Graphics Interface Conference 2015, 3–5 June, Halifax, Nova Scotia, Canada

The Performance of Indirect Foot Pointing using Discrete Taps and Kicks While Standing

William Saunders and Daniel Vogel*

Cheriton School of Computer Science, University of Waterloo

ABSTRACT

We investigate the performance of indirect foot pointing while standing using discrete taps and kicks. Two experiments show that left and right feet perform at similar levels, there is little difference in selection time across target configurations or directions, but targets with an angular size under 22.5° or radial size under 5cm should be avoided due to high error rates. There is a detectable advantage to tapping compared to kicking, but little practical difference. Although cursor feedback is optimal, we show that eyes-free foot pointing achieves an error rate of 27% for 45° angular targets. We translate our results into ten design guidelines and we illustrate their application by designing foot interaction techniques to control desktop applications at a standing desk.

Keywords: Foot input, pointing, tapping, kicking.

Index Terms: H.5.2. Information interfaces: User Interfaces.

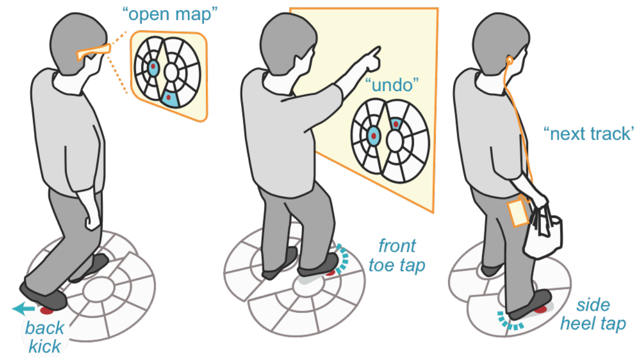
1 INTRODUCTION

People use their feet to drive cars and play instruments so it is not surprising that feet have been considered for cursor control [19]. Although the mouse proved better, foot-based interfaces have since been applied to, or proposed for, gaming [10,14], hands-free interaction for mobile devices [1,4,8,18], and large floor or wall displays [2,11]. Instead of *continuous* position control, these systems are designed around more *discrete* actions such as taps, kicks, and heel pivots. Some use a *direct input mapping* where a floor-based input sensor and display are coincident [2,11], but an *indirect input mapping*, where foot position is represented as a cursor on a display, provides greater flexibility (Figure 1).

This style of foot interaction while standing has broad applicability to situations when there is reason to avoid, reduce, or augment hand input. In mobile settings, foot input is useful with a head mounted display (Figure 1a) or when hands are occupied and smartphone input is burdensome [4] (Figure 1c). In large display settings, foot input could augment finger touches by triggering commands such as "undo" (Figure 1b). We are particularly interested in using foot input at "standing desk" to enable *foot input breaks* [3,13] for increased physical activity by occasionally using feet to control applications instead of the mouse and keyboard.

To develop new indirect, discrete foot interaction techniques for mobile, large display, standing desk, or other contexts, we need design guidelines based on empirical evidence. Researchers have examined indirect directional kicking with one foot while standing [1,8,14], but indirect tapping while standing has received little attention. Previous work has examined "0D" in-place floor tapping while seated [4], 1D pedal tapping while seated [6,9], and 2D interactive floor tapping emphasizing direct input issues such as perceived input point and occlusion [2].

wrsaunde@uwaterloo.ca, dvogel@uwaterloo.ca



(a) with feedback on HMD (b) with feedback on large display (c) without feedback on smartphone

Figure 1. Indirect foot pointing using discrete taps and kicks on virtual targets in semi-circular rings around feet, device and target examples: (a, b) indirect feedback using foot cursor (red dot) on high-density targets with head mounted display or large display; (c) indirect without feedback using low-density targets with smartphone.

We contribute empirical work for new interaction techniques based on indirect foot pointing using discrete taps and kicks while standing. We test a practical configuration of annular targets placed in semi-circular rings with-and-without cursor feedback. We extend kicking research by controlling for radial target size, testing both feet, and providing a direct comparison with tapping.

Based on the results of two experiments, we create a concise set of ten design guidelines. For example, left and right feet perform at similar levels, there is little detectable difference in time across target configurations or directions but targets with an angular size under 22.5° or radial size under 5 cm should be avoided due to high error rates. There is a small advantage to using tapping compared to kicking for pointing actions, but little practical difference. Pointing with foot cursor feedback is optimal, but we show that eyes-free pointing is feasible with a 27% lower bound error rate for 45° angular targets. To illustrate the applicability of our guidelines and demonstrate the utility of the investigated input space, we describe foot input techniques to control conventional applications at a standing desk for productive physical breaks.

2 RELATED WORK

We briefly survey previous studies of seated foot input and standing direct foot input, and then contrast our work with relevant studies examining standing indirect foot input.

2.1 Foot input While Seated

For seated desktop computing, researchers have not found a performance advantage for foot input. Kim and Kaber [12] were unable to show a clear benefit for using multiple foot pedals. Pearson and Weiser's rate control "foot joystick" [16] was slower and more error-prone than a mouse. Pakkanen and Raisamo's foot-controlled trackball [15] was slower, more error prone, and less preferred than a hand-controlled trackball. Other seated pedal studies [6,9] confirm that although foot motion follows Fitts' law, it is slower than comparable arm movements.

Copyright held by authors. Permission granted to CHCCS/SCDHM to publish in print and digital form, and ACM to publish electronically.

In mobile settings, foot input is more encouraging. Dearman et al. [5] found mobile text entry using two physical foot pedals could be as fast as touch, but more error prone. Crossan et al. [4] report that sequential toe tapping to select from an eyes-free menu is faster than retrieving the phone in some cases. They found an in-place "0D" tap took 1.2s.

However, using studies conducted while seated or with foot pedals is not appropriate to develop design guidelines for standing input given differences in balance and range-of motion.

2.2 Direct Foot Input

Augsten et al. [2] investigate foot input on an interactive floor display. This is *direct* input with a "fat foot" problem whereas our goal is to use *indirect* input where targets are represented on screen and foot position represented as a cursor. Their findings related to indirect input performance include minimum target sizes and selection of a "hotspot" (the point on a foot used to select a target). They found that 3.1 cm by 3.5 cm targets were needed for a reasonable (10%) error rate, while 5.3 cm by 5.8 cm targets achieved a low error rate of 3%. They found the perceived hotspot varies by individual from toe, offset from toe, and ball of foot. Indirect input also requires a hotspot to locate the cursor, but with no surrounding visible foot, the hotspot position is more arbitrary and universal (consider where the "hotspot" is on a mouse). Regardless, differences between direct and indirect touch input [17] are likely to translate to feet, making these findings less relevant.

