Interactions under the Desk:
A Characterisation of Foot Movements for Input in a
Seated Position

Eduardo Velloso[1], Jason Alexander[1], Andreas Bulling[2] and Hans Gellersen[1]

[1] School of Computing and Communications, Infolab21, Lancaster University,
Lancaster, LA1 4WA, United Kingdom
`{e.velloso, j.alexander}@lancaster.ac.uk, hwg@comp.lancs.ac.uk`
[2]Max Planck Institute for Informatics, Perceptual User Interfaces Group, Campus E1 4, 66123
Saarbrücken, Germany
`andreas.bulling@acm.org`

Abstract. We characterise foot movements as input for seated users. First, we built unconstrained foot pointing performance models in a seated desktop setting using ISO 9241-9-compliant Fitts's Law tasks. Second, we evaluated the effect of the foot and direction in one-dimensional tasks, finding no effect of the foot used, but a significant effect of the direction in which targets are distributed. Third, we compared one foot against two feet to control two variables, finding that while one foot is better suited for tasks with a spatial representation that matches its movement, there is little difference between the techniques when it does not. Fourth, we analysed the overhead caused by introducing a feet-controlled variable in a mouse task, finding the feet to be comparable to the scroll wheel. Our results show the feet are an effective method of enhancing our interaction with desktop systems and derive a series of design guidelines.

Keywords: Foot-based interfaces, Fitts's Law, Interaction Techniques

1 Introduction

Computer interfaces operated by the feet have existed since the inception of HCI [1], but such devices remained restricted to specific domains such as accessible input and audio transcription, being largely overshadowed by hand-based input in other areas. However, this overshadowing cannot be put down to lack of dexterity, as we regularly accomplish a wide variety of everyday tasks with our feet. Examples include the pedals in a car, musicians' guitar effect switches, and typists' use of transcription pedals. Recent technological advances renewed interest in foot-based input, be it for interacting with a touch-enabled floor [2], for hands-free operation of mobile devices [3], or for adding more input channels to complex tasks [4]. Despite this, we still lack a thorough understanding of the feet's capabilities for interacting in one of the most common computing setups—under the desk.

In particular, unlike previous work that used trackballs [5], pedals [6], and foot mice [7], we wished to explore unconstrained feet movements. This removes the need for a

physical device (as well as the related foot-to-device acquisition time) and provides a wide range of interaction possibilities (analogous to the ones available from a touch-screen over a mouse).

We envision numerous applications to arise from this greater understanding. These include using your feet to scroll a page while the hands are busy with editing the document, changing the colour of a brush while moving it with the mouse, or manipulating several audio parameters simultaneously (using both the hands and the feet) to create novel musical performances.

To address this gap we conducted a series of experiments exploring different aspects of foot-based interaction. In the first, we recorded 16 participants performing 1D and 2D pointing tasks with both feet to build the first ever ISO 9241-9 Fitts's Law models of unconstrained foot pointing for cursor control. This first study provided some evidence that side-to-side movement is faster than backwards and forwards. To confirm this hypothesis, we conducted a second experiment in which participants performed 1D serial pointing tasks in each direction. In the third experiment, we investigated the manipulation of multiple parameters using one and two feet. In the fourth and final experiment, we evaluated the use of the feet together with the hand.

In summary, (1) we built 1D and 2D ISO 9241-9 compliant movement time models for unconstrained foot pointing; (2) we found that unconstrained foot pointing is considerably slower than mouse pointing, but comparable to other input devices such as joysticks and touchpads; (3) we found no significant difference in performance between the dominant and non-dominant foot; (4) we found that left and right movement is easier than backwards and forwards; (5) the most comfortable movement for desktop foot interaction is heel rotation; (6) techniques that have a direct spatial mapping to the representation outperform the others; (7) when variables are shown separately, two feet work better than one; (8) we show that the feet perform similarly to the scroll wheel in tasks where the feet are used in conjunction with the mouse; and (9) we provide design guidelines and considerations based on our findings.

2   Related Work

Foot-Operated Interfaces: Early prototypes of foot-based interfaces aimed at reducing the homing time between switching from the keyboard to the mouse. Examples for such interfaces include English et al.'s knee lever [1] and Pearson and Weiser's moles [8]. Other works employed the feet for a variety of tasks including cursor control [1, 8], mode selection [9], spatial navigation [10], mobile phone control [3], command activation [11], gaming [12, 13], tempo selection [14, 15], user identification [16] and text input [6].

Tracking the Feet: Feet interfaces appear in the literature in different forms. Peripheral devices include foot mice, pedals, and switches. Wearable devices obtain input from sensors embedded in users' clothing [12], footwear [17] or mobile devices [3]. For tracking motion, inertial measurement units (accelerometer, gyroscopes and magnetometers) are commonly employed [3, 12], whereas pressure sensors [17] and textile switches [10] are used to detect button pressings and gestures such as heel and

toe clicking. Wearable tracking systems are usually more mobile and individualised than remote ones, but require some user instrumentation.

Conversely, remote tracking approaches rely on augmenting the environment where the system is going to be used, which usually means less instrumentation on the user and more versatility in cases where multiple users share the same system at different points in time. Sensing may be performed by conventional colour cameras [13], depth cameras [4, 16], optical motion capture [18], audio [15] and smart floors [2].

Quantitative Evaluations of Foot Interfaces: Analyses of feet performance date back to Ergonomics work by Drury [19] and Hoffman [20], who asked participants to tap on physical blocks on the floor. Springer and Siebes compared a custom-built foot joystick to a hand-operated mouse in an abstract target selection task [21]. Pakkanen et al. investigated the performance of trackballs operated by the feet and by the hand in common graphical user interface tasks [5]. Dearman et al. compared foot switches on a pedal to screen touch, device tilt, and voice recognition in text editing tasks [6]. Garcia et al. looked at how performance evolves as users learn to operate a foot joystick and a hand trackball [7]. Table 1 summarizes the results of these studies.

