Erschienen in: CHI '14 : Proceedings of the SIGCHI Conference on Human Factors in Computing Systems ; Toronto, ON, Canada, April 26 - May 01, 2014 / Matt Jones ... [General Chairs]. - New York : ACM, 2014. - S. 4127-4136. - ISBN 978-1-4503-2473-1

Bigger is not always better: Display Size, Performance, and Task Load during Peephole Map Navigation

Roman Rädle[1*], Hans-Christian Jetter[2*], Jens Müller[1], Harald Reiterer[1]
[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
[*] The first two authors contributed equally to this work.

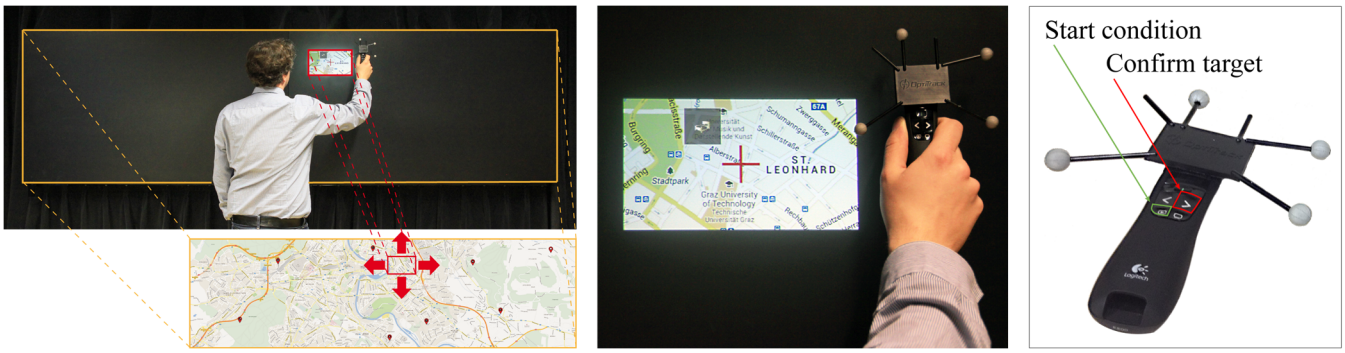

Figure 1. Dynamic peephole navigation of a map was simulated on a large vertical screen (left). The peephole was always displayed next to a handheld Presenter device (center) with buttons and a passive IR marker for 3D tracking (right).

ABSTRACT
Dynamic peephole navigation is an increasingly popular technique for navigating large information spaces such as maps. Users can view the map through handheld, spatially aware displays that serve as peepholes and navigate the map by moving these displays in physical space. We conducted a controlled experiment of peephole map navigation with 16 participants to better understand the effect of a peephole's size on users' map navigation behavior, navigation performance, and task load. Simulating different peephole sizes from 4″ (smartphone) up to 120″ (control condition), we confirmed that larger peepholes significantly improve learning speed, navigation speed, and reduce task load; however, this added benefit diminishes with growing sizes. Our data shows that a relatively small, tablet-sized peephole can serve as a "sweet spot" between peephole size and both user navigation performance and user task load.

Author Keywords
Peephole navigation; map navigation; display size; user study; navigation performance; experimentation

ACM Classification Keywords
H.5.2. User Interfaces: Input devices and strategies. Evaluation/methodology.

INTRODUCTION
Dynamic peephole navigation [11] is an increasingly popular technique for navigating large information spaces using small, spatially aware displays [4]. Typically, the display of a handheld computer [4,12,22,23], mobile phone [12,13,16,19], tablet [14], tangible display [21], or handheld projector [2,8,9,10,20] acts as a window or *peephole* to a much larger information space, such as a map [9,12,19,23]. Users can control the mobile display's content by physically moving it up, down, and sideways. By this, they can pan their view to move invisible off-screen content into the display and access the entire information space as if it was situated in physical space. This physical way of navigation provides users with more proprioceptive cues which are assumed to improve their orientation and understanding of the information space [4] and their spatial memory [9]. Ideally, peephole users are able to navigate quickly (short navigation time) and directly (short travelled path length) from their current location to any destination in the information space without an extensive task load, even if the location is off-screen or yet unknown.

We found that previous studies [9,13,14,19] have not sufficiently explored the peephole's size as an independent variable and how it affects navigation behavior, path lengths, navigation times, and user task load. This is surprising since it seems plausible that these aspects are all strongly dependent upon peephole size. A larger peephole reduces the need for slow physical panning and search in favor of a faster visual scanning of the display's content. It also allows for recognition rather than recall from spatial memory because it reveals more visual features that support user orientation all at once. However, a study of simulated tunnel-vision in front of large displays that included a task comparable to peephole navigation showed that the effect of a reduced peripheral vision and field of view is surprisingly small [1]. If this is also true for peepholes, it will open important design opportunities. In real-world systems, larger peepholes and displays increase cost, energy consumption, and weight, and the devices become more cumbersome. An alternative are small and lightweight handheld projectors which can produce a relatively large peephole. However, some practical problems (hand jittering, finding surfaces in the right size and lighting conditions for projection, privacy concerns when using projections in public spaces) come into play. Designers must make concessions due to these constraints. They want users to experience the benefits of larger peepholes while avoiding the many disadvantages that result from using and handling larger devices or mobile projections. Therefore answering the question of how small peepholes can become without overburdening their users during map navigation is of great practical relevance.

With this study, we wanted to find a good tradeoff, or "sweet spot", between peephole size and both user navigation performance and user task load. To do this, we conducted a controlled lab experiment during which 16 participants completed map navigation tasks on a large, vertical screen with physical navigation of simulated dynamic peepholes. The independent variable, peephole size, had 4 levels: smartphone, tablet, projector phone, and a control condition where the peephole was the size of the entire large, vertical screen. The dependent variables were travel path length, navigation time, and subjective workload. Results show that a relatively small, tablet-sized peephole can serve as the "sweet spot" mentioned above. In the following, we discuss related work, formulate our hypotheses, describe the experimental design, and report results before we conclude with a discussion of our findings.