2.3 Indirect Foot Input While Standing

Scott et al.'s [18] exploration of single foot, heel and toe pivoting is complementary to our work since pivoting can be combined with tapping and kicking. Relevant to tapping, their implemented technique used an in-place toe tap to demarcate interaction, but tapping was not included in their formative experiment.

The most relevant tapping study is Meyers et al. [13]. They used a Dance-Dance-Revolution (DDR) game mat (a 3 by 3 grid of footswitches) to enable physical breaks using foot-operated email and photo sorting applications. No controlled experiment was conducted, but a usability evaluation found people preferred to alternate or balance taps between their two feet. The Meyers et al. system serves to validate the discrete, indirect foot pointing input space we investigate and further motivates our interest in foot input for controlling desktop applications at a standing desk.

Although there is little previous work investigating tapping, there has been considerable interest in kicking. Han et al. [8] examined direction and velocity characteristics of forward kicks. They found people can use 5 distinct forward kick directions over a 120° arc (24° targets) and they can control two levels of kick velocity. Alexander et al.'s [1] elicitation study suggests people prefer spatial taps and kicks for certain tasks. They explore single-foot kick characteristics for controlling continuous map navigation and provide basic guidelines: backwards kicks are difficult and controlling kick direction is easier than kick distance. They do not investigate tapping performance beyond using an in-place foot tap like Crossan et al. [4] to stop navigation. Neither Han et al. or Alexander et al. evaluate kicking for discrete target selection or cover the larger two-level, semi circular target space we investigate.

3 Experiment 1: Tapping

The goal of our first experiment is to investigate discrete, indirect foot pointing using taps on a range of radial and angular target sizes. Once we establish a usable range of target sizes, our second experiment compares pointing using taps or kicks, and tests eyes-free indirect input with a no cursor condition. We use a circular array of targets around a central 'home position' – an extension of the front-facing array of kick directions used by Han et al. [8]. This layout allows for a consistent investigation of the effect of direction, reflects the biomechanical range of motion for legs while standing, and is a likely configuration for a foot-operated system.

3.1 Participants

Eleven people (3 female), ages 20 to 37 participated (13 people were recruited, but one had high tracking errors and one used an unanticipated slide-tapping strategy.) 11 reported they were right-footed (i.e. they kick a ball with their right foot) and 3 reported they had previously used a whole body input device. Participants were screened to exclude anyone with an injury or impairment that would interfere with their performance or lead to further injury.

3.2 Apparatus

Our aim is to establish an upper bound on performance. For this reason we use a Vicon motion tracking system for high-resolution, high frame rate (100 Hz), low latency data. The 3D position and orientation of both feet are tracked using infrared reflective markers on elastic bands wrapped around each foot (Figure 2a). A 17-inch display on a raised stand displayed all visuals (Figure 2b). Visuals were legible from 0.75 m away, the typical distance from participant to display. In addition to logging movement to Vicon capture files and logging all input events in our software, we also video recorded sessions for qualitative analysis (Figure 2a).

3.2.1 Foot Cursor Hotspot Calibration

We calibrated for each participant's shoe size by recording offsets from the tracked position of the band to the heel and toe using a floor registration point (Figure 2e). Augsten et al. [2] found that perceived hotspot varies by individual for direct input. Since there is more flexibility in hotspot location with indirect input and to be consistent across participants, we use the midpoint between heel and toe positions as the foot cursor and selection hotspot. The midpoint also avoids biasing the system towards a style of tapping: we wanted to observe if participants would tap with the front of the foot (toe or ball of the foot), the back of the foot (heel), or the whole foot. Visual feedback in the form of two red circles ("foot cursors") represented the real-time midpoint position of each foot (Figure 3b). The foot position in motor space is mapped to display space using a constant CD Gain of 8.5 px/cm.

3.2.2 Target Selection Action Detection

Targets were selected using thresholds determined in pilot experiments. When the height of either the toe or heel transitioned below 4 mm above the floor and momentary foot speed was less than 0.2 m/s, a selection event was triggered. To avoid hysteresis, the foot had to lift more than 8 mm above the floor or travel at a speed greater than 0.3 m/s before a previous selection event was exited. The speed threshold reduced false positives due to uncertainty from deformation of the shoe, but did not reduce the possibility to make rapid taps. These features made selections feel like tapping the floor, and allowed for tapping with the toe, heel, or whole foot.

3.3 Tasks and Stimuli

To complete each discrete foot-tapping task, participants lifted their foot off a center home target, moved it in the air until the foot cursor was over the task target, and tapped the floor. Then, they immediately returned their foot to the home position by lifting, moving, and tapping on the home target. This rapid cycle was repeated 3 times in succession for the same foot and task target. The accompanying video demonstrates the task and feedback.

3.3.1 Target Size and Distance

To vary target distance and test when target width is constrained in the direction of travel, our targets are shaped like pieces of a twodimensional ring, called “annular sectors” (Figure 3a), rather than the “circular sector” pie pieces used by Han et al. [8]. This creates a constant distance from the central home positions to any target and is more biomechanically compatible than a square grid. We chose this target shape to discover guidelines relevant to a system using two rings of targets (we discuss such a system later). All target sizes are actual size in motor space on the floor.

Each task was parameterized by 3 variables: angular size A, radial width R, and distance D (Figure 3a). Target SIZE is was fully defined by the pair (A, R). A is the angular size in degrees and R is the radial width of the target in cm. Informed by small pilot studies, we chose A and R values as whole number multiples of 11.25° for A and 5 cm for R. We use the concise target size notation A*i*R*j* for the set of target SIZES: A45R20 = (45°, 20 cm), A45R10 = (45°, 10 cm), A45R5 = (45°, 5 cm), A22R20 = (22.5°, 20 cm), A11R20 = (11.25°, 20 cm).

A fixed home position for each foot was calibrated near the center of the interaction area. Two home targets represent these positions each has a radius of 7.5 cm in the interaction area. Pilot tests determined this size as sufficiently constrained, but reliably selected. The home target represents a non-interactive area where the foot can rest between issuing commands. We use a pointing action going from home, to target, and back to home as a practical method of tapping which is consistent with kicks when tested later.