Table 1 - Previous studies on the performance of hand and feet pointing for seated users, summarising number of participants (N) and the ratios of task completion times and error rates between the feet and hands.

Foot Device	Hand Device	N	Foot Time / Hand Time	Foot Error / Hand Error	Ref
None	None	10	1.95[1]	-	[20]
None	None	10	1.7[2]	-	[20]
Joystick	Mouse	17	2.32	1.56	[21]
Trackball	Trackball	9	1.6	1.2	[5]
Pedals	Tilt	24	1.05[3]	1.20[3]	[6]
Pedals	Touch	24	0.98[3]	1.87[3]	[6]
Joystick	Trackball	16	1.58[4]	-	[7]

1 Ratio between the reported coefficients of the ID for *visually controlled* movements.
2 Reported ratio for *ballistic* movements.
3 Ratio between reported means for selection time and error rate in the text formatting task.
4 Mean ratio between reported task completion times for the foot joystick and the mouse.

These works evaluated several different foot interfaces, but the wide variety of experimental designs makes comparing results difficult. Further, these studies have only looked at 1st order devices (i.e. devices that control the rate of change of a value, rather than the value directly, such as the joystick and the pedals) and relative input devices (i.e. devices that sense changes in position, such as the mouse and the trackball) [22]. Hoffman investigated unconstrained absolute positioning, but with users tapping on physical targets rather than using the foot for cursor control. However, modern devices that take input from the feet, such as depth cameras [4] and interactive floors [2] use absolute positioning, sometimes without physical proxies, making it important to study this kind of interaction.

3 Study 1: Fitts's Law Performance Models

To fill this gap, we conducted an experiment in which 16 participants performed 1D and 2D pointing tasks with both feet, to build the first ever ISO 9241-9 Fitts's Law models of unconstrained foot pointing for cursor control. This allows us to compare our model to those of other input devices based on the same standard using the mean throughput for each condition. We also tested for effects of task and foot on user performance.

We recruited 16 participants (11M/5F), aged between 20 and 37 years (median = 27), with foot sizes ranging from 23 to 30cm (median = 26cm). Participants were inexperienced with foot interfaces and half of them were regular drivers. All participants were right handed, but one was ambidextrous. All participants were right footed.

The experiment was conducted in a quiet laboratory space, on a laptop with an 18-inch screen and 1920×1080 resolution. To track the feet, we used an implementation of Simeone et al.'s tracker [4]. This system uses a *Kinect* sensor mounted under the desk and a MATLAB program that subtracts the background, converts the coordinate system from the camera plane to the floor plane, isolates the feet and fits ellipses to the remaining data. The ellipses' foci are then used to approximate the position of the toes and heels. The tracker worked at 26 frames per second. We made sure that only one foot was visible to the camera at any point in time, by asking participants to keep the opposite foot under the chair, and that the cursor control was assigned to the toes of whichever foot was in view. Mouse clicks were performed using a conventional mouse with disabled movement tracking. The tracker was calibrated with a 1:1 CD gain, so that the cursor and foot movements matched exactly.

Participants performed 1D and 2D Fitts's Law tasks, for which we used Wobbrock et al.'s *FittsStudy* tool, an ISO-9241-9-compliant C# application to "administer conditions, log data, parse files, visualize trials and calculate results" [23]. The tool was configured to administer nine different combinations of A (amplitude) × W (width) defined by three levels of A {250, 500, 1000} crossed with three levels of W {20, 60,130}, yielding nine values of ID {1.55, 2.28, 2.37, 3.12, 3.22, 3.75, 4.14, 4.7, 5.67}.

We recorded all sessions using additional cameras pointed at participants' faces and feet, as well as the screen using the *Open Broadcaster Software* (see Figure 1).

3.1 Procedure

Participants first signed a consent form and completed a personal details questionnaire. The tasks were conventional ISO 9241-9 pointing tasks, in which targets appeared in blue on the screen. Participants selected targets by moving their feet so the cursor was above the target and by left-clicking the hand-held mouse. We chose this technique rather than foot tapping as we were interested in the time it takes to move the feet and the gesture time might delay the task unnecessarily.
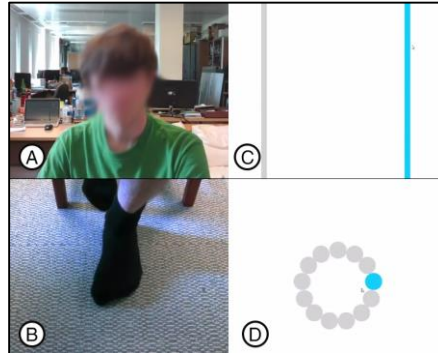
Figure 1 - We recorded participants faces (A) and feet (B), synchronized with the 1D (C) and 2D (D) tasks.

Participants performed both a 1D task (with vertical ribbons on either side of the screen) and a 2D task (with circular targets in a circular arrangement), with both their dominant and non-dominant foot. The order of tasks was randomized, but we ensured that the same foot was not used twice in a row. Each task comprised 9 *IDs* and was repeated in 13 trials (the first 3 discarded as practice). To summarize, each participant performed 2 feet × 2 tasks × 9 *IDs* × 13 trials = 468 movements. To make the friction with the floor uniform across all users, we asked them to remove their shoes and perform the tasks in their socks.

After completing the tasks we asked participants to fill in a questionnaire adapted from the ISO 9241-9 standard for the use with the feet (see Figure 2(B)). We also conducted an open-ended interview about participants' experience using the foot interface, what they liked and disliked about it, what strategies and movements they used to reach targets, etc. All interviews were transcribed and coded accordingly.

