Information Processing & Management Vol. 24, No. 6, pp. 665-680, 1988
Printed in Great Britain.
0306-4573/88 $3.00 + .00

SELECTION DEVICES FOR USER OF AN ELECTRONIC ENCYCLOPEDIA: AN EMPIRICAL COMPARISON OF FOUR POSSIBILITIES

DANIEL OSTROFF*

and

BEN SHNEIDERMAN†
Department of Computer Science, University of Maryland,
College Park, Maryland 20742, USA

(*Received July* 1987; *revised January* 1988)

Abstract—This study measured the speed, error rates, and subjective evaluation of arrow-jump keys, a jump-mouse, number keys, and a touch screen in an interactive encyclopedia. A summary of previous studies comparing selection devices and strategies is presented to provide the background for this study. We found the touch screen to be the fastest in time, the least accurate but the overall favorite of the participants. The results are discussed and improvements are suggested accordingly.

1. DESCRIPTION OF SELECTION SYSTEMS

1.1 *Introduction*

Selection as a means of interaction is ubiquitous in computer systems. Menus, in their various guises, are a common means of presenting items. An important aspect of using menus is the strategy of selection, a term that refers to the way a selection was made, such as by pointing or naming. When a selection strategy is employed on a particular device, it is termed a selection tactic. From the system view, a selection tactic is the implementation of strategy on a device; for the user, the tactic is how the selection is accomplished within a particular program. It is possible that the device and strategy could interfere with the selection process.

1.2 *Selection devices*

A selection device is an object that is used as the means for indicating a particular item given a set of selectable items. Noncomputer selection devices include a pointer on a blackboard, a finger on a restaurant menu, or even a nod of the head indicating which door is to be opened. Because of the technology, selection devices in computer applications can be expensive in addition to being difficult to learn.

The possible devices for a selection system are joystick, keyboard, light pen, mouse, touch screen, and trackball. Research into other means of input (cited in [28]) such as detecting foot, wrist, eye and head motion or voice recognition, continues. The effectiveness of a device is dependent on the task and the user, but there is a notion of virtual input devices which can be altered to the given user and task situation [5].

1.3 *Interpretation and summary of previous research*

The rapidly growing body of literature for selection devices is summarized in Table 1. Despite the range of tasks and experimental designs, some patterns emerge from these comparative studies. Most tasks involved positioning or locating, and some tasks included data entry as well. All studies involving the touch screen showed it to be the fastest device; it is most effective when it is on display employing the direct pointing strategy. The study by Karat *et al.* [18], however, pointed out that when combined with key-

*The Museum of Jewish Heritage, 342 Madison Avenue, New York, New York 10173.
†Author to whom correspondence should be addressed.

board entry, remaining on the keyboard was preferable to switching devices. Another consistent result is the more rapid use of the mouse over the joystick. Evaluating the light pen and the mouse, English [12] showed the mouse to be faster, but Haller *et al.* [16] showed the opposite to be true. No doubt, technological improvements in the light pen were a factor.

The keyboard itself has many variations, such as keypads, and many possible tactics such as text, letter, number and arrow keys. The study by Shinar *et al.* [35] showed the arrow keys to be the fastest and the letter keys the slowest; Beaumont [3] demonstrated the keypad to be faster than a keyboard. Both results showed that limited possibilities aid the novice user: the arrow keys were only up and down, the letter keys could range over the entire alphabet, and a keypad contained fewer keys than the keyboard.

Inconsistency appeared when keyboard and nonkeyboard devices were compared. The case of the touch screen and the keyboard mentioned above indicated that the comparison is highly task and user dependent. When either the mouse or light pen was compared to keyboard, the results were not clear. The type of task may be a factor in explaining the apparently inconsistent results between the mouse and other devices (cf. Card *et al.* [8] and Haller *et al.* [16] versus Karat *et al.* [18] and the study presented here. Note that the difference in users bears equal weight with that of task in explaining the results. Even though the mouse is quickly learned, novice users do not perform better or prefer an unfamiliar device, as shown in the study by Maclean *et al.* [21].

A comparative study by Ewing *et al.* [13]—in which these authors participated—examined the use of a ball mouse versus cursor or jump-arrow keys in the text selection environment of Hyperties (formally called TIES—The Interactive Encyclopedia System) (see later discussion). Performance and preference were the measure for the devices used in two types of tasks. Performance was measured by the time interval between target selection; preference was determined by a subjective evaluation given after completing the task.