Target distance was measured to the inner edge of the target. We asked participants to tap anywhere inside the target. Assuming they would tap with as little movement as possible, the distance to the inner edge of the target is a more representative distance than target center. We tested two values to investigate the effect of a distractor target in multiple target rings in a real system. A 7.5 cm distance positioned the task target right against the home target and a 15 cm distance created a 7.5 cm gap between the task target and home target like a target on an outer ring. The gap functions as a distractor target that must be avoided, allowing our results to generalize to a multi-layered ring of targets (e.g. Figure 7a,b).

While this design may seem similar to a Fitts’ Law style study (e.g. [6]), our goal is not cross-device comparison. Our focus is a formative study to guide interaction technique design similar to the approach of previous studies [8,18]. For this reason, we test different variables, including which foot is being used and direction (along with tapping, kicking, and level of feedback in Experiment 2). In addition, the spatial layout is chosen to match the physiology of leg and hip motion which are not core to traditional Fitts’ Law studies where difficulty of motion is more uniform in the interaction space, as with keyboard and mouse.

3.3.2 Target Directions

The targets for each foot were positioned at one of five directions (Figure 3b) 0 – forward, 1 – forward-diagonal, 2 – side, 3 – backward-diagonal, and 4 – backward. Using the largest angular target size of 45°, the edges of all targets in a real system would touch, making maximal use of the interaction area without overlapping.

3.3.3 Target Feedback

At the start of the task, a home target and task target appeared on the side of the display corresponding to the foot required for the task (Figure 3c). A purple border around the edge of the target indicated the target to tap next. The target was highlighted in bright blue when the system detected a foot cursor hotspot inside the target region and part of that foot (either the heel or toe) was touching the floor. If an error occurred, defined as tapping while the hotspot was outside of the task target, the purple border moved to the home target (no other feedback was given). Participants had to achieve three error-free repetitions to complete the task. Three repetitions allow an accurate calculation of ROUND TRIP TIME. Tasks alternated between feet to reduce fatigue.

3.4 Design and Protocol

The independent variables are FOOT (left or right), target SIZE, target DIRECTION, and target DISTANCE. Tasks were divided into 10 target configuration sets of 10 tasks. Each set covered all values of 5 target DIRECTIONS and FOOT for one target SIZE and DISTANCE, with a random ordering that always alternated between feet. All target configuration sets were presented in random order as one BLOCK, covering all 300 task settings. Participants completed three BLOCKS in order to test for learning effects.

A short instruction and demonstration block was presented at the beginning and rest breaks were provided at the end of each block. Participants were interviewed after the experiment for subjective feedback about fatigue and preference for toe or heel tapping. The experiment took 60 minutes on average.

In summary the design was:

3 BLOCKS ×
2 FEET × 5 target DIRECTIONS ×
5 target SIZES × 2 target DISTANCES ×
3 repetitions of serial selections
= 900 data points per participant

3.5 Analysis

The dependent variables are ERROR RATE, SELECTION TIME, and ROUND TRIP TIME.

ERROR RATE was calculated as the mean percentage of errors per repetition. Errors are defined as when the system detected that the participant tapped their foot while the foot hotspot was not on the desired target. To complete a selection and continue the task, the participant had to successfully tap on the target.

SELECTION TIME is defined as the time duration between the moment the participant had lifted their foot off the home target to

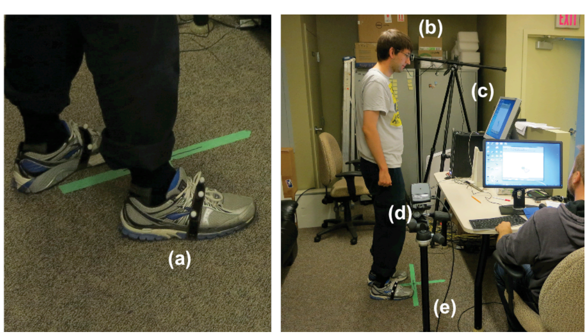


Figure 2: Apparatus: (a) Vicon tracking markers attached to both shoes; (b) video camera to record session; (c) desktop monitor on raised desk platform for experiment feedback; (d) Vicon tracking cameras; (e) marked calibration position on floor.

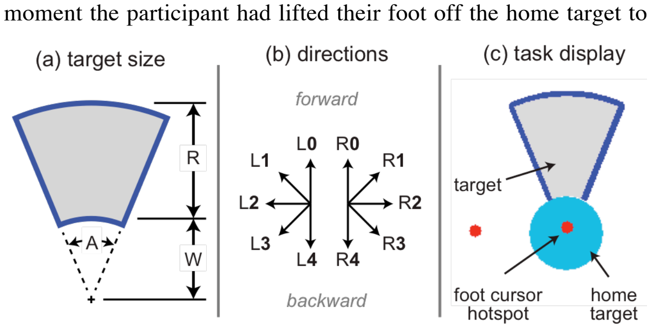


Figure 3: Task: (a) parameterization of target size and distance; (b) target directions for each foot, letter indicates foot, number indicates direction; (c) example task display stimuli for A45R20 (45°, 20cm) R0-forward target at distance 7.5cm, right foot is on home.

the moment when their foot touched down on the task target. SELECTION TIME is averaged over all repetitions in the task.

ROUND TRIP TIME is defined as the mean time to select the task target and return to the home target, including the stationary time at each target. This measurement captures the full time needed to select a target and return to the central neutral posture. ROUND TRIP TIME is averaged over the second and third error-free repetitions in each task to avoid effects of weight and attention shifting in the first repetition when the participant switches between targets. Only error-free repetitions are included in time measurements.

3.5.1 Outliers

Trials times more than 3 standard deviations from the mean for a target configuration were removed. This removed less than 2% of SELECTION TIME and ROUND TRIP TIME data points.

3.6 Results

Means over all conditions and participants were: SELECTION TIME: 309ms, ROUND TRIP TIME: 1203ms, ERROR RATE: 14.3%. ROUND TRIP TIME is 4 times longer than SELECTION TIME because it includes stationary time at each target. All main effects use a repeated measures ANOVA and all post hoc tests use a Bonferroni adjustment. A Greenhouse-Geisser correction is applied in the ANOVA where Mauchly's Test of Sphericity is significant, and corrected degrees of freedom are reported.

3.6.1 Learning Effect

A significant effect was found for BLOCK on ROUND TRIP TIME ($F_{1.074,10.74} = 24.451$, $p < .001$), and SELECTION TIME ($F_{1.077,10.77} = 8.917$, $p < .001$) but not ERROR RATE ($F_{2,20} = 3.073$, $p = 0.069$). Post hoc pairwise comparisons revealed a significant difference between all three blocks for both time measurements. SELECTION TIME decreased by 38ms (11%) from block 1 to 2 and by 19ms (6%) from block 2 to 3. ROUND TRIP TIME decreased by 259ms (18%) from block 1 to 2 and by 109ms (9%) from block 2 to 3. Garcia and Vu [7] suggest foot devices take time to learn, so this is expected. Given the decreasing learning trend, we only discard data in block 1 for the rest of the results.