RELATED WORK

Peephole navigation with handheld, spatially aware devices was originally conceived by Fitzmaurice et al. in 1993 [4]. It was then used in 2002 and 2003 for navigation and pen interaction in 3D [22] and 2D [23]. In 2004, Rapp et al. transferred this concept to handheld projectors for navigating the content of a general purpose UI (e.g., calendars, emails) [15]. From then on, many more peephole designs and systems were created, including augmented reality maps [6,12,19], peepholes using handheld projectors or projector phones [2,8,9,10,20], peephole navigation with smartphones, tablets, and tangible displays [6,13,14,21].

Comparative User Studies of Peephole Designs

Despite the popularity of peephole navigation, it took until 2006 for user studies to move beyond formative usability evaluations of individual systems and use controlled experiments to better understand the different design variants of peepholes more generally.

Mehra et al. [11] simulated a handheld peephole on a 15″ screen showing a 3.3″ peephole in two conditions: 1) *dynamic peephole navigation:* users move the peephole with the mouse across the screen to simulate physical navigation, 2) *static peephole navigation:* users use the mouse to scroll/pan the information space behind the peephole that always remains in the center of the screen to simulate virtual navigation. Results showed that dynamic peepholes improved users' speed and accuracy of discriminating lengths. Mehra et al. expect a substantial increase in users' situation-awareness and better estimation of spatial relationships when using handheld peepholes.

Rohs et al. used a phone as peephole for map navigation to compare virtual vs. physical navigation with and without visual context [19]. They found that physical navigation clearly outperforms virtual navigation with a joystick and that visual context (i.e. a map) behind the peephole does not substantially increase performance, potentially because of the costs of switching and refocusing between the two layers of information. These findings resonate with Henze & Boll who report that a simple off-screen visualization (i.e. arrow) can decrease the task completion time and that its effect is stronger than that of having visual context [6]. Rohs & Essl compared different peephole designs such as panning, zooming, and halo [16]. They report that the halo off-screen visualization is significantly faster and that only in complex situations zoom and halo show comparable performance, while the combination of halo and zooming is detrimental. In our study of peephole size, we therefore used only panning without zoom, no off-screen visualizations, and no visual context around the peephole to avoid confounding variables and to achieve better internal validity.

In 2013, three similar studies that compared physical vs. virtual touch-based peephole navigation were published: Kaufmann et al. conducted a study to find out if navigation performance and spatial memory performance during map navigation can be improved by using a projector phone with a peephole interface (54.7″) instead of a smartphone (4″) with a touchscreen. They report that users performed navigation in the zoomable map equally well, but that spatial memory performance was 41% better for projector phone users [9]. Rädle et al. compared physical vs. virtual navigation with a tablet (10.1″) in a zoomable landscape [14]. Opposed to Kaufmann et al., they report a

significantly better navigation performance (47% decrease in path lengths and a 34% decrease in task time), but no significant effect on users' spatial memory. Finally, Pahud et al. show for a map navigation task with a smartphone as peephole (4.3″) that physical navigation is significantly slower than virtual navigation unless navigation happens between a few known targets [13]. In the light of these contradicting results and 20 years after Fitzmaurice [4], Pahud et al.'s concluding remark that, our understanding of the subtleties of peepholes is still not sufficient, appears particularly true. Therefore, we designed our research to explore these subtleties by isolating the peephole sizes from above studies (54.7″/4″, 10.1″ and 4.3″) in a controlled map navigation experiment to understand their effect on navigation behavior, navigation performance, and task load.

Fitts' Law Peephole Target Acquisition Models

Another stream of related research concerns formulating models of peephole target acquisition based on Fitts' law and validating them with one-directional [3,7,8] or multi-directional pointing [18] or AR tasks [17]. While this work is of fundamental importance, we believe that for understanding the subtleties of real-world map navigation with dynamic peepholes these models are only a first step. They accurately model a subtask of navigation, namely the time and precision of pointing at a distant target. However, real map navigation is far more complex than only one-directional pointing between two targets, since it is a multi-directional task that involves recalling multiple different (off-screen) locations from a mental representation of a 2D map and navigating between them. Such map navigation also involves initial phases of learning the yet unknown locations and spatial features or, at least, reactivating them from memory. All these aspects of map navigation are not part of Fitts' law models, because Fitts' law does not consider them. Fitts' law models cannot help with finding design tradeoffs for peephole size since they propose that pointing performance always gets better with growing peephole size and thus assume that "bigger is always better…" They do not take limiting factors or boundaries into account. For instance, upper boundaries like the users' maximum field of view or the aforementioned practicalities of using large displays or projections or lower boundaries like the higher mental and physical demand when using small or very small peepholes. This is why we chose an experimental approach to measure the "sweet spot" for map navigation instead of attempting to approximate it using existing predictive models.

FORMULATION OF HYPOTHESES

We entered our experiment with the following basic assumption about the nature of map navigation with peepholes: A typical map navigation activity can be separated into two phases, a *learning phase* and a *navigation phase*.

The *learning phase* only takes place if the content of the map is unknown to the user or the spatial relations within the map are only partially remembered and must be reactivated from the users' memory. This is the case when users navigate an unknown map or a map they have not seen or used recently, for example in typical augmented map scenarios for tourist or cultural heritage sites. During this phase, users first have to scan the entire map by physically moving the peephole to get an overview and to memorize positions, map features, and their spatial relations before they then can navigate efficiently. As discussed above, a larger display size should facilitate learning by revealing more content and visual features at a time and reducing the amount of slower physical panning in favor of more visual scanning. Therefore our hypotheses for the *learning phase* were that a larger peephole size decreases 1) the travelled path lengths and 2) the times for completing the navigation tasks.

In the *navigation phase*, a mental representation of the actual map is already present in the users' memory. This is either the case when a mental representation of a map remains in a user's memory after the *learning phase* is completed or when they are already familiar with the map. In the *navigation phase* users can, in principle, navigate toward destinations in the map efficiently, even if they are currently invisible. They do not have to scan large parts of the map anymore to find their destination but can rely on their spatial memory (and proprioceptive cues and motor memory from physical navigation) to reach their targets faster and with a shorter travelled distance. In comparison to the *learning phase*, the *navigation phase* more resembles a pointing task without exhaustive scanning or searching and thus is less affected by peephole size. However, based on Fitts' law models of peephole target acquisition [3,7,8,18], there still should be differences between the peephole sizes. Therefore our hypotheses for the *navigation phase* were that a larger peephole size decreases 1) the travelled path lengths and 2) the navigation times for completing the navigation tasks.