3.2 Results

Our analysis had two objectives: to build a Fitts's Law performance model for each task and each foot and to check whether there was any difference in performance—as measured by the throughput—for different feet and tasks.
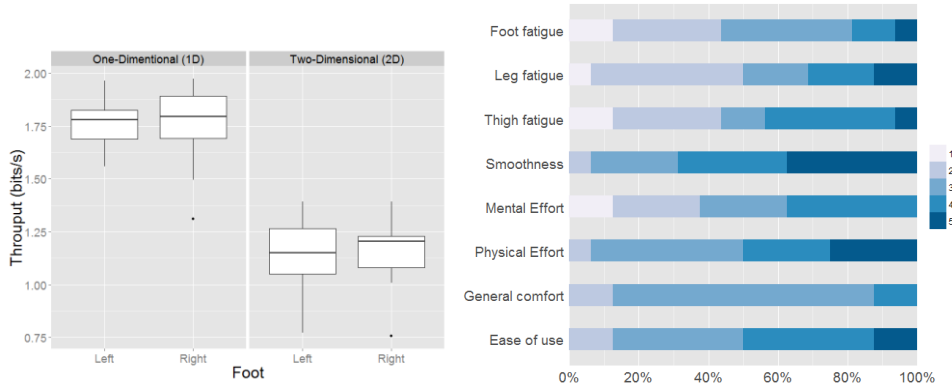
Figure 2 - Study 1 Results: (A) Mean throughput for each task and foot and (B) Subjective reactions to the interaction technique (1-Low, 5-High)

To build the performance models, we computed the mean movement time (*MT*) and the mean 1D and 2D effective indices of difficulty (*IDe*) for each participant and for each combination of $A \times W$. We then built the performance models using linear regression on these data, using the formulation described by Soukoreff and MacKenzie [24]. Table 2 summarizes the movement time models, as well as the R-squared and the mean throughput averaged over the individual mean throughputs for each participant. To test for differences in performance for each condition, we compared the mean throughput (see Figure 2(A))—a metric that takes into account both speed and accuracy of the movement performance [24]—using a factorial repeated-measures ANOVA. We found a significant main effect of the task on the average throughput ($F_{1,15}$=391, p<.001), with an average *TP* of 1.75 for the 1D task and 1.15 for the 2D task, but not of the foot ($F_{1,15}$=.75, p=.09) or the interaction between foot and task ($F_{1,15}$=.41, p=.12) (see Table 2). Participants' subjective ratings covered the whole scale (see Figure 2(B)), suggesting that some people like it, whereas other people hated it.

Table 2 – Performance model for each condition with its corresponding r-squared, mean throughput and error rate

Condition	Movement Time (ms)	R^2	TP (bit/s)	Error (%)
1D Right	$99 + 561 \times ID$	0.88	1.75	8.43
1D Left	$-56 + 609 \times ID$	0.96	1.75	8.49
2D Right	$423 + 739 \times ID$	0.85	1.16	7.64
2D Left	$372 + 789 \times ID$	0.75	1.14	8.60

After transcribing and coding the interviews, some consistent patterns of users' opinions emerged.

Movement Behaviours: In general, participants preferred to move around their hips and knees as little as possible, leaving as much of the movement as possible to the ankle joints. Participants reported five strategies for reaching targets on the screen: dragging the foot, lifting the foot, rotating the foot around the heel, rotating the foot around the toes and nudging the toes. At the beginning of the tasks, participants often started by

dragging the foot across the floor, but quickly realized that this was tiring *("was a bit uncomfortable", "I could instantly feel my abs working", "more taxing and not really natural")*. Four participants reported lifting the foot across the floor, but found that keeping the foot up was rather tiring (*"I'd have more control and I don't have the friction of the surface, but then I got very fatigued from keeping my whole leg up")*.

These strategies were used when targets were far apart; for shorter distances participants reached the targets by rotating the foot around the heels with the toes up, what they often referred to as "pivoting" (*"most of the time, I just tried to move around my heel")*. The reported advantages of heel rotation were the ease of movement, less fatigue, higher comfort and higher precision. Finally, for small adjustments and smaller targets, participants employed the toes in two ways: one participant reported rotating the foot around the toe and six participants reported bending and extending their toes, which would nudge the cursor towards the target (*"when I wanted to do a fine grained, on the smaller targets, I would crunch my toes")*.

Differences between Tasks: All participants but one found the one-dimensional task easier than the two-dimensional one, which is reflected in the quantitative difference in throughput. This can be explained by the fact that moving left and right could be accomplished with heel rotation (the easiest movement, as participants reported), whereas back and forth movements required knee flexion and extension, either by dragging the foot on the floor or lifting it above it, both strategies that were reported as being tiring.

Challenges: The biggest challenges reported by participants were the cognitive difficulty in reaching small targets (*"when the targets are smaller you need more precision so you need to focus")* and in coordinating the hands and feet (*"it was weird starting, because you'd have to coordinate your thought process, your clicking and your feet, but I think as you went on, It was pretty quick to adapt")*, fatigue (*"a little fatigue influenced the outcome")*, friction with the floor (*"I don't like this kind of rubbing with the floor")*, and overshooting (*"I knew that I was going to overshoot, so I just overshot and tried to click at the same time")*.

3.3 Discussion

Our regression models are in line with previous work as our one-dimensional model is very similar to Drury's ($MT = 189 + 550 \times ID$) [19]. Hoffman found a much lower coefficient ($MT = -71 + 178 \times ID$) [20], but both him and Drury conducted experiments with physical targets rather than cursor control. As Drury noted, this effectively increases the sizes of the targets by the size of the participant's shoe [19]. Also, whereas we use the Shannon formulation of *ID*, Drury used Fitts's original formulation and Hoffman used the Welford formulation.