3.6.2 Foot

No significant effect was found for FOOT on ERROR RATE ($F_{1,10} = 1.027$, $p = 0.335$), FOOT on ROUND TRIP TIME ($F_{1,10} = 2.816$, $p = 0.124$), or FOOT on SELECTION TIME ($F_{1,10} = 4.854$, $p = 0.052$). 95% confidence intervals indicate that if any difference, it is less than 4.7% for ERROR RATE, 46ms for ROUND TRIP TIME and 35ms for SELECTION TIME.

3.6.3 Target Size

A significant main effect was found for target SIZE on ERROR RATE ($F_{2.249,24.94} = 36.005$, $p < 0.001$). Pairwise comparisons showed that A45R20 had the lowest ERROR RATE of 3% compared to all other sizes, and A45R10 (9%) and A22R20 (6%) had a lower mean ERROR RATES compared to A11R20 (28%) and A45R5 (21%) (Figure 5c). Significant main effects were found for target SIZE on SELECTION TIME ($F_{1.738,17.38} = 6.617$, $p = 0.009$) and target SIZE on ROUND TRIP TIME ($F_{1.930,19.30} = 11.803$, $p < 0.001$). For SELECTION TIME, A45R20, A45R10, A22R20 had a significantly lower mean value (280ms), compared to A11R20 (334ms) (Figure 5b). For ROUND TRIP TIME. A45R20 (966ms) was significantly lower than A11R20 (1208ms) and A22R20 (1081ms); A22R20 (1081ms) and A45R10 (1045ms) were lower than A11R20 (1208ms) (Figure 5a).

3.6.4 Target Distance

A significant effect was found for target DISTANCE on SELECTION TIME ($F_{1,10} = 64.139$, $p < 0.001$) and target DISTANCE on ROUND TRIP TIME ($F_{1,10} = 49.938$, $p < 0.001$). Pairwise comparisons revealed a difference of 83ms [60ms, 107ms] (95% CI in quare brackets) or 25% for SELECTION TIME and a mean difference of 134ms or 11% [92ms, 177ms] for ROUND TRIP TIME (7.5 cm distance lowest for both). There was no effect for DISTANCE on ERROR RATE ($F_{1,10} = 0.105$, $p = 0.75$).

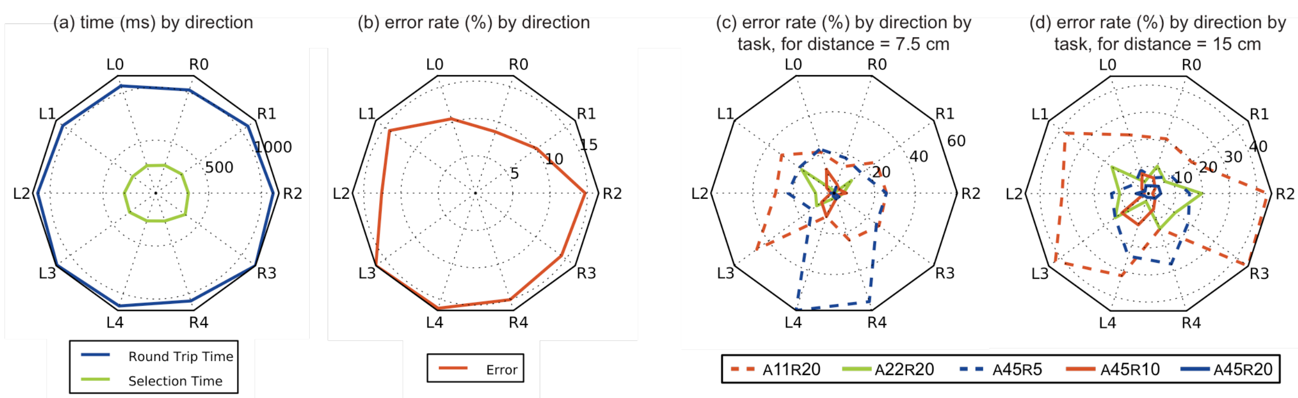


Figure 4: (a) time by direction for all tasks; (b) error rate by direction for all tasks;
(c) error rate by direction by target size for distance = 7.5 cm; (d) error rate by direction by target size for distance = 15 cm.

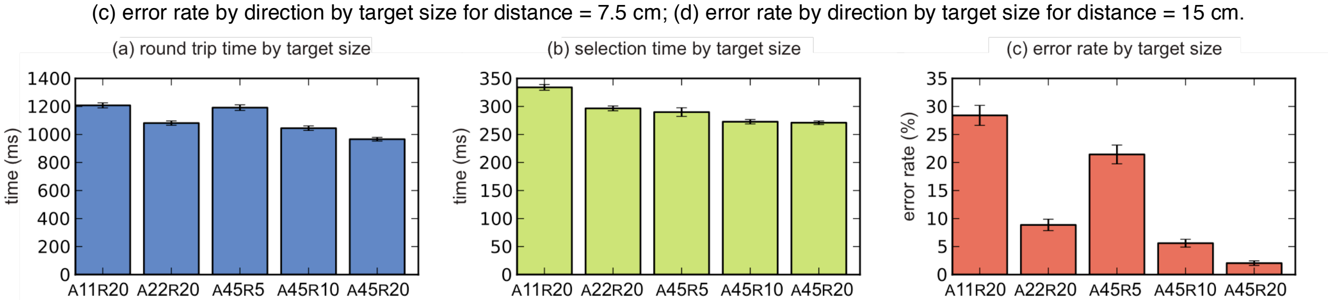


Figure 5. Effect of target SIZE on: (a) ROUND TRIP TIME; (b) SELECTION TIME; (c) ERROR RATE (all error bars are 95% CI).

3.6.5 Target Direction

A significant main effect was found for target DIRECTION on ERROR RATE ($F_{4,40} = 3.120$, $p = 0.025$) (Figure 4b). Pairwise comparisons found ERROR RATE for direction 0-forward, was 5.8% lower than direction 3-diagonal-backward, and direction 0-forward was 6.1% lower than direction 4-backward (all $p < 0.04$).

