# Assessing the efficacy of a touch screen overlay as a selection device for typical GUI targets

**Abstract**

In this paper we investigate the efficacy of a touch screen overlay compared to a mouse, when selecting typical graphical user interface (GUI) items in a desktop information system. A series of tests were completed involving multi-directional point and select tasks, and the results for both devices compared. The results showed that the touch screen overlay was not suitable for selecting GUI targets smaller than 4 mm. The touch screen overlay was slower and had a higher error rate than the mouse, but there was no significant difference in throughput. Testers rated the mouse easier to use and to make accurate selections, while the touch screen overlay resulted in greater arm, wrist and finger fatigue. These results suggest that a touch screen overlay is not a practical selection device for desktop interfaces with small GUI targets.

*Key words:* Touch screen overlay, Mouse, Selection device, Fitts' Law, Performance evaluation, GUI item

## 1 Introduction

Most modern information systems that run on desktop personal computers are designed to be used with a keyboard and mouse. While the combination of keyboard and mouse is the accepted method of interaction with such systems it does not necessarily suit all information systems. Information systems with limited data entry may be more usable through the use of a keyboard and touch screen. Touch screens require less physical space and thus the workstation environment in an office setting could be improved, allocating more space to the user and less to the computer.

The purpose of this paper is to investigate how effective a touch screen overlay is compared to a mouse, when selecting typical graphical user interface (GUI) items. The target types tested were buttons, check boxes, combo boxes and text boxes, which are typical of those found in an interface for an information system. Each target type was tested at three different sizes (see Section 2).

A typical touch screen device comprises a monitor enhanced with hardware for detecting touches on the screen surface. An alternative approach is to attach a discrete touch-sensitive surface to an existing conventional monitor. Sears and Shneiderman (1991) have previously assessed the efficacy of specialised touch screen hardware, but there appears to have been little research into the efficacy of touch screen overlays. We therefore chose to compare the performance of a touch screen overlay with that of a mouse.

The testing occurred in the context of a research project undertaken by the Department of Human Nutrition at the University of Otago. This project aims to improve complementary feeding diets for toddlers in New Zealand, by designing a computer program to help formulate population-specific food-based dietary guidelines for this high risk group. The program, which is based on a previously published linear programming approach (Ferguson et al., 2004), is a decision-making tool, designed specifically for nutrition programme planners to assist them in selecting appropriate and improved home-based complementary foods.

The remainder of the paper discusses our experiment and the results obtained. Section 2 briefly describes the types of targets used in the experiment, while Section 3 describes the measures used to evaluate the selection devices. The experimental design is described in Section 4. Section 5 describes how the data were analysed, and Section 6 presents the results of the experiment. Our conclusions are presented in Section 7.

## 2 GUI targets

Since the 1980's much work has gone into developing human computer interface guidelines. Today's interfaces are made up of a combination of different targets that include text boxes, check boxes, combo boxes, list boxes, buttons, labels, tool bars, etc. Sears and Shneiderman (1991) showed that touch screens can be successfully used as a selection device and can have advantages over a mouse, even for small targets. These results were, however, based on selecting arbitrary shapes and not the typical targets found in modern GUIs.

To accurately test the performance of each selection device within the experiment, three different sizes of GUI target were used, corresponding to small, medium and large GUI items. As our experiment took place in a Windows environment, we would have preferred to base these sizes on Microsoft's user interface guidelines. However, Microsoft's guidelines specify only a single standard size for most GUI items (Microsoft Corporation, 1999, pp. 448–450). Apple Computer's (2004) human interface guidelines specify three standard sizes (mini, small and large), but these proved to be rather small in our Windows-based testing environment. We therefore adjusted Apple's three sizes such that the "small" size was consistent with Microsoft's guidelines. The resultant target sizes are listed in Table 1. A screen resolution of 81 DPI was assumed.

Table 1

Target sizes (width × height) used in the experiment.

| **Target type** | **Large** | **Medium** | **Small** |
|---|---|---|---|
| Text box | 63 mm × 11 mm | 55 mm × 8 mm | 47 mm × 6 mm |
| Combo box | 63 mm × 11 mm | 55 mm × 8 mm | 47 mm × 6 mm |
| Button | 28 mm × 13 mm | 24 mm × 9 mm | 17 mm × 6 mm |
| Check box[a] | 9 mm × 9 mm | 6 mm × 6 mm | 4 mm × 4 mm |

[a]This refers to the size of the check box itself, not the associated text label.

## 3 Evaluation methods

Each selection device was assessed using a combination of performance and comfort measures. The performance measures were primarily taken from the ISO 9241-9 standard (ISO, 2000), while the comfort measures were derived from a questionnaire administered to test participants.

### *3.1 Performance*

ISO 9241 specifies standards for the ergonomic design of office computing environments. Part 9 of this standard describes different tests that can be used to evaluate one or more pointing devices (ISO, 2000). The standard describes a serial point and select task and recognises a dependent measure used with this test, known as *throughput*. The serial test comprises moving the cursor back and forth between two targets using the pointing device and selecting each target by pressing and releasing a button on the pointing device. One disadvantage of this approach is that only two targets are used in the test and therefore interactions between more than two targets—which often occur in a typical information system interface—are not studied.

MacKenzie and Jusoh (2001) note that throughput is a very important measure, as it reflects the efficiency of the user completing the task and is a measure of both speed and accuracy. Throughput is calculated by the formula $throughput = ID_e/MT$, where $MT$ is the movement time in seconds (defined as the time taken to successfully select a target) and $ID_e$ is Fitts' (1954) *index of difficulty* measured in bits. Throughput is thus measured in bits per second (bps).

