41, 1, Article 2 (Jan. 2009), 31 pages. DOI: http://dx.doi.org/10.1145/1456650.1456652
5. Sean Gustafson, Patrick Baudisch, Carl Gutwin, and Pourang Irani. 2008. Wedge: Clutter-free Visualization of Off-screen Locations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. ACM, New York, NY, USA, 787–796. DOI: http://dx.doi.org/10.1145/1357054.1357179
6. Sandra G. Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human mental workload* 1, 3 (1988), 139–183.
7. Kasper Hornbæk. 2006. Current Practice in Measuring Usability: Challenges to Usability Studies and Research. *Int. J. Hum.-Comput. Stud.* 64, 2 (Feb. 2006), 79–102.
8. Kasper Hornbæk, Benjamin B. Bederson, and Catherine Plaisant. 2002. Navigation Patterns and Usability of Zoomable User Interfaces with and Without an Overview. *ACM Trans. Comput.-Hum. Interact.* 9, 4 (Dec. 2002), 362–389. DOI:http://dx.doi.org/10.1145/586081.586086
9. Kasper Hornbæk and Effie Lai-Chong Law. 2007. Meta-analysis of Correlations Among Usability Measures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. ACM, New York, NY, USA, 617–626. DOI: http://dx.doi.org/10.1145/1240624.1240722
10. Mikkel R. Jakobsen, Yvonne Jansen, Sebastian Boring, and Kasper Hornbæk. 2015. *Should I Stay or Should I Go? Selecting Between Touch and Mid-Air Gestures for Large-Display Interaction*. Springer International Publishing, Cham, 455–473. DOI: http://dx.doi.org/10.1007/978-3-319-22698-9_31
11. Hans-Christian Jetter, Svenja Leifert, Jens Gerken, Sören Schubert, and Harald Reiterer. 2012. Does (Multi-)Touch Aid Users' Spatial Memory and Navigation in 'Panning' and in 'Zooming & Panning' UIs?. In *Proceedings of the International Working Conference on Advanced Visual Interfaces (AVI '12)*. ACM, New York, NY, USA, 83–90. DOI:http://dx.doi.org/10.1145/2254556.2254575
12. Bonifaz Kaufmann and David Ahlström. 2013. Studying Spatial Memory and Map Navigation Performance on Projector Phones with Peephole Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 3173–3176. DOI: http://dx.doi.org/10.1145/2470654.2466434
13. Daniel Klinkhammer, Jan Oke Tennie, Paula Erdoes, and Harald Reiterer. 2013. Body Panning: A Movement-based Navigation Technique for Large Interactive Surfaces. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces (ITS '13)*. ACM, New York, NY, USA, 37–40. DOI:http://dx.doi.org/10.1145/2512349.2512822
14. Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and Usability Engineering Group*. Springer, 63–76.
15. Svenja Leifert. 2011. The Influence of Grids on Spatial and Content Memory. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*. ACM, New York, NY, USA, 941–946. DOI: http://dx.doi.org/10.1145/1979742.1979522
16. Jens Müller, Roman Rädle, Hans-Christian Jetter, and Harald Reiterer. 2015. An Experimental Comparison of Vertical and Horizontal Dynamic Peephole Navigation. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1523–1526. DOI: http://dx.doi.org/10.1145/2702123.2702227
17. 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 *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1245–1249. DOI: http://dx.doi.org/10.1145/2858036.2858043
18. Jakob Nielsen and Jonathan Levy. 1994. Measuring Usability: Preference vs. Performance. *Commun. ACM* 37, 4 (April 1994), 66–75. DOI: http://dx.doi.org/10.1145/175276.175282
19. Kenton O'Hara, Abigail Sellen, and Richard Bentley. 1999. Supporting Memory for Spatial Location While Reading from Small Displays. In *CHI '99 Extended Abstracts on Human Factors in Computing Systems (CHI EA '99)*. ACM, New York, NY, USA, 220–221.
20. Henri Palleis and Heinrich Hussmann. 2016. Indirect 2D Touch Panning: How Does It Affect Spatial Memory and Navigation Performance?. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1947–1951. DOI: http://dx.doi.org/10.1145/2858036.2858334

21. Kathrin Probst. 2016. *Peripheral Interaction in Desktop Computing: Why It's Worth Stepping Beyond Traditional Mouse and Keyboard*. Springer International Publishing, Cham, 183–205. DOI: http://dx.doi.org/10.1007/978-3-319-29523-7_9

22. Roman Rädle, Hans-Christian Jetter, Simon Butscher, and Harald Reiterer. 2013. The Effect of Egocentric Body Movements on Users' Navigation Performance and Spatial Memory in Zoomable User Interfaces. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces (ITS '13)*. ACM, New York, NY, USA, 23–32. DOI: http://dx.doi.org/10.1145/2512349.2512811

23. Roman Rädle, Hans-Christian Jetter, Jens Müller, and Harald Reiterer. 2014. Bigger is Not Always Better: Display Size, Performance, and Task Load During Peephole Map Navigation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 4127–4136. DOI: http://dx.doi.org/10.1145/2556288.2557071

24. Joey Scarr, Andy Cockburn, and Carl Gutwin. 2012. Supporting and exploiting spatial memory in user interfaces. *Interaction* 6, 1 (2012), 1–84.

25. Amy Skopik and Carl Gutwin. 2005. Improving Revisitation in Fisheye Views with Visit Wear. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '05)*. ACM, New York, NY, USA, 771–780. DOI: http://dx.doi.org/10.1145/1054972.1055079

26. Desney S. Tan, Darren Gergle, Peter Scupelli, and Randy Pausch. 2003. With Similar Visual Angles, Larger Displays Improve Spatial Performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, USA, 217–224. DOI:http://dx.doi.org/10.1145/642611.642650

27. Desney S. Tan, Darren Gergle, Peter G. Scupelli, and Randy Pausch. 2004. Physically Large Displays Improve Path Integration in 3D Virtual Navigation Tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04)*. ACM, New York, NY, USA, 439–446. DOI: http://dx.doi.org/10.1145/985692.985748

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

29. F. Tyndiuk, G. Thomas, V. Lespinet-Najib, and C. Schlick. 2005. Cognitive Comparison of 3D Interaction in Front of Large vs. Small Displays. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST '05)*. ACM, New York, NY, USA, 117–123. DOI: http://dx.doi.org/10.1145/1101616.1101641

30. Norman G. Vinson. 1999. Design Guidelines for Landmarks to Support Navigation in Virtual Environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*. ACM, New York, NY, USA, 278–285. DOI: http://dx.doi.org/10.1145/302979.303062

31. Daniel Wigdor and Dennis Wixon. 2011. *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture* (1st ed.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.