

# Gaze+touch vs. Touch: What's the Trade-off When Using Gaze to Extend Touch to Remote Displays?

Ken Pfeuffer, Jason Alexander, Hans Gellersen

► **To cite this version:**

Ken Pfeuffer, Jason Alexander, Hans Gellersen. Gaze+touch vs. Touch: What's the Trade-off When Using Gaze to Extend Touch to Remote Displays?. 15th Human-Computer Interaction (INTERACT), Sep 2015, Bamberg, Germany. pp.349-367, 10.1007/978-3-319-22668-2\_27. hal-01599897

**HAL Id: hal-01599897**

**<https://hal.inria.fr/hal-01599897>**

Submitted on 2 Oct 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Gaze+touch vs. Touch: What's the Trade-off when Using Gaze to Extend Touch to Remote Displays?

Ken Pfeuffer, Jason Alexander, and Hans Gellersen

School of Computing and Communications, InfoLab21, Lancaster University,  
Lancaster, LA1 4WA, United Kingdom  
{k.pfeuffer, j.alexander, h.gellersen}@lancaster.ac.uk

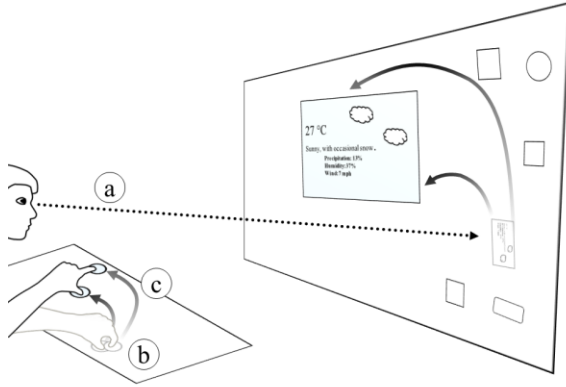
**Abstract.** Direct touch input is employed on many devices, but it is inherently restricted to displays that are reachable by the user. Gaze input as a mediator can extend touch to remote displays - using gaze for remote selection, and touch for local manipulation - but at what cost and benefit? In this paper, we investigate the potential trade-off with four experiments that empirically compare remote Gaze+touch to standard touch. Our experiments investigate dragging, rotation, and scaling tasks. Results indicate that Gaze+touch is, compared to touch, 1) equally fast and more accurate for rotation and scaling, 2) slower and less accurate for dragging, and 3) enables selection of smaller targets. Our participants confirm this trend, and are positive about the *relaxed finger placement* of Gaze+touch. Our experiments provide detailed performance characteristics to consider for the design of Gaze+touch interaction of remote displays. We further discuss insights into strengths and drawbacks in contrast to direct touch.

**Keywords:** Gaze Interaction, Eye-tracking, Multitouch, Multimodal UI.

## 1 Introduction

Multitouch gestures are now established on a variety of interactive surfaces such as phones, tablets, or tabletops, with much of their appeal based on the direct manipulation when touching the surface. It is of particular benefit to enable multitouch gestures on remote and large displays, as it is easy to learn, users are familiar with it, and users can easily manipulate content scattered across the large surface. However, the default touch is inherently restricted to interaction with direct surfaces in reach of the user. Although indirect interaction via a mediating mouse or touchpad can overcome this, the need for cursor dragging departs from the directness afforded by touch interaction.

A recent method to bring touch to remote displays is using the gaze modality as demonstrated in previous work ([21–25]). The control of Gaze+touch is simple: users look at the target on the remote surface for selection, and perform touch gestures on the close-proximity surface. For example, as illustrated in Figure 1: the user looks at the target (a), touches down on the close surface (b), then performs a gesture with the same touches to manipulate the target (c). Gaze is used to determine the target at touch down, after which 'touch' takes over and manipulates the target.



**Fig. 1.** Gaze can be used to bring multitouch to remote surfaces: (a) look at the target, (b) touch down, and (c) perform a multitouch gesture with the same touches.

From an input-theoretic standpoint Gaze+touch is a technique that ties characteristics of both direct and indirect touch together. Similar to direct touch, users can initiate manipulation the moment they touch down without prior cursor dragging as needed in common indirect touch techniques. Similar to indirect touch, users leverage a relaxed finger placement, use varying control-display gains, and avoid fat-finger and occlusion issues commonly associated with direct touch. For this reason, Gaze+touch should be considered as a hybrid technique rather than an instance of either direct or indirect category, which argues for a gradual characterisation of Gaze+touch through explorative evaluation as initiated in prior work [14, 21].

Yet, in light of real-world applicability, most present devices employ direct touch, making the similarities to direct touch particularly interesting. Considering a holistic perspective on 'multitouch', with regards to the entirety of direct manipulation rather than focusing on single gestures, Gaze+touch can be regarded as a method to bring all multitouch *as it is* to remote displays by using gaze as the mediator [21]. With this assumed, the question of trade-off becomes essential: what potential costs and benefits come with extending touch to remote displays? How does gaze selection affect the ease and familiarity of multitouch, and the quality of multi-finger and multi-hand interaction? Prior work compared Gaze+touch to head-movement based techniques [21, 22], to touch in a theoretical analysis [14], and between Gaze+touch variations for content transfer [24, 25] as well as large screens [23]. However, we are not aware of any work that empirically contrasted Gaze+touch to the default touch paradigm.

In this paper, we contribute four experiments that each compare remote Gaze+touch to standard touch interaction. The overall research question is what costs and benefits come with making touch indirect by gaze. The experiments regard task completion time, accuracy, and user feedback using the techniques. As 'multitouch' includes a broad range of gestures, we focus on the following commonly used gestures. Each gesture corresponds to one experiment:

1. Single-touch dragging of objects.
2. Two-touch rotation of objects across different sizes.
3. Two-touch rotation of objects of different orientations.
4. Two-touch scaling of objects.

Our experiments resulted in the following findings:

- **Completion Time:** Gaze+touch is as fast as touch for rotation and scaling, but slower for single-touch dragging.
- **Accuracy:** Gaze+touch is more accurate for rotation and scaling, but less accurate for dragging.
- **Fat-Finger Problem:** Gaze+touch allows to select smaller targets than touch.
- **User preference:** Gaze+touch is preferred for rotation and scaling, and touch is preferred for dragging.

These findings depict strength (e.g. accuracy) and shortcomings (e.g. dragging) that can be taken into consideration when designing Gaze+touch based user interfaces. While further validation is needed with real-world applicability and advanced technique design, our studies represent first groundwork in how Gaze+touch competes with and sets itself apart to standard touch interaction.

## 2 Related Work

The related work can be considered from two perspectives: research that evaluated gaze techniques, and research that evaluated touch techniques. Accordingly, we investigate the intersection of the domains by contrasting Gaze+touch to touch.

