Investigating Mid-Air Pointing Interaction for Projector Phones

Christian Winkler
University of Duisburg-Essen, Germany
christian.winkler@paluno.uni-due.de

Ken Pfeuffer
University of Duisburg-Essen, Germany
ken.pfeuffer@stud.uni-due.de

Enrico Rukzio
Ulm University, Germany
Lancaster University, UK
enrico.rukzio@uni-ulm.de

ABSTRACT
Projector phones, mobile phones with built-in projectors, might significantly change the way we are going to use and interact with mobile phones. The potential of combining the mobile and the projected display and further the potential of the mid-air space between them have yet to be explored. In this paper we assess these potentials by reporting two user studies: First, an experimental comparison of four techniques for target selection on the projection, including interaction on the touchscreen of the projector phone as well as performing pointing gestures in mid-air around the phone. Our results indicate that interacting *behind* the phone yields the highest performance, albeit showing a twice as high error rate. Second, a follow-up experiment where we analyzed the performance of the two best techniques of the first study within realistic mobile application scenarios such as browsing and gaming. The results show that mobile applications benefit from the projection, e.g., by overcoming the fat-finger problem on touchscreens and increasing the visibility of small objects. Our findings speak for the integration of a tracking camera at the bottom of the projector phone to enable mid-air pointing interaction.

Author Keywords
Projector phone; selection; target; pointing; gesture.

ACM Classification Keywords
H.5.2 [Information interfaces and presentation]: User Interfaces - Input devices and strategies.

INTRODUCTION
An inherent issue of mobile phones with touch screens is their small size, which is on the one hand essential for their mobility but on the other hand significantly limits the available space for input and output of information. While some years ago miniaturization of mobile devices seemed to be the ultimate goal of mobile phone manufacturers, nowadays we see mobile phone sizes and resolutions expanding again. The emergence of pico projectors and in particular projector phones, i.e. phones with built-in projectors, provides a versatile solution for this issue: users can project and interact with a large display almost anywhere and at any time. Such projector phones also support various forms of collocated media viewing, browsing and interactions which are not possible with conventional mobile phones.

Currently available projector phones (e.g. Samsung Beam, Sharp SH-06C) or accessory projectors (e.g. the SHOWWX+ from Microvision) that can be connected via TV-out to a conventional phone only mirror what is usually shown on the touch screen [18]. Projecting the touch screen user interface while maintaining the same interaction style must lead to suboptimal interactions as it requires many context switches during operation and because those interfaces were designed for high resolution screens with small dimensions operated through direct touch input.

Using a pointer as an intermediate that marks the current position on the projection is a basic way to overcome some of these problems. In particular, using the touchscreen of the mobile phone for indirectly controlling a mouse pointer on the projection requires no additional hardware and has been the focus of various research projects and products, e.g. [15]. The conceptual disadvantages are the indirectness and the unavailability of the touchscreen for interaction or as information display since it is occupied as a touchpad.

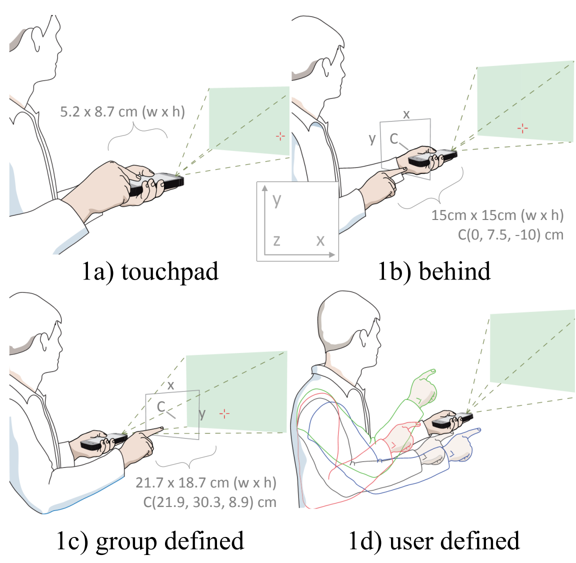


Figure 1. Compared pointing techniques. 'C' denotes the central point of the interaction area in relation to the projector phone position (0, 0, 0).

ITS'12, November 11–14, 2012, Cambridge, Massachusetts, USA.

Using mid-air finger-pointing techniques is an interesting alternative due to the more direct interaction and the possibility to use the mobile phone screen as secondary in-/output to the projection. Further, these techniques neither require the user to carry additional hardware nor do they require movement of the phone that interferes with the projection as accelerometer based interactions would do for instance. Thus they seem very suitable for typical ad-hoc mobile scenarios. However, the mid-air space around the user is quite large and unexplored considering the bimanual and interdependent control. So far it is unclear which interaction area will be optimal and how well users will be able to manage the dual-display, bimanual interaction.

To open this area of research we investigated the performance of three mid-air finger-pointing techniques leveraging different interaction areas (see Figure 1b-d) compared to the existing *touchpad* technique (Figure 1a). We compared the techniques through an experimental user study based on the ISO 9241-9 tapping task. Our results indicate that the interaction technique in which the user points *behind* the mobile phone to control a cursor on the projection performs significantly better than other mid-air techniques in all scenarios. Even more, it also performs ~15% faster than the *touchpad* option despite yielding ~2.5 times more errors. This makes it an interesting alternative interaction technique for a variety of application scenarios, even without considering its aforementioned advantage of keeping the touchscreen free.

In a follow-up experiment, we compared the interaction techniques *touchpad* and *behind* in common usage scenarios such as browsing, gaming, and drawing in order to analyze their performance in realistic contexts. Also, we included a standard smartphone without a projector in our comparison in order to analyze the performance of the projector phone interaction techniques in contrast to the current usage of a smartphone. Results of the second study confirm the familiarity of users with standard *touchscreen* phones but also highlight various advantages for the projector phone interaction techniques, e.g., in terms of not occluding targets on the screen, improved visibility, the usage in collaborative settings, and joy of use.

RELATED WORK

Interactions with a projector phone can be classified into four different categories: using controls on the phone, moving the phone, directly interacting with the projection and manipulation of the projection surface [18]. Most currently available solutions use the touchscreen of the mobile phone for input which requires no additional hardware but suffers from the separation of input (phone) and output (projection). This could lead to a large number of context switches as the user has to switch her focus constantly between the projection and the phone in case the phone screen is used for displaying information [6,8]. Moving the phone in order to perform interactions is another possibility but works mainly for simple commands and has the disadvantage that the projection moves as well [2,4]. Directly touching the projection of a projector phone or pico projector has recently been investigated but requires the user to be very close to the projection surface [9,22]. This limits the size of the projection through which collocated collaborations cannot be supported optimally.