Table 1. Comparative studies of selection devices

Study	Year	Device(s)	Results: speed
Ostroff	1986	KC, KN, MB, TS	TS > KC > KN > MB
Ewing *et al.*	1986	KC, MB	KC > MB
Maclean *et al.*	1985	KB, MB	KB > MB
Beaumont	1985	KB, KP, LP, TS	TS > KP > KB > LP
Shinar *et al.*	1985	KC, KL, KN	KC > KN > KL
Reinhart	1985	J, MB	MB > J
Haller *et al.*	1984	LP, GT, MB, TB, KC, VI	LP > GT > MB > TB > KC > VI
Karat *et al.*	1984	TS, MO, KL	TS > KL > MO
Whitfield *et al.*	1983	TS, TP, TP/K	TS > TP/K > TP
Albert	1982	DT, JF, JP, KC, LP, TB, TS	TS > LP > DT > TB > JF > JP > KC
Card *et al.*	1978	JP, KC, KT, MB	MB > JP > KT > KC
Goodwin	1975	KC, LP, LG	LP = LG > KC
Mehr and Mehr	1972	JP, TB	TB > JP
English	1967	LP, MB, JP	MB > JP > LP

Study	Results: accuracy	subjective evaluation
Ostroff	MB > KN > KC > TS	TS > KN > KC > MB
Ewing *et al.*	KC > MB	KC > MB
Maclean *et al.*	KB > MB	
Reinhart	MB > J	
Haller *et al.*	LP = KC > VI > GT > MB > TB	LP > MB > TB > CK > GT > VI
Karat *et al.*		KL > TS > MO
Albert	TB > DT > JF > JP > KB > LP > TS	
Card *et al.*	MB > KT > JP > KC	

Key:
DT -Data Tablet with puck
GT -Graphics Tablet
J -Joystick
JF -Force Joystick
JP -Position Joystick
KC -Cursor Keys
KL -Letter Keys
KN -Number Keys
KP -Key Pad
KT -Text Keys
LG -Light Gun
LP -Light Pen
MB-Mouse with ball
MO-Optical Mouse
TB -Trackball
TP -Touch Pad
TP/K-Touch Pad with key press
TS -Touch Screen

In part one of the experiment, the subjects rated the device according to comfort, control, enjoyment, frequency of looking at the device when manipulating, and ease of positioning. In part two, the subjects chose one device and were requested to explain why. Part three was allotted for comments.

For using the arrow-jump keys or the mouse, the task was to move a highlight bar (selector box) onto the target and to proceed until the path—a series of 28 targets—was completed. Besides device, the other condition was the type of path: short (targets that were on average less than two items away) and long (greater than two items). Note that distance is in terms of number of intervening items as experienced when using the arrow-jump keys, since the highlight bar was restricted to selectable items; the motion of the mouse cursor was continuous and unrestricted over the screen area.

Results showed that the arrow-jump keys were faster and preferred over the mouse, at the 5% level. The effect of target distance on the arrow-jump keys was significant, again at the 5% level, but not for the mouse. For the arrow-jump keys, where the selector box movement is discrete, distance is a function of how many items intervene, whereas the continuous motion of the mouse cursor, although also a function of distance, obviates having to traverse the intervening items. Although the fluid motion should result in faster time, contradicting the finding of the study, the unfamiliarity of the mouse and its implementation explains its lower rating. This is further confirmed by the subjective results, which showed consistent preference to the arrow-jump keys, with 88% of the participants favoring this strategy over the mouse.

The configuration of users, tasks, devices, and strategies in the previous studies reveals a complex interaction of issues. Judgment of selection tactics is most effective with a fixed task, homogeneous user population, and stated goals (speed, learnability, error rates).

2. THE STUDY

2.1 *Introduction*

This study compared four devices for selecting highlighted words or phrases in a textual database, Participants were asked to perform a series of selections using a touch screen, mouse, arrow-jump keys, and number keys. Performance was measured by time and number of errors. In addition, the subjective evaluation requested the participants to rate each device according to speed, comfort, and accuracy. They were also asked to rank the devices relative to each other. Data on the participant's age, sex, and computer usage was gathered and compiled. Oral and written comments were welcomed. Forty-six participants completed the subjective evaluation; data for the performance came from 24 participants.

2.2 *Method*

A repeated measures design with four treatment levels was used to measure the differences among the selection devices. To measure statistical significance, a repeated analysis of variance was employed. Tukey's Honestly Significant Difference was applied to determine pairwise significance for those effects that had a large enough F value.