### 2.1 Evaluation of Gaze-based Interaction

Gaze interaction is considered as fast and natural, but suffers from the Midas Touch problem (false positive activation of tasks, [9, 10, 22]). Early work therefore experimented using gaze pointing with dwell-time or button-press selection ([9, 27]). Sibert and Jacob compared gaze with dwell-time against mouse; resulting in gaze being faster than the mouse [19]. MAGIC uses gaze to replace most of the pointing of a manual device [29]. They compared MAGIC to a trackpad and found MAGIC to be faster while reducing physical effort and fatigue.

Touch received increased attention in recent time as a partner for gaze pointing ([21–25]). As touch is prime input for smartphones and tablets, it is particularly useful for interaction over distance where users point by gaze on remote displays, and confirm by touch on their local device. Stellmach and Dachsel first showed that Gaze+touch can improve selection accuracy that gaze-only usually lacks [21]. In a later work they showed that this combination also allows for dragging, scaling, and rotation of targets on distant displays [22]. In both works, Gaze+touch was compared to head based pointing techniques, and indicated performance benefit for using Gaze+touch. Turner et al. also investigated Gaze+touch with handheld and remote display, focusing on content transfer across devices ([24, 25]). They studied transfer techniques based on gaze pointing with varying touch actions, showing general user acceptance and the importance of visual feedback and eye-hand coordination. Turner et al. also investigated gaze-supported rotate, scale, and translate (RST) gestures on

large screens [23], indicating how subtle differences in the design of Gaze+touch techniques affect remote RST interaction. Pfeuffer et al. investigated Gaze+touch for direct, in-reach surfaces [14]. In their design space analysis, they theoretically discussed Gaze+touch in contrast to touch, arguing that using gaze can avoid typical pitfalls of touch, such as occlusion or the fat-finger issue. Our work is complementary in providing an empirical comparison of remote Gaze+touch to standard touch.

## 2.2 Comparative Studies with Direct Touch

Kin et al. compared touch on a tabletop to an indirect setting of input by mouse and output on a desktop screen [11]. The user study showed that touch can improve performance in a multi-target task, when more than one finger is used for selection. Forlines et al. also compared touch to mouse on the same tabletop device [6], indicating that touch is more appropriate for multi-finger tasks and for multiple users but less for single cases.

Cursor-based indirect touch based techniques were developed based on offset [16], multi-point [1, 3], and bimanual input [4, 28]. Potter et al. evaluated offset-based indirect touch to direct touch and found accuracy improved. Benko et al. compared indirect touch mice against direct touch [3]. They found that direct was faster as it allows explicit touch activation and resembles a single focus of interaction, contrasting the implicit cursor selection and dragging necessity. It is unclear whether these results apply to Gaze+touch since users can, like direct touch, manipulate the moment of touch down.

Other researchers proposed coupled drag gestures or consecutive touches to select small targets and increase mode selection [2, 7]. However, these techniques occupy specific touch gestures and require learning of additional techniques.

Indirect touch techniques without cursor dragging were developed by mapping touchpad and remote display 1:1 ([12, 13, 18, 26]). Schmidt et al. compared this indirect touch variant (input on horizontal surface, output on vertical surface), to direct touch (input and output on horizontal surface, [18]). Results showed lower performance with indirect touch, because users had difficulties to keep hands hovered over the surface while finding and selecting targets. With Gaze+touch, in contrast, users can touch down anywhere because selection is offloaded to gaze, eliminating the need to move and hover hands for long distances.

In summary, a variety of indirect touch techniques were developed in HCI literature and compared to direct touch as a baseline. Gaze+touch can be considered as a hybrid technique that inherits indirect characteristics without the drawbacks of cursor dragging nor absolute mappings. This makes it suited to transfer whole 'multitouch' to remote displays. Collectively, these reasons motivated us to take a more detailed look on this transfer, and to conduct a comparative study of Gaze+touch to touch.

### 3 Method and Design of the Experiments

**System.** The system (Figure 2) consists of a touch display mounted at  $30^\circ$  close to the user (Acer T272), a large remote display (Samsung SUR40), and an eye tracker (Tobii X300). Both screens' resolution is 1080p. The close-proximity screen supports capacitive touch input for up to ten fingers, and the eye tracker uses 60 Hz gaze tracking. The close-proximity screen is used for input and output of touch, while Gaze+touch uses it only for touch input (with visual output on the remote screen). The software runs on the SUR40 PC (2x2.9 GHz CPU, 4 GB RAM), implemented in Java using the MT4J framework (<https://code.google.com/p/mt4j/>, 18/01/2015).

**Eye Tracking Accuracy Mechanisms and Visual Feedback.** We implemented three mechanisms to cope with issues of eye tracking hardware (e.g. imprecise tracking, eye jitter, and calibration offset, see [10, 21]). First, we calibrate each time when users look at a target, but cannot select the target at all because of inaccuracy. We use Pursuit Calibration, a calibration method more flexible than standard procedures [15]. The calibration duration was 10 seconds, and on average users performed two to five calibrations during the study. Second, to cope with eye jitter, we average gaze samples for 150ms when only short eye movements occur ( $<2^\circ$  of visual angle). Third, a gaze cursor appeared after the user fixated on a point for 1 second and did not select a target during that time. Not being able to select a target is noticeable, as usually the target is highlighted yellow when users looked it. The appearing cursor allowed the user to reposition the cursor by slight head movement if the system's gaze estimate was slightly offset preventing target selection (similar technique to Look and Lean [20]), yet allowed cursorless interaction for the majority of the time.

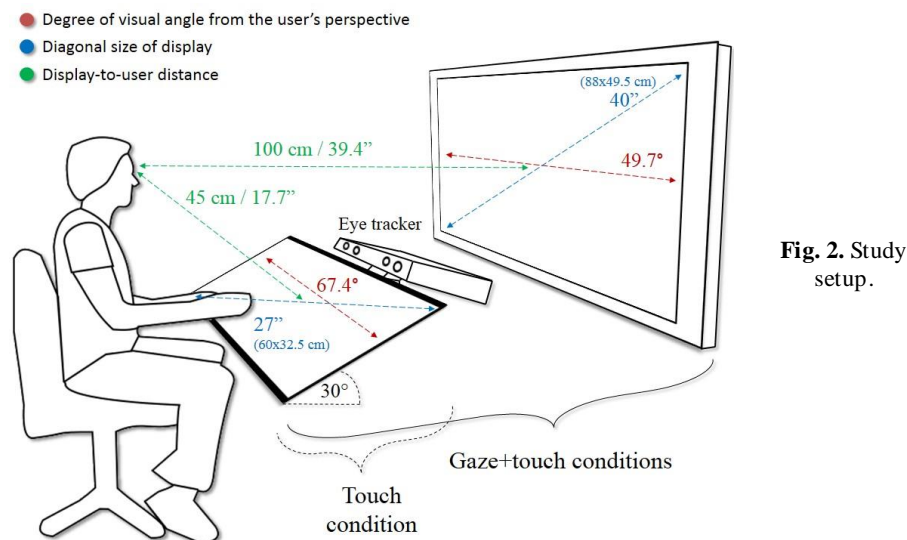
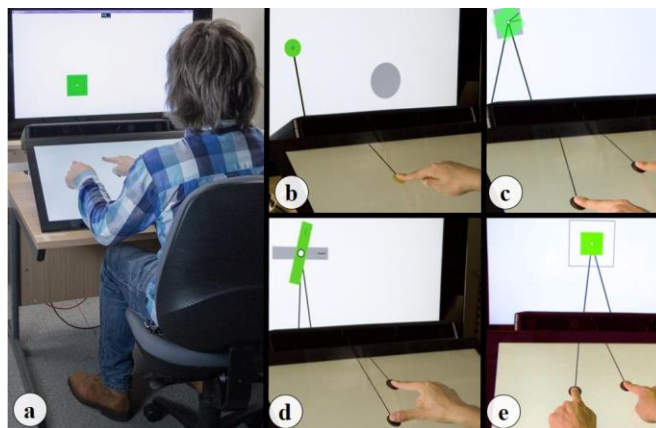