The usage of the touchscreen as a *touchpad* is an effective approach for controlling a cursor on a remote screen (Figure 1a) [13]. The advantage of this concept is that it is already very familiar from touchpads found on laptops. Conversely, it has the disadvantage that while using the screen as a touchpad, there is no easy alternative to interact with the content on the mobile phone screen at the same time. One possible solution would be a hardware button on the side of the device to toggle between touchscreen and touchpad mode. While this seems feasible it still would not allow for interactions where both displays are simultaneously active, e.g. for seamless dragging of pictures from phone display to projection.

Recently researchers started to look at the usage of finger pointing and hand gestures for controlling interactions due to improved tracking technology. *SideSight* explored gestural interactions around a mobile phone to increase the input space on small devices [3]. Similarly, researchers looked at direct interactions with personal projections. Shadow interactions as described e.g. by Cowan and Li [7] are not applicable for single user scenarios as depicted in Figure 1, though, as the fingers are very close to the projector and this leads to very large shadows occluding a large area of the projection.

In the first study we aimed for investigating how well simple pointing tasks and target selections can be performed on the projection from a projector phone. Remote pointing has extensively been researched on large *fixed* projections. The *Pointable* facilitates remote interaction with distant targets on large tabletop displays through perspective pointing and ray-casting [1]. Pointing on vertical displays has been researched in regard to the influence of effects like parallax and control type, different ray pointing techniques [12] and different devices like laser pointers [17] or bare hands [20] including various mid-air selection techniques [1,20]. However, the findings from this research strand can only partially be applied within the context of personal projectors as the mobile scenario is substantially different: The projection is constantly moving with the device. The user has to hold the projector phone during the whole interaction, which introduces jitter to the projection and the interaction, limits the possible movement area per hand and makes the interaction bimanual by nature. Further, mobile users usually do not want to carry or use additional hardware like a laser pointer or air mouse, why mobile interaction techniques have to get by with the user's bare hands. Since interaction happens in unaltered environments, the gaze of the user cannot easily be made available why image-plane ray-pointing techniques are unpractical and pointing must usually be based on the relation to the projector alone.

Remote mid-air pointing nevertheless has shown good performance why we decided to compare the usage of the touchscreen as a *touchpad* against mid-air pointing techniques. Both have much potential to enrich interaction in various situations and do not interfere with the projection.

INTERACTION TECHNIQUES

The aim of our research is a first exploration of mid-air pointing for projector phones. Since the area around the user is quite large, we considered different spaces around the device including behind, before, above, below, and to both sides of the device. All spaces have substantial implications on the usability (cf. [11]) and technical feasibility of the approach.

Through two preliminary user studies we discovered that interacting in front of the projector phone is not a well-suited space. While this might work for projectors worn around the neck such as in [14], interacting with the right hand in front of the projector that is held with the left hand requires that the right hand must be held very far away from the body. Additionally, the shadow on the projection created by the finger close to the projection occludes large parts of the projection. In contrast, when pointing with the index finger *behind* the projector phone (Figure 1b) to control a cursor on the projection, the user does not interfere with the projection. Further, it might allow for a convenient posture as the user is able to rest the upper arm of the pointing hand on the upper part of the body. Also, this technique is more independent of the user's girth.

In contrast, interacting to the right side of the device (respectively left side for left handed users) as well as interacting above or below the device poses a more difficult challenge for a real implementation: the necessity for maintaining an input space that is planar to the projection surface ($x \times y$ in Figure 1) assumed, the device would require a depth camera facing to the side of the device, providing the finger's horizontal x-movement via depth sensing. Similarly, an upward facing depth camera would have to provide vertical y-movement via depth sensing for interaction above the device. Despite mobile depth sensing is gaining ground as recently presented by Omnitouch [9], commodity hardware with sufficiently precise depth sensing cannot be expected to become available on mobile devices soon. Therefore, we refrained from *specifically* designing input spaces above/below or beside the device.

Instead we wanted to learn and also lay more stress on participant's own likings for input (which could include any desired space around the device) following a user-elicited approach. Therefore, we derived the *group defined* technique from a separate previous assessment where we tracked pointing preferences of 27 people. Participants (7 female, 20 male) of this study were undergraduates with an average age of 23 and have not had any prior experience with our work. Each participant was asked to define their preferred input space by showing the pointing gestures they would perform when selecting the four corners of the pro-

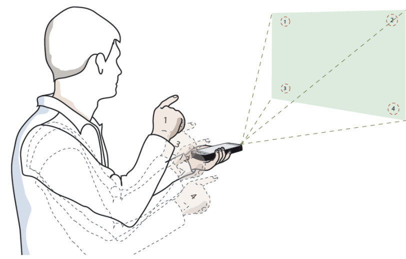


Figure 2. Pointing gestures performed for the definition of the input space for *group defined* and *user defined*.

jection by pointing at each corner three times while holding the projector phone (see Figure 2). Those pointing interactions were observed and measured by an optical tracking system. We calculated the average of those readings that led to an input space as specified in Figure 1c, which is on the top right side of the projector phone.

The *user defined* technique (Figure 1d) is similar to *group defined* but here only the currently tested user defines her preferred input space by pointing three times at the four corners of the projection. *Group defined* and *user defined* follow the approach described by Nielsen et al. [16] in which users show how they would perform a certain interaction. As each user defined a different input area there was no common input space for *user defined* that we can provide. However, the average of users chose a 16.0 cm in width (*SD*=7.5) and 14.2 cm in height (*SD*=7.3) interaction space with its center lying at $\{6.8\text{ cm x}, 20.3\text{ cm y}, 9.7\text{ cm z}\}$ (*SD*=9.8) away from the phone. In terms of size this would be similar to *behind* whereas the position would rather resemble *group defined*. Based on the different input sizes, the four techniques have slightly different control-display (C-D) gains. However, findings of Casiez et al. [5] indicate that C-D gain has a less important impact in studies modeled after Fitts' Law which our results will confirm as *behind* and *user defined* involved a very similar C-D gain but yielded significantly different results. Moreover, *touchpad*, in spite of having the highest C-D gain, was the slowest technique.

The actual selection of a target shown on the projection is in *all four* interaction techniques performed by a tap on any position of the touchscreen of the smartphone.

