Keyboard and Mouse Errors Due to Motor Disabilities

Shari Trewin and Helen Pain
University of Edinburgh
Department of Artificial Intelligence
80 South Bridge
Edinburgh
EH1 1HN
Scotland
shari@aisb.ed.ac.uk

November 19, 1996

Abstract

There are many people who find the standard computer input devices – the keyboard and mouse – difficult to use due to a motor disability. A number of keyboard and mouse configuration options designed to overcome physical difficulties exist. However, their development has tended to be based on personal experience and intuition rather than real user data. There is, in fact, little data available on the precise nature of physical difficulties with input devices. Hence it is difficult to gauge the adequacy of existing access provision.

This paper presents a empirical study of the keyboard and mouse errors encountered in a sample of twenty computer users with motor disabilities, and six without disabilities. It describes how this data was gathered and analysed, and summarises the nature and frequency of the problems experienced by these users. It is hoped that these results will help to inform the development of more accessible software and hardware.

1 Introduction

There are many computer users with motor disabilities who find that the keyboard and mouse, while not ideal, are the fastest and most convenient input devices for them to use. *Performance errors*, those made due to a physical problem in manipulating the keyboard or mouse, occur frequently for many such users. Examples of performance errors are pressing keys for too long producing repeated letters, striking adjacent keys in addition to the one aimed for, moving the mouse while double clicking, and dropping the mouse button while dragging.

In response to these problems, many mechanisms for reducing or eliminating performance errors have been developed (Brown, 1992) (Lazzaro, 1995) (CALL Centre, 1994). In particular, modern operating systems include a number of features designed to make standard input devices easier for people with motor disabilities to use. One example is *Sticky Keys*, a software facility which causes modifier keys to latch, so that the user need never press two keys at once.

These mechanisms are for their users perhaps the most crucial feature of the computer interface, aside from the hardware devices themselves. Without them many would find the keyboard or mouse unusable. Their design and use is, therefore, an important human-computer interaction (HCI) problem.

While features intended to improve keyboard and mouse accessibility have become a standard part of most popular operating systems over the past decade (CALL Centre, 1994) (Novak & Vanderheiden, 1993) (Lee, 1989) (Novak *et al.*, 1991), for the majority, no formal evaluation is reported in the literature. As McMillan has observed, most work in computing for people with disabilities is:

> "carried out by professionals in education, rehabilitation and communication disorders, usually in isolation from more theoretical research in the field of HCI" (McMillan, 1992)

When the computing needs of people with physical disabilities are assessed, their keyboard and mouse skills are established by observation, and usage data is not recorded (Broadbent & Curran, 1992) (Lee & Thomas, 1990). In fact, the authors are not aware of any detailed input data recorded for physically disabled users of keyboards and mice.

There is, on the other hand, a body of HCI literature assessing the usability of input devices among the general population (Greenstein & Arnaut, 1987) (Hargreaves *et al.*, 1992). Unfortunately, these results are often not relevant to people with disabilities (August & Weiss, 1992).

Historically, HCI research has either examined expert, error-free performance (Card *et al.*, 1987) (Roberts & Moran, 1983), or concentrated on cognitive errors and their causes (Egan, 1988) (Miller & Swain, 1987) (Norman, 1983), despite evidence that 'keyboarding errors' are important and significant, particularly in large databases (Peterson, 1980).

Because there is little or no quantified research on the actual problems that users with motor disabilities experience with standard keyboards and mice, it is difficult to assess the adequacy of the existing provision, and to improve upon it.

An additional motivation for gathering data on keyboard and mouse usage is to allow investigation of the possibilities for automatic recognition of performance errors. If this were achieved, then dynamic support for keyboard and mouse configuration would become possible. This would relieve users of the burden of investigating and activating the relevant features in the existing set of keyboard and mouse access facilities, and could improve the uptake of such facilities among those who would benefit from them.

This paper attempts to bring rigorous empirical HCI research techniques to interface design for people with motor disabilities. It describes a study of twenty people who find keyboards and mice difficult to use, and six who do not, and examines the performance errors occurring and the frequencies of different types of performance error.

The data gathering methodology is described in Section 2 and the subjects are described in Section 3. Section 4 gives an overview of the process by which the data was analysed, arriving at the set of keyboard performance errors described in Section 5 and mouse performance errors described in Section 6. Final conclusions are drawn and directions of current work are described in Section 7.

2 Experimental Design

Data was gathered from twenty computer users with motor disabilities (the main group) and six subjects with no motor disability (referred to as the comparison group). All subjects were asked to perform two typing tasks, two mouse-based tasks, and one editing task. The experimental methodology was tested by the first four subjects (all with some motor disability), who formed a pilot for the remainder (Trewin, 1996b). After the first two subjects, the mouse and editing tasks were revised. The revised version was then carried forward unchanged into the main study.

2.1 Materials

The tasks are based on Apple Macintosh computers, and the ClarisWorks[1] word processing package. Three different venues were used.

Both typing tasks involve copying out the same 100 word passage, the first without correcting errors, and the second with error correction. The former provides easily analysable data, while the latter is a more realistic sample of typing, and introduces problems that occur when errors are made in corrections. It also gives some indication of the time spent correcting errors. The passage is constructed so as to test the user's ability to reach all parts of the keyboard, and to use the shift key in conjunction with keys in a variety of positions. It requires a minimum of 547 key presses, including 25 uses of the *Shift* key.

The two mouse-based tasks are identical. They require the user to perform a set of specified pointing, clicking, multiple clicking and dragging operations on a text passage in which all but the target words are obscured. Targets vary in size from 3 pixels wide to the whole width of the text window. Again, the tasks deliberately involve targets covering the majority of the screen.

The editing task requires the use of both the mouse and the keyboard. A pre-typed passage is edited, the editing tasks covering the same set of basic skills as are examined in the typing and mouse tasks, with the addition of scrolling and selection from hierarchical menus. The editing task provides a more realistic context for performing the operations.

2.2 Procedure

Prior to recordings being taken, each subject was made as comfortable as possible, in order to minimise the effort required to operate the computer, and reduce fatigue. This involved making use of facilities such as wrist rests and adjustable tables, which were available for the majority of subjects. In some cases further adaptions were made as the experiment

[1] ClarisWorks is a registered trademark of Claris Corporation.

progressed. The only external aid that was not permitted was the keyguard.[2] Recordings were made using the default system configuration. The only access option permitted was *Repeat Keys*, as it does not affect the input events recorded.