Fig. 2. Study setup.

**Visual Angle as a Size Metric.** To normalize measures of the close-proximity and the remote surface, we use degree of visual angle as a metric so that targets appear as the same size from the user's view. In the Touch condition, users were approximately 45cm in front of the screen's center. In the Gaze+touch condition, users were approximately 100cm from the screen (Figure 2). Thus absolute measures have a different size from the user's view. To normalize pixel measures, we use degree of visual angle from the user's perspective as a metric for distance and size. For example,  $1^\circ$  of visual angle represents 25.1px on the direct condition, and 36.2px on the Gaze+touch condition. Figure 3 shows exemplary task designs for the Gaze+touch case:  $3^\circ$  (b),  $2^\circ \times 10^\circ$  (d), or  $4^\circ$  (c and e).

**Study Procedure.** After completing a consent form, participants were given an introduction to the study, and then conducted the four experiments. They were instructed to perform each task as quickly and as accurately as possible, with speed taking priority. Before each experiment  $\times$  technique block, the experimenter explained how the interaction technique is used, followed by a first trial with assistance. Each block was repeated five times. Notably, the whole first block was training (with assistance when necessary), and excluded from the final results. Each session lasted approximately 95 minutes.



**Fig. 3.** Example user setup (a) and design of each experiment: dragging (b), rotation across sizes (c), rotation across orientations (d), and scaling (e). Lines show the mapping between finger and object (not visible during study).

**Task Procedure.** Each task begins when the user touches the required number of fingers on the centre of the touchscreen, and (for Gaze+touch) looks at the centre of the remote screen. Based on a similar object dragging study [22], the object's starting position was always placed toward a random screen corner. For rotation and scaling tasks it was  $10^\circ$  from the screen's center; dragging tasks involved specific distances (described below). The user completed a task by pressing a button, although the actual finishing time is taken as the last time the user manipulated the object. After each experiment  $\times$  technique block, the participant filled out a questionnaire. The categories were rated on a scale between 1 (strongly disagree) and 5 (strongly agree), and included 6 categories: ease of use, speed, accuracy, learning effort, eye tiredness, and

physical effort. In addition, after each experiment users selected their preference of technique.

**Experimental design.** We used a within-subjects experimental design. Each user conducted the four experiments sequentially as presented in this paper. Within each experiment, the tested technique conditions were counterbalanced. All additional factors were randomized.

**Participants.** 16 paid participants completed the study, aged 19-31 years ( $M=25.8$ ,  $SD=3.3$ , 7 female). Seven wore contact lenses and none wore glasses. 15 participants owned a smartphone and were right-handed. On a scale between 1 (None) to 5 (Expert), users rated themselves as experienced with digital technology ( $M=4.5$ ,  $SD=0.8$ ), multitouch input ( $M=3.4$ ,  $SD=1.1$ ), and less experienced with eye tracking ( $M=2.4$ ,  $SD=1.1$ ).

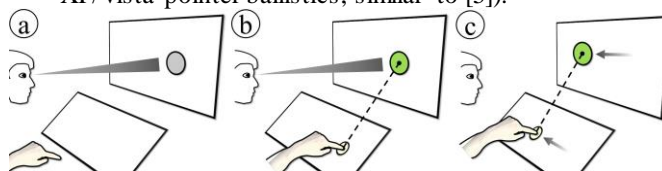
**Data Analysis.** For the statistical analysis of completion time and accuracy, we used a factorial repeated-measures ANOVA (Greenhouse-Geisser corrected if sphericity violated), and post-hoc pairwise comparisons with Bonferroni corrections. For the analysis of Likert ratings, we used a Friedman test with a post-hoc Wilcoxon Signed Rank test (Bonferroni corrected if necessary).

## 4 Experiment 1: Object Dragging

We first investigate an object dragging task. Dragging is commonly used on touchscreens, such as for moving files into a folder, or when positioning icons. In our experimental task, users move objects from a start to a destination location.

### 4.1 Compared Interaction Techniques

- **Touch:** Select the object with direct touch, and drag the object to the destination. The object starting and destination position are on the close, direct display.
- **Gaze+touch<sub>1:1</sub>:** Look at the object, touch down anywhere, indirectly drag object to the destination (Figure 4). The task is performed on the remote screen, while touches are issued on the close display. The finger movement translates 1:1 to object movement on remote screen).
- **Gaze+touch<sub>dynamic</sub>:** Same procedure (Figure 4), but the movement translates relatively using a dynamic control-display (CD) gain. This amplifies dragging i.e. it is more precise at slow and faster at fast finger movement (based on Windows XP/Vista pointer ballistics, similar to [5]).



**Fig. 4.** Drag with Gaze+touch: look at the object (a), touch down (b), and drag (c).



## 4.2 Experimental Design

The object's size is set to  $3^\circ$  which is 31 mm on close ( $31_{\text{mmc}}$ ) and 49 mm on remote display ( $49_{\text{mmr}}$ ). It appears towards a corner of the screen, and the dragging direction is always toward the opposite diagonal corner. The target of the dragging was displayed as a grey circle (see Figure 3b). The independent variables were:

- Dragging distance:  $17.5^\circ$  ( $185_{\text{mmc}}$ ,  $292_{\text{mmr}}$ ), and  $35^\circ$  ( $1210_{\text{mmc}}$ ,  $1307_{\text{mmr}}$ ) (between start and destination).
- Destination size: 110% ( $34_{\text{mmc}}$ ,  $54_{\text{mmr}}$ ), 220% ( $69_{\text{mmc}}$ ,  $109_{\text{mmr}}$ ), and 330% ( $103_{\text{mmc}}$ ,  $164_{\text{mmr}}$ ) of the object size.