2.2.1 *Participants.* There were a total of 46 participants, 24 of whom completed all four paths. Participants were undergraduate and graduate students drawn from history, education, and library courses and the psychology subject pool. Data about the individual's background were collected but not correlated to the timing results or to the subjective questionnaire.

Of the 46 participants who filled out the evaluation, 20 were female, 26 male. As a group, the average age was 22.33 with a standard deviation of 5.60. There was no significant difference between the sexes according to age.

As for computer experience, participants were asked to choose among six possibilities: once a day, at least three times a week, once a month, few times a year, and never. Scoring a once a day as 1 and never as 6, the average and standard deviation of frequency of computer use was 3.78 (1.55) overall. As expected of college students, participants, on average, used a computer between once a week and once a month. The proximity of the average to the median coupled with a small standard deviation (1.55) is encouraging since it shows that the participants, as a whole, were neither so expert as to render the task trivial or so naïve as to make computer anxiety a significant factor.

2.2.2 *Materials.*

System Configuration. The experiment was conducted on an IBM PC with 384K memory and two floppy disk drives. The IBM PC was interfaced with a Sony Trinitron monitor, a Carroll Touch Technology controller and panel, and a Microsoft mouse. In an IBM PC expansion slot was a gl512 graphics board from On-Line Products, Inc., which controlled the image on the screen. Mounted onto this gl512 board was the board controlling the touch panel. Another slot was occupied by the board controlling the mouse.

Software: Hyperties. The study was run on an early version of Hyperties (formerly, The Interactive Encyclopedia System). Hyperties (available from Cognetics Corp., 55 Princeton-Hightstown Rd., Princeton Junction, NJ 08550) is a software tool for authoring and browsing a group of interrelated articles on a given topic or theme. It was designed for a museum environment with the casual user in mind. Users select terms that are of interest and a brief definition appears at the bottom of the screen. If there is full text for the selected term, an option to read the complete article is presented. The top of the screen contains the header, identifying the article, along with the current and total number of pages. A user of Hyperties may page forward or backward, return to the previous article if possible, or exit the system. Figure 1 shows a sample screen from the study.

All selectable items, whether words or phrases, appeared in yellow on the screen. Each article had a heading that appeared in gray characters at the top of the screen, along with current and total number of pages, also in gray. The remaining text was displayed in white.

The conception for Hyperties is due to Shneiderman with Ostroff. It was designed by Shneiderman, Morariu, and Ostroff, and the various versions have been implemented by Ostroff. The software for Hyperties in the Ewing *et al.* study was written in STSC APL. For the second study, Hyperties had been rewritten in DeSmet C.

Screen description. The size of the visible screen on the Sony Trinitron Monitor was 9 in. wide and 7 in. high. The resolution of the monitor was 512 pixels for both width and length. The screen allowed for a maximum of 65 characters per row and 30 rows. Due to the limitations of the touch panel and the screen, only 24 rows and 55 characters were used. The background color was blue, with white, yellow, or gray lettering as indicated above.

Features of the selection devices

Arrow-jump keys. The arrow-jump keys, also known as cursor control, selector, or jump keys, are located on the numeric keypad to the right side of the IBM keyboard. Only the left and right arrows were used.

The arrow-jump keys caused a selector box, outlined in yellow, to move among the selectable items. When the selector box was on the desired target, the RETURN key was

PLACE: AUSTRIA PAGE 1 OF 4

Austria holds a special place in the history of the **Holocaust**. Straddled between Eastern and Western Europe, possessing a vibrant and culturally creative Jewish community on the eve of World War II, Austria had also provided the young **Adolf Hitler**, himself an Austrian raised near **Linz**, with important lessons in the political uses of **antisemitism**.

ADOLF HITLER—Nazi Dictator of Germany 1933–45; sought the "Final Solution to the Jewish Problem"	**NEXT PAGE**
FULL ARTICLE ON HITLER, ADOLF	
RETURN TO ASSIMILATION	**EXIT**

Fig. 1. Sample screen used in the study. Selectable items that appear here in boldface were in yellow on the monitor.

pressed to complete the selection. The selector box wrapped around the screen whenever the top-leftmost or bottom-rightmost option was passed. The arrow-jump keys were labeled for the two directions in which the selector box could move, and the RETURN key had a diamond prominently drawn on top of it to facilitate its use.

Mouse—"Jump" or "Blind"? The Microsoft Mouse is 2 in. wide, 3.5 in. long and 1.25 in. high, a little taller than a deck of cards. Two green oblong buttons are on the front of the mouse. The surface area for moving the mouse was approximately 4 sq ft on a formica desk top.