For the overall navigation task including both phases, we assumed that the cognitive load and the amount of physical panning increases with a smaller peephole size. Therefore, we hypothesized that the users' reported task load (based on the mental and physical demand items of NASA-TLX [5]) increases for smaller peepholes.

Finally, we predicted that the smaller the peephole, the greater the likelihood that users built an unreliable or incorrect mental spatial representation of the map and thus, when exposed to similar maps, they might not be able to recognize the one they navigated in the experiment. Therefore our final hypothesis is that the number of errors in a post-navigation map recognition task should increase for smaller peephole sizes.

EXPERIMENT

To better understand the role of peephole size during both phases of a map navigation task, we designed a controlled laboratory experiment with high internal validity. We

isolated peephole size from other possible confounding variables, such as: existence of off-screen visualizations, design of navigation gestures, and ergonomic aspects or device-specific properties (weight of the device, readability from different viewing angles, resolution, or latency). The study was conducted as a 4×4 within-subjects design and systematically counterbalanced using a balanced Latin Square. The independent variable, peephole size (see Figure 2), had four within-subjects factors: control condition (S1), projector-sized peephole (S2), tablet-sized peephole (S3), and smartphone-sized peephole (S4). We used the four different maps A, B, C, D (Figure 4) to control for systemic errors and to avoid learning effects. The navigation path length, the navigation time, task load, and the post-navigation map recognition were the dependent variables.

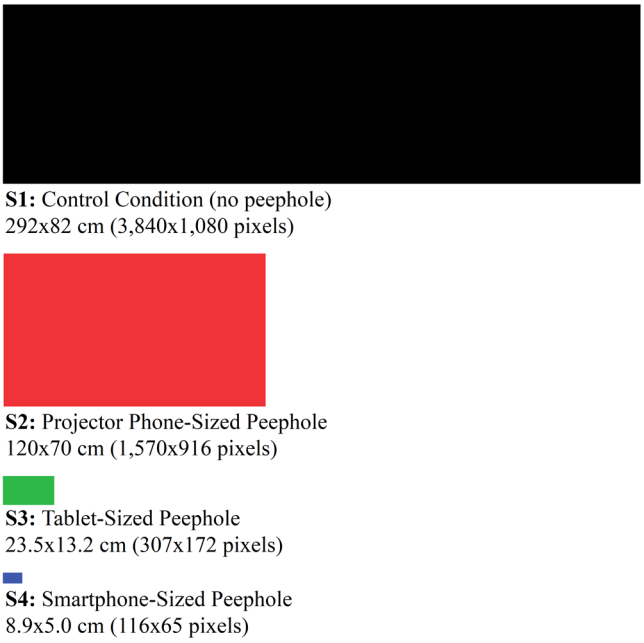

Figure 2. Overview of peephole sizes S1-S4 used as independent variable in the experiment.

In order to achieve a high degree of internal validity, we simulated the peepholes on a large display (rather than using the actual devices) so that the only variation from condition to condition was the peephole size itself. We initially discussed using different real-world devices instead of simulations, so that users would experience all device-specific properties such as different weight, resolution, or latency. However, we decided against this for following reason: Our overall goal is to understand the subtleties of peephole navigation as suggested by Pahud et al. [13] and the different factors that contribute to navigation performance. As a first step, in this study, we wanted to focus only on the effect of peephole size which arguably is the most important property and ideally arrive at generalizable results. Comparing devices would have led to recommending a certain device instead of a "sweet spot" peephole size without being able to isolate the role of peephole size from other device-specific properties (e.g. weight, resolution). We still would not truly understand the role of peephole size because other device properties such as weight and resolution would have come into play. Also, the recommendation would have been short-lived since such properties change with each new device generation.

Participants

16 participants (8 female, 8 male) were recruited to take part in the experiment. The mean age was 26.6 years (SD = 6.2, min = 19 years, max = 37 years) with a skewness of 0.64 (SE = 0.56) and kurtosis of -1.15 (SE = 1.10). One participant was left-handed. To get a realistic sample of participants, we excluded participants from the computer science department or with a background in computer science. 12 of the participants were students, 1 was a lecturer in linguistics, 2 were administrative staff, and 1 was a construction worker.

Apparatus

We used a large, vertical high resolution screen (size 292x82 cm, 120″ diagonal, 3,840x1,080 pixels resolution) to simulate peephole sizes of typical device displays at a constant resolution of 13.1 pixel/cm (or 33.5 ppi). This resolution was lower than that of actual mobile devices, but the display quality was more than sufficient for our purposes (Figure 3). The maps used in the experiment covered the entire screen, but users were only shown a rectangular section the size of the simulated peephole while the surrounding screen was black.

Participants used a wireless Logitech Professional Presenter R800 device (total weight 102 grams, Figure 1) to move the peephole on the screen. The Presenter was equipped with passive markers and continuously tracked in space using an OptiTrack 3D motion capturing system (18 cameras) with a tracking mean error of less than .5 mm and a tracking rate of 100 Hz. Participants held the Presenter in their preferred hand. A Kalman filter was used to reduce jittering caused by hand tremor and the noise or inaccuracies of the OptiTrack motion capturing system.

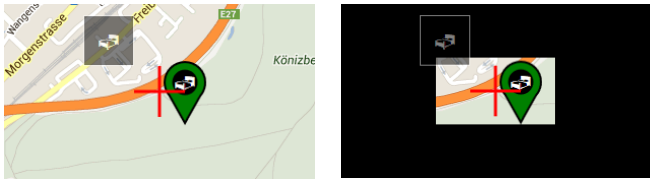

Figure 3. The tablet-sized peephole S3 with 307x172 pixels (left) and the smartphone-sized peephole S4 with 116x65 pixels (right). Note that the visual features and symbols stay recognizable and have sufficient detail.