FIRST EXPERIMENT: TARGET SELECTION

The main goal of this experiment was to investigate whether finger pointing based techniques (controlled and user-elicited types) provide a similar performance in terms of target selection times and error rate when compared with *touchpad*. In addition, the experiment should clarify whether users perceive these techniques as beneficial. This experiment compares the previously described four techniques through a two-dimensional target selection task based on the ISO 9241-9 tapping test.

Participants

12 right-handed participants (6 female) took part in the experiment and were rewarded 10€ afterwards. All were undergraduate students and aged between 15 and 27 (mean = 23 yrs.). Their academic backgrounds were humanities, economics, and computer science.

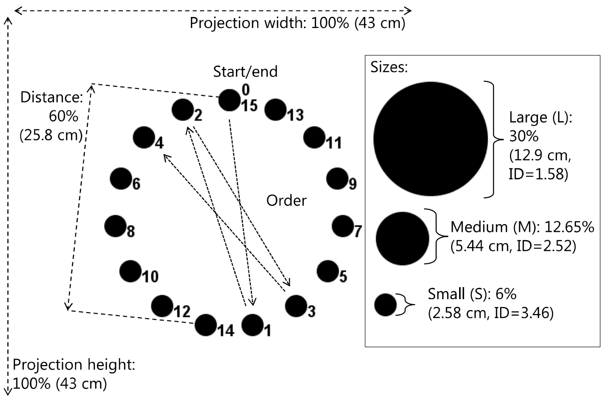


Figure 3. ISO 9241-9 task. Visualization of size and height of projection in relationship to the three target sizes.

Experimental Design

The experiment used a within-subjects design, i.e. all participants participated in all conditions of the experiment (counterbalanced). The first independent variable *technique* contained four levels: *touchpad*, *behind*, *group defined* and *user defined*. The second independent variable *size* of targets contained three levels: small, medium, and large (see Figure 3). The smallest target size was defined through a preliminary test were we looked for the smallest size that could be comfortably selected with *touchpad*.

Prototype and Setup

We assume that the three finger pointing based interaction techniques can be realized through an additional camera on the bottom and / or side of the projector phone. Corresponding algorithms and approaches like coarse-grained depth tracking for background removal or IR-camera sensing have been reported previously, e.g. [9,14,21]. We used an external optical tracking system (OptiTrack V100:R2, 100Hz from NaturalPoint) and infrared markers attached to the user's finger and the projector phone in order to support accurate tracking of the index finger in relation to the phone (Figure 4). With this approach it is possible to compare the interaction techniques independently from a potentially inaccurate tracking solution. A SHOWWX pico laser projector from Microvision connected to a Samsung Galaxy S was used as no projector phone has been commercially available in the country where the study was conducted.

The software used for conducting the study was written in Java and executed on the Android phone. Apart from running the study tasks and logging phone properties such as acceleration sensor values, the software also performed the pointer calculations based on the input from the tracking system in real-time. Pointer movement worked instantly without any noticeable delay. For the *touchpad* technique we implemented pointer acceleration similar to the algorithm used in Microsoft Windows [24]. Thus, and because the screen size was notably bigger than the farthest distance between targets, clutching was not required with *touchpad* in the first study.

Procedure

The experiment was conducted in a light dimmed laboratory room. The position of the participant was marked with an X on the floor, facing a wall 100cm away, resulting in a projection screen size of 43x43cm (see Figure 3 and 4). A user standing in front of a nearby wall is considered as typical scenario for mobile usage of projector phones. Participants were asked to stay at this location throughout the study. Participants were holding the projector phone with their left (non-dominant) hand (see Figure 1 and 4), and pointing with the other hand. Participants were allowed to freely move the projector and their finger as only their spatial relation defined the position of the pointer.

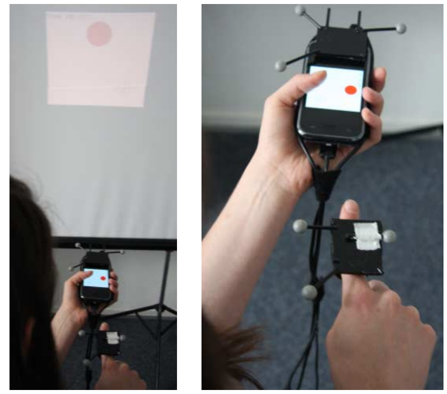


Figure 4. Study setting (for *behind*) and used hardware (participants did not look at phone display).

Participants took part in the study individually. Initially, to define the *user defined* technique, the participant was instructed to point three times at each corner of the projection as they would want to point at them in the subsequent experiment. After that the experimenter explained and demonstrated the four interaction techniques and asked participants to rate each interaction technique on a 10-point Likert scale (1 – very bad to 10 – very good) based on their sole expectation. Then, each technique was tested with three different target sizes. For each of the 12 possible combinations of interaction technique × size the participant performed 1 test and 3 consecutive study rounds, each including 15 targets (see Figure 3). In each round, the user started with a click (tap any position of the phone screen) on the circle in the middle, then went to the first circle at the top from which on the time taken to every subsequent target was measured. After the user's click the target turned to green (hit) or red (miss) for 150ms and only one trial per target was allowed. After each round the user was shown their time taken and the percentage of hit and missed targets. In addition we logged hit locations and jittering of the phone using the built-in accelerometer. After each technique, participants rated the technique regarding perceived speed, precision, satisfaction, difficulty and fatigue. Finally, participants were asked to rate each interaction technique again on the 10-point Likert scale from before based on their actual experiences.

Results of first experiment

Movement times and error rates measured are depicted in Figure 5 and 6. Movement time (*MT*) is defined by the duration between the occurrence of the target on the projec-

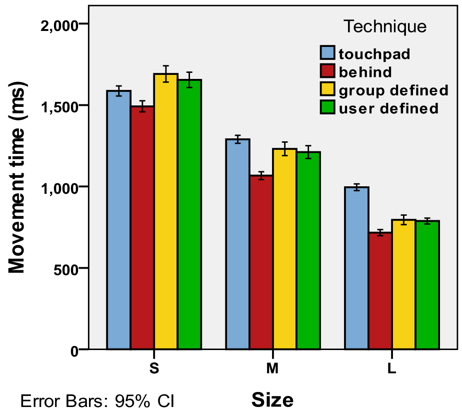


Figure 5. Movement times.