The index of difficulty is calculated by the formula $ID_e = \log_2((D/W_e) + 1)$, where $D$ is the distance to the target and $W_e$ is the *effective width* of the target. The effective width reflects spatial variability in a sequence of trials, and thus differs from the actual width of the target. The effective width of a target is calculated by the formula $W_e = 4.133 \times SD_x$, where $SD_x$ is the standard deviation in the selection coordinates measured along the path to target $x$.

ISO 9241-9 does not provide any guidance on the range of index of difficulty values to use in testing. Douglas et al. (1999) recommend using a range from 2 to 6 bits. They also recommend calculating the *error rate* as a separate dependent measure of accuracy. The error rate is defined as the ratio of incorrect to correct selections made on a target, so an error rate of 100% implies that there were as many errors made as correct selections. Error rate is not included in ISO 9241-9, but has been used in several other studies (Sears and Shneiderman, 1991; Sears et al., 1993; Harada et al., 1996; Bender, 1999; Douglas et al., 1999; MacKenzie and Jusoh, 2001; Po et al., 2004). Computing both throughput and error rate gives a more detailed performance analysis for the selection device in question.

### *3.2 Comfort*

ISO 9241-9 argues that to fully evaluate a selection device requires assessment of user effort and comfort in addition to performance measurements. Comfort is subjective and can be assessed by means of questionnaires, while effort can be evaluated objectively by measuring the biomechanical load on users as they use a device. Unfortunately, such measurements require reasonably sophisticated equipment (Douglas et al., 1999) that was not available to us. We therefore omitted effort measurements from our experiment.

A questionnaire was used to assess comfort and user satisfaction for each selection device in our experiment. The selection device assessment questionnaire comprised sixteen questions, eight of which were taken from the ISO "Independent Questionnaire for Assessment of Comfort" (Douglas et al., 1999). The remaining eight questions related specifically to the target types and target sizes that were tested. In particular, the questionnaire aimed to assess the participants' comfort in using the selection device, the difficulty in accurately selecting each of the target types and the preferred size of each target type using the selection device.

The responses to twelve of the questions were based on a five point ordinal scale. The remaining four questions referred to the participant's preferred size for each target type and were based on a three point response corresponding to the target sizes tested—small, medium and large (see Table 1). There was also a space for participants to provide additional general feedback about the testing process.

### *3.3 Other considerations*

Douglas et al. (1999) also note that ISO 9241-9 does not take into account any possible effects of learning, which can affect movement time and accuracy. For example, MacKenzie et al. (1991) found that the movement times from the first of five testing sessions were significantly higher than in later sessions. This can be explained as a result of learning and shows that input device studies should take learning into account and test for it; indeed, Douglas et al. (1999) recommend applying a repeated measures paradigm and testing for learning effects.

One interesting aspect of using typical GUI items as targets is the variation in selection behaviour for different target types, compared to earlier studies that used simple rectangular targets. The button, check box and text box target types can be said to exhibit a "one-step" selection behaviour, because they require only single action (i.e., the user clicks on them) in order to be selected. A combo box is different, however, because it exhibits a "two-step" selection behaviour: first the combo box must be selected in order to show the list of items, and then an item must be selected from the displayed list, as illustrated in Figure 1. To complicate matters further, users may execute this two-step behaviour using either a "one-click" or a "two-click" approach. In the former approach, the user clicks on the combo box, drags down to the desired list item, then releases. In the latter approach, the user clicks once on the combo box, then clicks again on the desired list item. If the drop-down list were longer than what could be displayed on screen, this could even lead to a "multiple-click" approach, where the user clicks multiple times on the downward scroll arrow in the list. We have, however, not considered this possibility in our experiment.

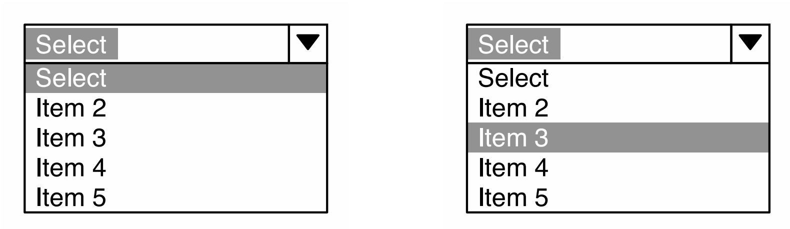


Fig. 1. The "two-step" action required to select a combo box.

## 4 Method

An experiment was carried out to test the effect of size for different GUI target types with different selection devices. The experiment involved participants completing a series of simple point and select tasks. Small, medium and large sizes (see Table 1) were tested for combo boxes, text boxes, check boxes and buttons, using either a touch screen overlay or a mouse. The test was multi-directional, meaning the targets appeared in multiple directions from the initial starting point. A variety of different sizes, angles and distances were used for each target position.

The test itself comprised a screen containing a button in the centre and a target for the participant to select as illustrated in Figure 2. When a participant clicked on the centre "Go" button, a trial was started and a target appeared on the screen. The trial ended when the participant successfully clicked the target, which then disappeared. The time taken between clicking the "Go" button and successfully clicking on the target was recorded as well as the number of errors made during the trial. The final coordinates of the successful click on the target were recorded in order to calculate the effective width of the target.


Fig. 2. Screenshot of the test environment with a target in the top left of the screen and the "Go" button in the centre.