During the experiment, the peephole travelled left or right of the Presenter (depending on handedness) to simulate physical navigation with a handheld dynamic peephole. By movement of their hands and lateral movement of their bodies, participants could move the rectangular peephole in the XY-plane of the display to view any location on the map, similar to the augmented maps in [6,19]. To minimize occlusion by hands, the anchor point was adjusted to right-

and left-handed users. The ratio between physical movement of the hand in control space and the peephole's XY-movement in display and map space was always 1:1. To constrain the distance between hand and screen to realistic holding and viewing of mobile devices, the peephole only appeared on the screen when the hand was within a range of 15 cm to the display. Except this, participants were free to choose their preferred head, body, and arm position during navigation and thereby set the optimal viewing distance to the display as it is the case when using an actual handheld device. However, they could not use rotation around the X-, Y-, or Z-axis as it is possible in AR see-through scenarios [12,17,22]. There was a red crosshair in the center of the peephole for selecting targets. We showed the target that the user searched for above and to the left of the crosshair.

Task Design

There was one condition for each peephole size. Due to the balanced Latin Square design, display sizes and maps were counterbalanced. Each map had 4 target pins that acted as navigation goals and 4 distractor pins. Maps were taken from Google Maps but were all unknown to the participants. All maps had similar visual features and complexity, such as a city with roads and a river (Figure 4).

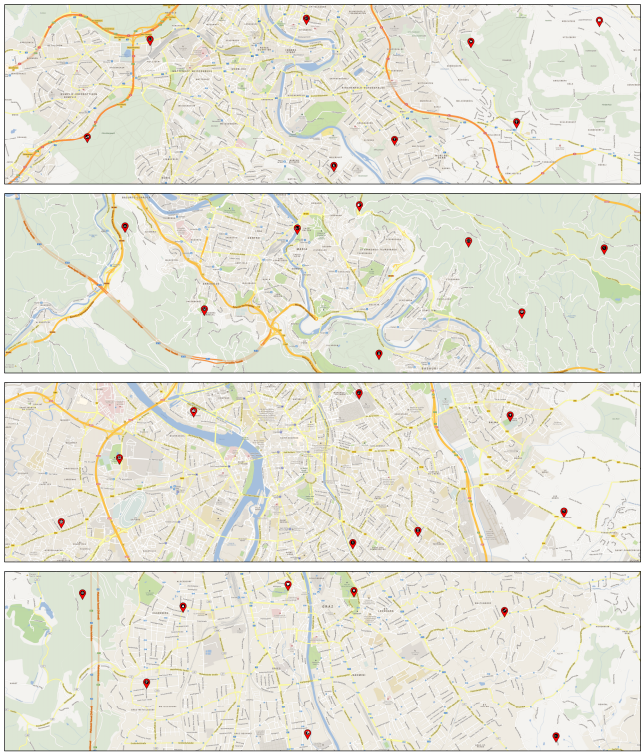

Figure 4. The four maps A-D used for the navigation tasks.

During the task, participants were asked to navigate with the peephole to a target pin in the map that shows a certain symbol, e.g., a bed (Figure 3). They were asked to navigate as quickly and as precisely as possible and to select the target with the peephole's crosshair by pressing the confirm button on the Presenter (Figure 1). The next target was presented to the participants if the correct target was selected. Otherwise the trial continued. The navigation path and time travelled between presenting the new symbol and its selection with the crosshair was recorded. The recording of a trial started immediately after completion of the previous trial and at the last position of the peephole. All targets were systematically placed on each map to ensure comparable target distances between the maps.

In each condition, participants had to navigate to 4 targets in the same order for 8 times (blocks). This added up to 16 participants × 4 conditions × 8 blocks × 4 targets = 2048 trials with 128 trials per participant.

After each condition participants reported their subjective workload ratings using the NASA-TLX questionnaire [5]. After this, they chose the map they navigated from a selection of three maps. Two maps served as distractors. The purpose of this task was to test if the participants could recognize the map they had just used based on the mental representation of the map that they created during the navigation tasks.

Procedure

Each participant was first asked to fill out a demographic questionnaire and was asked about their dominant hand. After this, participants were introduced to holding the Logitech Presenter device with their dominant hand, its two buttons, and how to move it with their hand. To avoid learning effects of handling the peephole during the actual data collection, they then could familiarize themselves with the task, the technique for moving the peephole, and the different peephole sizes during a training phase that lasted as long as they wanted.

After this preparation phase, the actual data collection started with the four conditions. After each condition they reported their task load using NASA-TLX and choose their map in the post-navigation map recognition task. The entire experiment lasted approximately 30 minutes per participant and each participant was rewarded with 8 Euros for their time.

RESULTS

For each peephole size, Figure 5 (top left) shows the mean path length that the participants travelled during each block. Path lengths were normalized by dividing them by the shortest possible path length, so that 1.0 is the minimum. Figure 5 (top right) shows the mean navigation time per block in milliseconds. Additional plots are provided for blocks 4-8 where the data points are too close together on the Y-axis to discriminate them.

Path length and movement time analyses were done using repeated measures ANOVAs (Greenhouse-Geisser corrections are marked as GGc) with post-hoc pairwise comparisons. Table 1 shows their *p*-values for the mean of each individual block, the mean of blocks 1-4 (B1-B4), the mean of blocks 5-8 (B5-B8), and the mean of blocks 1-8 (B1-B8). All post-hoc tests were Bonferroni corrected.