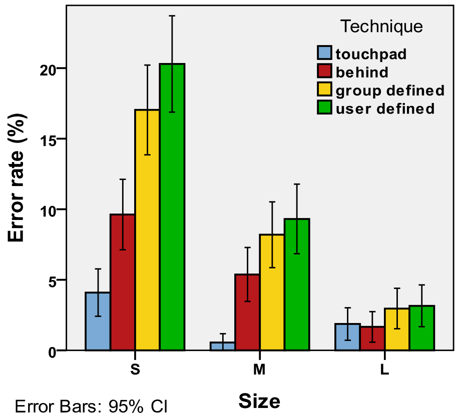


Figure 6. Error rates.

		Movement time	Error rate
Technique	η^2	161.850	162.925
	$\varepsilon_{G.-Geisser}$	.864	.857
	df	2.591	2.570
	df_{Error}	1386.343	1374.708
	F	73.587	38.147
Size	η^2	52.744	106.886
	$\varepsilon_{G.-Geisser}$	.914	.846
	df	1.828	1.693
	df_{Error}	978.017	905.705
	F	2251.249	95.670
Tech. x size	η^2	445.843	489.437
	$\varepsilon_{G.-Geisser}$	.767	.777
	df	4.603	4.663
	df_{Error}	2462.472	2494.854
	F	20.886	10.029

Table 1. ANOVA and post-hoc analysis of measured data.

tion and the selection of the target. An error is defined as click outside the target area. The results reveal *behind* to require a 15.4% shorter average movement time than *touchpad* when considering all sizes. The results also reveal a 2.55 times lower error rate of *touchpad* compared to the second best error rate of *behind* that we will discuss later.

Movement times (MT) and error rates (ER) were analyzed using a factorial repeated-measures ANOVA. Since sphericity had been violated for all effects, degrees of freedom were corrected using Greenhouse–Geisser estimates of sphericity (Table 1). According to this the main effects and the interaction effect were reported as significant ($p < .001$). The main effect *technique* and the interaction effect *technique* × *size* (split by *size*) were further post-hoc analyzed using pairwise comparisons of means with Bonferroni correction (for 6 and 18 comparisons respectively):

- $MT \times technique$: There were significant differences in movement time between all techniques ($p < .01$) except for *group defined* vs. *user defined*. Hence, users performed fastest with technique *behind* and slowest with *touchpad* ($MT_{touchpad} = 1291ms$, $MT_{behind} = 1092ms$, $MT_{group\ defined} = 1239ms$, $MT_{user\ defined} = 1217ms$).
- $ER \times technique$: The error rate significantly differed ($p < .001$) between all techniques except for *group* vs. *user defined*, revealing that users made the most errors with *group* and *user defined*, less with *behind*, and least with *touchpad* ($ER_{touchpad} = 2.2\%$, $ER_{behind} = 5.6\%$, $ER_{group\ defined} = 9.4\%$, $ER_{user\ defined} = 10.91\%$).
- $MT \times technique \times size$: No significant differences were found between *group* vs. *user defined* (S/M/L), *touchpad* vs. *group defined* (M) and *touchpad* vs. *user defined* (M). *Touchpad* vs. *user defined* on sizes S, M were reported significantly different ($p < .05$). Remaining differences were reported as significant ($p < .01$).
- $ER \times technique \times size$: On target size S, all pairs revealed significant differences ($p < .01$) except *group defined* vs. *user defined*. On size M, only *touchpad* vs. all other techniques showed significant differences ($p < .01$). Size L revealed no significant differences.

For further evaluation of the results we used the Fitts' Law model and calculated throughputs (TP) as described in [19,23]. First all measurements of the circular tapping task were rotated to horizontal 0º and 16 of 6480 targets (0.25%) were filtered out as spatial outliers. Then we calculated the effective index of difficulty (ID_e) individually for each subject and condition (technique and target size) based on the users trials (successful or not) over all test rounds of the condition (3 rounds × 15 targets) using equation

$$ID_e = \log_2(A_e/W_e + 1) \quad (1)$$

where A_e is the average actual movement distance over all rounds for a particular combination [19] and W_e reflects the standard deviation of endpoints as

$$W_e = 4.133 \times SD_{x,y} \quad (2)$$

where $SD_{x,y}$ is the bivariate endpoint deviation calculated as the spread of hits $\langle x_i | y_i \rangle$ around the center of mass $\langle \bar{x} | \bar{y} \rangle$

$$SD_{x,y} = \sqrt{\frac{\sum_{i=1}^{n} \left(\sqrt{(x_i - \bar{x})^2 + (y_i - \bar{y})^2}\right)^2}{N-1}} \quad (3)$$

Having ID_es for each subject, technique and target size, we calculated the individual throughput for each subject and technique using the mean-of-means approach [19], and the grand throughput by averaging individual throughputs. The grand throughputs, depicted in Figure 7, show a similar picture as the movement times. *Behind* outperformed other techniques, especially showing a 28.5% higher TP than *touchpad*. Our measured throughput of 1.957 for touchpad is in line with measured throughputs of traditional touchpad usage in the literature, which agrees on values between 0.99 and 2.9 [19]. As pointing on movable displays has not been studied before we can relate the throughput of *behind* to fixed pointing only. The *fixed-origin* pointing described by Jota et al. [12] shares with *behind* the similarity that the pointing ray depends on the user's finger *and* another point in space, which albeit is fixed. They measured throughput of ~3.4b/s for *fixed-origin* pointing – for one-dimensional tasks only, though. In this light, the throughput of *behind* pointing might be slightly smaller than similar pointing on fixed projections, which can be explained by the increased complexity of the bimanual control.

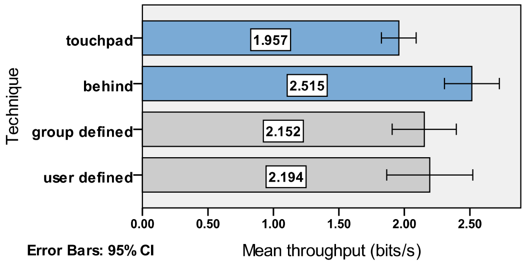


Figure 7. Grand throughput of interaction techniques (since *group* and *user defined* yielded a comparably high error rate close to or above 10% their calculated throughput values may be less meaningful).