A significant main effect was found for target DIRECTION on SELECTION TIME ($F_{1.689,16.89} = 5.911$, $p = 0.014$) and ROUND TRIP TIME ($F_{1.881,18.81} = 6.384$, $p = 0.009$) (Figure 4a). For ROUND TRIP TIME, direction 0-forward was significantly lower than 4-backward, but a small 51ms difference [2ms, 101ms], $p = 0.042$). For SELECTION TIME, no pairwise differences were found.

3.6.6 Subjective Feedback

When interviewed, the majority of participants (11) reported using their toe to tap, 2 participants reported using their heel for at least some of the targets, and 2 only participants reported using the whole foot on some targets. No participants reported significant fatigue or discomfort. 8 experienced some minor fatigue or discomfort at some point and 3 reported no discomfort at all.

3.7 Discussion

Although task time is significantly affected by all variables except FOOT, the effect size is small – differences in time were generally less than 25%. With many repetitions, these small time differences may add up, but error rate has the largest effect on usability due to additional costs from user frustration and mistaken actions. We consider error rate to be the most important factor.

The most significant factor influencing error rate is target size. Targets with angular size less than 22.5° or radial size less than 5cm had error rates in excess of 20%, and should be avoided in a real system. The best target size considering both ERROR RATE and ROUND TRIP TIME was the largest (radial 20cm, angular 45°) with an error rate of 3%. Note this translates Han et. al.'s finds for directional kicking to target tapping [8]: they found participants could reliably kick in directions spaced 24° with 88% accuracy.

Our results show tapping forwards is easiest and the backwards and backwards-diagonal directions somewhat more difficult. There were moderate differences in ERROR RATE (about 5%) and very small differences in ROUND TRIP TIME (about 50ms). This translates Alexander et al. [1] results for directional kicking to tapping, they also report backwards kicking most difficult.

Distance has the greatest effect on SELECTION TIME and ROUND TRIP TIME, increasing both on the order of 100ms for a 7.5cm increase in distance.

There was no significant main effect for feet for the right-footed participants in this study, and 95% confidence intervals indicate the possible effect size is small. Foot dominance is not an important consideration in foot interaction.

With no reports of significant fatigue or discomfort in this 60-minute rather intensive experiment, 60 minutes may be a reasonable upper bound for continuous discrete foot input.

4 EXPERIMENT 2: KICKING AND FEEDBACK

Building on Experiment 1, the goal of our second experiment is to compare tapping and kicking and test performance with no cursor feedback. A direct comparison with kicking contextualizes different interaction options and previous related directional kicking work from Han et al. [8] and Alexander et al. [1], and tests whether the target size recommendations for tapping apply to kicking. Including a no feedback condition tests the feasibility of eyes-free foot input, where indirect cursor feedback is not available or the user's visual attention is focused elsewhere. To accommodate these additional factors, we reduce the number of target sizes by eliminating the lowest performing sizes from Experiment 1.

4.1 Participants

Twelve participants were recruited (5 female), ranging in age from 20 to 30. 12 reported they were right-footed, and 7 reported they previously used a whole body input device.

4.2 Apparatus

The same apparatus was used as Experiment 1, but the pointing action detection algorithm was modified to permit both tapping and kicking. To accomplish this, the algorithm ignored height and used speed and direction of travel only. Specifically, a pointing action was triggered when foot speed fell below 0.2 m/s, or direction of foot travel reversed along a vector from home target to task target. To avoid hysteresis, foot speed had to be greater than 0.3m/s and the foot had to move away from the home target before a pointing action was triggered. This enabled rapid taps and kicks.

Using an under-constrained detection algorithm has multiple benefits. First, it simplified the system and reduced unnecessary system errors. Second, it allowed participants to adopt a wider range of movements and pointing strategies that could inform system design. Third, we gather more representative data of tap and kick actions that could be mined in the future to tune the design of a tap or kick specific sensing algorithm. During the experiment participants were instructed to perform either taps or kicks and the experimenter monitored their adherence.

4.3 Tasks and Stimuli

The task and stimuli were the same as Experiment 1, but with a reduced subset of target size and distance variations to accommodate the new extra factors of POINTING ACTION and FEEDBACK.

4.3.1 Pointing Action

Two types of POINTING ACTIONS were tested: TAP and a mid air short KICK. To complete the task using the KICK action, the participant lifted one foot off the center home target, moved their foot in the air until the foot cursor hotspot was over the task target, and reversed direction to select it. They immediately returned their foot to the home position, tapping the floor with the foot cursor inside the home target. This cycle was repeated 3 times in rapid succession for the same foot and task target. The new detection algorithm also permitted the exact same TAP pointing action as Experiment 1 with either heel or toe taps.

4.3.2 Feedback

The FEEDBACK condition used the same red dot foot cursor as Experiment 1, but in the NO FEEDBACK condition this cursor was hidden. NO FEEDBACK was tested with both TAP and KICK pointing actions. Targets were shown with post-selection feedback in both conditions, and the change in color when a target was activated faded out over a brief period of time, rather than disappearing immediately. Error feedback was also made clearer with a soft error sound, and sounds for target selection.

Hiding the cursor in the NO FEEDBACK condition establishes if it is possible to have foot interaction occur without a person looking at a feedback display. They would only need prior knowledge of the position of targets they wish to activate and receive feedback resulting from the system action they selected.

4.3.3 Target Size and Distance

For the FEEDBACK condition, target SIZE was limited to three (A, R) pairs used in Experiment 1: A22R20 = (22.5°, 20 cm), A45R20 = (45°, 20 cm), and A45R10 = (45°, 10 cm). These were used in 4 combinations with two DISTANCES (7.5 cm and 17.5 cm): A45R20 at a distance of 7.5cm, A22R20 at distance of 7.5cm, A45R10 at a distance of 7.5cm, and A45R10 at a distance of 17.5cm. These combinations correspond to tasks from Experiment 1 with reasonable error rates. The far DISTANCE was increased to exactly simulate a system with two concentric, non-overlapping rings of 10cm radial size targets (since the 5cm radial size was eliminated due to high error rate).

For the NO FEEDBACK condition, we only used A45R20 at a distance of 7.5cm, as this task had the lowest error in the previous experiment and is the easiest target to select.

4.4 Design and Protocol

The independent variables are FOOT (left or right), target SIZE, target DIRECTION, target DISTANCE, POINTING ACTION (TAP or KICK), and FEEDBACK (FEEDBACK or NO FEEDBACK).