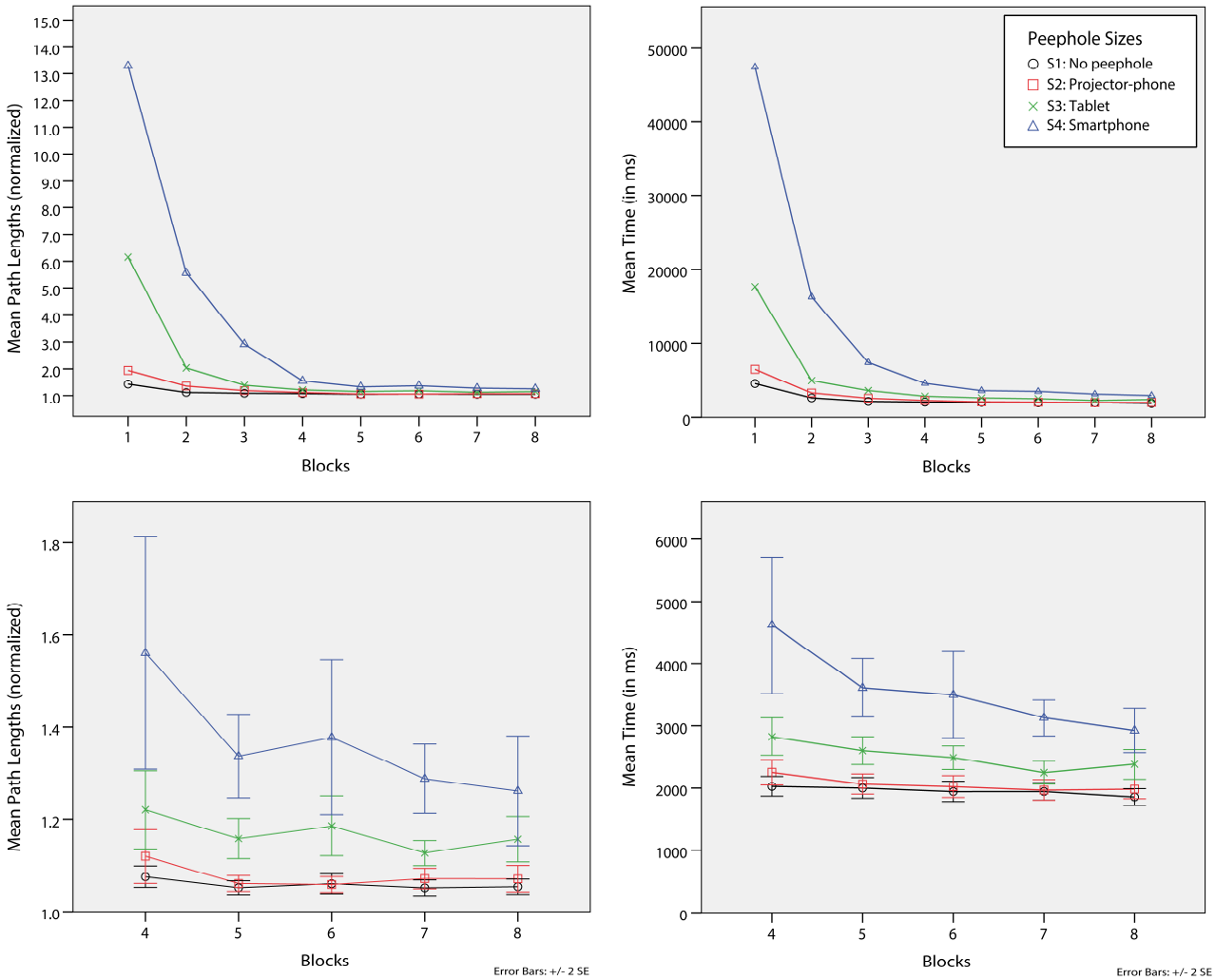

Figure 5. Plots of mean paths lengths and mean times.

		B1	B2	B3	B4	B5	B6	B7	B8	B1-B4	B5-B8	B1-B8
S1 vs. S2	Length	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	Time	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
S1 vs. S3	Length	**.023***	1.000	1.000	.862	**.023***	.330	**.043***	.133	**.017***	**.000***	**.008***
	Time	.061	1.000	1.000	.279	**.001***	.178	.054	**.004***	**.014***	**.000***	**.036***
S1 vs. S4	Length	**.000***	**.000***	**.008***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***
	Time	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***
S2 vs. S3	Length	.057	1.000	1.000	1.000	**.049***	.310	.291	.340	.064	**.000***	**.038***
	Time	.168	1.000	1.000	.904	**.004***	.379	.095	.057	.078	**.000***	.146
S2 vs. S4	Length	**.000***	**.001***	**.015***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***
	Time	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***
S3 vs. S4	Length	**.000***	**.008***	**.045***	**.004***	**.000***	**.020***	**.000***	.119	**.000***	**.000***	**.000***
	Time	**.000***	**.000***	**.010***	**.000***	**.000***	**.000***	**.000***	**.002***	**.000***	**.000***	**.000***

Table 1. The *p*-values (* means significant) for pairwise comparisons of lengths and times.

Results NASA-TLX Task Load Index

The mean results of the NASA-TLX questionnaire about task load are S1: 13.07, S2: 18.49, S3: 25.51, and S4: 47.19 (scale from 0 to 100). An ANOVA with repeated measures revealed a statistically significant main effect of the subjective workload ratings on the peephole sizes, GGc: $F_{1.82,27,23} = 37.642$, $p < .001$, *partial* $\eta^2 = .715$. Pairwise comparisons (Bonferroni corrected) of subjective workload of peephole sizes revealed statistically significant differences between S1 vs. S3 ($p < .001$), S1 vs. S4 ($p < .001$), S2 vs. S4 ($p < .001$), and S3 vs. S4 ($p < .001$). All other comparisons were not significant. The individual subscales of NASA-TLX such as mental demand or physical demand are shown in Figure 6.

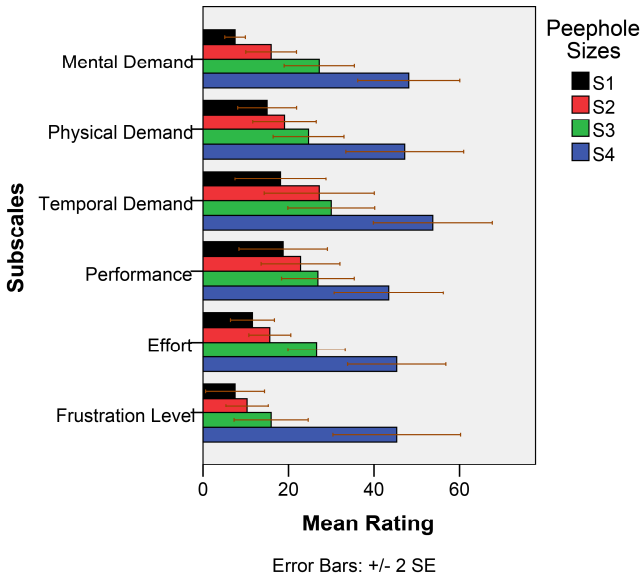

Figure 6. Subscales of the NASA-TLX questionnaire.