Overall, each user performs 90 trials: 3 techniques  $\times$  2 distances  $\times$  3 target sizes  $\times$  5 repetitions. All tasks were successfully completed by all users.

## 4.3 Results

- **Task Completion Time:** Task completion times are presented in Table 1, line 1-3. A main effect of technique ( $F(2,30)=47.9$ ,  $p<.001$ ) showed that Touch is significantly faster than both Gaze+touch techniques ( $p<.001$ ), and Gaze+touch<sub>dynamic</sub> was faster than Gaze+touch<sub>1:1</sub> ( $p<.05$ ). Participants were slower with increasing distance ( $F(1,15)=86.26$ ,  $p<.001$ ). An interaction effect between technique and distance ( $F(1.3,19.3)=9.58$ ,  $p<.01$ ) showed Touch was the fastest technique across both distances ( $p<.01$ ). Further, Gaze+touch<sub>dynamic</sub> was faster than Gaze+touch<sub>1:1</sub> on  $35^\circ$  distance ( $p<.001$ ), which can be accounted to CD gain.
- **Accuracy:** Accuracy is the distance between the object's final position and center of the destination, normalized to degree of visual angle to allow direct comparison. Results are listed in Table 1, line 4-6. A main effect of technique ( $F(1.9,28.8)=5.17$ ,  $p<.05$ ) showed that users were overall more accurate with Touch than both Gaze+touch techniques ( $p<.05$ ). Further, users were more accurate with shorter distances ( $F(1,15)=23.63$ ,  $p<.001$ ) and larger target sizes ( $F(2,30)=23.91$ ,  $p<.001$ , all pairs differ,  $p<.01$ ).

**User Feedback:** Participants' preferences were split during the dragging task, with 9 of the 16 users preferring Touch. The users' rationale for Touch was less mental demand ("*my attention is not completely absorbed [with Touch]*"), familiarity with this technique ("*I have been using touchphones for a while, so it is a familiar technique*"), ease, and speed of Touch ("*it was easy and quick to perform the task*"). Arguments in favour of Gaze+touch were no occlusion ("*Touch had the problem that you obscured the object*"), the speed of the eyes ("*the eyes were faster than the fingers on the screen*"), and less perceived effort ("*[with Touch] it felt there was more effort because you saw the hand moving*"). Gaze+touch<sub>dynamic</sub> was preferred over Gaze+touch<sub>1:1</sub> by 13 of the 16 participants due to less physical movement with the hands.

Experiment	Line	Dependent Variable	Interaction Technique		
			Gaze+touch (1:1)	Gaze+touch (dynamic)	Touch
			Mean (SD)	Mean (SD)	Mean (SD)
Dragging	1	Time (17.5°)	3.12 (1.35)	2.9 (1.07)	2.11 (.93)*
	2	Time (35°)	4.24 (1.56)*	3.64 (1.56)*	2.57 (1.01)*
	3	Time (all)	3.68 (1.56)*	3.27 (1.39)*	2.34 (1)*
	4	Accuracy (17.5°)	0.44 (.55)	0.54 (.66)	0.38 (.37)
	5	Accuracy (35°)	0.64 (.76)	0.54 (.56)	0.46 (.58)
	6	Accuracy (all)	0.54 (.67)	0.54 (.61)	0.42 (.61)*
Rotation (Varying Object Size)	7	Skipped tasks	0%		19%
	8	Time (2°)	4.36 (1.59)*		6.23 (2.74)
	9	Time (4°)	4.25 (1.68)		4.24 (1.41)
	10	Time (8°)	4.43 (1.56)		4.04 (1.22)
	11	Time (all)	4.35 (1.6)*		4.8 (2.13)
	12	Accuracy (2°)	1.08 (.94)*		1.71 (1.43)
	13	Accuracy (4°)	0.81 (.75)		0.98 (.88)
	14	Accuracy (8°)	0.55 (.52)		0.59 (.52)
	15	Accuracy (all)	0.81 (.78)*		1.08 (1.1)
Rotation (Orientation)	16	Time (all)	5.79 (2.16)		5.81 (2.51)
	17	Accuracy (all)	0.74 (1.07)*		0.87 (.88)
Scaling	18	Skipped tasks	0%	0%	24.7%
	19	Time (2°)	3.64 (1.47)*	4.31 (1.79)	4.52 (1.82)
	20	Time (4°)	3.75 (1.39)	3.98 (1.67)	3.68 (1.5)
	21	Time (8°)	4.6 (1.76)	4.65 (1.94)	3.77 (1.4)*
	22	Time (all)	4 (1.6)	4.32 (1.85)	3.99 (1.62)
	23	Accuracy (2°)	0.072 (.05)	0.073 (.042)	0.105 (.074)
	25	Accuracy (4°)	0.074 (.059)	0.078 (.059)	0.098 (.07)
	26	Accuracy (8°)	0.082 (.064)	0.099 (.077)	0.102 (.086)
	27	Accuracy (all)	0.076(.058)*	0.083(.062)*	0.102 (.07)*

**Table 1.** Quantitative results of the experiments ('green' denotes higher performance, 'light green' denotes higher performance than 'white', asterisks denote significance, 'time' measured in seconds, 'accuracy' measured in visual angle).

- **Questionnaire:** The questionnaire's result is presented in Table 2, line 1-6. Statistical tests showed that for the category of perceived ease ( $X^2(2,16)=12.1, p=.002$ ), Touch was perceived easier than the Gaze+touch<sub>1:1</sub> variant ( $Z=-2.7, p<0.05$ ). Regarding speed ( $X^2(2,16)=12.1, p=.002$ ), both Touch ( $Z=-2.76, p<0.05$ ) and Gaze+touch<sub>dynamic</sub> ( $Z=-2.49, p<0.05$ ) were perceived as faster than Gaze+touch<sub>1:1</sub>. Regarding learning effort ( $X^2(2,16)=13.6, p=.001$ ), users find Touch easier to learn than Gaze+touch<sub>dynamic</sub> ( $Z=-2.8, p<0.05$ ) and Gaze+touch<sub>1:1</sub> ( $Z=-2.8, p<0.05$ ).