Tasks were divided into 10 target configuration sets of 10 tasks. Each set covered all values of 5 target DIRECTIONS and FOOT for one combination of SIZE, DISTANCE, POINTING ACTION, and FEEDBACK. All target configuration sets were presented in random order as one BLOCK. Participants completed three BLOCKS in order to test for learning effects. After the experiment, participants were asked to demonstrate their comfortable range-of-motion for tapping and kicking and complete a post-experiment questionnaire for subjective ratings for tapping and kicking and feedback and no feedback. The experiment took 60 minutes on average.

In summary the design was:
3 BLOCKS ×
2 POINTING ACTIONS ×
2 FEET × 5 target DIRECTIONS ×
(1 SIZE and DISTANCE with NO FEEDBACK
+ 4 SIZE and DISTANCE combinations with FEEDBACK) ×
3 repetitions of serial selections
= 900 data points per participant

4.5 Analysis

The primary dependent variables are the same as Experiment 1: ERROR RATE, SELECTION TIME, and ROUND TRIP TIME. A secondary dependent variable characterizes pointing action characteristics: DWELL TIME is the time that the task target was activated, i.e. the mean time that foot speed is below the 0.3 m/s threshold.

4.5.1 Outliers

Outliers were removed as in Experiment 1, removing less than 2% of SELECTION TIME, ROUND TRIP TIME, and DWELL TIME data points.

4.6 Results

All main effects use repeated measures ANOVA, all post hoc tests use a Bonferroni adjustment.

4.6.1 Learning Effect

A significant effect was found for BLOCK on SELECTION TIME ($F_{2,20}$=7.530, p=.003), ROUND TRIP TIME ($F_{2,20}$ = 14.947, p < .001) and ERROR RATE ($F_{2,20}$ = 7.390, p = 0.004). Post hoc pairwise comparisons revealed a significant (p < 0.03) difference in all three conditions between the first and third block. Block 1 was discarded as in the first experiment.

4.6.2 Foot

No significant main effect was found for FOOT on ERROR RATE ($F_{1,11}$ = 1.758, p = 0.212) or for FOOT on SELECTION TIME ($F_{1,11}$ = 0.956, p = 0.349), but a significant main effect was found for FOOT on ROUND TRIP TIME ($F_{1,11}$ = 7.624, p = 0.019). Pairwise comparisons showed that the left foot was slower than the right foot, but only by 24ms [5ms, 42ms].

4.6.3 Pointing Action

Analysis of POINTING ACTION is only applicable to tasks with FEEDBACK. A significant main effect was found for POINTING ACTION on SELECTION TIME ($F_{1,11}$ = 17.642, p = 0.001). Pairwise comparisons showed TAP is 34ms [16ms, 52ms] faster, taking 247ms compared to 284ms for KICK. There was also a significant difference for POINTING ACTION on DWELL TIME ($F_{1,11}$ = 42.964, p < 0.001). The mean difference was 187ms [124ms, 250ms], with KICK having the lowest dwell time of 121ms vs. TAP with a DWELL TIME of 308ms. There were no significant effects of POINTING ACTION on ROUND TRIP TIME ($F_{1,11}$ = 0.961, p = 0.348) or POINTING ACTION on ERROR RATE ($F_{1,11}$ = 0.152, p = 0.704).

4.6.4 Feedback

We compare using target A45R20 at distance 7.5cm for FEEDBACK and NO FEEDBACK. A significant main effect was found for FEEDBACK on ERROR RATE ($F_{2,11}$ = 28.859, p<0.001). NO FEEDBACK had a mean error rate of 27.5%, while FEEDBACK had a mean error rate of 4.4% (mean difference 23.5%, [32.6%, 13.7%]). No significant effect was found for FEEDBACK on SELECTION TIME ($F_{2,11}$ = 0.274, p = 0.611), or on ROUND TRIP TIME ($F_{2,11}$ = 3.604, p = 0.084) and no significant interaction effects were found involving FEEDBACK.

4.6.5 Heel or Toe

Since our target selection algorithm tests if either the front or back of the foot is near the floor, we can examine the data to see how participants naturally tap. For this analysis, we classify tap type using the angle Θ of the heel-to-toe vector above the floor plane at selection time. If Θ > 5°, we consider it a heel tap; if Θ < -5° it is a toe tap; otherwise a whole foot tap. Using this metric, participants tapped with their heel in 17% of the trials, their whole foot in 29%, and their toe in 54%. The distribution of tap type by foot

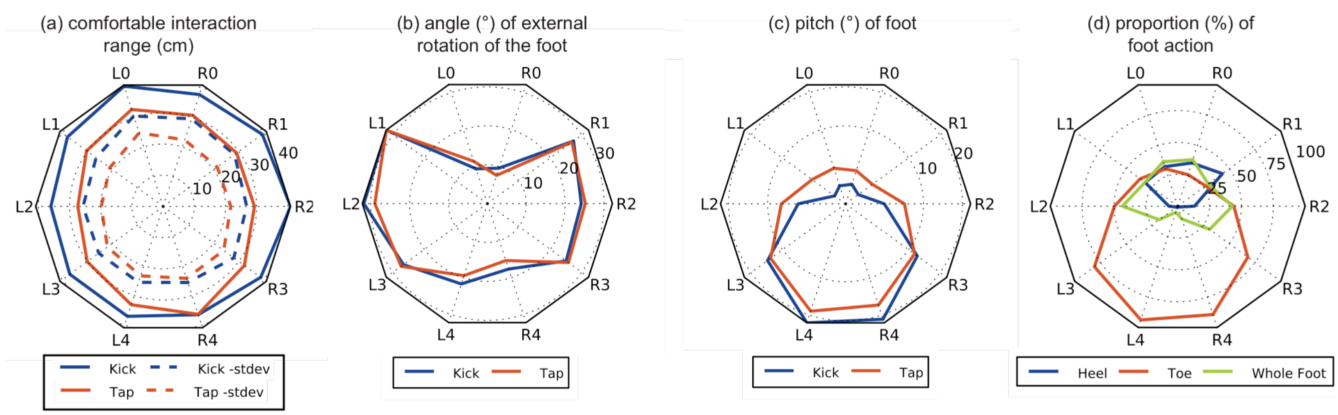


Figure 6. (a) Comfortable range-of-motion in cm for tapping and kicking as demonstrated by participants; (b) angle of external rotation of the foot in degrees; (c) pitch of foot in degrees; (d) proportion of heel, toe, or whole foot actions based on Vicon log analysis

is near-symmetric, but there are proportionally more toe taps to backwards targets (Figure 6d)

4.6.6 Comfortable Interaction Range

At the end of the experiment, we asked participants to demonstrate the maximum range they would be comfortable interacting at with both of the techniques. Using the Vicon tracker, we processed these demonstrations into range of motion by distance (Figure 6a). We include the mean value and the mean value less one standard deviation as a more conservative estimate. The comfortable range of motion is 30cm for TAP on average, and for KICK it is roughly 40 cm on average, falling to 35 cm in the backwards direction. Reducing the area by one standard deviation yields a conservative estimate of 20cm for TAP and 30 cm for KICK.