The particular implementation of the jump-mouse was complicated by the use of a monitor and controlling board that did not use display memory of the PC. Instead, a mapping between the 25 × 80 characters or 200 × 640 pixels on the IBM display monitor and the 523 × 512 pixels on the Sony Trinitron monitor was required. Any cursor on the Sony had to be software emulated.

Due to this mapping, the selector box appeared around the item that was closest to the hardware cursor, after translating it to the 512 by 512 coordinates. There was no cursor on the Sony; instead, the selector box appeared to "jump" from one selectable item to another, hence the term "jump" mouse. When the selector box was on the target, a click of either mouse button selected the item. Advocates of this approach suggest that seeing the selector box and not the cursor avoids confusion.

Unfortunately, the location of the cursor was often distant from the selectable item, and this disparity caused frustration for the user in navigating the selector box. Moreover, if a definition of an article was selected, the selector box would then move from that reference to appear on the option FULL ARTICLE, but the true location of the mouse cursor remained on the reference. Once the jump-mouse was moved again, the selector box would "jump" off the FULL ARTICLE and move to the item closest to the cursor location. The lack of precise knowledge of the cursor position, the only feedback being the location of the selector box, caused one observer to term this implementation, appropriately, a "blind" mouse. The nature of the "blind" mouse was pointed out to participants during the practice period.

Number keys. The number keys were located along the top row of the keyboard. To avoid confusion with the arrow-jump keys, the numeric keypad located on the right was not used.

The highlighted options were shown on the screen with numerical superscripts. When an appropriate number was entered, a yellow selector box appeared around that item. Pressing the RETURN key effected the selection, turning the yellow selector box red. If a number were entered incorrectly, pressing the Escape (ESC) key removed the selector box, and a new choice could be made.

If there were more than nine selectable items on the screen, the following protocol was used. Entering the first digit would cause the corresponding item to be boxed in yellow. If another digit were entered, the selector box would move to the new number. Thus, pressing the "1" would cause the selector box to appear in yellow around the first item. If the "1" were followed by a "0" and there were 10 items, the yellow box would be removed from the first item and appear around the tenth. If there were not at least 10 selectable items, then the selector box would disappear from the screen. Pressing any illegal combination, or Escape, resulted in clearing the selector box.

Touch Screen. A Carroll Touch Technology controller box and panel was used. This panel, mounted on the Sony Trinitron Monitor, is an optomatrix frame of 48 × 36 discrete points whose resolution is 0.20 in. × 0.20 in. (0.51 cm × 0.51 cm). It consists of infrared light-emitting diodes (LEDs) and phototransistor detectors. When the point in the grid of infrared light beams is broken, the coordinates are transmitted to the CPU.

For each selectable item, a corresponding touch zone was established. If the touch was within this set of defined zones, the appropriate highlighted option was boxed in red, the selection having been effected. When the computer processing was completed, the red box was removed and a new selection was ready to be made. If the touch was outside this set of zones, nothing would happen.

As with many touch screens, the lack of feedback, the coarse grid, and the parallax problems meant users had to wiggle their fingers to produce a selection, and they often hit the wrong selection. Newer touch screens strategies are far more effective [30].

Paths. A path is defined to be a series of selections or a list of target items. Four paths with 37 selections each were printed on paper and presented to the participant. Each path was designed to be equivalent in time yet different in target items. In terms of time, they were equal by the total number of selections, of page turns (defined to be either a NEXT PAGE or BACK PAGE), of total pages formatted, of definitions appearing, and the number of article accesses.

2.3 *Experimental procedure*

The task was to complete a path traversal, 37 target selections, for each device.

2.3.1 *Administration.* The study was conducted in a quiet room where participants were given an introduction to the interactive encyclopedia. They were allowed to practice on each one of the four selection devices until they felt comfortable, typically 10 to 15 minutes. Participants were tested individually.

After the practice session, the participants were presented with the first of the four paths. Path 1 was always given first regardless of the selection device and so on for the remaining three paths. All the permutations were presented for the order of devices.

During the traversal of the path, if the participant lost track of the path or of how to get back, he or she was guided by the experimenter back to the correct path. For this reason, the error rate was not measured according to the number of incorrect screens per path, but by the number of times the participant strayed from the path. Therefore, the time spent returning to the path was not counted.

After the four paths were completed, the participants filled out the subjective evaluation form and were interviewed.