A factorial repeated-measures ANOVA on throughput revealed a significant main effect of *technique* ($F_{3,33} = 6.219$, $p < .01$, $\eta^2 = 7.104$). Post-hoc pairwise comparisons with Bonferroni corrections showed no significant differences except for *touchpad* vs. *behind* ($p < .001$). Finally, we created Fitts' Law models of the form

$$MT_{Technique} = a + b \cdot ID_e$$

using linear regression. The average model fits (Pearson *r*) and parameters (*a*, *b*) are given in Table 2 and fit the measured results well: In particular, it shows the lower initial time required to start moving in mid-air (*a* of *behind*) as well as the smaller slope *b* of *touchpad* that indicates faster movement on the touchscreen for targets with a higher *ID*.

Technique	*a*	*b*	*r*
Touchpad	442.33	333.5	.953
Behind	-222.17	492.67	.937
Group defined	-735.42	795	.862
User defined	-563.25	741.58	.938

Table 2. Fitts' Law parameters and model fits.

After the study we asked participants to rate each interaction technique again on the 10-point Likert scale (1 – very bad to 10 – very good interaction) from before. Here *touchpad* performed best (average rating prior experiment 7.31, after the experiment 7.38) directly followed by *behind* (6.31, 7.31) that increased an entire point. Conversely, decreasing differences were found between *user defined* (7.46, 6.31) and *group defined* (6.46, 5.85).

We collected participants' ratings (Likert scale 1 – 7) after completing the tasks with each technique. Participants rated perceived speed, accuracy, fatigue of different body parts and selected questions from the Nasa TLX [10]. Participants' feedback delivered an overall similar picture to quantitative results in terms of speed, precision [*touchpad* (*MED* 6) slightly better than *behind* (*MED* 5)], difficulty and user satisfaction. As expected, overall fatigue was the lowest and almost non-existent with *touchpad*. *Behind* was rated the second best on fatigue scales overall (left/right finger, hand, wrist, and shoulder) but was rated one point worse than other pointing techniques for left hand, left arm, and left shoulder fatigue – the body parts involved in holding the projector. This can be explained by the fact that the projector had to be held slightly further away from the upper body to not interfere with the pointing right hand. The latter is supported by our analysis of phone jitter (Figure 8) which shows *behind* to cause the highest jitter among the mid-air pointing techniques. For right-sided body parts fatigue was rated as almost non-existent in contrast.

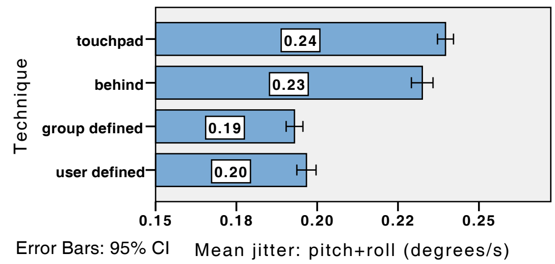


Figure 8. Mean phone jitter (pitch+roll) measured as the sum of differences over time.

Discussion of first experiment

Contrary to our initial expectations, the experiment revealed a significant difference between *behind* and the other techniques. The difference between *behind* and *group defined / user defined* can mainly be explained by the fact that the independent group of 27 people who provided the information for the input space of *group defined* and the participants of our study preferred on average an area on the right top side of the projector phone. Users seem to choose this area because it allows them to move the right arm freely, unrestricted by the upper body or the projector phone. The negative implication of this area is that upper arm, lower arm, and finger have to be controlled simultaneously. Based on our results it seems that most participants were not able to control the attitude of their pointing arm exactly and steadily enough in those two interaction techniques. This caused pointing jitter, inaccurate pointing and arm fatigue.

When using the *behind* technique participants were able to rest their upper arm of the pointing hand on the upper part of their body. Therefore, they had to control only their lower arm and index finger, which allowed accurate and steady pointing and led to lower arm fatigue. The results show that those advantages outweigh the disadvantage of *behind* that is the slightly limited input space. For instance it is more difficult to select areas on the bottom-left of the input space (see Figure 9), especially for corpulent and female users.

Compared to *touchpad* the interaction technique *behind* has the significantly lower movement times because the user needs less time to start moving in the air whereas *touchpad* requires to place the finger on the screen and overcome the initial resistance on the surface. However, *behind* is more vulnerable to errors for small and medium sized targets because of hand jitter and arm fatigue. This is less an issue with *touchpad* because it is easier to brake or rest the finger on the touchscreen surface. In real usage scenarios it will likely depend on the type of application whether the faster movement time or the higher error rate will have the higher impact. For instance, *behind* will likely perform worse than *touchpad* for text entry on the projection as errors are very

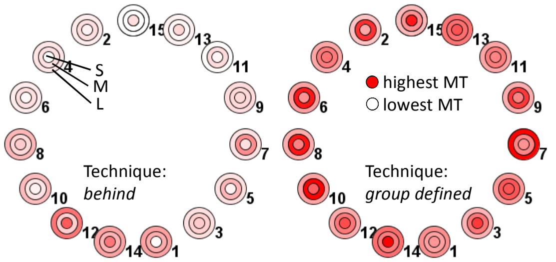


Figure 10. Target heat maps of *behind* and *group defined* averaged over all users showing movement times (MT) for the target sizes (S, M, L) and the overall benefit of *behind* despite the problematic area at the bottom left.

frustrating for the user in this scenario. During browsing a website on the other hand, being 15% quicker in general might easily compensate for missing every 18^{th} link (5.6% error rate). Furthermore, if the application made good use of the dual-display setup enabled by mid-air techniques like *behind*, e.g. a browser showing an overview of open tabs on the touchscreen and the currently active tab on the projection, the user interface could benefit further in terms of speed, clarity and user satisfaction.

SECOND EXPERIMENT: APPLICABILITY

Before we can study dual-display mobile applications with projector phones, though, we need to test how the superior mid-air technique *behind* compares to the *touchpad* technique in real world application scenarios with *unaltered* mobile applications. We further added the standard mobile touchscreen usage as third technique that would allow participants and us to distinguish between the impact of the projection and the interaction technique.

Participants

For the second experiment we recruited the same 12 participants from the first experiment to ensure they had the same amount of practice with the projection techniques. All of them owned a laptop and were hence familiar with touchpads and all but P4 and P7 (10 of 12 participants) owned a touchscreen phone themselves (only a few featured multi-touch or VGA+ resolution screens, though).