4.6.7 Foot Rotation and Pitch

We calculated the external rotation and pitch of the foot using Vicon logs. External rotation is the outwards angle of foot rotation around the heel axis relative to the foot pointing forward (if rotated inward, it would be negative external rotation). Pitch is similar to plantar flexion but the angle is measured relative to the floor. Lowering the toe relative to the heel increases pitch (raising the toe relative to the heal leads to negative pitch). Participants externally rotated the foot more on side and diagonal targets (Figure 6b), and lifted the heel more on backwards targets (Figure 6c).

4.6.8 Post-Experiment Questionnaire

The post experiment questionnaire asked for a preference between TAP and KICK pointing action with FEEDBACK on a 5 point Likert scale. 8 participants preferred or strongly preferred TAP over KICK as an interaction technique. However, 9 participants agreed or strongly agreed that KICK was easy to perform physically, 11 participants agreed or strongly agreed that KICK was easy to learn, and all participants agreed or strongly agreed that KICK was easy to perform mentally. Similar to Experiment 1, 5 participants reported minor discomfort, 3 of them attributing the discomfort to either KICK or KICK with NO FEEDBACK.

4.7 Discussion

We did not find a large quantitative difference between tapping and kicking pointing actions in time or error rate measures. Participants showed an overall preference for tapping, but kick was not rated poorly on an absolute scale.

The lower selection time for tapping may be due to how the foot rapidly decelerates when it contacts the floor, while kicking requires the foot to decelerate using only leg muscles. The faster tap selection time occurs in spite of tap having a larger DWELL TIME. This can be explained by tapping requiring more time to accelerate after stopping and touching the floor. This difference in tapping and kicking motion characteristics might be exploitable in conjunction with foot height for robust discrimination between them.

With an error rate above 20%, the NO FEEDBACK condition as tested may be unfeasible for use in a real system. However, given the difficulty of this task, this result is encouraging. An acceptable error rate is within reach by increasing the radial target size.

5 IMPLICATIONS FOR DESIGN

We apply our results to create design guidelines, illustrated with example foot interaction technique designs.

5.1 Design Guidelines

Based on experiment results, we propose ten design guidelines for indirect foot pointing using discrete taps and kicks while standing:

G1 Tapping and kicking are both feasible, but users have a slight preference for tapping: use tapping for more frequent actions.

G2 People use both feet equally well; any effect of foot dominance is small. This means the user preference to alternate feet [13] is supported without increased time or errors.

G3 When tapping, people prefer toe taps. Use toe taps for most common actions, then whole foot taps, then heel taps.

G4 All of the investigated target directions are feasible. Forward movement is less error prone to use, and backwards and backwards-diagonal interaction are hardest to use.

G5 For indirect cursor feedback, target angular size should be at least 22.5°; two target levels is feasible with radial size 10cm.

G6 Without cursor feedback, target angular size should be much greater than 45°: a conservative recommendation is 90°.

G7 Increasing distance of targets within reasonable limits increases interaction time, but does not increase error.

G8 A conservative estimate for an appropriate interaction radius is 20 cm for tap interaction, and 30 cm at the front and 25 cm in radius at the back for kick interaction.

G9 60 minutes of continual foot interaction, with occasional breaks, is feasible for users to do with only minor discomfort.

G10 Sensing techniques must be robust to changes in foot pitch and external rotation of the feet with sideways motion.

5.2 Example Design: Foot Input at a Standing Desk

To illustrate our design guidelines and demonstrate the utility of an indirect, discrete tapping and kicking interaction space, we describe a design for a standing desk foot input system (Figure 7a). People use standing desks to avoid health problems caused by sedentary computing [3], so such a system could enable foot input breaks for increased physical activity [13] by occasionally using feet to control applications instead of a mouse and keyboard. This would require a foot input technique resilient to false positives. We imagine tasks such as reading documents, browsing web pages, or even code debugging. Targeting short, occasional usage keeps continuous foot input well below 60 minutes, as per G9.

Based on our design guidelines, we describe how two target layouts, low-density and high-density, can be combined with foot action and feedback. In addition, deployable methods to sense foot action and location are briefly described.

5.2.1 Low-density Target for Invoking Commands

Application control, such as scrolling a document, should be eyes-free (not requiring constant feedback of foot input). We propose that taps and kicks on low-density targets can be used for this purpose (Figure 7b). As G6 suggests, these targets must be larger than 45° in angular size and 20cm in radial size. False positives can be reduced with explicit activation of foot input using the keyboard and mouse, or by adjusting the target layout to be centered on the foot to compensate for changes in static stance. To help learn the input space, the system should show foot cursors and target locations in a small, but always-visible location such as a side-bar with most recently sensed foot actions clearly indicated.

With 4 actions (toe tap, heel tap, whole foot tap, or kick) and 6 virtual target locations (using both feet in all directions by G4), 24 commands can be accessed at one time. Following G1, we would map less frequent commands to kicks such as application switching or swapping the command set. G4 suggests mapping frequent commands to forward actions and G3 suggests most frequent to toe taps. For example, close document could be a whole-foot back tap and scrolling the document up could be a forward toe tap. Since G2 suggests people can use both feet equally well and peo-

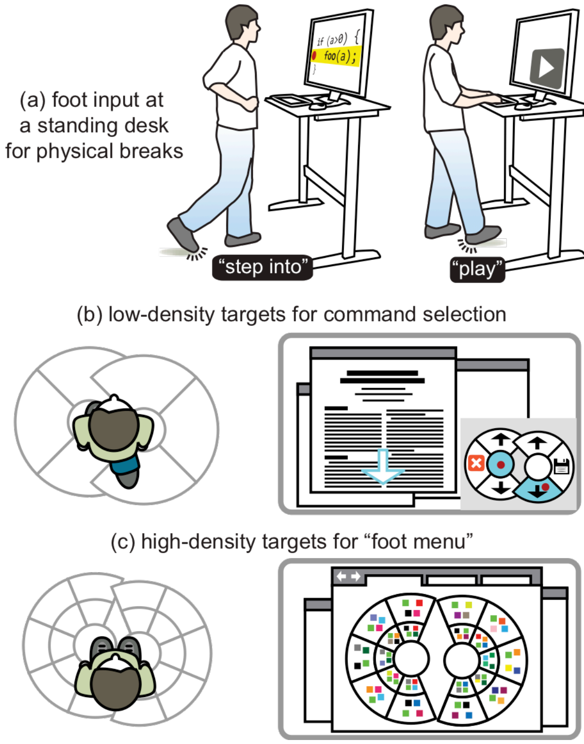


Figure 7. (a) the example foot input system for controlling conventional applications at a standing desk: (b) low-density targets for command invocation with minimal feedback; (d) high-density targets for "foot menu" with full indirect pointing feedback.

ple prefer to alternate feet, we would duplicate command mappings to be symmetrical across feet. For example, scroll up should be mapped to both left and right forward taps.