### *4.1 Participants*

A participant sample size of twenty-four was used for the experiment. Each participant was allocated to one of two groups with each group using one selection device in testing. The allocation of groups was based on the results of a questionnaire completed by each participant prior to testing. The purpose of the pre-test questionnaire was to establish the level of computer, mouse and touch screen experience of each participant. Participants were then allocated to a selection device group based on which device they had the least experience with.

Due to the testing being done within the nutrition program environment mentioned in Section 1, the participants were all nutritionists (i.e., typical users of the program). There were twenty-one female and three male participants, all with a university level of education. All participants were unpaid volunteers.

### *4.2 Apparatus*

The test environment was implemented in Visual Basic.NET using Microsoft Studio 2003, and is illustrated in Figure 2. During each test, data corresponding to the relevant measures (movement time, number of errors and selection coordinates) were captured by the software and automatically written to a Microsoft Excel worksheet.

The touch screen used in testing was a 17" Magic Touch USB overlay Model KTMT-1700-USB-M (KEYTEC, 2005). This device uses a take-off touch strategy, that is, a selection is not confirmed until the user's finger is removed from the screen. An important property of touch screen overlays is that they are placed over a conventional monitor and the touch surface is thus not coincident with the display surface. This can cause a slight discrepancy or parallax effect between where the user touches the overlay and where the cursor is positioned on the screen.

The touch screen overlay was fitted to a Dell 15" Flat Panel Model E151FPb monitor. A flat panel monitor was chosen because it was noticed during pre-testing that typical CRT monitors with curved screens produced a variable gap between the overlay and the display surface, thus potentially leading to a greater parallax effect than with a flat display surface.

The mouse used in testing was a standard Dell PS/2 Optical Mouse Model M071KC. Both devices were connected to a Dell Inspiron 7500 laptop computer that ran the testing software.

### *4.3 Design*

A mixed design experiment was used with the selection device as a between-subjects factor. The independent (between-subject) variables were:

- Target type (text box, combo box, button and check box)
- Target size (large, medium and small—see Table 1)
- Target distance (40 mm, 80 mm and 160 mm—see below)
- Target angle (45°, 135°, 225° and 315°—see below)
- Trial (1 to 144)
- Block (1 to 6)

The dependent variables within the experiment were throughput, movement time and error rate.

The entire test was divided into six blocks. Each block contained every possible combination of target type (four combinations), size (three combinations), angle from initial starting point (four combinations) and distance from initial starting point (three combinations). Consequently there were 144 trials in each block and the entire experiment per participant comprised a total of 864 trials (six blocks of 144 trials each). Combinations of target type, distance and angle were presented to the participant in random sequence with no repetition. Target size was deliberately set to large for the first forty-eight trials in each block, followed by medium for the next forty-eight trials, and finally small for the remaining trials, in order to compensate for learning effects.

The combination of distance and angle from the initial starting point yielded twelve possible target positions for each trial, as illustrated in Figure 3. Three distances were used that represented target positions ranging from close to the initial starting point to very far away from the initial starting point. Four angles were chosen so that targets could be tested in ninety degree blocks and to provide a good range of screen positions for the target. The first angle was set to 45° with 90° increments thereafter, in order to mimic real life user interface target selection, where targets are situated in different areas of the screen and therefore selections are made in multiple directions that are neither simply horizontal nor vertical.

The index of difficulty ($ID_e$) was ascertained for each possible task using the combination of distance and non-adjusted target width. This showed that the test had a range of $ID_e$ values from 0.7 bits (63 mm width and 160 mm

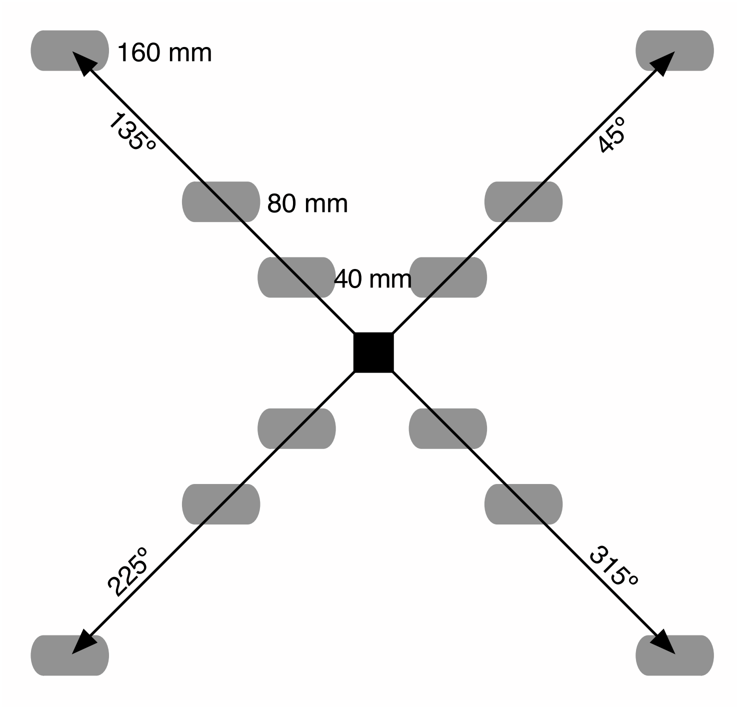


Fig. 3. Positions of targets tested. The black box represents the initial starting point and the rounded rectangles represent the target positions.

distance) to 5.4 bits (4 mm width and 40 mm distance). It is important to note that the combo box distance values were adjusted in these calculations to reflect the two-step selection behaviour of this target type (as discussed in Section 3.3). That is, we need to consider not just the distance from the initial starting point to the target, but also the extra distance from the main combo box to the selected list item. In our experiment, participants were told to always select the third list item in combo boxes (as illustrated in Figure 1), so the adjusted distance for a combo box was equal to the normal distance from the initial starting point to the target, *plus* the additional distance to the third list item. Additional data about the selections made on combo boxes were recorded in order to account for the selection approach of the participant, whether it be a “one-click” or “two-click” approach.