Results Map Recognition

The results of the map recognition task revealed the following error rates for each peephole size: S1: 5 errors (31.25%), S2: 4 errors (25%), S3: 4 errors (25%), and S4: 4 errors (25%). Since there are only marginal differences in the error rates for the different peephole sizes, we have not used the error rates in the further data analysis.

DISCUSSION

Resonating with our previous assumption about an initial *learning phase* followed by a *navigation phase*, the first blocks in Figure 5 (e.g. blocks 1-4) show very long path lengths and navigation times with great standard deviations. During these first blocks users still had to scan the information space to memorize locations and to build up a spatial mental representation of the map. This *learning phase* initially lead to a rapid fall of path lengths and times until the values stabilized and stayed roughly constant which indicates the beginning of the *navigation phase*. In the following, we discuss both phases in greater detail.

Evidence for a Learning Phase

Noticeable improvements in peephole navigation occurred during blocks 1-4. This can be explained by users' improving mental spatial representation of the map that they achieved by systematically scanning the map for targets with the peephole. The nature of this initial scanning process becomes evident when plotting peephole movements. Figure 7 shows two examples of such a scanning process by participant 1 for B1 to B4 using the tablet-sized peephole S3 (top) and the smartphone-sized peephole S4 (bottom). The blue dots show the movement of the peephole's anchor point on the screen. The red dots show the locations of the navigation targets. The figure illustrates characteristic scanning patterns with vertical scanning movements that are repeated horizontally or vice versa. They also visualize the potential benefit of a greater peephole size during this *learning phase*. Since a greater peephole reveals more visual information, it is possible to choose larger distances between the repeated movements, thus shortening the overall scanning path.

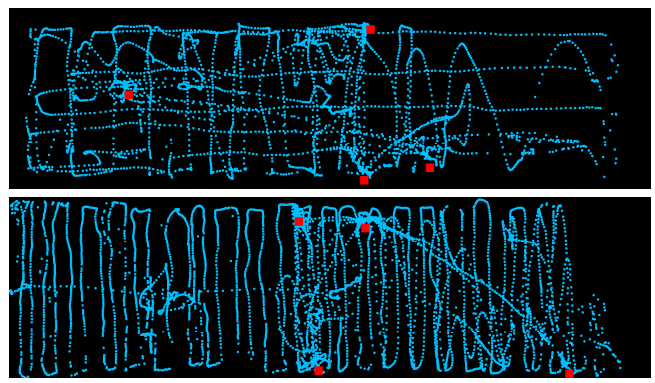

Figure 7. Example of peephole movement by participant 1 during B1 to B4 with S3 (top) and S4 (bottom).

A 4×4 (peephole size × repetition) ANOVA with repeated measures revealed a statistically significant main effect of peephole size in terms of travelled path lengths, GGc: $F_{1.12,16.82} = 26.79$, $p < .001$, *partial* $\eta^2 = .641$, as well as a significant effect of repetition, $F_{1.97,29.49} = 30.69$, $p < .001$, *partial* $\eta^2 = .672$. There was also an interaction between peephole size and repetition, GGc: $F_{2.39,35.91} = 13.69$, $p < .001$, *partial* $\eta^2 = .477$. This indicates that the effects of peephole size on path lengths and navigation time is depended on the number of repetitions. We consider this as evidence of a learning process during B1 to B4. This is further supported by the fact that the same interaction effect cannot be found in the assumed *navigation phase* during B5 to B8 as we discuss below.

Moreover, the results show that a larger peephole facilitates this learning process and leads to better initial performance. The initial performance in B1 for path length ($M_{S1} = 1.44$ with SE = .105, $M_{S2} = 1.94$ with SE = .152, $M_{S3} = 6.16$ with SE = .736, $M_{S4} = 13.29$ with SE = 2.096) shows significant differences between and the smartphone-sized peephole (S4) and all other peephole sizes (S1, S2, S3). Clearly, the

smartphone-sized peephole was outperformed (Table 1, column B1). Interestingly, there are no significant differences for S1 vs. S2 and S2 vs. S3, a fact that we discuss below in a dedicated section on peephole sizes.

Looking at the entire learning phase B1-B4 reveals similar characteristics: The mean path lengths for B1-B4 (M_{S1} = 1.18 with SE = .031, M_{S2} = 1.41 with SE = .063, M_{S3} = 2.70 with SE = .240, M_{S4} = 5.84 with SE = .853) have significant differences between S1, S2, and S3 vs. the smartphone size S4, which is clearly outperformed again (Table 1, column B1-4). There are no significant differences for S1 vs. S2 and S2 vs. S3.

Evidence for a Navigation Phase

The different nature of the *navigation phase* compared to the *learning phase* becomes immediately visible when looking at the plots of peephole movement in Figure 8 that show the same tasks as Figure 7 but this time for B5 to B8. The navigation trajectories show direct navigation movements between the targets without scanning. This illustrates how participants successfully applied their mental spatial representation and proprioceptive cues of the physical peephole navigation to efficiently move between invisible but known targets without a need for scanning.

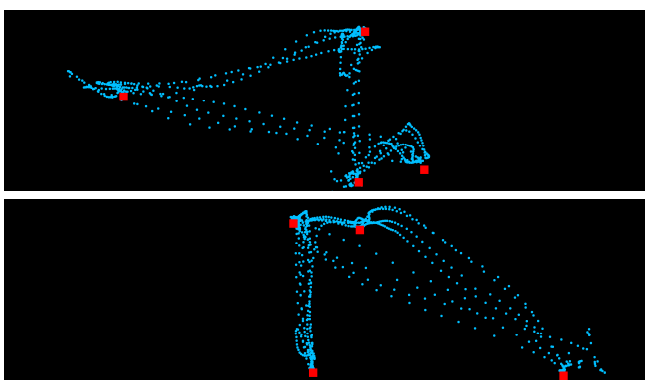

Figure 8. Example of peephole movement by participant 1 during B5 to B8 with S3 (top) and S4 (bottom).

As discussed, the navigation performance in terms of path lengths and navigation time substantially improved between B1 and B5. After this, as is visible in the plots of mean path lengths and mean time for B4-B8 in Figure 5, the navigation performance in B5, B6, B7, and B8 stayed almost constant, however at different levels depending on the peephole size. These results indicate gradual transition from the end of the *learning phase* to the beginning of the *navigation phase*.