5.2.2 High-density Targets for Special-purpose Foot Menu

By focusing on indirect feedback with the foot cursor, a high-density layout of targets could be used to choose from browser bookmarks (Figure 7c). Based on G5, targets could have an angular size of 45° (twice the tested minimum) and be positioned in two rings, from 10 to 20cm, and from 20 to 30cm (a compromise of comfortable ranges for tap and kick from G8).

With 4 actions (toe tap, heel tap, whole foot tap, or kick), 10 virtual target locations (using both feet in all directions by G4), and 2 target bands, 80 bookmarks could be accessed. Frequent bookmarks could be placed to optimize preferred directions and actions from G3 and G4. Our results suggest this layout would have an error rate of about 6%.

5.2.3 Sensing

There are practical sensing techniques that can be deployed in real settings to enable this kind of interaction. Position tracking can be achieved with a depth sensing or conventional camera or with accelerometers mounted on the leg or on footwear. Differentiation between different foot actions can be accomplished using pressure sensors embedded in the sole of footwear, or by making use of other characteristics (such as dwell time). Sensing technology must handle the movement characteristics noted by G10.

6 Conclusion and Future Work

Our work contributes fundamental empirical results summarized in practical design guidelines for foot-based discrete pointing while standing. Our example standing desk input system shows how these guidelines can be incorporated into new foot input vocabularies enabling new foot interaction techniques. This style of discrete foot pointing while standing can be used in other contexts, such as with watches, smartphones, large private and public displays, tabletops, and head-mounted displays.

As future work, we plan to build a fully working system following the design above and evaluate this style of foot interaction in a real setting. In addition, although we cover a useful subset of the discrete foot pointing while standing input space, we see avenues for more formative work: such as examining eyes-free foot input more closely; controlling for toe, heel, or whole foot tapping; combining discrete taps and kicks with Scott et al.'s [18] heel and toe pivoting gestures; incorporating continuous input; perhaps even adding physical floor mounted props. The identification of these future directions, and methodology for their exploration, are enabled by the results presented here.

7 References

1. Alexander, J., Han, T., Judd, W., Irani, P., and Subramanian, S. Putting Your Best Foot Forward: Investigating Real-world Mappings for Foot-based Gestures. *Proc. CHI*, ACM (2012), 1229–1238.
2. Augsten, T., Kaefer, K., Meusel, R., et al. Multitoe: high-precision interaction with back-projected floors based on high-resolution multi-touch input. *Proc. UIST*, ACM (2010), 209–218.
3. Berque, D., Burgess, J., Billingsley, A., Johnson, S., Bonebright, T.L., and Wethington, B. Design and evaluation of persuasive technology to encourage healthier typing behaviors. *Persuasive* (2011), 9:1–9:10.
4. Crossan, A., Brewster, S., and Ng, A. Foot tapping for mobile interaction. *Proc. BCS Interaction Specialist*, (2010), 418–422.
5. Dearman, D., Karlson, A., Meyers, B., and Bederson, B. Multi-modal text entry and selection on a mobile device. *Proc. GI* (2010), 19–26.
6. Drury, C.G. Application of Fitts' Law to foot-pedal design. *Human Factors: 17*, 4 (1975), 368–373.
7. Garcia, F.P. and Vu, K.-P.L. Effectiveness of hand- and foot-operated secondary input devices for word-processing tasks before and after training. *Computers in Human Behavior 27*, 1 (2011), 285–295.
8. Han, T., Alexander, J., Karnik, A., Irani, P., and Subramanian, S. Kick: investigating the use of kick gestures for mobile interactions. *Proc. Mobile HCI*, (2011), 29–32.
9. Hoffmann, E.R. A comparison of hand and foot movement times. *Ergonomics 34*, 4 (1991), 397–406.
10. Hoysniemi, J. International survey on the Dance Dance Revolution game. *Comput. Entertain. 4*, 2 (2006).
11. Jota, R., Lopes, P., Wigdor, D., and Jorge, J. Let's Kick It: How to Stop Wasting the Bottom Third of Your Large Screen Display. *Proc. CHI* (2014), 1411–1414.
12. Kim, S.-H. and Kaber, D.B. Design and Usability Evaluation of Foot Interfaces for Dynamic Text Editing. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 52*, 6 (2008), 581–585.
13. Meyers, B., Brush, A.J.B., Drucker, S., Smith, M.A., and Czerwinski, M. Dance your work away: exploring step user interfaces. *CHI Ext. Abstracts*, ACM (2006), 387–392.
14. Paelke, V., Reimann, C., and Stichling, D. Foot-based mobile interaction with games. *Proc. SIGCHI International Conference on Advances in computer entertainment technology*, ACM (2004), 321–324.
15. Pakkanen, T. and Raisamo, R. Appropriateness of foot interaction for non-accurate spatial tasks. *CHI Ext. Abstracts*, (2004), 1123–1126.
16. Pearson, G. and Weiser, M. Exploratory Evaluations of Two Versions of a Foot-Operated Cursor-Positioning Device in a Target-Selection Task. *SIGCHI Bull. 19*, 3 (1988), 70–75.
17. Schmidt, D., Block, F., Gellersen, H. A comparison of direct and indirect multi-touch input for large surfaces. *INTERACT*, 2009, 582–594.
18. Scott, J., Dearman, D., Yatani, K., and Truong, K.N. Sensing foot gestures from the pocket. *Proc. UIST*, ACM (2010), 199–208.
19. Doug Engelbart Discusses Mouse Alternatives. ftp://ftp.cs.utk.edu/pub/shuford/terminal/engelbart_mouse_alternatives.html.