### *4.4 Procedure*

The participant was given an introduction to the test by the research observer, which included a brief summary of the aims of the study and what the test involved. The participant was also given an instruction sheet that they had access to throughout the duration of the test. After reading the instruction sheet the participant had the opportunity to ask questions or raise any issues.

Participants were instructed to complete each block of trials as quickly as possible without losing accuracy. Participants were given the opportunity to rest for as long as they wished between blocks. It was made clear to participants that a task was only complete once the target was successfully selected.

Before the test began, participants were instructed to complete a practice session involving fifteen random trials of the same point and select tasks used in the test. This brought all participants up to a minimal level of experience with their selection device. This also meant that each participant knew how to correctly select each target type including the combo box.

At the conclusion of the test the participant was asked to fill out a questionnaire regarding comfort and user satisfaction with the selection device used.

## 5 Analysis

The data collected from the software included movement time, error rate and throughput and was used to evaluate selection device performance. A mixed design repeated measures analysis of variance model (MANOVA) was used for movement time and throughput to examine within subject differences in target type and size, as well as between subject differences in selection device. A Greenhouse and Geisser correction of the F-ratio was used whenever the Mauchly's test results showed that assumptions of sphericity were violated.

Post hoc tests, for multiple comparisons, were made using the Bonferroni method. Due to the skew observed in the error rate data (see Section 6.1), inter-device differences in error rates were assessed using the Mann-Whitney U Test.

The comfort questionnaire was based on a five point ordinal scale, where five generally indicated a poor rating. Because of the small data size, a Mann-Whitney (non-parametric) test was used.

All statistical analyses were performed using SPSS version 11.0. A $p$-value of $< 0.05$ was considered statistically significant.

## 6 Results and discussion

### *6.1 Adjusting for learning*

Douglas et al. (1999) have shown that the effects of learning can affect movement time and accuracy. They therefore recommend that input device studies should apply a repeated measures paradigm and test for learning effects.

From analysing the results of movement time and throughput over each test block, it was clear for the combo box and check box target types that learning occurred from the first to the second block with the touch screen overlay, as seen in Figure 4. Due to prior participant experience, no learning was observed with the mouse. No learning occurred with the text box or button, most likely due to their relatively large size and simple selection behaviour.

Statistical analysis using a simple repeated measure ANOVA was carried out on movement time for both the check box and combo box. For movement time of the combo box, the effect of block × device was significant ($F(1.549, 1335.219) = 4.373, p < 0.05$). Helmhert contrasts showed that the differences between blocks became non-significant after block 1 ($p > 0.05$), which implies that learning occurred during block 1.

For movement time of the check box, the effect of block × device was significant ($F(1.608, 1385.960) = 4.763, p < 0.05$). Helmhert contrasts again showed that the differences between blocks became non-significant after block 1 ($p > 0.05$), implying that learning occurred during block 1.

To account for learning with the combo and check box, results from block 6 only were used to calculate the movement time and throughput measures, as

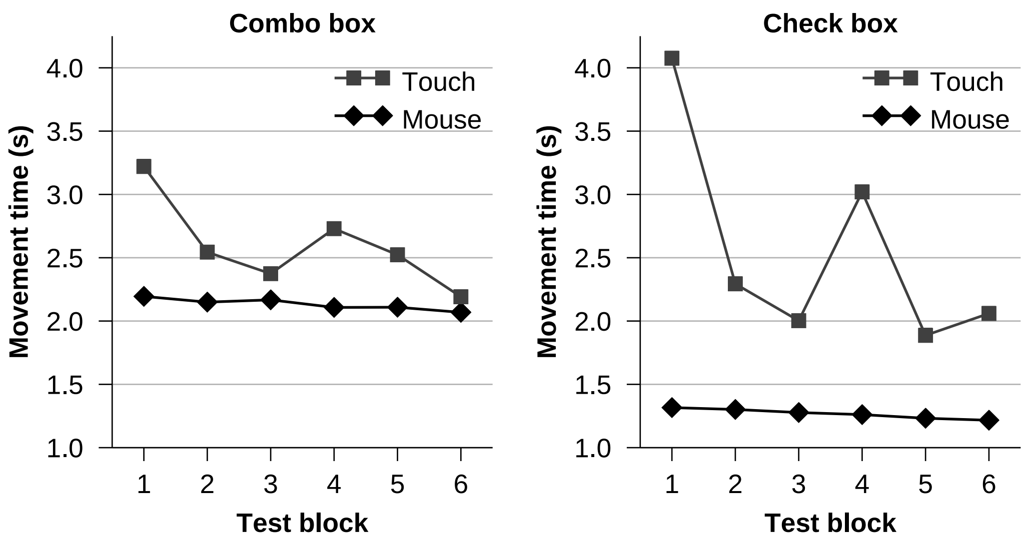


Fig. 4. Learning is displayed for movement time by selection device and test block for the combo box and check box.

the results from block 6 alone gave a good measure of performance. However, the error rate results were highly skewed for both target types, with two participants accounting for almost 90% of the errors. Error rates for the check box and combo box were therefore calculated using results from all six test blocks.