Since our model is compliant with the ISO standard, we can compare our throughputs with other studies reported in literature. The typical range of throughput for the mouse is between 3.7 and 4.9 bit/s, considerably higher than the 1.2–1.7 range we found for the feet, but expected given users' experience and practice with it [24]. The values we found, however, fall into the range for other input devices such as the

isometric joystick (1.6–2.55) [24], the touchpad (0.99–2.9) [24] and video game controllers (1.48–2.69) [25].

By allowing participants to choose how to reach the targets, we obtained valuable insights into the most comfortable ways of using the feet. Although heel rotation was perceived as the most comfortable movement, most foot-operated interfaces do not use this movement (an exception is Zhong et al.'s *Foot Menu* [26]). Our results are also in line with Scott et al.'s in which users also reported that heel rotation was the most comfortable gesture, followed by plantar flexion, toe rotation and dorsiflexion [3]. The use of heel rotation is suitable for radial and horizontal distributions of targets. This kind of interaction could be used in a discrete (e.g. for foot activated contextual menus) or in a continuous fashion (e.g. controlling continuous parameters of an object while the hands perform additional manipulations).

We investigated foot performance for seated users so our results apply to foot-only (e.g. mice for people with hand disabilities), and foot-assisted (e.g. driving simulators, highly-dimensional applications) desktop interfaces. It remains to be seen how these results apply to standing users (e.g. using a touch-enabled floor). A second limitation is that our participants were not familiar with this kind of input device, which might affect the predictive power of our models if the device is used more frequently.

4   Study 2: Effects of Foot and Direction of Movement

One possible use for the feet in a seated position is to provide one-dimensional input, be it discrete (e.g. selecting an option in a menu) or continuous (e.g. changing the music volume). To better understand how to design such interfaces, it is important to understand if there is a significant difference in the movement times and comfort between different directions of movement. In this study, we tested the effects of the direction in which the targets were distributed (horizontal vs. vertical) and the foot (dominant vs. non-dominant) on the movement times and error rates.

For this experiment, we recruited 12 participants (8M/4F), aged between 19 and 31 years (mean 27), with posters on campus and adverts on social networks. All participants were right handed and one was left footed; seven participants were car drivers. Foot sizes ranged from 22 cm to 33 cm (mean 27.1 cm). None of the participants had ever used a foot mouse or similar foot-operated pointer before. The experimental setup was the same as in Study 1.

4.1   Procedure

To begin, participants signed a consent form and filled in a questionnaire. The task in our experiment was analogous to other Fitts's Law experiments. The user was presented with a green and a red bar, in either horizontal or vertical orientation, with a certain width (W) and separated by a certain distance (A) on the screen. For each trial, the user had to select the green bar, at which point the colours of the bars switched.

To select a target, participants used their feet to position an on-screen cursor over the green bar and press the space bar. We selected 14 combinations of W and A to yield

exact indices of difficulty from 1 to 7, using the Shannon formulation. Each ID combination was executed with each foot in both horizontal and vertical configurations. We balanced the order of the feet and the direction of the bars among participants but and ensured the task was not repeated with the same foot twice in a row so as to reduce fatigue. The order of difficulty was randomised. The complete procedure was repeated ten times. To summarise, each participant performed 2 feet ×2 directions×14 ID combinations×10 repetitions = 560 movements.

The system continually logged the position of the feet and cursor and a video camera placed under the desk recorded participants' leg movement. At the end of the experiment participants filled in another questionnaire on the perceived difficulty and speed of the target selection on the top, right, bottom, left and centre of the screen for each foot. We also asked for suggested applications of foot-operated interfaces.

4.2   Results

To compare the conditions we computed the mean of means of the throughput of each case. We compared the throughputs using a factorial repeated-measures ANOVA. We found that horizontal movements had a significantly higher throughput (2.11 bit/s) than vertical ones (1.94 bits/s), $F_{1,11} = 14.06$; $p < .05$, but dominance of the foot used had no significant effect, $F_{1,11} = 4.62$, $p = .055$, on the task completion time. We also found no significant interaction effect between the foot and the direction of movement, $F_{1,11} = 4.72$, $p = .052$, indicating that both feet perform roughly the same in both directions.

4.3   Discussion

The results from this study confirm our hypothesis that it does not matter which foot is used, but moving it horizontally is faster than moving it vertically. We attribute this phenomenon to the possibility of pivoting the foot when moving it horizontally and the necessity to drag or lift the foot when moving it vertically.

Suggestions for tasks that could be improved by the use of the feet together with traditional input modalities pointed to the fact that it is not suitable for fine positioning, but it would be useful for **mode switching** ("switching tasks", "switching between colours when drawing", "changing tabs in a browser"), **navigation** ("scrolling", "game exploration", "Google maps", "navigating a document"), and **selection between a reduced number of options** ("anything where you have a limited number actions to do", "two or three big buttons", "if there were large quadrants, it would be useful")

For real-world use, one participant said that he "(...) would not want the tracking to be always on. To toggle this mode, I would suggest holding down a key".

5   Study 3: Simultaneous Manipulation of Two Parameters

The previous experiment focused on how the feet can be used to control one parameter. However, the feet have a greater bandwidth than one parameter as their positions and orientations in space can have meaning for input. In this experiment, we aimed to understand how people can use their feet to control two parameters at the same time. Is it better to use one foot to control multiple parameters or distribute these parameters across the two feet? Further, does the visual representation of the control of parameters affect the interaction?

For this experiment, we recruited a group of 12 participants (8M/4F), aged between 19 and 42 years (mean 28) using posters on campus and adverts on social networks. Two of the participants were left handed and footed and nine were drivers. Foot sizes ranged from 22 cm to 32 cm (mean 26.7 cm). None of the participants had ever used a foot mouse or similar foot-operated pointer. The experimental setup was the same as for the previous experiment.