Experiment	Line	Dependent Variable	Interaction Technique		
			Gaze+touch absolute Mean (SD)	Gaze+touch relative Mean (SD)	Touch Mean (SD)
Dragging	1	Ease	3.81 (.83)	4.25 (.68)	4.75 (.45)*
	2	Speed	3.63 (.81)	4.25 (.68) *	4.44 (.73)*
	3	Accuracy	3.81 (.98)	4.06 (.57)	4.25 (.58)
	4	Learning	4.19 (.66)	4.06 (1.06)	4.88 (.34)*
	5	Eye tired.	3.63 (.81)	3.81 (1.05)	4.62 (.50)*
	6	Physical	3.81 (.91)	4.13 (.96)	3.69 (.79)
Rotation (Varying Object Sizes)	7	Ease	4.06 (.57)*		2.75 (1)
	8	Speed	3.94 (1)*		2.69 (.87)
	9	Accuracy	4.31 (.79)*		3.00 (.73)
	10	Learning	4.38 (.50)		4.13 (.72)
	11	Eye tired.	3.50 (.97)		4.25 (1)*
	12	Physical	3.50 (.97)		3.00 (.97)
Rotation (Orientations)	13	Ease	3.94 (.77)*		3.38 (.81)
	14	Speed	3.88 (.96)		3.50 (.63)
	15	Accuracy	4.00 (.82)		3.81 (.66)
	16	Learning	4.25 (.68)		4.06 (.93)
	17	Eye tired.	3.38 (1.31)		4.44 (.51)*
	18	Physical	3.50 (.97)*		2.75 (1)
Scaling	19	Ease	4.44 (.51)*	4.13 (.72)*	2.75 (.86)
	20	Speed	4.44 (.51)*	4.13 (.72)*	3.06 (.85)
	21	Accuracy	4.25 (.58)*	4.44 (.51)*	2.81 (.75)
	22	Learning	4.50 (.52)	4.38 (.72)	3.94 (.85)
	23	Eye tired.	3.19 (.98)	3.31 (1.01)	4.25 (.68)*
	24	Physical	3.69 (.95)	3.94 (.57)	3.25 (.93)

**Table 2.** Qualitative results of the experiments (1 = strongly disagree, 5 = strongly agree, 'green' denotes better rating, asterisks denote significance).

For eye tiredness ( $X^2(2,16)=14.4$ ,  $p=.001$ ), users find Touch less tiring than Gaze+touch<sub>dynamic</sub> ( $Z=-2.57$ ,  $p<0.05$ ) and Gaze+touch<sub>1:1</sub> ( $Z=-2.86$ ,  $p<0.05$ ).

#### 4.4 Discussion

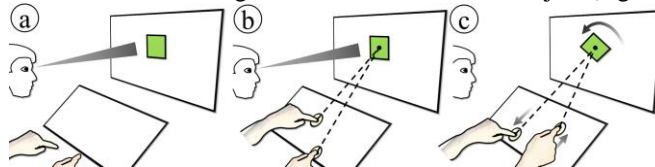
Users were faster and more accurate with Touch, and Gaze+touch<sub>dynamic</sub> was faster than Gaze+touch<sub>1:1</sub>. User preferences did not show clear results, however they indicate that more users prefer using Touch. Further analysis of the questionnaire showed that users perceive Touch to be beneficial across most categories (ease, speed, accuracy, etc.). Regarding both Gaze+touch techniques in comparison, users preferred Gaze+touch<sub>dynamic</sub> over Gaze+touch<sub>1:1</sub>, for requiring less hand movement.

## 5 Experiment 2: Object Rotation (Varying Object Sizes)

The previous study investigated a single-touch context; we now study multitouch interaction beginning with rotation of objects. We also investigate the fat-finger problem: we hypothesise that Touch is affected, but not Gaze+touch as fingers can be placed more freely. The task is a two-finger rotation tasks. Participants had to select an object, and then rotate it to a specific orientation. In each task, the object's size and rotation angle is varied.

### 5.1 Compared Interaction Techniques

- **Touch:** The user directly puts two fingers on an object, then rotates it to a target orientation. Again, touches and manipulation occur on the close and direct display. After users selected a target, users can also expand their fingers to manipulate the target more freely (as in all the following rotation and scaling experiments).
- **Gaze+touch:** The user looks at the target, indirectly touches down two fingers, then rotates these fingers to rotate the selected object (Figure 5).



**Fig. 5.** Gaze+touch: look at the object (a), touch down two fingers (b), and rotate (c).

### 5.2 Experimental Design

The target orientation was displayed underneath the object to visually indicate the rotation direction and angle (grey box, Figure 3c). The object and the target both displayed a line, which the users had to match to finish the rotation. The independent variables were:

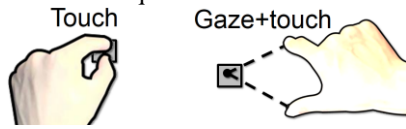
- Object size (visual angle):  $1^\circ$  ( $10_{\text{mmc}}$ ,  $16_{\text{mmr}}$ ),  $2^\circ$  ( $20_{\text{mmc}}$ ,  $33_{\text{mmr}}$ ),  $4^\circ$  ( $41_{\text{mmc}}$ ,  $66_{\text{mmr}}$ ), and  $8^\circ$  ( $83_{\text{mmc}}$ ,  $123_{\text{mmr}}$ ).
- Rotation angle:  $10^\circ$ ,  $50^\circ$ , and  $90^\circ$ .

The smallest size of  $1^\circ$  is chosen as a realistic lower limit of the used eye tracker. If users struggled with the acquisition of this target with Touch (as  $1^\circ=25\text{px}$ ), they could skip the task. Overall, each user performs 120 trials: 2 techniques  $\times$  4 sizes  $\times$  3 rotation angles  $\times$  5 repetitions.

### 5.3 Results

- **Error:** All tasks were successfully completed with Gaze+touch, and users skipped 19% of tasks with Touch. In particular, 76% of tasks with a size of  $1^\circ$  were skipped, demonstrating the effect of the fat-finger problem. The error is illustrated

in Figure 6. For the following statistical analysis, we excluded all trials with 1° to ensure an equal numbers of conditions.



**Fig. 6.** Users skipped more tasks with Touch, as the fat-finger issue became apparent with small targets.

- **Task Completion Time:** Table 1 line 8-11 summarise task completion times. A main effect of technique ( $F(1,15)=6.12$ ,  $p<.05$ ) showed that overall users were significantly faster with the Gaze+touch technique ( $p<.05$ ). A main effect of object size ( $F(1.3,19.3)=40.77$ ,  $p<.001$ ) showed that object size affected task completion time: tasks with 2° sized objects were significantly slower than 4° and 8° ( $p<.001$ ), yet no difference was found between 4° and 8°. Similarly with rotation angle ( $F(2,30)=17.23$ ,  $p<.001$ ), task completion time decreased at 10° ( $p<.01$ ), yet 50° and 90° did not show significant differences. An interaction between technique and size ( $F(1.5,21.9)=35.68$ ,  $p<.001$ ), revealed that participants performed faster with Gaze+touch at the 2° tasks than with Touch ( $p<.001$ ).
- **Accuracy:** Accuracy is measured in degrees as the difference between the final angle of the object and that of the destination. The results are presented in Table 1, line 12-15. A main effect of technique ( $F(1,15)=20.32$ ,  $p<.001$ ) showed that Gaze+touch was significantly more accurate than Touch ( $p<.001$ ). Accuracy increased with increasing object size ( $F(2,30)=129.59$ ,  $p<.001$ , all pairs differ at  $p<.001$ ), but not significantly with rotation angle. An interaction between technique and size ( $F(2,30)=12.47$ ,  $p<.001$ ) showed that for target size 2° Gaze+touch was significantly more accurate than Touch ( $p<.01$ ).
- **User Feedback:** 15 of 16 users preferred Gaze+touch. Most users' reasons were based around the fat-finger problem (*"I wasn't constrained by finger size"*).
- **Questionnaire:** The questionnaire's result is shown in Table 2, line 7-12. The following statistical differences were revealed. For category ease ( $X^2(1,16)=12$ ,  $p=.001$ ), Gaze+touch was perceived as easier than Touch. For category speed ( $X^2(1,16)=13$ ,  $p=.0$ ), Gaze+touch was perceived as faster than Touch. For category accuracy ( $X^2(1,16)=16$ ,  $p=.0$ ), Gaze+touch was perceived as more accurate than Touch. For category eye tiredness ( $X^2(1,16)=7.4$ ,  $p=.007$ ), Gaze+touch was perceived as more eye tiring than Touch.