### *6.2 Movement time*

The results showed that the mouse had an overall movement time of 1.32 s across all target types compared to 1.57 s for the touch screen overlay. We can thus conclude that the touch screen overlay was on average 1.2 times slower than the mouse. This is interesting because Sears and Shneiderman (1991) found that the movement times for mouse and touch screen monitor were similar for rectangular targets larger than 2 mm. Therefore the nature of the two types of touch screen (overlay versus monitor) may affect the movement time associated with the type of touch screen. It is also likely that due to the loss of accuracy found with the overlay during testing, the touch screen monitor will have faster movement times compared to the touch screen overlay.

The movement times for each target type showed that the text box had the fastest movement time, followed by the button, the check box and the combo box. These results are illustrated in Figure 5 and are consistent with Fitts' Law (Fitts, 1954), in that the largest target (the text box) had the fastest movement time. It may appear that the combo box results violate Fitts' Law, as the combo box is the same size as the text box, yet is over twice as slow. In this case however, the slow movement times are not a function of the target size, but rather a result of the more complex two-step behaviour required to successfully select a combo box (which is not considered by Fitts' Law). The extra movement of selecting an item from the drop-down list clearly dramatically increases the movement time for the combo box. As the additional distance from the main combo box area to the list item is relatively short, the significant increase in movement time is therefore most likely due to users making more errors.

The touch screen overlay was found to have similar movement time to the mouse for the medium and large targets, but for the small targets, the touch screen overlay was on average 1.5 times slower than the mouse. The only time that the touch screen overlay was found to be faster than the mouse was with the largest target type (the large text box), but the difference was less than one percent.

The movement time for the small check box on the touch screen overlay was about 2.7 times slower than that for the mouse. The small check box was the smallest item tested, with dimensions of 4 mm × 4 mm. We can conclude

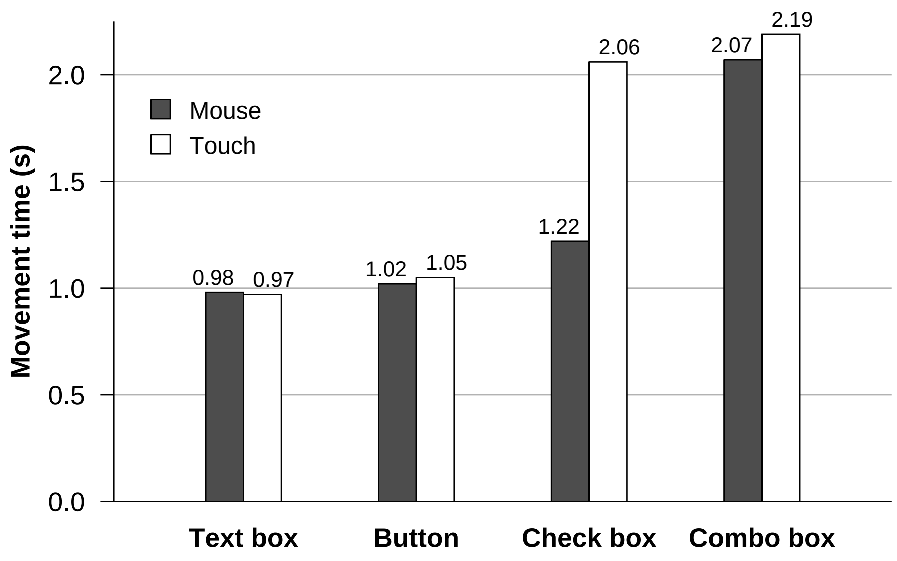


Fig. 5. Movement time by target type and device, averaged across all target sizes.

that the touch screen overlay was not efficient for selecting targets as small as 4 mm. Compare this with Sears and Shneiderman (1991), who showed that a touch screen monitor has similar movement time to a mouse for targets as small as 2 mm. While a touch screen monitor can be used with targets as small as 2 mm, a touch screen overlay should only be used for targets with a size of greater than 4 mm. The results from the error rate analysis also support this conclusion (see Section 6.4).

### *6.3 Throughput*

The average throughput for the mouse was 1.242 bps, slightly higher than the 1.214 bps average throughput for the touch screen overlay. The selection device by itself was shown not to have a significant effect on throughput ($F(1, 22) = 0.02, p > 0.05$). Throughput did not vary for size but throughput did vary depending on target type ($F(2.07, 45.55) = 4.77, p < 0.001$), as shown in Figure 6. Check boxes had the highest throughput rate of 1.967 bps (sd = 0.720). This is interesting as the check box had the second worst movement time and the worst error rate. Upon further investigation it was seen that the movement time for the check box was in fact in the middle range of all targets, and due to its small size it had a high index of difficulty. These two factors are the most likely reason for the check box having such a high throughput rate.

The combo box had the worst average throughput of 0.501 bps (sd = 0.213). The index of difficulty was not very high for the combo box, so its low throughput rate could be attributed to its high movement time.

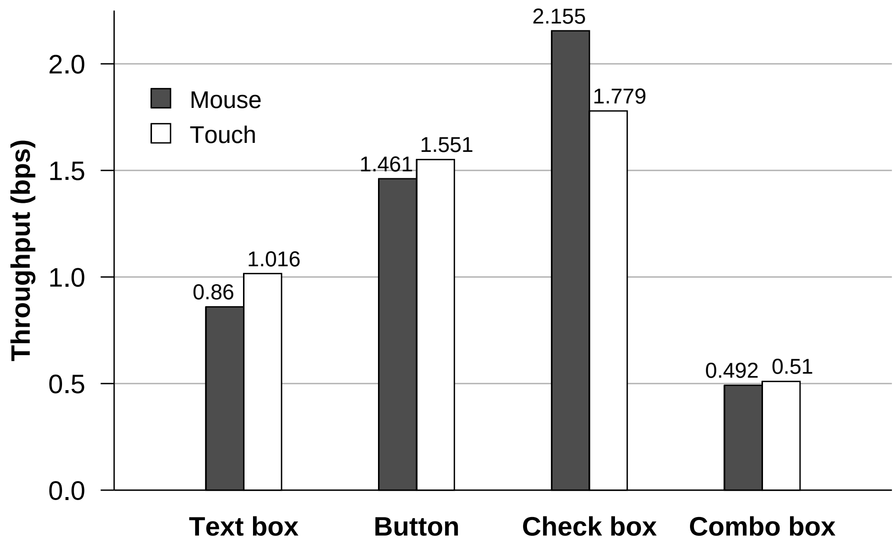


Fig. 6. Throughput by target type and device, averaged across all target sizes.