A statistical indicator for the end of the learning phase and the beginning of the navigation phase is the absence of the interaction between peephole size and repetition that we witnessed for B1-B4: A 4×4 (peephole size × repetition) ANOVA with repeated measures on B5-B8 revealed a statistically significant main effect of peephole size on travelled path lengths, GGc: $F_{1.27,19.10}$ = 27.11, p < .001, *partial* η^2 = .644 but no difference for the repetition, $F_{1.88,28.21}$ = 1.87, p = .175, *partial* $\eta 2$ = .111. As stated above, there was no interaction between peephole size and repetition, GGc: $F_{3.47,51.98}$ = 1.04, p = .388, *partial* $\eta 2$ = .065. Also the small standard deviations in Figure 5 allow the conclusion that participants reached almost constant performance levels for the different peephole sizes.

A post-hoc pairwise comparison with Bonferroni corrections of peephole sizes for B5-B8 (Table 1) revealed statistically significant differences between peephole sizes S2 vs. S3, S2 vs. S4, and S3 vs. S4. However, it did not show a difference for the comparison of S1 vs. S2. The mean values for path lengths for B5-B8 are M_{S1}= 1.05 (SE = .006) for S1, M_{S2} = 1.07 (SE = .006) for S2, M_{S3} = 1.16 (SE = .017) for S3, and M_{S4} = 1.32 (SE = .043) for S4.

Peephole Sizes: Is Bigger Always Better?

Up to now, the results were largely reflecting our initial assumptions about the existence of a learning phase, a navigation phase, and the benefits of larger peepholes that we formulated above. However, there are some unexpected observations that shed light on the question, "Is bigger always better?"

Control Condition (S1) vs. Projector Phone-Size (S2)

Table 1 shows that for all blocks in B1-B8, each individual block, the learning phase (B1-B4), and navigation phase (B5-B8), there was no significant difference between control condition S1 and the peephole S2. This is clearly a case for projector phones since there were no significant differences in performance between S2 and a 120″ large screen without any peephole. Also the NASA-TLX questionnaires did not report a significantly different workload with S2 compared to S1. Therefore, when comparing S1 vs. S2, bigger is not better. To expand this conclusion, peephole sizes greater than a projector phone do not pay off in terms of navigation performance or task load when used in a map navigation scenario that is similar to our experiment.

However, in our study, S2's size of 54.7″ covers a greater field of view than might be typical in real-world uses of projector phones. Participants stood within a close range (approx. 40 cm) to the screen resulting in covering approx. 127° of the users' typical field of view. In [9], users stood at a distance of 200 cm, so that the projection covered approx. 33.4° of the users' field of view. Interestingly, in the light of this size of S2 in our study, it is therefore even more surprising that the tablet-sized peephole S3 achieved an almost comparable performance as we discuss in the following.

Projector Phone-Size (S2) vs. Tablet-Size (S3)

The comparison of the projector phone condition S2 vs. the tablet condition S3 in Table 1 reveals that there are no significant differences in both devices except for B5, B5-8 (*navigation phase*), and, as a result, also for the overall performance B1-B8. S2 outperforms S3 only during the *navigation phase*, but not during the *learning phase*. A

comparison of the absolute differences in terms of path lengths and times during the navigation phase shows an 8.4% longer navigation path length and 419 ms longer navigation time per target.

While statistically significant, these differences have to be seen in relation to the afore-mentioned disadvantages and practicalities of mobile projections vs. tablets. In our interpretation, the only moderately increased performance during *navigation phase* cannot outweigh the many disadvantages of mobile projection and the many advantages of using off-the-shelf tablets. Furthermore, there are no significant differences between S2 and S3 in terms of the reported subjective workload. By this, we do not imply that a tablet-sized peephole should be considered as an equivalent to a projector-phone-sized peephole in every respect. However, designers of peephole navigation systems should carefully balance the specifics of both technologies. We therefore suggest for use cases that are similar to our experiment that a tablet-sized peephole is more suitable than a larger one.

Smartphone-Size (S4)

Our results clearly show that a peephole with the size of a smartphone is outperformed by all other peephole sizes. This is particularly interesting with respect to tablets which are natural competitors to smartphones in peephole navigation scenarios due to their great availability, popularity, price, and mobility. The tablet-size S3 outperforms S4 in blocks B1 to B7, during the learning phase B1-B4, the navigation phase B5-B8, and the overall performance B1-B8. This is also reflected in the report of subjective workload from the NASA-TLX questionnaires, which is 82% higher for S4 than for S3 and also higher in all subscales (MD = 72%, PD = 85%, TD = 76%, P = 53%, E = 69%, and F = 75%).

These findings about better navigation performance and less workload with S3 compared to S4 (13.8% in path lengths and 864 ms in time) could be helpful for revisiting the study of Pahud et al. [13]. Replacing the 4.3″ device in their study with a tablet should lead to better navigation performance and reduced task load in their physical navigation condition. This could possibly lead to different results for their comparison between virtual and physical navigation. These findings are also relevant for Kaufmann and Ahlström's study of spatial memory and map navigation performance with projector phones vs. a smartphone [9]. It would be interesting to see if the reported significant differences in spatial memory still exist when replacing the smartphone with a tablet-sized peephole.

Limitations of the Study

As this comparative study of the effect of peephole size on navigation performance is the first of its kind, it has some limitations that we are aware of and that could be addressed in future work.

First, controlling device-specific properties such as weight or resolution and only using peephole size as independent variable increases the internal validity but also decreases the external validity. It would be interesting to repeat the experiment using real-world physical devices to see if the same results can be replicated or if the differences in device-specific properties such as resolution, weight, or latency outweigh differences in peephole size.

Second, in future studies, the peephole size S2 in the projector phone condition should be reduced to reflect more realistic scenarios like in Kaufmann and Ahlström [9] where the peephole covers only a much smaller field of view. However, it is very likely that such an adjustment would have led to an even greater similarity between our conditions S2 and S3, thus further strengthening our recommendation of tablet-sized peepholes over projector phone-sized peepholes for many scenarios.