## 5.4 Discussion

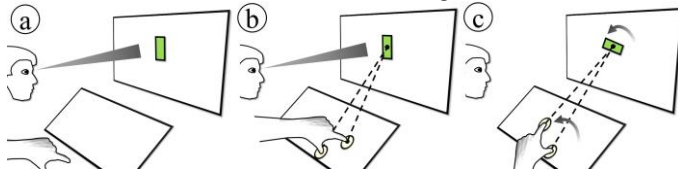
Users skipped more tasks when selecting small targets with Touch, a result that was expected considering the fat-finger issue. Gaze+touch did not show any errors across all target sizes, indicating that a gaze-selection is more accurate than a raw Touch-selection for two-touch gestures. Users were faster with Gaze+touch for small targets as they were easier to acquire ( $\text{size} \leq 2^\circ$ ). In case of accuracy, users performed more accurate than Touch in the rotation tasks with small targets, which can be accounted to the more relaxed placement of fingers. User feedback confirms these results, most users preferred Gaze+touch for being less constrained to finger size.

## 6 Experiment 3: Object Rotation (Varying Object Orientation)

The previous study investigated rotation of objects with varying object sizes. Another important factor for rotation is the initial orientation of targets which can affect the user's performance [8]. We therefore investigate this context in our third experiment, where users perform unimanual rotation gestures. In each task, the object's initial orientation and the rotation direction are varied.

### 6.1 Compared Interaction Techniques

- **Touch:** The user directly puts two fingers of the same hand on an object, then rotates it to a target rotation.
- **Gaze+touch:** The user looks at the target, indirectly touches down two fingers of one hand, then rotates these fingers to rotate the selected object (Figure 7).



**Fig. 7.** Rotate with Gaze+touch: look (a), touch down two fingers of one hand (b), and rotate (c).

### 6.2 Experimental Design

The target object is a rectangle with a width of  $2^\circ$  ( $20_{\text{mmc}}$ ,  $33_{\text{mmr}}$ ) and a height of  $10^\circ$  ( $104_{\text{mmc}}$ ,  $166_{\text{mmr}}$ ), and the angle of required rotation is fixed at  $90^\circ$ . We chose  $2^\circ$  as it allowed users to comfortably place their fingers on the object. The target orientation was indicated underneath the object with a rectangle and with an arrow to visually indicate the rotation direction and angle (Figure 3). Independent variables were:

- Object starting orientation:  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$  relative to the screen's x-axis.
- Rotation direction: clockwise or anticlockwise.

Overall, users perform 90 trials: 2 techniques  $\times$  4 initial orientations  $\times$  2 directions  $\times$  5 repetitions. The participants completed all tasks with both techniques.

### 6.3 Results

- **Task Completion Time:** No significant effects were found for task completion times; users performed similarly fast with both techniques (Table 1, line 16).
- **Accuracy:** A significant main effect of technique for accuracy (again measured in angle difference,  $F(1,15)=19.43$ ,  $p<.01$ ) revealed that Gaze+touch was more accurate than Touch (Table 1, line 17).
- **User Feedback:** 12 of 16 users preferred rotation with Gaze+touch, with the reason that Touch, with differently oriented objects, made some objects hard to acquire ("You do not need to put the hand on the object").

- **Questionnaire:** Results are shown in Table 2, line 13-18. In category ease of use, users perceived Gaze+touch as easier than Touch ( $X^2(1,16)=5.4$ ,  $p=.02$ ). For category eye tiredness ( $X^2(1,16)=6.4$ ,  $p=.011$ ), Gaze+touch was perceived as more eye tiring than Touch. For category physical effort ( $X^2(1,16)=6.4$ ,  $p=.011$ ), Gaze+touch was perceived as less physically demanding than Touch.

## 6.4 Discussion

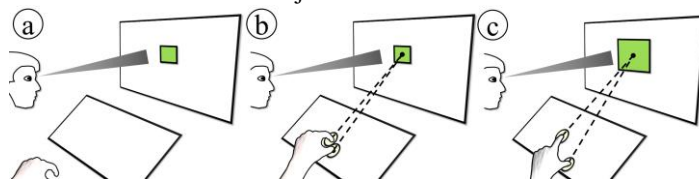
Overall, users performed equally fast with both techniques, but more accurately with Gaze+touch. Contrasting the first rotation study (Experiment 2), no difference in speed is found because the target size remained fixed for all tasks in this experiment. However, users were more accurate with Gaze+touch, again to be accounted to the relaxed finger placement, as users indicated objects are difficult to acquire with Touch. The questionnaire showed that users perceived Gaze+touch as easier and as less physically demanding, although more eye tiring.

## 7 Experiment 4: Object Scaling

Our previous experiments investigated dragging and rotation tasks. Another prevalent gesture for multitouch is pinch-to-scale, often used to scale images, maps, or other objects. This experiment assesses users' performance of Gaze+touch in a scaling task. Object size and scaling amplitude is varied. We let users choose between uni- or bi-manual interaction.

### 7.1 Compared Interaction Techniques

- **Touch:** The user directly puts two fingers on an object, then pinches them to fit an outline of a differently scaled target.
- **Gaze+touch<sub>1:1</sub>:** The user looks at the target, indirectly touches down two fingers, then pinches them to scale (Figure 8). The scaling translates absolutely, thus finger movement is 1:1 applied to the object's scaling.
- **Gaze+touch<sub>dynamic</sub>:** Same operation (Figure 8). However, the scaling translates with a different control-display gain, the distance between both fingers is relative to the width of the object.