The average throughput rate of 1.242 bps for the mouse is much lower than that found by previous research. A study by Douglas and Mithal (1994) showed that a mouse had a throughput rate of 4.15 bps. MacKenzie et al. (1991) compared three devices (mouse, tablet and trackball) using four target sizes (8, 16, 32 and 64 pixels) over two different types of tasks (pointing and dragging), and found that the throughput for the mouse was 4.5 bps. This may indicate that the level of selection difficulty in our experiment is higher than in previous research. This could be due to the selection of GUI targets instead of arbitrary rectangular targets.

### *6.4 Error rate*

The average error rate for the mouse was only 2.7% which is consistent with previous studies. The touch screen overlay, on the other hand, had an average error rate of 60.8%. Sears and Shneiderman (1991) found that a touch screen monitor had an average error rate of 49% but this was across much smaller targets. This suggests that there is a loss of accuracy when using a touch screen overlay as opposed to a touch screen monitor.

The check box had a significantly high average error rate of 78.5% across all sizes and both devices; in particular, the small check box on the touch screen overlay had an error rate of 393%. The touch screen overlay incurred the majority of the errors. For all sizes of the check box the mouse produced an average error rate of 6.7%, while the touch screen overlay produced an average error rate of 178.5%. The distinguishing factor of the check box compared to the other target types is its small size, so we can conclude from this that a touch screen overlay is more inaccurate for selecting small targets (4 mm or less).

The button and text box had much lower error rates than the check box and combo box (as seen in Figure 7). As the button and text box also exhibited low movement times, we can conclude that these two target types have very good overall performance.

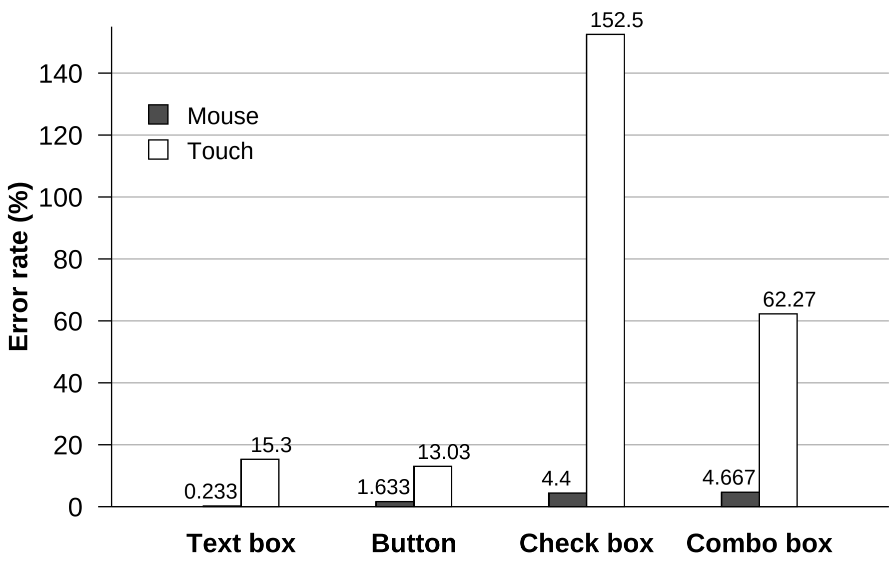


Fig. 7. Error rate by target type and device, averaged across all target sizes.

Note that the error rate calculation for combo boxes assumes a "two-click" selection approach, as only two of the twenty-four participants used the "one-click" approach. Both of these participants used the mouse.

### *6.5 Comfort*

In terms of accurate pointing the mouse (2.083) was rated easier than the touch screen overlay (3.000). These differences were statistically significant ($p < 0.01$). The responses regarding the question on neck, wrist and arm fatigue showed that the touch screen overlay had a high rating (4.083), whereas the mouse was rated in the midpoint range (3.167). These differences were statistically significant ($p < 0.05$). The final question rated the overall difficulty in using the selection device. The mouse (4.250) was rated easier to use than the touch screen overlay (3.333). These differences were statistically significant ($p < 0.05$).

Participants using the touch screen overlay rated both the text box and button as easy to accurately select, with the large and medium sizes being most preferred. This feedback is consistent with the data collected in that the text box and button have low movement times and low error rates (i.e., they are easy to accurately select). The combo box was rated in the middle of the range, and the check box was rated as very hard to select. Three quarters of the touch screen overlay participants preferred to select large combo boxes and check boxes, which is consistent with the high error rates and movement times associated with these two target types on the touch screen overlay.

Participants using the mouse also rated the text box and button as easy to accurately select, with the large and medium sizes being most preferred. Both the combo box and check box were rated harder to select than the button and text box with the check box having the worst rating. As with the touch screen overlay, participants preferred large combo boxes and check boxes.