Experimental Design

The second experiment comprised two independent variables *technique* and *application*. Techniques consisted of *touchscreen* (the application was used on the mobile touchscreen without projection), and *touchpad* and *behind* (application was used on the projection, controlled via a cursor). We decided to test three specific applications that are likely to benefit from the larger projection or the different input technique instead of fielding the projection in tasks that are optimized for and advantageous (like private text entry) on the touchscreen. The three applications (Figure 10) and reasons for choosing them were as follows:

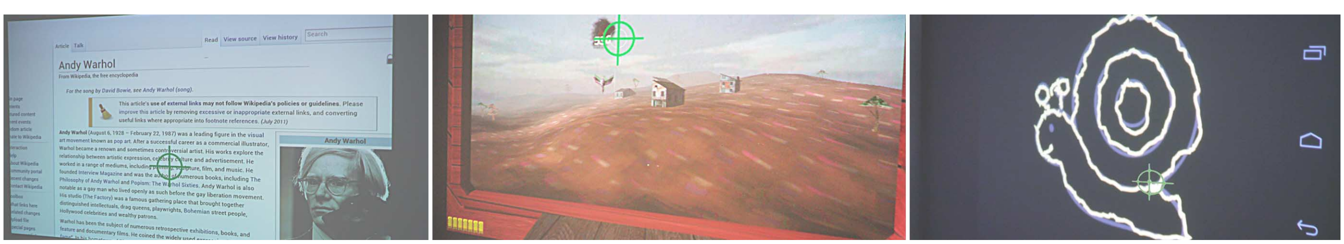
1. Browsing

Browsing has become one of the most common tasks performed on smartphones. With mobile phones reaching display resolutions comparable to laptops, websites can be used in “full site” or “desktop view” mode instead of their usually very restricted mobile versions. However, due to the small physical display size, this requires several zooming and panning operations by the user. In contrast, on the projection even small text can easily be read without zooming.

We used the standard Android browser in full screen mode (Figure 10a) in all three techniques. The participant always started with a Wikipedia article about San Francisco. Starting on this web page the experimenter asked the participant to follow one of three predefined paths (counterbalanced). On every path the participant had to scroll down to the table of contents of the article, and then navigate to one of three predefined sections (e.g. museums). Then, the participant had to perform twice: following a link (e.g. to the Wikipedia article of the Museum of Modern Art) and finding a certain piece of information (e.g. when the museum was established). All tasks required roughly the same amount of scrolling, reading, clicking and time.

2. Gaming

As games are ultimately diverse we acknowledge that a single game cannot be representative. However, it can provide a preliminary sense for a particular group of mobile gaming applications. Since a shooting game resembles much of the Fitts’ Law tapping task, yet in a completely different setting, the popular app “Drunken Hunting” seemed to be a reasonable candidate. The goal in this game (see Figure 10b) is simply to shoot flying ducks by touching or pointing and clicking on them respectively. In contrast to other similar simple shooting games, it features targets at different sizes with shooting smaller ones yielding more points than larger ones. We anticipated that smaller targets would be easier to see and hit while displayed on the (large) projection than on the (small) screen because of the bigger size and the eliminated fat-finger problem. Every participant played two levels with each level comprising 10 shots.



a) Fullscreen Android Browser **b) Drunken Hunting game** **c) Painting (post-painting outlines)**

Figure 9. Applications used in the 2nd experiment.

3. Painting

With mobile phones taking over increasingly more traditional PC tasks, accurate pointing *and* steering gains importance. Painting combines both requirements very well and the huge number of downloads of painting applications in the app stores shows their increasing distribution. One obvious problem with painting, though, is the lacking accuracy caused by the fat-finger problem. With this application we want to research if the usage on the projection with the presented techniques increases the accuracy during the task.

We used the – at the time of writing top ranking – application "Paint Joy" from the Android market store. The task of the participant was to post-paint the outlines of a snail with house (Figure 10c). This image was chosen because it combines horizontal, vertical, and circular lines – the basic subset of every more complex painting task.

Prototype and Setup

Hardware Setup

For the second experiment we employed a different prototype as we wanted to maximize the user experience of different display sizes between phone and projected display. We therefore used a Samsung Galaxy Nexus Android phone featuring 720p HD resolution. This phone has the highest physical display size and resolution currently available on mobile phones. Hence it seemed to be the strongest competitor against a projection. Similarly, we wanted to provide a large, bright, and high quality projection. Since none of the currently available pico projectors supported HD resolutions or a brightness beyond 50 Lumens, we opted for the palm-sized projector Qumi Q2 from Vivitek. This projector provides the same 720p HD resolution and a brightness of 300 Lumens while still only weighing 617g. Phone and projector thus weighed 742g together. While this was possible to hold in one hand and use for a short time, we decided to additionally uphold the projector from a rod affixed to a tripod moving freely in all directions. Thereby we equilibrated the weight of the projector to some extent, but it still had to be uphold and steered by the user as it would have without equilibration (Figure 11b). The phone was attached in landscape mode (Android's default when connected to a projector) on a flexible plastic attached to the bottom of the projector. This construction allowed the user to hold the "projector phone" with one hand in *behind* mode (Figure 11a) and two hands in *touchpad* mode (Figure 11c). Participants could stand and hold the device comfortably while looking on an almost leveled projection, yet were required the typical balancing to preserve the levelness and position of the projection and cope with hand jitter as with a real projector phone.

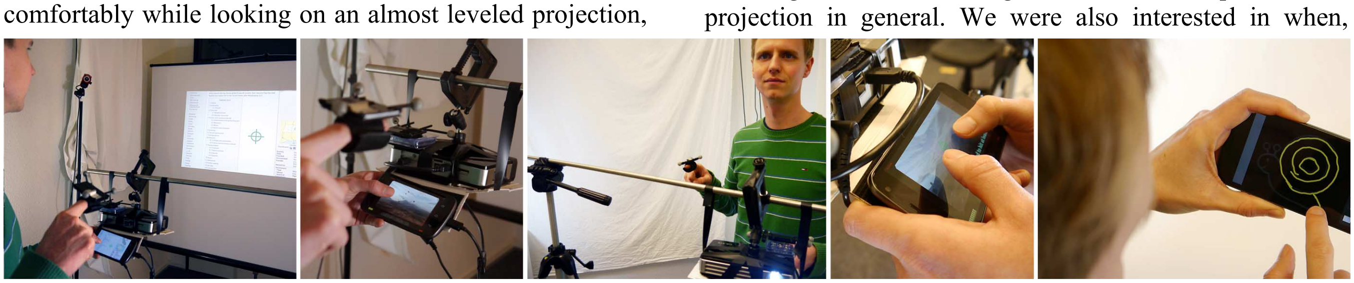
Software

The pointing software was realized as an Android background service, which showed a shiny green cursor on top of all other Android windows and applications and intercepted all user touch events. Our background service processed these events and based on the current mode of interaction (*touchpad* or *behind*) sent them as new touch events to Android's input system. The latter was accomplished using Android's built-in monkey service, which we hijacked on our rooted device to send arbitrary touch events to the system. Additionally we attached to the native Linux events from the touchscreen. Overall, this gave us full control over Android's touch input handling to send our own events to the Android system and its built-in applications.