5.1   Procedure

Participants were first asked to sign a consent form and complete a personal information questionnaire. They were then given time to familiarise themselves with the interface. The goal of the study was to investigate how interaction technique and visualisation influence task completion time and error rate. To this end, participants were asked to manipulate two variables, within a certain threshold, while we varied the following two factors: *Interaction Technique* (3 levels): The two input values were manipulated by (1) XY position of 1 foot (1F); (2) X position of both feet (XX) and; (3) X position of one foot and Y position of the other (XY); and *Visualisation* (2 levels): Rectangle resizing and slider adjustment (described below).

In the first visualisation, the task was to fit the dimensions of an adjustable rectangle to those of a target rectangle (see Figure 3(A)). The target values were the width and height of the destination rectangle while the threshold was represented by the thickness of the rectangle's stroke. In the second visualisation, participants were asked to set two sliders along a scale to different target values marked by red tags (see Figure 3(B)). Here, the target values were the centres of the tags and the threshold was represented by their thickness. We chose these two visualisations because in the first the two degrees of freedom are integrated (as the corner of the rectangle) while in the second they are independent (as separate sliders). We hypothesised that these different visualisations might influence the performance depending on the number of feet used in the interaction. For each task, we measured the task completion time and error rate.
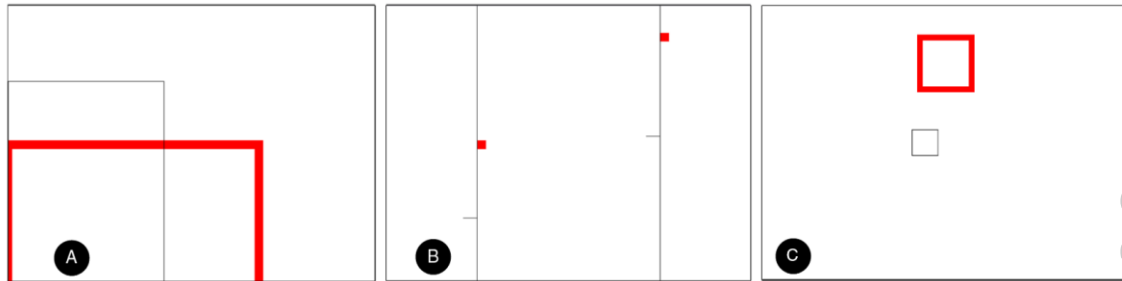
Figure 3 - Tasks in Study 2: (A) Resizing a rectangle and (B) Setting sliders and (C) task in Study 3

5.2 Results

We computed the mean task completion time and error rates for each condition (see Table 3). We considered an error when users clicked the mouse outside the target bounds. We compared the task completion times using a factorial repeated-measures ANOVA, testing the assumption of sphericity with Mauchly's test where appropriate. All effects were reported as significant at $p < .05$. There was a significant main effect of the technique, $F_{2,22} = 14.82$, and of the visual representation, $F_{1,11} = 50.46$, on the task completion time. There was also a significant interaction effect between the technique and the visual representation, $F_{2,22} = 34.10$, indicating that the interaction technique influence on participants' speed was different for the rectangle and slider representations of the task.

Table 3 - Time to select a target and error rate for each technique and visualization in Study 3

	Rectangle			Sliders		
	1F	**XX**	**XY**	**1F**	**XX**	**XY**
Time (ms)	3428	5068	4161	9098	4625	4902
Error (%)	5.5	3.7	8.0	8.2	8.1	8.0

Bonferroni post-hoc tests revealed that using one foot is significantly different than all other conditions in the slider representation ($p<.05$), but not in the rectangle one, as in this condition, it was not significantly different than using one foot horizontally and the other foot vertically ($p=0.38$). The two conditions in which participants used both feet were not significantly different in any combination of techniques and representations at $p<.05$.

We also compared accuracy using a factorial repeated-measures ANOVA. Mauchly's test indicated that the assumption of sphericity had not been violated neither by the effects of the technique (W = 0.86, $p = 0.47$) nor by the effects of the interaction between technique and visual representation (W = .91, $p = .72$). Our results showed no significant effect of the technique, of the visual representation or the interaction between them.

5.3 Discussion

Our results show that when manipulating multiple variables with the feet the visualisation strongly affects performance. The best performances amongst all conditions were interaction techniques 1F and XY in the rectangle representation, which were not significantly different at $p<.05$. In these two conditions, there was a direct spatial mapping between the technique and the task, since in technique 1F, the foot moved together with the corner of the rectangle and in technique XY, the feet moved together with its edges.

Users were confused when this spatial mapping was broken. The worst performing condition was using technique 1F for the slider task. Even though the underlying task was exactly the same, the change in visualisation caused the mean completion time to increase over twofold. This can be explained by how users would complete the task. In the slider task, participants would often set one slider at a time and in technique 1F, this meant moving the foot in one direction and then in the other. The problem is that users find it hard to move the foot in only one direction at a time. As we discovered in our previous study, when moving the foot horizontally, users tend to pivot their feet, rather than drag them, and this movement causes the cursor to move in both directions at the same time, resulting in users setting one slider, then setting the second one and having to go back and forth between them to make final adjustments. This was not a problem when controlling each value by a different foot. Regardless of whether the user tried to set both values at the same time or in sequence, moving one foot did not affect the other, so the visual representation was not an issue when using two feet.

An interesting effect we observed was in technique XY in the rectangle representation. Even though only one axis of the movement of each foot was being used to control the size of the rectangle, some participants would move both feet diagonally and symmetrically. One participant was even conscious of this, but kept on using this strategy: "I knew that each foot controlled only one dimension, but I found myself moving each one in both directions." This suggests that symmetrical movements might be more comfortable than independent ones when using two feet.