In the general feedback, one participant noted the lack of arm support when selecting targets at the top of the screen. This is an interesting comment, because the nature of using a touch screen means the user's arm might be raised off the desk, and thus be self-supporting when selecting items towards the top of the screen.

Another suggestion was making the target change colour when the cursor is located above it. This is a similar concept to that of interactive rollover items commonly used in web pages. Auditory feedback has been shown to affect speed and accuracy when making a selection (Bender, 1999), so it is likely that the visual feedback received from GUI targets will affect the selection performance. All of the target types tested provide some form of immediate visual feedback, from the button being visually depressed to a tick appearing in the check box. Further study is needed to assess how visual feedback affects selection performance of GUI items and what is the most effective method of providing feedback.

### *6.6 Other findings*

Two interesting patterns emerged when we calculated the standard deviation of the final selection coordinates on the screen. First, when selecting text boxes with the mouse, there was greater variation in final selection coordinates on the right side of the screen (45° and 315° target angles) than on the left side of the screen (90° and 225° target angles). This behaviour was not apparent when selecting text boxes with the touch screen overlay, as illustrated in Figure 8. This would mean that participants using the mouse were more careful when making selections with text boxes on the left side of the screen than on the right.

Second, and even more interesting, was the observation that while selections made on the combo box with the mouse exhibited similar behaviour to the text box, the behaviour on the touch screen overlay was the opposite (although less pronounced), as shown in Figure 9. That is, selections of combo boxes on the left side of the touch screen overlay had greater variation than selections made on the right.

These effects were most pronounced for the large size of both the text box and the combo box. The variation of selection coordinates for the smaller targets (the button and check box) was consistent across both the mouse and touch

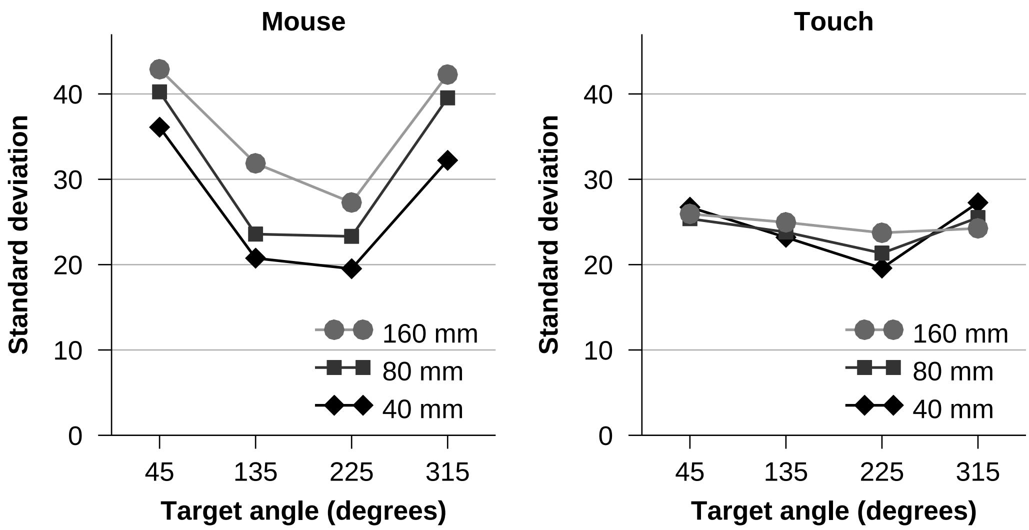


Fig. 8. Standard deviation of final selection coordinates for the text box, averaged across all target sizes.

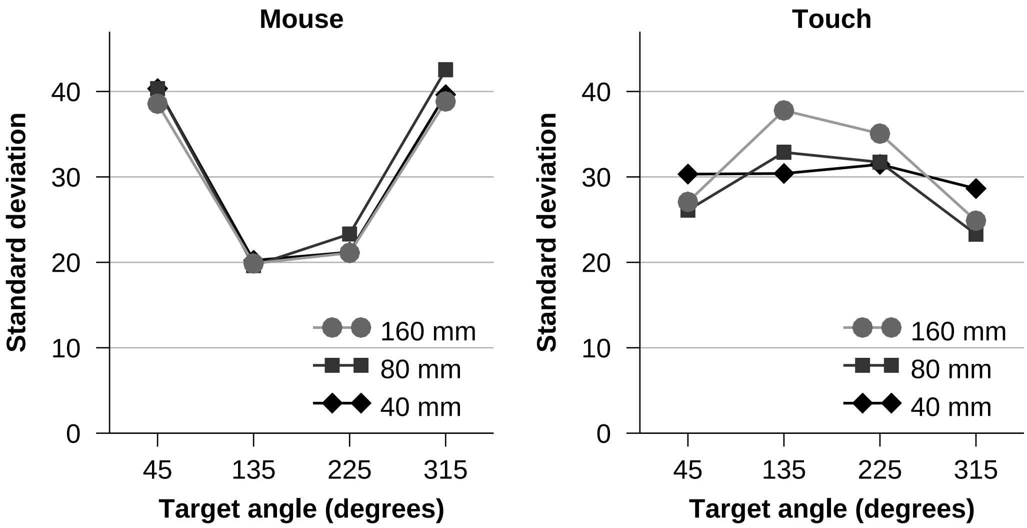


Fig. 9. Standard deviation of final selection coordinates for the combo box, averaged across all target sizes.

screen overlay and for all target sizes.