In both projector interaction modes clicking anywhere on the device resulted in a click at the current position of the cursor. In *touchpad* mode the cursor position was changed relatively to movement of the finger on the device (same as in the first experiment). Scrolling in the browser application and painting in the paint application was initiated with a double click from where on movement of the finger was passed through to the application until the finger was lifted up again (in *browsing* the cursor position remained fixed during scrolling). In *behind* mode the cursor was moved by moving the finger in mid-air just as in the first experiment. In this case, scrolling and painting was executed while the finger was down on the touchscreen, i.e. the website was "grabbed" with the left hand's finger and moved up or down by moving the right hand's finger in the air. The game only required positioning and clicking to shoot.

Procedure

We employed a within-subjects design as in the first experiment. Each participant tried each of the three techniques with each of the three applications (counterbalanced). Each application was used with each technique between 2 and 3 minutes. We followed a qualitative analysis approach that would reveal differences that have not become apparent in the first study. We instructed participants to think aloud during all interactions, which we recorded for later analysis. After having tried all 9 combinations we asked participants for their feedback about speed, accuracy, liking, joy of use, advantages and disadvantages of each technique and the projection in general. We were also interested in when,



a) *Behind* with browsing and gaming b) Device equilibration c) *Touchpad* d) Painting with *touchscreen*

Figure 11. 2nd study prototype and techniques in use.

where, and for which applications participants would favor using the projection over using the touchscreen alone.

Results

Overall, the projection techniques were liked much by participants and more fun to use than *touchscreen* as reported by 9 participants. Partly, this has to be attributed to the novelty effect. Nevertheless, it indicates a positive user experience with both projection techniques, albeit being highly dependent on the application type.

In the browsing task *touchpad* was perceived as slower than other techniques by at least four participants since *touchpad* required a double-click to initiate scrolling which four participants perceived to slow down the interaction. In contrast, *behind* was reported to be very fast (P4, P5, P9) and precise (P3, P6, P10) as was *touchscreen* (7 and 8 participants respectively) despite the required zooming and panning steps. Especially in this task, those participants owning a high class smartphone and therefore being trained on getting by with the small screen for browsing performed much better with *touchscreen* while for novice smartphone users both systems seemed to perform equally well.

In the gaming task participants scored most successfully with the *behind* technique, which also felt intuitive (P1, P2, P8), but also became more aware of the freehand pointing jitter. P8 and P9 said “it was difficult to keep still”. *Touchpad* was more affected by clutching than in other scenarios, as moving the pointer over long distances from a previous shooting target to the next required more than one movement. 4 participants said they felt constricted by the small touchscreen compared to the large projection (P6 said “I didn’t know where I was on the screen with my finger”). But *touchpad* was on the other hand perceived as the most precise (6 participants) since targets that didn’t move too fast could be hit more accurately than with any other technique. The *touchscreen* also performed fast but showed the problems that very small ducks could not be recognized on the small screen and that the finger occluded the targets at the expense of accuracy.

In the painting task 9 participants reported the fat-finger problem to hinder accurate painting on the *touchscreen*. Yet, *touchscreen* (Figure 11d) performed much better than *behind*, which was very unsatisfactory to use because of the comparably high jitter. Despite *behind*’s lower caused jitter compared to other mid-air techniques, the jitter is still too high for the technique to be qualified for steering tasks. Here, *touchpad* showed its huge advantage in that it, as P10 said, “combines the advantages of projection and touch-screen”, namely the elimination of the fat-finger problem on one side and the haptic affordance of the touchscreen on the other that improves precision.

Further comments, independent of application, included that *behind* is an interactive performance like playing Wii (P5), which can be liked or disliked (as by P3 and P10 in our case). 3 participants also stated they would like to perform the click in the air, too, which we had thought about before but decided to stick to bimanual input as this will likely be the standard use case in future projector phone interaction. With *touchpad* participants liked that it feels familiar from laptops (P1, P3, P7) and requires little space (P2, P4) as well as little effort (P3, P5, P10) and therefore is more versatile in its application than *behind*. But it also requires a lot of movement on the touchscreen surface, which got uncomfortable over time for P7 and P9.

Discussion second experiment

The second experiment has shown that mobile applications indeed can benefit from a mobile projection. Despite private or public media broadcasting and collaboration, the projection can even enhance unaltered mobile applications that originally have been designed for touchscreens. Further, the advantages of the projection are very co-dependent on the usage scenario and can for instance be very useful to overcome the fat-finger problem on touchscreens or to increase the visibility and ease the selection of small objects on the display. Based on these findings we predict that new application-specific interaction techniques that sensibly integrate touch and mid-air interaction on both displays will largely enrich the projector phone experience.

LIMITATIONS

Even with comfortable arm postures such as with the *behind* technique, mid-air interaction might lead to higher fatigue than traditional solutions. Luckily, mobile situations rarely entail longsome series of interactions. Further investigation of fatigue, especially on larger mobile projections, also in respect to different C-D gains that might affect speed, accuracy, and fatigue, is required. In our studies we did not experiment with different C-D gains: the lower bound of C-D gain was set by the physical size of the touchscreen that we did not want to exceed to maintain comparability. Higher C-D gains in contrast might have decreased the accuracy of mid-air techniques further.

When testing the applicability of the techniques in common mobile scenarios we did not include all mobile factors such as interacting on the go or sudden breaks. In contrast to touch input, the mid-air techniques forbid pausing of the cursor as long as the user’s hand is within the input area. Furthermore, we only evaluated existing mobile applications specifically designed for touch input. Studying applications designed for dual-display mid-air interaction will deliver further interesting results. Finally, people used to multi-touch performed the tasks of the second study quicker on the touchscreen, albeit acknowledging many advantages of the projection. However, the majority of participants were unacquainted with multi-touch for why we implemented *touchpad* interaction similar to laptop touchpads.