Third, compared to the size of S3 and S4, we used a large map and a large physical navigation space with a 1:1 control-display ratio. This could have penalized S3 and S4 and been in favor of S1 and S2. However, similar to [14], our intention was to employ frequent physical navigation and strong proprioceptive cues to support users' spatial memory. The absolute values for the NASA-TLX (all below 50 on a scale to 100) and the absence of significant differences between task loads of S2 vs. S3 reflect that the tasks were still solvable by the users even with a comparably small, tablet-sized display.

CONCLUSION

With this study of peephole map navigation, we wanted to find a "sweet spot" between peephole size and both user navigation performance and user task load. By simulating different peephole sizes from 4″ (smartphone size) up to 120″ (control condition), we found that a smartphone-sized peephole is outperformed by all other sizes and that larger peepholes significantly improve learning speed, navigation speed, and reduce task load. However, this added benefit diminishes with growing sizes, and peephole sizes greater than a projector phone do not pay off in terms of navigation performance or task load anymore. Our data shows that a relatively small, tablet-sized peephole can serve as a "sweet spot" in terms of both user navigation performance and user task load.

We have also shown that for understanding the subtleties of real-world map navigation with dynamic peepholes, existing models of peephole target acquisition based on Fitts' law [3,7,8,18] are only a first step. They were not intended to model different phases of map navigation such as a *learning phase* and a *navigation phase* whose existence we have shown using a statistical and visual analysis of the users' navigation paths in our study. By this, we have contributed to the better understanding of the subtleties of peephole navigation as motivated by Pahud et al. in [13].

ACKNOWLEDGMENTS
This work was partially supported by DFG Research Training Group GK-1042 and by the Ministry for Science, Research, and Art Baden-Wurttemberg under the project Blended Library.

REFERENCES

1. 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.
2. Cao, X. and Balakrishnan, R. Interacting with dynamically defined information spaces using a handheld projector and a pen. *In Proc. of UIST '06*, ACM Press (2006), 225–234.
3. Cao, X., Li, J.J., and Balakrishnan, R. Peephole Pointing: Modeling Acquisition of Dynamically Revealed Targets. *In Proc. of CHI '08*, ACM Press (2008), 1699–1708.
4. Fitzmaurice, G.W. Situated information spaces and spatially aware palmtop computers. *Communications of the ACM 36*, 7 (1993), 39–49.
5. Hart, S. and Staveland, L. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human mental workload*, North Holland Press (1988).
6. Henze, N. and Boll, S. Evaluation of an off-screen visualization for magic lens and dynamic peephole interfaces. *In Proc. of MobileHCI '10*, ACM Press (2010), 191–194.
7. Huber, J., Steimle, J., and Mühlhäuser, M. A model of embodied dynamic peephole pointing for hidden targets. *In Proc. of BCS-HCI '11*, British Computer Society (2011), 315–320.
8. Kaufmann, B. and Ahlström, D. Revisiting Peephole Pointing: A Study of Target Acquisition with a Handheld Projector. *In Proc. of MobileHCI '12*, ACM Press (2012), 211–220.
9. 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.
10. Löchtefeld, M., Gehring, S., and Krüger, A. Distributed User Interfaces for Projector Phones. In J.A. Gallud, R. Tesoriero and V.M.R. Penichet, eds., *Distributed User Interfaces: Designing Interfaces for the Distributed Ecosystem*. Springer London, London, 2011, 113–123.
11. Mehra, S., Werkhoven, P., and Worring, M. Navigating on Handheld Displays: Dynamic versus Static Peephole Navigation. *ACM Transactions on Computer-Human Interaction 13*, 4 (2006), 448–457.
12. Morrison, A., Oulasvirta, A., Peltonen, P., et al. Like bees around the hive: a comparative study of a mobile augmented reality map. *In Proc. of CHI '09*, ACM Press (2009), 1889–1898.
13. Pahud, M., Hinckley, K., Iqbal, S., Sellen, A., and Buxton, B. Toward Compound Navigation Tasks on Mobiles via Spatial Manipulation. *In Proc. of MobileHCI '13*, ACM Press (2013), 113–122.
14. Rädle, R., Jetter, H.-C., Butscher, S., and Reiterer, H. The effect of egocentric body movements on users' navigation performance and spatial memory in zoomable user interfaces. *In Proc. of ITS '13*, ACM Press (2013), 23–32.
15. Rapp, S., Michelitsch, G., Osen, M., et al. Spotlight Navigation: Interaction with a handheld projection device (Video paper). *In Proc. of Pervasive '04*, (2004).
16. Rohs, M. and Essl, G. Which one is better?: information navigation techniques for spatially aware handheld displays. *In Proc. of ICMI '06*, ACM Press (2006), 100–107.
17. Rohs, M., Oulasvirta, A., and Suomalainen, T. Interaction with magic lenses: real-world validation of a Fitts' Law model. *In Proc. of CHI '11*, ACM Press (2011), 2725–2728.
18. Rohs, M. and Oulasvirta, A. Target acquisition with camera phones when used as magic lenses. *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*, ACM Press (2008), 1409–1418.
19. Rohs, M., Schöning, J., Raubal, M., Essl, G., and Krüger, A. Map Navigation with Mobile Devices: Virtual versus Physical Movement with and without Visual Context Michael. *In Proc. of ICMI '07*, ACM Press (2007), 146–153.
20. Rukzio, E., Holleis, P., and Gellersen, H. Personal Projectors for Pervasive Computing. *IEEE Pervasive Computing 11*, 2 (2012), 30–37.
21. Spindler, M., Stellmach, S., and Dachselt, R. PaperLens: Advanced Magic Lens Interaction Above the Tabletop. *In Proc. of ITS '09*, ACM Press (2009), 69–76.
22. Tsang, M., Fitzmaurice, G.W., Kurtenbach, G., Khan, A., and Buxton, B. Boom Chameleon: Simultaneous capture of 3D viewpoint, voice and gesture annotations on a spatially-aware display. *In Proc. of UIST '02*, ACM Press (2002), 111–120.
23. Yee, K. Peephole displays: pen interaction on spatially aware handheld computers. *In Proc. of CHI '03*, ACM Press (2003), 1–8.