All subjects used a standard design mouse, with a single button. The mouse tracking was set to sensitive, but reduced if necessary, and the mouse double click speed was on the middle setting.

All concepts used in the experiment (e.g. scroll bars) were explained prior to recording, and subjects who were unfamiliar with computers, keyboards, mice or ClarisWorks were given initial practice time. The effect of unfamiliarity with computers, or of familiarity with alternative systems, is discussed in Section 4.

The tasks were administered by the same observer for each subject. She explained each task as it was presented, and provided verbal help where the subjects required it. Subjects performed each task in their own time, and could rest whenever they chose. Tasks were performed in the order: typing without corrections, mouse 1, editing, mouse 2, typing with corrections. Sessions lasted for up to two hours, extended only if the subject chose to continue. Consequently, many subjects did not complete all the tasks.

For each subject, the following data was recorded:

- *An automatically generated log of input events*, produced by *InputLogger* (Trewin, 1996a). This consists of time-stamped input events, including key up and mouse movement events.
- *A video of the subject performing the tasks.* This is useful in establishing the actual performance errors that occurred. For the typing tasks, the video is focussed on the keyboard. For the mouse and editing tasks, the video is focussed on the screen.
- *Observations made during the word processing tasks.* For each subject, the same observer recorded impressions and particular examples of the keyboard and mouse difficulties experienced by the subject.
- *Background information about the subject.* This includes the nature of their disability, their previous experience with computers and word processors, the set of configuration options they usually use (if known), and their self-reported levels of fatigue and ease of performance of the tasks.

3 Subjects

A total of 26 subjects (12 female and 14 male) aged 25 to 72 took part. Six of these had no physical disability, and provided data for comparison with the main group of subjects. All were volunteers. Some were recruited through personal contacts, some responded to an advertisement seeking people with motor disabilities affecting their hands or arms, and the remainder were contacted through organisations providing computing for people with disabilities. As a result, the subjects' computing experience ranged from expert to none at all. Table 1 summarises each subject's experience, disability, typing technique and hand used to operate the mouse.

[2] Keyguards prevent the majority of performance errors but slow down the rate at which a user can type. This research investigates those same performance errors for users typing at their natural speed.

Sub-ject	Experience	Disability	Typing Technique	Mouse Hand
1	moderate	Stroke	right hand only	right
2	good	Stroke	mainly right hand, prodder	right
3	expert	Spasticity, weakness	both hands, several fingers	left
4	expert	Incomplete tetraplegia	mainly right hand	right
5	moderate	Proprioceptive disorder	left hand, mainly 1st finger	left
6	none	Radial palsy	mainly left hand	left
7	none	Wrist stiffness	both hands, several fingers	right
8	none	Muscle wastage	right hand, mainly 1st finger	right
9	moderate	Shoulder replacement	both hands, several fingers	left
10	moderate	Stroke	left hand only	left
11	a little	Stroke/myelitis	left hand, mainly 1st finger	left
12	moderate	Cerebral palsy	mainly right hand, several fingers	right
13	moderate	Stroke, spasms	both hands, 1st 2 fingers	right
14	a little	Shakiness	mainly right hand	right
15	expert	Cerebral palsy	left hand only, several fingers	left
16	expert	Muscle loss, weakness	both hands, middle fingers and thumbs	left
17	a little	Spina bifida	left hand, 1st 2 fingers + thumb	left
18	a little	Muscular weakness	both hands, any digit	right
19	expert	Cerebral palsy, RSI	both hands, any digit	right
20	expert	Cerebral palsy	both hands, any digit	right
C1	none	none	both hands, 1st 2 fingers	right
C2	expert	none	both hands, any digit	right
C3	expert	none	both hands, several fingers & thumbs	right
C4	moderate	none	both hands, several fingers & thumbs	right
C5	a little	none	both hands, any digit	right
C6	moderate	none	both hands, any digit	left

Table 1: Disability and Previous Computer Experience

4 Analysis

Analysis of the recorded data involves identifying all performance errors. This is done by examining the log data, observations and video evidence, and annotating the log file. Performance errors are annotated according to their type.

Errors which are not performance errors are placed in a single error class – deliberate wrong inputs. Examples include spelling errors, misunderstanding of the task, and errors caused by external events such as the subject being nudged, or distracted in some way.

Seven of the subjects had little or no computer or word processing experience. For these subjects, the practice session provided was vital. After this session, three of these subjects showed no major effects due to inexperience. For four others (Subject 6, Subject 8, Subject 11 and Subject 18) the practice session was not enough, and some difficulties attributable to lack of experience were observed, particularly in mouse dragging tasks. For example, subjects sometimes abandoned a drag operation because they did not understand how to complete it. Where observed, these errors have been classified as deliberate wrong inputs, and not performance errors.

Some deliberate wrong inputs are also caused by a subject having word processing experience on a different application. Ten subjects (including three of the non-disabled subjects) had previous experience that may have conflicted with the operational requirements of ClarisWorks. An example is a subject who clicks on a menu instead of holding down the mouse button to keep the menu in view. It is important not to mistake this habit for an inability to hold down the mouse button.

An additional seven subjects were used to using differently designed mice, or configuration options such as *Sticky Keys*. These differences also caused errors, even though practice was given. Where identifiable, errors due to conflicts with previous experience have been classified as deliberate wrong inputs, and not performance errors.

The annotated log files are then processed to produce summary statistics, and to transform the data into formats appropriate for visualisation and further statistical analysis.

4.1 Keyboard Errors

Keyboard errors are the easiest to identify and classify. A keyboard error occurs wherever an intended key is missed, an unintended key is hit, a key is pressed for the wrong period of time, or keys are pressed in the wrong order. The typing and editing tasks specify what keys should be pressed, and so identifying errors is simply a matter of looking at places where the input differs from that dictated by the task. The observations and video are used to classify these errors according to their underlying cause. The classes of keyboard performance error are described in Section 5.

4.2 Mouse Errors

When using the mouse, the boundary between correct performance and an error is often dependent on the context in which the action is being performed. For example, when dragging the box in the scroll bar, the end of the drag may be a little distance away from the scroll bar itself, and the operation will still be successful. When dragging to select a piece of text in the middle of the screen, on the other hand, the end of the drag must be specified very accurately.

The basic mouse operations are clicking, pointing and dragging. All of these usually have a target. Targets vary greatly in size. The volume of mouse movement acceptable within a click is dependent on the target size and starting position within the target. Even one pixel of movement may cause an error. Consequently, this analysis looks at the movement within all mouse clicks, whether or not an error actually occurred. This is justified by the observation that no click movement is deliberate.