In terms of technical feasibility, the setup of both experiments used a precise external tracking system for position tracking during the mid-air techniques. A current implementation based on an embedded mobile camera with less hardware capabilities will likely not achieve the same accuracy/performance to the detriment of the mid-air techniques while touchpad performance could remain the same.

CONCLUSION

Projector phones raise various questions regarding their interaction design due to the large remote display, availability of various sensors, and movement of the projection. Most currently available commercial projector phones only mirror the information displayed on the phone display on the projection, involving frequent context switches and unsuited user interfaces. Using the *touchpad* interaction technique already provides a significant advantage as the user can focus primarily on the projection. But in real world scenarios the indirectness, the effect of clutching and the occupation of the screen diminish its applicability.

Our first study showed that more direct pointing using *behind* provides distinct advantages in terms of movement time and throughput when compared with *touchpad*, in particular when considering medium and large targets. The notably higher error rate of *behind* however makes it more suitable to application scenarios such as browsing and gaming and less to painting or text input. Interestingly, *behind* performed better than *group* and *user defined* although the latter two were gathered through a user-elicited approach.

The second study analyzed the user experience of *behind* and *touchpad* in relationship to conventional *touchscreen* usage. Here, we have seen that projection-based techniques (*behind, touchpad*) received overall equally good feedback as *touchscreen* despite not having been explicitly designed for the chosen applications. Furthermore, participants saw various disadvantages of touchscreens such as the fat-finger problem and numerous advantages of using a projection, e.g., for collaboration and application areas such as gaming.

Overall, our findings showed many advantages of the mid-air technique *behind*, which speaks for the integration of a camera at the bottom of the projector phone. Further, our research confirms that the interaction technique *touchpad* is an interesting option for future projector phones in particular as it comes for free. In our future research we will investigate *behind* further with dual-display mobile applications.

ACKNOWLEDGEMENTS

The presented research was conducted within the Emmy Noether research group “Mobile Interaction with Pervasive User Interfaces” funded by the DFG. We thank Markus Löchtefeld for helping proof-read the paper.

REFERENCES

1. Banerjee, A., Burstyn, J., Girouard, A., and Vertegaal, R. Pointable: an in-air pointing technique to manipulate out-of-reach targets on tabletops. *Proc. ITS 2011*, ACM (2011), 11–20.
2. Blasko, G., Feiner, S., and Coriand, F. Exploring Interaction with a Simulated Wrist-Worn Projection Display. *Ninth IEEE International Symposium on Wearable Computers (ISWC’05)*, 2–9.
3. Butler, A., Izadi, S., and Hodges, S. SideSight: multi-“touch” interaction around small devices. *Proc. UIST 2008*, ACM (2008), 201–204.
4. Cao, X., Forlines, C., and Balakrishnan, R. Multi-user interaction using handheld projectors. *Proc. UIST 2007*, (2007), 43.
5. Casiez, G., Vogel, D., Balakrishnan, R., and Cockburn, A. The Impact of Control-Display Gain on User Performance in Pointing Tasks. *HUMAN–COMPUTER INTERACTION 23*, 3 (2008), 215–250.
6. Cauchard, J.R., Löchtefeld, M., Irani, P., et al. Visual separation in mobile multi-display environments. *Proc. UIST 2011*, ACM (2011), 451–460.
7. Cowan, L.G. and Li, K.A. ShadowPuppets: supporting collocated interaction with mobile projector phones using hand shadows. *Proc. CHI 2011*, ACM Press (2011), 2707–2716.
8. Greaves, A. and Rukzio, E. Evaluation of picture browsing using a projector phone. *Proc. Mobile HCI 2008*, ACM (2008), 351–354.
9. Harrison, C., Benko, H., and Wilson, A.D. OmniTouch: wearable multitouch interaction everywhere. *Proc. UIST 2011*, ACM (2011), 441–450.
10. Hart, S.G. and Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology 52*, (1988), 139–183.
11. Hinckley, K., Pausch, R., Goble, J.C., and Kassell, N.F. A survey of design issues in spatial input. *Proc. UIST 1994*, ACM (1994), 213–222.
12. Jota, R., Nacenta, M.A., Jorge, J.A., Carpendale, S., and Greenberg, S. A comparison of ray pointing techniques for very large displays. *Proceedings of Graphics Interface 2010*, Canadian Information Processing Society (2010), 269–276.
13. McCallum, D.C. and Irani, P. ARC-Pad: absolute+relative cursor positioning for large displays with a mobile touchscreen. *In Proc. UIST 2009*, ACM Press (2009), 153–156.
14. Mistry, P., Maes, P., and Chang, L. WUW-wear Ur world: a wearable gestural interface. *27th CHI extended abstracts*, ACM (2009), 4111–4116.
15. Myers, B.A., Stiel, H., and Gargiulo, R. Collaboration using multiple PDAs connected to a PC. *Proc. CSCW 1998*, ACM Press (1998), 285–294.
16. Nielsen, M., Störring, M., Moeslund, T., and Granum, E. A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI. *Proc. GW 2003*, Springer (2004), 409–420.
17. Olsen,Jr., D.R. and Nielsen, T. Laser pointer interaction. *Proc. CHI 2001*, ACM (2001), 17–22.
18. Rukzio, E., Holleis, P., and Gellersen, H. Personal projectors for pervasive computing. *IEEE Pervasive Computing 11*, 2 (2012), 30 – 37.
19. Soukoreff, R.W. and MacKenzie, I.S. Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts’ law research in HCI. *International Journal of Human-Computer Studies 61*, 6 (2004), 751–789.
20. Vogel, D. and Balakrishnan, R. Distant freehand pointing and clicking on very large, high resolution displays. *Proc. UIST 2005*, ACM (2005), 33–42.
21. Wilson, A.D. PlayAnywhere: A Compact Interactive Tabletop Projection-Vision System. *In Proc. UIST 2005*, ACM Press (2005), 83–92.
22. Winkler, C., Reinartz, C., Nowacka, D., and Rukzio, E. Interactive phone call: synchronous remote collaboration and projected interactive surfaces. *Proc. ITS 2011*, ACM (2011), 61–70.
23. Wobbrock, J.O., Shinohara, K., and Jansen, A. The effects of task dimensionality, endpoint deviation, throughput calculation, and experiment design on pointing measures and models. *Proc. CHI 2011*, ACM (2011), 1639–1648.
24. Pointer Ballistics for Windows XP. http://msdn.microsoft.com/en-us/windows/hardware/gg463319.aspx.