We can only speculate as to the reasons for this variation. In the case of the combo box, one possibility is its asymmetrical appearance compared to the other target types, which may encourage participants to try to click specifically on the drop-down arrow of the combo box, rather than treating the entire combo box as a target (this could also be another factor in the slow movement times for this target type). However, this does not explain the variation between left and right sides of the screen, nor why the same behaviour was observed with the completely symmetrical text box.

Another possibility is the handedness of the participants, which in the case of a touch screen might affect how difficult it is to select targets on different sides of the screen. Unfortunately we did not ask participants whether they were right- or left-handed, and thus can draw no conclusions on this point.

## 7 Conclusion

The goal of our study was to assess the efficacy of a touch screen overlay compared to a mouse, when selecting the typical GUI targets commonly presented to users in desktop information systems. This was achieved by an experiment measuring movement time, throughput and error rate for various combinations of target type, size, angle and distance. Comfort and user satisfaction were assessed by means of a questionnaire.

The results showed that the touch screen overlay was both slower and less accurate than the mouse. The touch screen overlay was found to have reasonable performance with large GUI items but poor performance with smaller GUI items. The touch screen overlay did have comparable movement times to the mouse for medium and large sized targets. Throughput did not vary across device or target size but did vary across target type. Both selection devices had the same user preference except with respect to the smallest target type (check boxes), in which the mouse had a higher preference. The mouse was rated easier to make accurate selections with than the touch screen overlay. The touch screen overlay also had worse arm, wrist and finger fatigue compared to the mouse. From these results we can conclude that the mouse had higher user satisfaction than a touch screen.

An unusual variation in final selection coordinates was noted for both text boxes and combo boxes. Further studies are required to establish why this variation occurs.

In general we can conclude that a touch screen overlay with no external device (e.g., a pen) is not an effective selection device for targets with dimensions of

4 mm or smaller. When designing interfaces that will be used with a touch screen overlay, selection within the interface will be more efficient if the GUI items are larger than 4 mm.

Although the results showed that the touch screen overlay was not efficient and usable for selecting GUI items with a size of 4 mm or less, this may not be the case when a pen or some external device is used in conjunction with the touch screen overlay. In general there seems to be a lack of research in device assessment of touch screen overlays with pens or other external devices. Further testing on touch screen overlays used with an external device such as a pen may well show that a touch screen overlay is adequate and efficient for selecting small GUI items.

## References

Apple Computer, May 2004. Apple Human Interface Guidelines. Apple Computer, Inc.

Bender, G. T., 1999. Touch screen performance as a function of the duration of auditory feedback and target size. Phd thesis, Wichita State University, Wichita, Kansas.

Douglas, S. A., Kirkpatrick, A. E., MacKenzie, I. S., May 15–20 1999. Testing pointing device performance and user assessment with the ISO 9241, Part 9 standard. In: The CHI is the Limit: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99). ACM Press, Pittsburgh, Pennsylvania, pp. 215–222.

Douglas, S. A., Mithal, A. K., Apr. 24–28 1994. The effect of reducing homing time on the speed of a finger-controlled isometric pointing device. In:

Celebrating Interdependence: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '94). ACM Press, Boston, Massachusetts, pp. 411–416.

Ferguson, E. L., Darmon, N., Briend, A., Premachandra, I., 2004. Food-based dietary guidelines can be developed and tested using linear programming. Journal of Nutrition 134, 951–957.

Fitts, P. M., Jun. 1954. The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology 47 (6), 381–391.

Harada, H., Katsuura, T., Kikuchi, Y., Dec. 1996. Fundamental study on the size and inter-key spacing of numeric keys for touch screen. Journal of Human Ergology 25 (2), 167–174.

ISO, 2000. Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs)—Part 9: Requirements for Non-keyboard Input Devices. International standard ISO 9241-9:2000, International Organisation for Standardisation (ISO), Geneva, Switzerland.

KEYTEC, 2005. KEYTEC add-on magic touch screen for 16–17" LCD and CRT monitors. Web document, KEYTEC, Inc., Richardson, Texas, accessed 3 May 2005.
URL `http://www.magictouch.com/KTMT-1700.html`

MacKenzie, I. S., Jusoh, S., May 11–13 2001. An evaluation of two input devices for remote pointing. In: Little, M. R., Nigay, L. (Eds.), Proceedings of the 8th IFIP International Conference on Engineering for Human-Computer Interaction (EHCI 2001, revised papers). Vol. 2254 of Lecture Notes in Computer Science. Springer, Toronto, Canada, pp. 235–250.

MacKenzie, I. S., Sellen, A., Buxton, W. A. S., Apr. 27–May 2 1991. A comparison of input devices in element pointing and dragging tasks. In: Reaching Through Technology: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '91). New Orleans, Louisiana, pp. 161–166.

Microsoft Corporation, 1999. Microsoft® Windows® User Experience: Official Guidelines for User Interface Developers and Designers. Microsoft® Professional Editions. Microsoft Press.

Po, B. A., Fisher, B. D., Booth, K. S., Apr. 24–29 2004. Mouse and touchscreen selection in the upper and lower visual fields. In: Dykstra-Erickson, E., Tscheligi, M. (Eds.), Proceedings of the 2004 Conference on Human Factors in Computing Systems (CHI 2004). Vienna, Austria, pp. 359–366.

Sears, A., Revis, D., Swatski, J., Crittenden, R., Shneiderman, B., Jan.–Feb. 1993. Investigating touchscreen typing: The effect of keyboard size on typing speed. Behavior & Information Technology 12 (1), 17–22.

Sears, A., Shneiderman, B., 1991. High precision touchscreens: Design strategies and comparisons with a mouse. International Journal of Man-Machine Studies 34 (4), 593–613.