When dragging, for some targets it is acceptable to raise the mouse some distance outside the target area, the scroll box is one common example. Such targets are referred to here as *loose*. Although the mouse up position is outwith the target, the result is not considered an error, since visual feedback is provided to indicate to the user that their action will have the desired effect.

Another parameter of mouse ability is the ease with which a user can position the mouse. Ease can be measured not only in terms of the accuracy of the final position, but also the length of time taken to reach that position, and the directness of the path taken by the mouse. Long times and erratic mouse paths are not errors, but may indicate performance difficulties. However, these phenomena could have many causes besides difficulty with mouse manipulation. The user may have been nudged, or run out of table/mat to move the mouse over.

When dragging, timings are more useful, since it is less likely that the user will pause in the middle of a dragging task. The path taken while dragging is not available for analysis, due to technical restrictions in the *InputLogger* software.

5 Keyboard Performance Errors

Only one of the twenty subjects with disabilities, Subject 2, had no difficulty in using the keyboard. He is excluded from this analysis. Of the remaining subjects, including the subjects without disabilities, 18 attempted both tests and 7 only one of the tests. All tests were completed except for the second typing test for Subject 11, who stopped due to fatigue.

Table 2 summarises the two typing tests, labelled T1 and T2, and includes data for the comparison group (subjects C1-C6). For each subject, the total number of keystrokes they made and the number of minutes they took to complete each task are given. The times varied from 53.2 minutes for Subject 13's second task to 3.8 minutes for Subject 3's first task, the average being 14.0 minutes for the first task and 16.8 minutes for the second. There was some relation between the times taken and experience level.

Among the comparison group the average time was 4.9 minutes for the first and 6.0 minutes for the second task, with a strong relationship between experience level and time taken. The difference in times for the two tasks is largely due to subject C1, who was much slower than the others, and did not do the first task.

Also shown is the time spent correcting performance errors (Perf. Errs) and other errors (Other Errs), given as a percentage of the total time taken. Other errors are all of the deliberate wrong inputs described in the previous section. Error correction times should

Subject	Keystrokes T1	Keystrokes T2	Time (mins) T1	Time (mins) T2	Perf. Errs (% time) T1	Perf. Errs (% time) T2	Other Errs (% time) T1	Other Errs (% time) T2
1	588	1348	15.5	30.7	0.0	18.1	0.0	14.1
3	550	637	3.8	5.1	0.0	18.2	0.0	5.4
4	565	559	4.6	4.5	0.0	0.0	0.0	0.0
5	547	567	8.0	7.7	0.0	3.8	0.0	4.8
6	556	590	27.5	24.0	0.0	3.0	0.0	0.7
7	561	594	17.5	14.9	2.0	16.2	0.0	6.3
8	660	-	36.6	-	0.0	-	0.0	-
9	572	608	6.7	7.6	0.9	8.8	0.7	5.3
10	-	782	-	49.3	-	8.9	-	5.3
11	647	325	8.2	4.4	0.0	0.0	0.0	0.0
12	593	556	16.2	13.2	1.7	2.1	1.9	1.4
13	-	766	-	53.2	-	10.8	-	8.7
14	552	-	17.1	-	0.0	-	3.3	-
15	605	575	14.8	14.7	4.4	0.6	1.2	0.1
16	582	583	5.7	5.5	0.0	0.6	0.0	4.8
17	597	-	28.0	-	0.0	-	0.0	-
18	571	-	14.9	-	0.0	-	0.0	-
19	574	620	5.2	5.9	0.0	9.8	0.0	2.2
20	624	626	16.6	11.3	3.4	8.2	3.6	2.7
Ave.	585	649	14.0	16.8	0.7	7.3	0.6	4.1
C1	-	551	-	11.8	-	0.0	-	8.0
C2	557	566	3.0	3.2	0.0	0.6	0.0	0.0
C3	562	568	3.8	3.8	0.0	2.0	0.0	4.2
C4	554	566	3.9	4.2	0.0	1.5	0.0	15.8
C5	537	593	9.7	8.8	0.0	4.5	0.0	3.8
C6	555	569	3.9	4.4	0.0	0.0	0.0	3.6
Ave.	566	569	4.9	6.0	0.0	1.7	0.0	5.9

Table 2: Summary of the Typing Tasks

be zero for task one for all subjects, but some subjects did make corrections. Subject 4 and Subject 11 made no corrections in the second typing task, having failed to notice the few errors they did make, while Subject C1 and Subject C6 made no performance errors in their second typing tasks. Where errors were corrected, the average time spent correcting performance errors was greater than the average time spent correcting other errors for the main group, and less for the comparison group.

In the typing tests for the main group, examples of 7 different performance errors were observed.[3] In order of frequency, these were:

1. *Long Key Press Errors*: An alphanumeric key is unintentionally pressed for longer than the default key repeat delay.
2. *Additional Key Errors*: A key adjacent to the intended key is activated.
3. *Missing Errors*: The intended key is missed entirely.
4. *Dropping Errors*: The subject fails to press two keys simultaneously (e.g. use of the *Shift* key).
5. *Bounce Errors*: The subject unintentionally presses the intended key more than once.
6. *Remote Errors*: A key not adjacent to any intended key is pressed (e.g. the subject accidentally leans on a key).
7. *Transposition Errors*: Two keys are transposed.

For all performance errors except long keypress errors, Table 3 gives the number of errors of each type observed in each test. The numbers of long keypress errors given are projected values showing the errors that would have occurred had the default key repeat delay been used.

Subjects were also asked to rate how difficult they found it to press two keys at once (Shift), reach all the keys on the keyboard (Reach), aim accurately at a key (Aim), only press a single key (Isolate) and to press keys quickly (Fast). Answers were on a scale ranging from *easy*, through *some difficulty*, *moderate difficulty*, *hard*, *very hard*, *extreme difficulty* up to *impossible*. Table 4 gives each subject's answer for each category.

These responses provide some indication of the subjects' personal opinion of their keyboard difficulties. In general, a subject's opinions were reasonable indicators of the relative numbers of performance errors of different types that were observed. In some cases, subjects reported difficulty but made no errors, illustrating that error rates are not always a good indicator of the ease of performance of a given task. Between subjects, there is less correlation between reported levels of difficulty and relative numbers of errors, since each individual has a different awareness of tolerance of their own keyboard errors.