2.3.2. *Grading.* For all screens viewed, the time per screen and the selection made were logged onto disk. The time for each selection was the time between appearances of the screens until the end of the path. The time, therefore, included the selection time and the execution of that selection. The times for the correct screens were ten averaged over the path. The mean time per correct screen was the statistic used to measure speed of performance.

Analysis of the response times proceeded along three lines: target location, number of items, and distance to the target. The target location imparts semantic meaning within TIES. Potential targets were divided into three types: page turn (NEXT PAGE or BACK PAGE), definitions, and article accesses (FULL ARTICLE ON or RETURN TO). The system response time was shortest for page turning and longest for an article access.

Screens were grouped by the number of selectable items: 3 to 5 items, 6 to 8 items, and 9 to 23 items. They were also grouped according to distance to the target. Distance to a target was defined to be the number of intervening items between the default location of the selector box and the target item.

Error count was based on the number of times the participant strayed from the path. This measure did not capture how far off the participant strayed but the propensity to stray.

The evaluation sheets had a scale of 1 to 7 for each attribute: speed, comfort, and accuracy. The wording for accuracy was in terms of fewer or more errors. Desirable attributes, such as fastest, most comfortable and fewest errors, were given the value 7, while undesirable attributes rated a one. Computer usage was on a scale of 1 to 6, the lower the number the more experience. The ranking of the devices restricted the range from one to four.

2.4. *Results*

2.4.1. *Performance measurements.*

Time. Time was measured as the mean time per correct screen per path. The touch screen was the fastest device with a mean time of 5.30 secs, a STD of 0.86, and the arrow-jump keys were next, followed by the number keys and the jump-mouse (Table 2).

Table 2. Performance results: Mean time per correct screen (seconds) (with standard deviations)

	Arrow keys	Jump-mouse	Number keys	Touch screen
Time (secs)	6.23 (1.70)	6.86 (1.48)	6.70 (1.41)	5.30 (0.86)

Table 3. Analysis of variance for mean time per correct screen

	SS	DF	MS	F
Total	222.77	95		
Device	35.58	3	11.86	18.04*
Subject	141.82	23	6.17	9.30
Residual	45.37	69	0.66	

*Significant $p < .01$.

Table 4. Pairwise comparison for all screens. HSD coefficient is 3.741. HSD Value is .619 (* indicates significance)

Device	Jump-mouse	Number keys	Touch screen
Arrow	0.637*	0.473	0.928*
Mouse		0.164	1.565*
Number			1.401*

Table 5. Performance results mean time per correct screen (seconds) by target location as given by type of screen (with standard deviations)

	Arrow keys	Jump-mouse	Number keys	Touch screen
Page turn	6.18 (1.69)	6.59 (0.95)	6.66 (1.57)	5.18 (1.06)
Definition	7.00 (2.00)	7.89 (3.00)	6.78 (1.82)	5.20 (0.95)
Article access	5.15 (1.76)	5.75 (1.46)	6.45 (1.64)	5.53 (1.15)

The main effect for devices ($F(3,69) = 18.04$) was significant at the $p < 0.1$ level. With the touch screen faster, the HSD between the touch screen and the jump-mouse was the greatest (1.57) followed by the value between the touch screen and the number keys (1.40) and between the touch screen and arrow-jump keys. The arrow-jump keys were significantly faster than the jump-mouse (Tables 3 and 4).

Target Location. Table 5 displays the times for the correct screens for each screen type. Except for the touch screen, accessing an article was fastest and accessing a definition was slowest; page turn was the fastest for the touch screen and article access the slowest; and the times for the touch screen and number keys were more consistent over type of screen.

At the 1% level, there was a significant F value for device, location, and their interaction (Table 6). Inspection of the cursor device shows best the interaction between target location and device. Since definitions are scattered throughout the top portion of the screen, there is always more search time. The default location of the cursor had a greater effect on the time. For article accesses, the cursor was at most one item away from the tar-

Table 6. Analysis of variance target location as given by type of screen

	SS	DF	MS	*F*
Blocks	438.02	23	19.04	
Treatments	194.04	11		
Device	93.10	3	31.03	21.10*
Location	48.81	2	24.40	16.59*
Interaction	52.14	6	8.70	5.91*
Residual	372.17	253	1.47	
Total	1004.23	287		

*$p < .01$.