6   Study 4: Parallel Use of Feet and Hands

The previous experiments investigated interactions using the feet alone. In this experiment we wanted to investigate the overhead caused by using the feet in parallel with one hand. More specifically, we wanted to test whether there is an effect of resizing technique (scrolling with the mouse wheel, the position of one foot, or the distance between two feet) on the completion times and accuracy of the task, while the hand repositions the same square.

For this experiment, we recruited a group of 12 participants (10M/2F), aged between 19 and 32 years (mean 26), with posters on campus and adverts on social networks. Two participants were left-handed and -footed while 10 participants were drivers. Foot sizes ranged from 22 cm to 34 cm (mean 28 cm). None of the participants had ever used a foot mouse or similar foot-operated pointer. The experimental setup was exactly the same as the one for the previous experiments.

6.1 Procedure

Upon arrival, participants signed a consent form and completed a personal information questionnaire. They were then given some time to familiarise themselves with the interface. The task consisted of resizing and positioning a square to match a destination square at a different place on the screen. In all experimental conditions, the positioning was done with the mouse but the size of the square would be manipulated by one of three controls: the scroll wheel of the mouse, the horizontal coordinate of one foot or the horizontal distance between the two feet. We chose the scroll wheel as it is widely used for manipulating continuous variables. When the size and position of the two squares were matched, the user would click with the mouse and the button would reappear in the centre of the screen. We considered an error when users clicked the mouse outside the target bounds. Each participant repeated this task 40 times for each condition, with the target square in different positions and with different sizes. We measured the task completion time and the error rate. In the end of the study participants were asked to rank their preference of interaction techniques.

6.2 Results

The mean task time was similar across all conditions: 2.50s for the scroll wheel, 2.63 for the two feet condition, and 2.95 for the one foot condition. We first compared the task completion times in each condition with a one-way repeated measures ANOVA. Mauchly's test indicated that the assumption of sphericity had not been violated, W=0.55, p=0.05. Our results showed a significant effect of the technique used for resizing on the task completion time, $F_{2,22} = 5.08$, $p < .05$. Post-hoc tests showed that using one foot was significantly slower than the other two conditions ($p < .05$), but no significant difference was found between using two feet and scrolling ($p = .062$).

We also computed the error rates for each technique: 0.15 for the scroll wheel, 0.089 for the two feet and 0.081 for the one foot. We then compared the error rates for the conditions using a one-way repeated measures ANOVA. Mauchly's test indicated that the assumption of sphericity had not been violated, W=0.96, p=0.86. Results show a significant effect of the technique on the error rate, $F_{2,22} = 20.03$, $p < .05$. Bonferroni *post-hoc* tests showed that using the scroll wheel was significantly more accurate than the other two conditions ($p < .05$), but no significant difference was found between using one or two feet ($p = 1.00$).

The most preferred technique was the scroll wheel, chosen as the top technique by 85% of participants. Participants were divided between the feet techniques, with eight preferring two feet and four preferring one foot.

6.3 Discussion

We chose an increment value for each step of the scroll wheel so that users would not overshoot the thickness of the stroke of the target rectangle, but it also caused the

scroll wheel to be slower, so it might have fared better with adjustments in its sensitivity. In terms of task completion times, the feet performed similarly to the hands, showing little overhead for the task being performed, but with a significant decrease in accuracy. Taking into account that users are more familiar with the scroll wheel and none of our participants had any experience with foot-operated interfaces, from these results we speculate that with training, the feet could match (if not outperform) the scroll wheel as a means of providing continuous input to applications.

Our results show that using two feet was significantly faster than using one. We suggest two explanations for this. First, because what mattered was the relative distance between the feet, users could place their feet wherever they felt most comfortable within the tracked area. Because in the one foot condition, what mattered was the absolute position of the foot, depending on how the user was seated, this position might not have been ideal, causing a decrease in performance. Second, as both conditions used the same calibration, moving two feet simultaneous would cause a twofold change in the size of the rectangle, as compared with moving just one foot, increasing the overall speed of the interaction. Despite being faster, almost 40% of participants still preferred one foot, citing that moving two feet was more tiring than moving just one.

7   Guidelines and Design Considerations

Based on previous work, the quantitative and qualitative results from our experiments and our own experience while investigating the subject, we suggest a set of guidelines and considerations for designing desktop interactive systems that use feet movements as input.

Resolution: Our findings confirm the observations of Raisamo and Pakkanen that pointing with the feet should be limited to low fidelity tasks, in which accuracy is not crucial [5]. For example, when compared to using only the hands in experiment 3, the feet were significantly less accurate.

Visibility & Proprioception: In a desktop setting, the desk occludes the feet, which prevents the direct manipulation of interfaces, such as the floor-projected menus in Augsten et al. [2]. Moreover, foot gestures suffer from the same problems as other gestural interactions (see Norman [27] for a discussion of such problems), which are amplified by this lack of visibility of the limbs. Our second study showed that when designing such interactions, on-screen interfaces should provide a direct spatial representation of the movement of the feet. However, the lack of visibility of the feet is somewhat compensated by the user's proprioception: the inherent sense of the relative positioning of neighbouring parts of the body. Therefore, even though users are not able to see their feet they still know where they are in relation to their body.

Fatigue: Similarly to mid-air gestures, users report fatigue after extended periods of time using leg gestures. In all of our studies, participants reported that, in order to minimise fatigue, they preferred pivoting the foot around the heel to dragging the feet across the floor. Fatigue must also be taken into account when designing interactions where any foot is off the floor. In our experiment, when moving the feet across the floor, users preferred dragging the foot to hovering it over the floor.