5.1 Long Keypress Errors

On a Macintosh, the default delay before a key will repeat is 16 ticks (1 tick = 1/60 sec). For many people, this is too short. Unwanted extra copies of a letter are long keypress errors.

[3] One example of an eighth performance error – that of holding down the *Shift* key for too long – was observed in the comparison group, but is not discussed here.

Subject	ave. key len. (ticks)	no. of shifts	long key press (proj.)		additional		miss		drop		bounce		remote		transpose	
			T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2
1	7	8	0	0	38	35	11	6	0	3	0	0	0	0	0	0
3	9	56	30	35	9	4	10	3	0	1	0	0	0	1	0	2
4	4	56	0	0	4	0	2	1	8	6	0	0	2	0	0	0
5	6	56	13	39	6	4	4	4	0	0	0	0	0	0	0	0
6	10	22	31	27	2	2	12	18	2	3	8	12	2	0	1	0
7	11	56	37	23	11	3	8	3	4	0	0	0	10	2	0	0
8	12	28	114	–	5	–	4	–	0	–	0	–	2	–	0	–
9	9	56	16	39	13	6	11	5	0	2	0	1	0	0	0	0
10	17	11	–	377	–	3	–	0	–	0	–	0	–	0	–	0
11	4	11	1	0	4	1	1	0	0	0	0	0	0	0	0	0
12	16	56	311	171	4	1	1	1	4	2	2	1	0	1	0	0
13	20	28	–	510	–	4	–	4	–	3	–	3	–	0	–	0
14	10	28	14	–	0	–	0	–	0	–	0	–	0	–	0	–
15	10	22	14	44	9	8	10	8	11	2	3	3	1	7	0	0
16	5	56	0	0	3	1	4	1	0	0	0	0	0	0	0	0
17	10	0	45	–	4	–	1	–	0	–	0	–	0	–	0	–
18	9	11	9	–	0	–	0	–	0	–	0	–	0	–	0	–
19	16	56	299	357	28	21	7	8	0	0	3	0	2	2	0	0
20	10	56	36	18	25	14	20	13	2	2	5	3	5	0	0	1
Total			2610		272		179		55		44		37		4	
C1	8	28	0	–	0	–	0	–	0	–	0	–	0	–	0	–
C2	5	56	0	0	2	2	1	1	0	0	0	0	0	0	0	0
C3	4	56	0	0	0	3	0	0	0	0	0	0	0	0	0	0
C4	5	56	0	0	2	1	1	1	0	0	0	0	0	0	0	0
C5	4	56	0	0	0	1	1	5	0	1	0	0	0	0	0	1
C6	5	56	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Total			0		11		10		1		0		0		2	

Table 3: Performance Errors in the Typing Tasks

Subj.	Shift	Reach	Aim	Isolate	Fast
1	impossible	easy	some diff	moderate	easy
2	easy	easy	easy	easy	easy
3	easy	easy	some diff	some diff	easy
4	moderate	easy	some diff	some diff	easy
5	hard	easy	some diff	some diff	easy
6	impossible	moderate	easy	easy	easy
7	easy	easy	easy	some diff	some diff
8	easy	easy	easy	easy	easy
9	some diff	some diff	some diff	easy	easy
10	very hard	easy	easy	easy	hard
11	impossible	easy	easy	easy	easy
12	some diff	easy	easy	easy	easy
13	easy	easy	easy	easy	very hard
14	easy	easy	easy	some diff	hard
15	hard	moderate	some diff	some diff	hard
16	easy	easy	some diff	easy	easy
17	easy	easy	easy	easy	easy
18	easy	easy	easy	easy	easy
19	moderate	easy	moderate	moderate	extreme
20	moderate	some diff	some diff	easy	easy
C1	easy	easy	easy	easy	easy
C2	easy	easy	easy	easy	easy
C3	easy	easy	easy	easy	easy
C4	easy	easy	easy	easy	easy
C5	easy	easy	easy	easy	easy
C6	easy	easy	easy	easy	easy

Table 4: Reported Typing Difficulties

To investigate a user's ideal key repeat delay, all that is required is knowledge of the lengths of their key presses. Consequently the first six subjects performed the experiment with the key repeat switched off. The recordings made indicated their most natural keypress length, in the absence of any requirement to lift keys quickly. Switching off the repeat facility, however, made it impossible to record time spent correcting long keypress errors. To provide some data on correction of these errors, the key repeat facility was set to repeat after 24 ticks (a medium long delay) for twelve of the remaining subjects, while Subject 15 and Subject 20 both chose to have the repeat facility disabled. A long value was chosen because the use of a short delay such as 16 ticks would have caused so many long keypress errors for some subjects that correction of those errors in the second typing task would have been a daunting task, and would have taken an unreasonably long time. The comparison group all used the minimum repeat delay, which was 10 ticks.

Because of the differences in repeat delay settings used by the subjects, this analysis does not report the actual number of errors that occurred, but the number of errors that would have occurred had each subject been using the default key repeat delay of 16 ticks. The actual numbers of errors were zero for subjects 1-6, 15 and 20, since the key repeat facility was disabled. For subjects 7-14 and 16-19, a long repeat delay was in force, so the actual numbers of errors occurring were much smaller than the numbers in the table (the maximum being 43 for Subject 13 and Subject 10).

Table 3 details, for each subject, the average time for which they pressed keys down, and the number of key presses longer than the default key repeat delay. These values are measured using alphanumeric and punctuation keys only. The *Delete* key, the arrow keys, *Return* and all modifier keys are excluded from the calculation.

Among the comparison group, the average key press length was 5 ticks, with the longest for any individual being 8 ticks. No key presses longer than 16 (or even 10) ticks were observed. Among the main group of subjects, Subject 13 had the longest average key press length, at 20 ticks, and reported that he found it very difficult to press keys quickly. Subject 10, Subject 12 and Subject 19 also had long average key press lengths. Subject 19 reported that she found short key presses extremely difficult, while Subject 10 also found them difficult. Subject 12 reported no difficulty, having the key repeat facility disabled for his test. He has word processing experience on a PC and reported no problem with long key presses there. It is not known what key repeat setting is in force on his usual machine.

In general, there seems to be little correlation between a subject's reported ease of pressing keys quickly and their average key press length. A user's perceptions may well be more dependent on the settings of the system they usually use than the actual time for which they press keys.

The large variability in average key press lengths between individuals, and the high upper values found indicate that key repeat settings are perhaps the most important issue in keyboard accessibility.