**Fig. 8.** Gaze+touch scaling: look (a), touch down two fingers (b), and pinch (c).

### 7.2 Experimental Design

Users had to scale an object to a specific size which was visually indicated by a black rectangle (see Figure 3). The independent variables were:

- Initial object size (visual angle): 1° (10<sub>mm</sub>, 16<sub>mm</sub>), 2° (20<sub>mm</sub>, 33<sub>mm</sub>), 4° (41<sub>mm</sub>, 66<sub>mm</sub>), and 8° (83<sub>mm</sub>, 123<sub>mm</sub>).
- Scaling factor: 50%, 90%, 120%, and 200% of the initial object's size.

Overall, a user did 240 trials: 3 techniques × 4 sizes × 4 scale factors × 5 repetitions.

### 7.3 Results

- **Error:** Participants completed all tasks with Gaze+touch, and skipped 24.7% tasks with Touch (c.f. Fig. 6). In particular, users skipped with Touch within 1° tasks 100% of both downscaling conditions, 68.8% upscaling-to-120% tasks, and 56.3% upscaling-to-200% tasks; and within 2° tasks 29.7% downscaling-to-50% tasks, and 40.6% downscaling-to-90% tasks. Like in Experiment 2, we exclude skipped-task conditions from statistics for an equal number of trials between techniques.
- **Task Completion Time:** Task completion times results are summarised in Table 1, line 19-22. There was no significant difference for technique at task completion time. A significant main effect for object size ( $F(2,30)=7.68$ ,  $p<.01$ ) showed that users performed faster with 4° than 8° objects ( $p<.01$ ). A main effect of target size ( $F(1,15)=34.98$ ,  $p<.01$ ) revealed that users were faster with 120% than 200% upscaling ( $p<.05$ ). An interaction between technique and the object's size ( $F(2.3,33.8)=7.6$ ,  $p<.01$ ) showed that within 2° sized object tasks, users performed faster with Gaze+touch<sub>1:1</sub> than with Gaze+touch<sub>dynamic</sub> ( $p<.05$ ), and within 8° tasks, users were faster with Touch than with either Gaze+touch technique ( $p<.05$ ).
- **Accuracy:** Accuracy (the difference between radius of the object and radius of the target, normalized to degree of visual angle) showed a main effect of technique ( $F(2,30)=9.8$ ,  $p<.01$ ) with users more accurate with both Gaze+touch techniques than with Touch ( $p<.05$ , Table 1, line 23-27).
- **User Feedback:** All users preferred using one of the Gaze+touch techniques. The user's reason were based around the fat-finger problem ("*I could not get both fingers within the box*"), occlusion, and hand fatigue ("*my arms do not get so tired because I can rest on the desk*"). Between both Gaze+touch techniques, 8 users preferred Gaze+touch<sub>dynamic</sub>, and the other 8 users Gaze+touch<sub>1:1</sub>. Reasons were more precision with Gaze+touch<sub>dynamic</sub> ("*I was more precise*"), and speed for Gaze+touch<sub>1:1</sub> ("*[Gaze+touch<sub>dynamic</sub>] is slower [than Gaze+touch<sub>1:1</sub>], because your fingers need to make a bigger movement*").
- **Questionnaire:** Results are presented in Table 2, line 19-24. For category ease ( $X^2(2,16)=26$ ,  $p=.0001$ ), Gaze+touch<sub>dynamic</sub> ( $Z=-3.4$ ,  $p<0.05$ ) and Gaze+touch<sub>1:1</sub> ( $Z=-3.5$ ,  $p<0.05$ ) were perceived as easier than Touch. For category speed ( $X^2(2,16)=19.4$ ,  $p=.0001$ ), Gaze+touch<sub>dynamic</sub> ( $Z=-2.9$ ,  $p<0.05$ ) and Gaze+touch<sub>1:1</sub> ( $Z=-3.3$ ,  $p<0.05$ ) were perceived as faster than Touch. For category accuracy ( $X^2(2,16)=24.6$ ,  $p=.0001$ ), Gaze+touch<sub>dynamic</sub> ( $Z=-3.4$ ,  $p<0.05$ ) and Gaze+touch<sub>1:1</sub> ( $Z=-3.3$ ,  $p<0.05$ ) were perceived as more accurate than Touch. For category eye tiredness ( $X^2(2,16)=16.8$ ,  $p=.0001$ ), Gaze+touch<sub>dynamic</sub> ( $Z=-3$ ,  $p<0.05$ ) and Gaze+touch<sub>1:1</sub> ( $Z=-3.1$ ,  $p<0.05$ ) were perceived as more tiring than Touch.



## 7.4 Discussion

Similar to experiment rotation (size), users had difficulties acquiring small targets with Touch (fat-finger issue). Despite small targets, users performed similarly fast with Gaze+touch and Touch, while more accurate with Gaze+touch. Only for large targets, Touch was faster than Gaze+touch. Users' reasons were about the fat-finger problem (occlusion, fatigue). Considering both Gaze+touch variants, users can scale smaller objects faster with Gaze+touch<sub>1:1</sub>. User opinion was split, half of them preferred Gaze+touch<sub>1:1</sub> and the remaining users preferred Gaze+touch<sub>dynamic</sub>.

## 8 Overall Discussion

Our experiments investigated how Gaze+touch combines properties of direct (same gesture operation) and indirect touch (relaxed finger placement), which we follow up with a high level discussion on the following points:

**Advantages of Gaze+touch.** The relaxed finger placement can reduce effects of the fat-finger and occlusion issues. This was particularly beneficial for rotation and scaling tasks, where more than one finger is used for manipulation. Our experiments showed that with the design techniques, users performed similarly fast, and more accurately than with Touch. Manipulation of objects with a size of  $<2^\circ$  was not feasible with Touch, however with Gaze+touch, users were able to manipulate the  $1^\circ$  objects (although completion time increased).

**Shortfalls of Gaze+touch.** Gaze+touch was slower and less accurate for dragging tasks. Based on observations and feedback, we believe this is accounted to two aspects. First, the 'leave-before-click' issue [17]: users can already look away from a target before they touch down, or touch down before they look at the target; both will void a Gaze+touch selection. Second, users could 'lose' a target during dragging, e.g. when the system wrongly detected a 'touch up' event or the user's finger briefly hovered. With direct touch, the finger is on the target, thus the system will immediately receive a 'touch down' event and the dragging continues. However, with Gaze+touch, it is indirect touch, and to reselect users have to look to the target before they touch down. Both advantages and shortfalls correlated with the users' feedback.

**Control-Display Gain.** The relaxed finger placement allowed us to experiment with dynamic CD gains (Gaze+touch<sub>dynamic</sub>). During dragging, this led to faster performance than a 1:1 mapping of Gaze+touch. It eliminated the need to clutch when the touch screen's size did not suffice, confirming previous studies of cursor acceleration [5]. CD gain in scaling tasks did not use acceleration, but rather input relative to the fingers' distance. Users were faster with the absolute scaling, but more precise with relative scaling.