Balance: Foot gestures performed whilst standing up only allow for one foot to be off the floor at the same time (except when jumping). While sitting down, the user is able to lift both feet from the floor at the same time, allowing for more complex gestures with both feet. To prevent fatigue, such complex gestures should be limited in time and potentially also space. In this work, even though we tracked the feet in three dimensions, we only took into account their two-dimensional position in relation to the floor. It remains an open question how adding a third dimension could affect the interaction.

Chair & Spatial Constraints: The kind of chair where the user is seated may influence the movement of the feet. For example, the rotation of a swivel chair might help with moving the foot horizontally. Further, when both feet are off the floor, swivel chairs tend to rotate as the user moves which may hamper interaction. The form factor of desks, chairs and clutter under the desk also affect the area in which the user can perform gestures. This also offers opportunities for interaction, as physical aspects of the space can help guide the movement of the feet or serve as reference points. Another aspect that needs to be taken into account are the properties of the floor, which might influence the tracking (shiny floors will reflect the infra-red light emitted by the Kinect, creating additional noise) and interaction (floors covered in carpet or anti-slip coating may slow down feet movements, while smooth flooring may speed them up).

Rootedness: Mid-air gestures often suffer from the problem of gesture delimiters, similar to the classic Midas touch problem, as it is hard to tell specific actions and gestures from natural human movement [28]. This is less of a problem for feet gestures in a seated stance, because when sat down, most leg and foot movement consists of postural shifts, reducing the number of movements that might be recognised as false positives in gesture recognition systems. We addressed this problem in our studies by defining an area on the floor where the feet would provide input for the system, but in applications where it would be desirable to track the feet at all times, it is necessary to pay special attention to designing gesture delimiters that are not part of users' normal lower limb behaviour.

Footedness: In the same way that people favour one hand they also favour one foot and, even though they are often correlated, there are exceptions to this rule, with approximately 5% of the population presenting crossed hand-foot preference [29]. Our findings indicate there is no significant difference between the dominant foot and the non-dominant one. These results, however, reflect the performance of users with no experience with feet-based interfaces. It is not clear if this similarity in performance still holds for experienced users. Further, it is necessary to consider which foot will be used in the interaction, as crossing one foot over the other to reach targets on the opposite side might be too uncomfortable.

Hotspot: Touch-based interfaces, despite suffering from the phenomenon of 'fat-fingers', can still provide a high resolution of input due to the small relative size of the contact area between the finger and the touch-sensitive area. Feet, however, provide a large area of contact with the floor. The designer can then opt for reducing the foot to a point or using the whole contact area as input. The former has the advantage of providing high resolution input, but users' perceptions of the specific point on the foot that should correspond to the cursor is not clear, as demonstrated by Augsten et al. [2]. Using the whole of the foot sole makes it easier to hit targets (as shown by Drury's modification of Fitts's Law [19]), but increases the chance of hitting wrong targets.

Hence, if using this approach, the designer needs to leave enough space between targets as to prevent accidental activation.

8   Limitations

In this work, we described four experiments that attempt to characterise some fundamental aspects of the use of foot movements for interacting with desktop computers. These experiments, however, have some limitations. We collected data from a relatively small number of participants, so more precise estimates of the real value of the times and error rates presented here can certainly be achieved in experiments with larger pools of participants. Also, our participant pool was not gender-balanced in every study and did not cover a wide age range. We present results using only one tracking system that has several limitations of its own. For example, our prototype was implemented in Matlab, achieving a frame rate of 25 fps, but the tracking speed could be improved by porting the system to a faster language, such as C++. While our results are in line with the ones in related work, further work is necessary to assess whether they translate to other foot interfaces.

9   Conclusion

In this work we took a bottom-up approach to characterising the use of foot gestures while seated. We implemented a foot tracking system that uses a Kinect mounted under a desk to track the users' feet and used it to investigate some fundamental characteristics of this kind of interaction in three experiments.

First, we presented ISO 9241-9 performance models for 1D and 2D foot pointing in a sitting position. Our results suggest little difference in performance between the dominant and non-dominant foot and that horizontal foot movements are easier to perform than vertical ones. We identified five strategies that participants used to reach targets and found that the preferred one was rotating the foot around the heel. We also found that the biggest challenges for foot-based interaction in a desktop setting are difficulties in reaching small targets, hand-feet coordination, fatigue, friction with the floor, and overshooting targets. These findings are important because they help us complete our understanding of the potential of foot-operated interfaces and provide guidance for future research in this emerging domain.

Second, we studied the performance of each foot in controlling a single parameter in a unidimensional task. Our results showed no significant difference between the dominant and non-dominant foot, but it showed that horizontal movement on the floor is significantly faster than vertical. Also, users showed a preference for pivoting their feet rather than dragging them. Third, we looked at controlling two variables at once, comparing the use of one foot against the use of two (each foot using the same movement axis or different ones). Our results showed that the visual representation of the variables do matter, with the performance for techniques that have a direct spatial mapping to the representation outperforming the others. It also showed that when the variables being manipulated are shown separately (such as in independent sliders), it is

preferable to use two feet rather than one. Fourth, we analysed the use of the feet in parallel to the hands, showing that the feet perform similarly to the scroll wheel in terms of time, but worse in terms of accuracy, suggesting that with training and more accurate tracking systems, the feet could be used to support hand based interaction in a desktop setting.

Future work will focus on using these insights to design and implement techniques that can possibly enhance the interaction by supporting the hands in everyday computing tasks. While we provide some guidelines for design, it is still an open question as to which tasks can effectively be supported by the feet and the size of the cognitive overhead of adding such an interactive modality.