5.2 Additional Key Errors

As Table 3 shows, the next most common error, made by 18 of the subjects, was that of pressing down keys adjacent to the intended key. All subjects who made more than 5 such errors on either task reported some or moderate difficulty in isolating keys or in reaching all the keys on the keyboard. Two other subjects, both with previous experience, reported some difficulty in isolating keys, but made few errors of this type. It may be that the text passage was too short to show up their additional key errors. Users are very aware of these errors, and these results suggest that they are a significant problem for many users with disabilities.

For those who make many additional key errors, keyguards are often suggested as a way of reducing these errors. A keyguard is a piece of hard clear plastic which fits over the keyboard, with holes through which the keys are pressed.

Only one of the subjects, Subject 20, usually uses a keyguard. For the experiment she used a wrist rest and no keyguard, which she found a more comfortable configuration, of comparable speed to using a keyguard. The number of additional key errors she made was surprisingly low, particularly in the second test, when she was more used to the wrist rest.

Subject 1 and Subject 19, who were the most prone to additional key errors, have both tried using keyguards but prefer the keyboard bare. Subject 1 had trouble getting his fingers through the holes, while Subject 19 felt that it would slow her down too much.

Additional key errors were the most common performance error among the comparison group. Interestingly, 10 of the 11 examples observed were made by the three most experienced computer users. Among the main group, the subjects making the most additional key errors had a variety of experience levels.

5.3 Missing Key Errors

Seventeen of the subjects sometimes missed the key they were aiming for entirely – a total of 179 examples of this error were observed. Neither of the two subjects who made the most missing key errors reported much difficulty in hitting the key they were aiming for, but both of them did report some or moderate difficulty in reaching all the keys on the keyboard. In general, however, subjects were reasonably aware of their missing key errors. Again, two experienced subjects reported difficulty that was not reflected in their error rates. This could be due to the experimental conditions – subjects being more careful than usual, or to the short length of the text passage.

Missing key errors were also one of the major performance errors among the comparison group, particularly for Subject C6.

5.4 Dropping Errors

The keyboard operation rated as the most difficult by the subjects was that of pressing two keys at once. This was rated *impossible* for three subjects, *hard* or *very hard* for three, and *quite* or *moderately difficult* for five. There were 28 modified key presses in the passage, 17 of which were capital letters and 11 punctuation marks.

A dropping error occurs when the *Shift* key is raised before the key to be modified has been pressed. Table 4 gives each subject's rating of difficulty of pressing two keys at once, while Table 3 shows the number of dropping errors they made, and the number of uses of *Shift* they attempted. Despite the avoidance of use of *Shift* by several subjects, 55 examples of dropping errors were observed. Not all subjects who found this operation difficult made dropping errors – some were able to use *Shift* accurately, but with great effort.

Nine subjects used the *Shift* key for all modification, and three used *Shift* most of the time, but sometimes switched to *Caps Lock*. Five subjects used the *Caps Lock* key for all capital letters and *Shift* for all punctuation. One subject omitted all capital letters and punctuation and one used *Caps Lock* for all modification, even though it didn't work for punctuation.

All subjects who rated the use of *Shift* as very difficult or above used the *Caps Lock* key for all capital letters, and one also used it for punctuation.

5.5 Bounce Errors

Seven subjects made bounce errors, which occur when a key is pressed more than once, perhaps because the user's finger twitched while releasing the key. For most subjects, these errors occurred perhaps once in every 200/300 keystrokes. For Subject 6, however, approximately one in sixty keystrokes bounced. The observation of 44 instances of bounced keys in 32 short typing tests is notable – this performance error is reasonably common.

5.6 Remote Errors

A remote error occurs when a user accidentally presses a key while reaching for another key in a different part of the keyboard. Accidentally pressing down a key by leaning on it is also a remote error. Of the 37 remote errors observed, 25 were on the bottom row of the keyboard. Only Subject 7 and Subject 15 seemed particularly prone to remote errors, and Subject 7's total decreased dramatically the second time she performed the typing task – inexperience may have played a part in her initial high count.

Subject 20, who usually uses a keyguard, made five remote errors. While this is higher than most subjects, it is low enough that a keyguard is not necessary to prevent remote errors.

5.7 Transposition Errors

When two keys are typed in the wrong order, a transposition error has occurred. Transposition errors have been reported as significant in at least one study of spelling errors (Damerau, 1964), although it is not clear whether they occurred through human error or machine malfunction in this case. Certainly, two examples of transposition errors were observed in the comparison group, which would seem to indicate that they are worthy of consideration among the general population. However, they were not common among the main group of subjects, probably due to the large number of one finger typists and slow typists among the subjects.

It is debatable whether the majority of transposition errors are true performance errors, since they could be attributed to timing misjudgements rather than difficulty in controlling the timing of movements of different fingers. The data gathered in this study suggests that whatever the definition of transposition errors, they are not an important source of keyboard difficulty for those with motor disabilities.

6 Mouse Performance Errors

Subjects were asked to perform the same set of pointing, clicking, multiple clicking and dragging operations twice. They were also asked how difficult they found each of these operations, and how difficult it was to pick up the mouse and reposition it on the table. Their responses are given in Table 5, along with the minutes taken to perform each of the mouse tasks (M1 and M2).

The data for Subject 18, who had great difficulty in understanding how to use the mouse, has been excluded. Some mouse data is not available for Subject 1 and Subject 2, who performed a pilot version of the mouse tasks, but they are included where possible.

Of the 19 subjects, 13 performed the mouse task twice and 6 only once, due either to fatigue or lack of time.

All but one of the subjects used a one button curved mouse where the button occupied the whole of the upper part of the mouse. Subject 19, who performed the experiment on her own computer, used a one button flat mouse where the button was rectangular and positioned in the middle of the upper part of the mouse.

12 subjects had tried alternative mouse designs, usually a trackerball. Of these, 4 preferred the trackerball and 8 preferred the standard mouse.

The times taken by the main group of subjects varied from 2.9 to 28.6 minutes. These times are influenced by their experience level and in some cases by cognitive impairments. Most subjects performed the task more quickly the second time around, and these second times give a more reliable estimate of the extent of each subject's mouse difficulties. The comparison group took between 2.3 and 7.4 minutes to perform the tasks, with little difference between the first and second tasks.

The difficulties observed in pointing, clicking, multiple clicking and dragging are each discussed in the following sections.

6.1 Pointing

Pointing is the most fundamental mouse operation, and yet also one of the most difficult. Seven of the subjects with disabilities rated pointing as being as hard or harder than any other mouse operation, and only four rated it as easy.