**Observations.** We observed users exploiting the relaxed finger placement as a strategy to improve their performance: for long rotations, users had their fingers drawn together to manipulate faster, while for short rotations, users set them further apart to be more accurate. Although technically normal touch allows the same operation, it is restricted by a lower limit, as for small targets users have little room to adjust their fingers on it, and an upper limit, defined by the object's size (although users can expand their fingers once they selected an object).

**Limitations.** Users were slower with Gaze+touch for dragging tasks. One method to overcome this could be using gaze for target dragging: after a touch selection, the target follows the user's gaze ([22, 24, 25]). Future studies can investigate how this advanced technique would compare to standard touch dragging. In addition, our experiments compared Gaze+touch against the raw default of direct touch. This is apparently prone to issues such as the fat-finger problem or occlusion. These issues can be overcome, e.g. by indirect touch techniques ([1, 4]), and should be considered in future evaluation.

## 9 Conclusion

Gaze as a mediator can bring touch to remote displays, but at what cost and benefit? As a first step, we compared Gaze+touch to touch in four experiments across dragging, rotation, and scaling tasks. Our experiments provide detailed performance characteristics of both input modes, and indicate that while Gaze+touch is slower and less accurate for dragging tasks, users are equally fast and more accurate in rotation and scaling tasks. This can support the design of Gaze+touch interaction for remote displays, such as combined close and remote surfaces consistently controlled by multitouch. In light that further evaluation beyond abstract tasks may be required to validate the real-world applicability, our experiments provide empirical groundwork in the exploration of how Gaze+touch sets itself apart from touch interaction.

## References

1. Albinsson P-A, Zhai S: High precision touch screen interaction. CHI '03. ACM, pp 105–112 (2003).
2. Au OK-C, Su X, Lau RWH: LinearDragger: A Linear Selector for Target Acquisition on Touch Screens. CHI'14. ACM, pp 2607–2616 (2014).
3. Benko H, Izadi S, Wilson AD, Cao X, Rosenfeld D, Hinckley K: Design and Evaluation of Interaction Models for Multi-touch Mice. GI'10. CIPS, pp 253–260 (2010).
4. Benko H, Wilson AD, Baudisch P: Precise selection techniques for multi-touch screens. CHI '06. ACM, pp 1263–1272 (2006).
5. Casiez G, Vogel D, Balakrishnan R, Cockburn A: The Impact Control-Display Gain on User Performance in Pointing Tasks. *Human-Computer Interact* 23:215–250. (2008).
6. Forlines C, Wigdor D, Shen C, Balakrishnan R: Direct-touch vs. Mouse Input for Tabletop Displays. CHI '07. ACM, pp 647–656 (2007).

7. Gutwin C, Cockburn A, Scarr J, Malacria S, Olson SC: Faster Command Selection on Tablets with FastTap. CHI'14. ACM, pp 2617–2626 (2014).
8. Hoggan E, Williamson J, Oulasvirta A, Nacenta M, Kristensson PO, Lehtiö A: Multi-touch Rotation Gestures: Performance and Ergonomics. CHI '13. ACM, pp 3047–3050 (2013).
9. Jacob RJK: What you look at is what you get: eye movement-based interaction techniques. CHI '90. ACM, pp 11–18 (1990).
10. Jacob RJK: Eye Movement-Based Human-Computer Interaction Techniques: Toward Non-Command Interfaces. *Adv. Human-Computer Interact.* Vol. 4, Ablex Publishing, pp 151–190 (1993).
11. Kin K, Agrawala M, DeRose T: Determining the Benefits of Direct-touch, Bimanual, and Multifinger Input on a Multitouch Workstation. GI '09. CIPS, pp 119–124 (2009).
12. Malik S, Laszlo J: Visual touchpad: a two-handed gestural input device. ICMI '04. ACM, pp 289–296 (2004).
13. Moscovich T, Hughes JF: Indirect Mappings of Multi-touch Input Using One and Two Hands. CHI '08. ACM, pp 1275–1284 (2008).
14. Pfeuffer K, Alexander J, Chong MK, Gellersen H: Gaze-touch: Combining Gaze with Multi-touch for Interaction on the Same Surface. UIST '14. ACM, pp 509–518 (2014).
15. Pfeuffer K, Vidal M, Turner J, Bulling A, Gellersen H: Pursuit Calibration: Making Gaze Calibration Less Tedious and More Flexible. UIST '13 (2013).
16. Potter RL, Weldon LJ, Shneiderman B: Improving the Accuracy of Touch Screens: An Experimental Evaluation of Three Strategies. CHI '88. ACM, pp 27–32 (1988).
17. Salvucci DD, Anderson JR: Intelligent Gaze-added Interfaces. CHI'00. ACM, New York, NY, USA, pp 273–280 (2000).
18. Schmidt D, Block F, Gellersen H: A Comparison Direct and Indirect Multi-touch Input for Large Surfaces. INTERACT '09. Springer, pp 582–594 (2009).
19. Sibert LE, Jacob RJK: Evaluation of eye gaze interaction. CHI '00. ACM, pp 281–288 (2000).
20. Špakov O, Isokoski P, Majaranta P: Look and Lean: Accurate Head-assisted Eye Pointing. ETRA'14. ACM, New York, NY, USA, pp 35–42 (2014).
21. Stellmach S, Dachselt R: Look & touch: gaze-supported target acquisition. CHI '12. ACM, pp 2981–2990 (2012).
22. Stellmach S, Dachselt R: Still looking: investigating seamless gaze-supported selection, positioning, and manipulation distant targets. CHI '13. ACM, pp 285–294 (2013).
23. Turner J, Alexander J, Bulling A, Gellersen H: Gaze+RST: integrating Gaze and multitouch for remote Rotate-Scale-Translate tasks. CHI'15. ACM, p to appear (2015).
24. Turner J, Alexander J, Bulling A, Schmidt D, Gellersen H: Eye Pull, Eye Push: Moving Objects between Large Screens and Personal Devices with Gaze & Touch. INTERACT'13. Springer, pp 170–186 (2013).
25. Turner J, Bulling A, Alexander J, Gellersen H: Cross-device Gaze-supported Point-to-point Content Transfer. ETRA'14. ACM, pp 19–26 (2014).
26. Voelker S, Wacharamanotham C, Borchers J: An evaluation state switching methods for indirect touch systems. CHI '13. ACM, pp 745–754 (2013).
27. Ware C, Mikaelian HH: An Evaluation of an Eye Tracker As a Device for Computer Input. CHI '87. ACM, pp 183–188 (1987).
28. Wigdor D, Benko H, Pella J, Lombardo J, Williams S: Rock & Rails: Extending Multi-touch Interactions with Shape Gestures to Enable Precise Spatial Manipulations. CHI'11. ACM, pp 1581–1590 (2011).
29. Zhai S, Morimoto C, Ihde S: Manual and gaze input cascaded (MAGIC) pointing. CHI '99. ACM, pp 246–253 (1999).