References

1. English, W.K., Engelbart, D.C., Berman, M.L.: Display-selection techniques for text manipulation. Transactions on Human Factors in Electronics. 5–15 (1967).
2. Augsten, T., Kaefer, K., Meusel, R., Fetzer, C., Kanitz, D., Stoff, T., Becker, T., Holz, C., Baudisch, P.: Multitoe: high-precision interaction with back-projected floors based on high-resolution multi-touch input. UIST, 209–218. ACM (2010).
3. Scott, J., Dearman, D., Yatani, K., Truong, K.N.: Sensing foot gestures from the pocket. Proceedings of the 23nd annual ACM symposium on User interface software and technology. pp. 199–208. ACM, New York, NY, USA (2010).
4. Simeone, A., Velloso, E., Alexander, J., Gellersen, H.: Feet Movement in Desktop 3D Interaction. Proc. of the 2014 IEEE Symposium on 3D User Interfaces. (2014).
5. Pakkanen, T., Raisamo, R.: Appropriateness of foot interaction for non-accurate spatial tasks. CHI'04 EA. pp. 1123–1126. ACM (2004).
6. Dearman, D., Karlson, A., Meyers, B., Bederson, B.: Multi-modal text entry and selection on a mobile device. Proceedings of Graphics Interface 2010. pp. 19–26. Canadian Information Processing Society (2010).
7. Garcia, F.P., Vu, K.-P.L.: Effects of Practice with Foot-and Hand-Operated Secondary Input Devices on Performance of a Word-Processing Task. Human Interface and the Management of Information. Designing Information Environments. pp. 505–514. Springer (2009).
8. Pearson, G., Weiser, M.: Of moles and men: the design of foot controls for workstations. Proc. of CHI, pp. 333–339. ACM, New York, NY, USA (1986).
9. Sellen, A.J., Kurtenbach, G.P., Buxton, W.A.: The prevention of mode errors through sensory feedback. Human-Computer Interaction. 7, 141–164 (1992).
10. LaViola Jr, J.J., Feliz, D.A., Keefe, D.F., Zeleznik, R.C.: Hands-free multi-scale navigation in virtual environments. Proceedings of the 2001 symposium on Interactive 3D graphics. pp. 9–15. ACM (2001).
11. Carrozza, M.C., Persichetti, A., Laschi, C., Vecchi, F., Lazzarini, R., Vacalebri, P., Dario, P.: A Wearable Biomechatronic Interface for Controlling Robots with Voluntary Foot Movements. Transactions on Mechatronics, 12, 1–11 (2007).
12. Han, T., Alexander, J., Karnik, A., Irani, P., Subramanian, S.: Kick: investigating the use of kick gestures for mobile interactions. Proceedings of the 13th

International Conference on Human Computer Interaction with Mobile Devices and Services. pp. 29–32. ACM Press, New York, New York, USA (2011).

13. Paelke, V., Reimann, C., Stichling, D.: Foot-based mobile interaction with games. Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology. pp. 321–324. ACM (2004).

14. Hockman, J.A., Wanderley, M.M., Fujinaga, I.: Real-time phase vocoder manipulation by runner's pace. Proc. Int. Conf. on New Interfaces for Musical Expression (NIME) (2009).

15. Lopes, P.A.S.A., Fernandes, G., Jorge, J.: Trainable DTW-based classifier for recognizing feet-gestures. Proc. of RecPad, (2010).

16. Richter, S., Holz, C., Baudisch, P.: Bootstrapper: recognizing tabletop users by their shoes. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1249–1252. ACM (2012).

17. Paradiso, J.A., Hsiao, K., Benbasat, A.Y., Teegarden, Z.: Design and implementation of expressive footwear. IBM Systems, 39, 511–529 (2000).

18. Kume, Y., Shirai, A., Sato, M.: Foot Interface : Fantastic Phantom Slipper. ACM SIGGRAPH 98 Conference abstracts and applications. p. 114 (1998).

19. Drury, C.G.: Application of Fitts' Law to foot-pedal design. Human Factors: The Journal of the Human Factors and Ergonomics Society. 17, 368–373 (1975).

20. Hoffmann, E.R.: A comparison of hand and foot movement times. Ergonomics. 34, 397–406 (1991).

21. Springer, J., Siebes, C.: Position controlled input device for handicapped: Experimental studies with a footmouse. International Journal of Industrial Ergonomics. 17, 135–152 (1996).

22. Hinckley, K., Jacob, R., Ware, C.: Input/output devices and interaction techniques. In: Tucker, A.B. (ed.) CRC Computer Science and Engineering Handbook. pp. 20.1–20.32. CRC Press LLC: Boca Raton, FL. to appear (2004).

23. Wobbrock, J.O., Shinohara, K., Jansen, A.: The effects of task dimensionality, endpoint deviation, throughput calculation, and experiment design on pointing measures and models. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1639–1648. ACM (2011).

24. Soukoreff, R.W., MacKenzie, I.S.: Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. International journal of human-computer studies. 61, 751–789 (2004).

25. Natapov, D., Castellucci, S.J., MacKenzie, I.S.: ISO 9241-9 evaluation of video game controllers. Proceedings of Graphics Interface 2009. pp. 223–230. Canadian Information Processing Society (2009).

26. Zhong, K., Tian, F., Wang, H.: Foot Menu: using heel rotation information for menu selection. Wearable Computers (ISWC), 2011 15th Annual International Symposium on. pp. 115–116. IEEE (2011).

27. Norman, D.A., Nielsen, J.: Gestural interfaces: a step backward in usability. interactions. 17, 46–49 (2010).

28. Benko, H.: Beyond flat surface computing: challenges of depth-aware and curved interfaces. Proc. of Multimedia. pp. 935–944. ACM (2009).

29. Dargent-Paré, C., De Agostini, M., Mesbah, M., Dellatolas, G.: Foot and eye preferences in adults: Relationship with handedness, sex and age. Cortex. 28, 343–351 (1992).