Of these subjects, Subject 20 made the most errors in pointing and clicking on a given target: 47.0% of her clicks and drag starting positions were off target. A total of 8 of the subjects had error rates over 20%, and 14 were over 10%. This is in stark contrast to the results for the comparison group, where the maximum error rate recorded was 6.3%. The three highest error rates in the comparison group were recorded by the three most experienced subjects. No link between experience and pointing accuracy can be determined for the main group.

Another indicator of difficulty in pointing is the time taken to position the mouse. This is difficult to separate from time spent getting the next task, locating the target, or pausing to ask questions. However, to give a rough idea of the time spent pointing, two tasks are presented here. These are "click on the Apple menu" and "double click on the word 'so' ". They occur consecutively in the mouse test. For the main group, the average time between completing the previous task and clicking on the Apple menu was 15.6 seconds,

Subject	Point	Click	Multi Click	Drag	Pick up	time (mins)	
						M1	M2
1	easy	easy	some diff	easy	easy	n/a	n/a
2	some diff	easy	easy	moderate	impossible	n/a	n/a
3	some diff	easy	easy	moderate	easy	3.0	2.9
4	some diff	easy	easy	some diff	some diff	4.7	3.1
5	some diff	easy	easy	hard	easy	4.6	7.3
6	very hard	some diff	moderate	impossible	moderate	16.6	8.4
7	some diff	some diff	easy	some diff	easy	28.6	14.7
8	some diff	easy	easy	moderate	easy	9.7	-
9	hard	easy	easy	moderate	easy	6.8	5.2
10	easy	some diff	easy	hard	easy	4.7	9.3
11	very hard	moderate	moderate	moderate	moderate	13.7	23.0
12	hard	hard	easy	very hard	easy	14.3	11.1
13	hard	hard	moderate	some diff	easy	21.7	-
14	moderate	easy	easy	extreme	easy	23.0	-
15	moderate	moderate	hard	hard	moderate	10.7	-
16	easy	easy	easy	easy	easy	4.1	3.5
17	easy	easy	easy	hard	easy	16.4	12.6
19	some diff	some diff	very hard	some diff	hard	3.6	2.9
20	hard	easy	easy	hard	easy	15.8	-
C1	easy	easy	easy	easy	easy	7.4	-
C2	easy	easy	easy	easy	easy	4.5	2.6
C3	easy	easy	easy	easy	easy	3.2	2.8
C4	easy	easy	easy	easy	easy	2.7	2.3
C5	easy	easy	easy	easy	easy	3.2	2.4
C6	easy	easy	easy	easy	easy	3.7	2.3

Table 5: Summary of Mouse Difficulties and Tasks Performed

and the average time between the click on the Apple menu and double clicking on ‘so’ was 24.0 seconds. The distance travelled to each target from the previous target is similar. The difference in time taken may be explained by the observation that the Apple menu is easier to find and its position in the corner of the screen reduces time wasted through overshooting. For the comparison group, the average time to click on the Apple menu was 6.4 seconds, and to reach the target ‘so’ was 11.7 seconds, showing a similar proportional difference in the time taken to reach each target.

Examination of the path taken by the mouse in reaching these targets showed little difference between the groups. In both groups, the mouse path was sometimes very direct, sometimes overshot the target, or sometimes took an indirect route to the target.

Most subjects performed the mouse tasks with the mouse at its most sensitive setting. This had the advantage that it reduced the range of motion required to operate the mouse, and the frequency of needing to lift the mouse off the mouse mat and reposition, but it also made smaller targets more difficult to pinpoint. One subject found it impossible to lift the mouse and reposition it on the table, while five others reported some difficulty. Several subjects reported that they found the mouse more sensitive than they were used to, but in only two cases was this sufficiently problematic for them to require a slower setting.

Another difficulty that some subjects experienced was in positioning the mouse without activating the mouse button. Accidental clicks can be difficult to differentiate from deliberate but random clicks, and where there was doubt, clicks were assumed to be deliberate. Nevertheless, examples of 116 unintentional clicks were observed, the maximum for any individual being 19, for Subject 6. Many of these had unwanted side effects such as bringing the Finder to the front or changing the current text view, and recovery could take some time.[4]

6.2 Clicking

Difficulty in controlling the mouse can also manifest itself in the mouse clicks themselves. It is particularly interesting to examine any movement between the mouse down and mouse up events. Such movement may or may not cause an error, depending on whether the mouse up event lies on the same target as the mouse down event. Even one pixel of movement could potentially cause an error. The median value for the percentage of clicks that moved was 28.1% in the main group, compared to 6.3% in the comparison group. The maximum value observed was 75%, for Subject 1, and the minimum 7.1%, for Subject 3. In the comparison group, the maximum click movement was 15.3%, and the minimum was 1.8%.

Interestingly, some subjects showed strong bias in the direction of mouse movement while clicking. Subject 1 and Subject 6 are good examples. Their click movements are graphed in Figure 1. The axes of the graphs measure pixels, and are oriented as they would be on a Macintosh screen. Taking the origin as the mouse down position, the graphs show the number of times the mouse was raised at each position.

Subject 1 tends to slip up and right or down and left, he is using his right hand to

[4] Genuine recovery times are not known, as the observer provided instructions on how to recover from such errors.

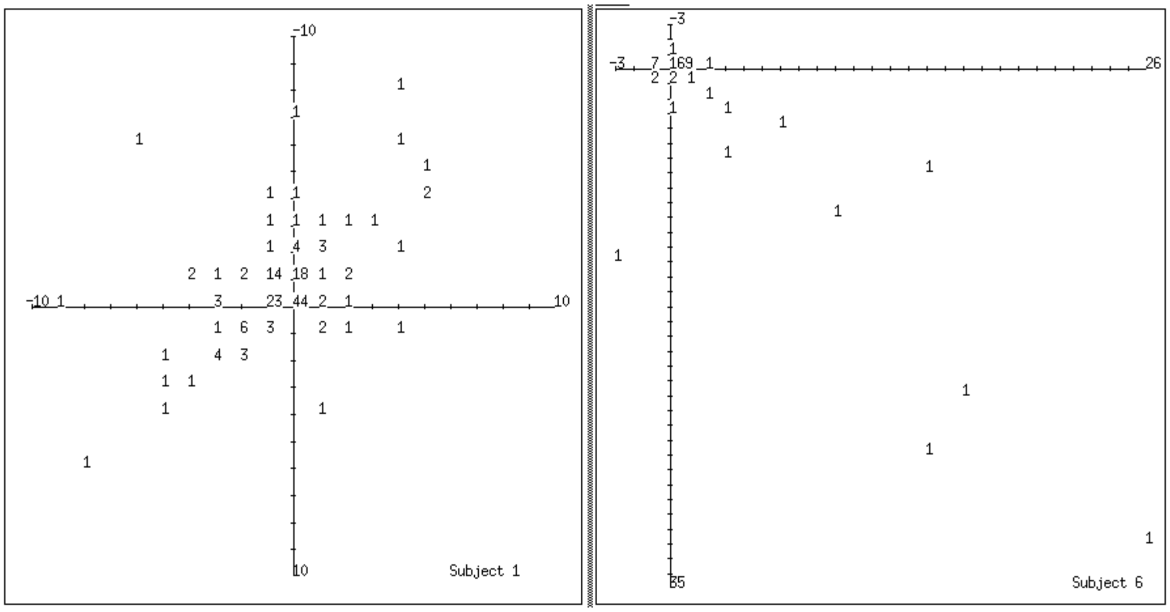

Figure 1: Click Movements for Subject 1 and Subject 6

operate the mouse. Subject 6 slips down and to the right, and uses the mouse in his left hand. Directional biases such as these were observed in 11 subjects. The direction of movement appeared to be independent of the hand in which the mouse was held.

6.3 Multiple Clicking

For those who find clicking difficult, double, triple or other multiple clicks pose even more problems. The mouse task required the user to perform three double clicks, and one triple click. The editing task contained at least two double clicks, more if subjects chose to use that technique for selecting words in the text.

A total of 164 multiple clicks on a known target were observed. 233 attempts to multiple click on these targets were made, resulting in 141 successful attempts. Of the unsuccessful attempts, 25 missed the target entirely, 29 involved too much movement within or between clicks, and 13 were too slow to be recognised as multiple clicks. In addition, in 23 cases the wrong number of clicks were made. Not all of these were performance errors, sometimes subjects deliberately added extra clicks because earlier clicks had moved, or been too slow, or because the system was slow to respond.

As part of the mouse tasks, subjects were also asked to click the mouse quickly many times while keeping it still. This provided additional multiple click data. The average time between connected clicks for each subject in the main group varied between 6 (Subject 10, Subject 16) and 20 (Subject 6) ticks. The default maximum time between clicks on a Macintosh is 16 ticks. The average for five of the subjects was on or over this value. In the comparison group, average gaps were between 3 and 6 ticks.

Overall, multiple clicking problems were fairly evenly divided between positioning difficulties, difficulties in keeping the mouse still while clicking, and click timing difficulties. There was little or no relation between subjects' reported difficulty in multiple clicking and the numbers of errors observed.

6.4 Dragging

Dragging is an extremely common operation in word processing, and the one rated most difficult by 12 of the subjects. HCI research has also reported dragging to be more difficult and error-prone than pointing (MacKenzie *et al.*, 1991) (Bewley *et al.*, 1990). This is hardly surprising, since a drag operation requires the ability to point, press the mouse for a prolonged time, and move the mouse accurately with the button pressed down.

Positioning and pressing down the mouse button at the start of the drag have been discussed in Section 6.1. The following discussion examines the drag operations for which the end target of the drag is known. 54 drags deliberately abandoned for reasons such as misunderstanding the task are excluded. Drags abandoned because of physical difficulties are included.

Once a drag had been started, the most common difficulty observed was in raising the mouse button in the correct position at the end of the drag. While 215 correctly completed drags were observed in the main group, in a further 106 examples, a subject deliberately completed a drag, but the text selected was wrong. Often, they had been careful to position the mouse correctly, but it had slipped as they raised the button.

Another common problem was difficulty in holding down the mouse button while moving the mouse. In 89 cases, the subject accidentally raised the mouse button while dragging.

A third difficulty, of which 38 examples were observed, was that some subjects could get stuck and have to abandon a drag. Sometimes this was because they had run out of space to move the mouse and could not lift it up while holding down the mouse button. In other examples, they reached a position where they were physically unable to make the required movement in order to complete the drag.

Finally, 30 drags were abandoned because the subject had moved the pointer off the text window causing it to scroll. This commonly happened with targets at the bottom of the text window.

The comparison group made no dropping errors and did not abandon any drags for physical reasons, or because the screen had scrolled. Of 138 deliberately completed drags, 131 were accurately targeted.

The average time taken to complete a drag varied greatly between subjects. The lowest average was 181 ticks for Subject 3, while the highest was 2117 ticks for Subject 7 (excluding an extreme value of 5273 ticks for Subject 6, for whom only one successful drag was observed). The median value was 555 ticks. Among the comparison group, the maximum average was 567 ticks, the minimum 171 ticks and the median 197 ticks, noticeably faster.

It is difficult to relate subjects' reported difficulty in dragging to either the time taken or the number of failed drags observed. Factors such as each subject's experience, patience and desire for accuracy cloud the relationship. Three subjects (Subject 6, Subject 11 and Subject 20) found dragging so frustrating that they switched to alternative selection mechanisms. As a result, the number of drags recorded for these subjects is small, and the figures reported here downplay the actual difficulty of dragging tasks.

7 Conclusions

The experiments described have gathered data recording the keyboard and mouse usage of twenty computer users with physical disabilities, and six comparison subjects with no disability affecting their use of these input devices. In both groups, the computing and word processing experience of the subjects ranged from none at all to expert. From this data, performance errors related to physical control have been extracted and examined.

Seven keyboard performance errors have been identified, the most common of which was that of pressing keys for longer than the default key repeat delay. Other common errors were in pressing keys adjacent to the intended key, and missing the intended key. Difficulties in pressing two keys at once, bouncing on the intended key and activating keys remote from any intended key were also fairly common. In contrast, the comparison group made no key press length errors. For them, additional key presses and missing letters were the most common errors.

It is interesting to note that the inexperienced subjects in the comparison group made less keyboard errors than the experienced subjects. This is partly because they were typing very carefully, and also partly because the experienced subjects were used to using different keyboards. Among the main group of subjects, there was no discernible effect of experience on the number of errors, as the effect of disability dominated.

The proportion of time spent correcting performance errors was significant for many of the subjects, even though long keypress errors were suppressed for the majority of them. This emphasises the importance of appropriate keyboard configuration in order to provide maximum accessibility.

The majority of the subjects observed found the mouse difficult to use, and many generally try to avoid mouse operations as much as possible. This was reflected in high error rates for positioning the mouse in clicking and dragging operations, and long times spent positioning and dragging relative to the comparison group. The subjects also found difficulty in holding the mouse still while clicking, and in performing multiple clicks successfully.

Despite problems with the mouse, most subjects who had tried alternative mouse devices such as the trackerball preferred the standard mouse, in some cases because they were more used to it.

This last point is important. Many people have access to, or are required to use, standard computers with ordinary keyboards and mice. Many people with physical disabilities can use these devices, but find that performance errors frequently arise. Although specialised input devices are necessary for some users, others prefer the convenience of being able to use the standard devices, and tolerate the frustration caused by performance errors.

Keyboards and mice can be configured in order to reduce performance errors. These results provide some data by which the adequacy of existing configuration options can be assessed. For an analysis of keyboard configuration options see (Trewin & Pain, 1996). In addition, the authors are currently developing automatic recognisers for specific performance errors, which will lead to the development of dynamic support for the configuration of keyboards and mice. By enabling such work, the data gathered in this study will, it is hoped, contribute to the development of improved access features in the future.

8 Acknowledgements

The authors would like to thank the University of Edinburgh for funding this research; all the volunteer subjects and organisations who helped in contacting potential subjects, particularly Craighall Day Centre, Stevenson College and the Thistle Foundation; Lothian Regional Council and the Hands-On-Technology project for providing access to the computers and equipment used for the majority of the subjects; and Annalu Waller, Nik Whitehead, Simon Kelly, Mike Ramscar, Richard Cox and Catriona Smith for helping out with feedback, time and equipment.

References

August, Sharon and Weiss, Patrice. (1992). A human-factors approach to adapted access device prescription and customization. *Journal of Rehabilitation Research and Development*, 29(4):64–77.

Bewley, William, Roberts, Teresa, Schroit, David and Verplank, William. (1990). Human factors testing in the design of Xerox's 8010 'Star' office workstation. In Preece, Jenny and Keller, L., (eds.), *Human-Computer Interaction: Selected Readings*. Prentice-Hall, Hemel Hempstead.

Broadbent, Steven and Curran, Sandra. (1992). *The assessment, disability and technology handbook*. North West Regional ACCESS Centre and Oldham Education Department, Oldham.

Brown, Carl. (May 1992). Assistive technology computers and people with disabilities. *Communications of the A.C.M.*, 35(5):36–45.

CALL Centre. (July 1994). Alternative access to the Apple Macintosh: Using in-built features. University of Edinburgh, 4 Buccleuch Place, Edinburgh.

Card, Stuart, Moran, Thomas and Newell, Allen. (1987). The keystroke-level model for user performance time with interactive systems. In Baecker, Ronald M. and Buxton, William A. S., (eds.), *Readings in human-computer interaction : a multidisciplinary approach*, chapter 5, page 192. Morgan Kaufmann, Los Altos, Calif.

Damerau, F.J. (1964). A technique for computer detection and correction of spelling errors. *Communications of the A.C.M.*, 7(3):171–6.

Egan, Dennis E. (1988). Individual differences in human-computer interaction. In Helander, Martin, (ed.), *Handbook of Human-Computer Interaction*, chapter 24, pages 543–580. Elsevier Science Publishers B.V.

Greenstein, Joel and Arnaut, Lynn. (1987). Human factors aspects of manual computer input devices. In Salvendy, Gavriel, (ed.), *Handbook of Human Factors*, chapter 11.4, pages 1450–1489. John Wiley and Sons, New York.

Hargreaves, William, Rempel, David, Halpern, Nachman, Markison, Robert, Kroemer, Karl and Litewka, Jack. (1992). Toward a more humane keyboard. In *Proceedings of CHI*, pages 365–368, New York. ACM.

Lazzaro, J. J. (1995). *Adapting PCs for disabilities.* Addison Wesley.

Lee, K. S. and Thomas, D. J. (1990). *Control of Computer-based technology for people with physical disabilities: an assessment manual.* University of Toronto Press, Canada.

Lee, Charles. (1989). Access to microsoft windows 2.0 for users with physical disabilities. In *Resna '89: Proceedings of the 12th Annual Conference*, pages 23–24, Washington D.C. Resna Press.

MacKenzie, Scott, Sellen, Abigail and Buxton, William. (1991). A comparison of input devices in elemental pointing and dragging tasks. In *Proceedings of CHI*, pages 161–166, New York. ACM.

McMillan, William W. (1992). Computing for users with special needs and models of computer-human interaction. In *Proceedings of CHI*, pages 143–148, New York. ACM.

Miller, Dwight and Swain, Alan. (1987). Human error and human reliability. In Salvendy, Gavriel, (ed.), *Handbook of Human Factors*, chapter 2.8, pages 219–250. John Wiley and Sons, New York.

Norman, Donald A. (April 1983). Design rules based on analyses of human error. *Communications of the A.C.M.*, 26(4):254–258.

Novak, Mark and Vanderheiden, Gregg. (June 1993). Extending the user interface for X windows to include persons with disabilities. In *Resna '93: Proceedings of the 16th Annual Conference*, pages 435–436, Washington D.C. Resna Press.

Novak, Mark, Schauer, Joseph, Hinkens, Jay and Vanderheiden, Gregg. (1991). Providing computer access features under DOS. In *Resna '91: Proceedings of the 14th Annual Conference*, pages 163–165, Washington D.C. Resna Press.

Peterson, J. (1980). Computer programs for detecting and correcting spelling errors. *Communications of the A.C.M.*, 23(12):676–687.

Roberts, T. L. and Moran, T. P. (1983). The evaluation of text editors: methodology and empirical results. *Communications of the A.C.M.*, 26:265–283.

Trewin, Shari and Pain, Helen. (1996). On the adequacy and uptake of keyboard access facilities for people with motor disabilities. DAI research paper. submitted to *chi97*, AI Dept, University of Edinburgh.

Trewin, Shari. (1996a). *InputLogger*: General purpose logging of keyboard and mouse events on a Macintosh. DAI research paper. submitted to *behaviour research methods, instruments and computers*, AI Dept, University of Edinburgh.

Trewin, Shari. (April 1996). A study of input device manipulation difficulties. In *Proceedings of the Second Annual ACM Conference on Assistive Technologies*, pages 15–22, USA. ACM. ACM ISBN: 0-89791-776-6. Also available as DAI Research Paper 812, Dept. of Artificial Intelligence, University of Edinburgh.