Table 7. Pairwise comparisons target location as given by type of screen. HSD coefficient is 3.741 (* indicates significance)

Type of screen	Device	Jump-mouse	Number keys	Touch screen
All	Arrow-jump keys	.637	.473	.928*
Page		.412	.483	.994*
Def		.889	.222	1.802*
Article		.601	1.300*	.377
All	Mouse		.164	1.565*
Page			.072	1.406*
Def			1.111	2.692*
Article			.699	.224
All	Number keys			1.401*
Page				1.477*
Def				1.581*
Article				.923*

get. The response times for the noncursor positioning tactics remained fairly constant across target locations. Although the effect of target location and its interaction with the devices were shown to be significant, the significance is due to the distance to the target rather than any additional scanning time by the user or processing time by the computer.

The touch screen was significantly faster than all other devices for all types of screens, except for the two cursor positioning devices—the jump-mouse and arrow-jump keys—when accessing an article. The results of the analysis of variance for each type of screen were significant at $p < 0.01$. For page turn, $F(3,69) = 14.12$, for definitions, $F(3,69) = 10.26$, and for article access $F(3,69) = 7.46$. The corresponding HSD values were .677, 1.309, and .749 (Table 7).

Number of Items. We expect that response time increases with the number of items, as the reader has more items to scan. Within the confines of this study, the number of options may not have varied enough to draw conclusions. Nevertheless, the pattern for all the devices seems counterintuitive: as the number of items increased, the response times decreased, but not dramatically, except for the touch screen. For the group of screens with the largest number of items, the touch screen had the lowest response time for all other screen groups and all other devices. Moreover, the screens with only 3 to 5 items on them had a relatively high value for this device. The explanation may have little to do with the number of items on the screen, but the physical location of the target on the screen. For the screens with few items, the target was likely to be at the bottom of the screen, where most touches were ineffective. The screens with 9 to 12 items had a higher probability of the target being somewhere in the middle of the screen, a location where there are fewer misses (Table 8).

The results of an analysis of variance to test for device and number of targets (Table 9) showed significance at the 1% level for the effect of both these factors, but their interaction was not significant. Unlike analysis by target location or distance, both of which showed a significant interaction effect, the number of items on a screen is a func-

Table 8. Performance results mean time per correct screen (seconds) by number of items (with standard deviations)

Number of items	Arrow keys	Jump-mouse	Number keys	Touch screen
3–5 items	6.98	7.00	7.70	6.20
	(1.92)	(1.27)	(1.53)	(1.14)
6–8 items	6.18	6.91	6.77	5.39
	(1.47)	(1.69)	(1.44)	(0.95)
9–12 items	6.04	6.79	6.13	4.69
	(2.23)	(1.99)	(1.95)	(1.19)

Table 9. Analysis of variance number of items on the screen

	SS	DF	MS	*F*
Blocks	412.49	23	17.93	12.47
Treatments	140.42	11		
Device	105.20	3	35.07	24.38*
Number	22.26	2	11.11	7.72*
Interaction	13.00	6	2.17	1.51
Residual	363.83	253	1.47	
Total	916.73	287		

*$p < .01$.

Table 10. Performance results mean time per correct screen by distance to target (with standard deviations)

Distance in items	Arrow keys	Jump-mouse	Number keys	Touch screen
Zero items	5.14	5.45	6.56	5.43
	(1.71)	(1.46)	(1.58)	(1.03)
One item	6.08	6.81	6.52	5.12
	(1.65)	(1.29)	(1.54)	(0.93)
Two items	7.28	7.78	7.10	5.41
	(2.06)	(2.46)	(1.76)	(1.04)
Three+ items	8.06	9.50	7.21	5.87
	(2.69)	(4.12)	(2.70)	(1.48)

tion of the user's perception and the ability to scan and match. Since that process must occur for all devices, it is not surprising to find no significant effect of interaction.

Distance to Target. Upon inspection of Table 10, the difference between the direct pointing and cursor positioning tactics illustrates interaction between device and distance: response time for the touch screen did not change as the distance increased whereas for the arrow-jump keys and the jump-mouse the response time noticeably increased. By the same logic, the number keys did not show much change over distance, because there is no default cursor location. In fact, there was a slight increase in response time for the number keys as the number of items increased. The effect of distance, device and their interaction was significant at the 1% level (Table 11). The data in Table 11 show the sharp rise in time as the distance increased for the cursor positioning devices but that distance probably has little or no effect on command or direct pointing strategies.

Nevertheless, the effect of the device when the cursor defaulted to the target item was significant, $F(3,69) = 10.44$. For a distance of one item, $F(3,69) = 18.90$, for two items, $F(3,69) = 10.45$, and three or more items, $F(3,69) = 7.96$, for $p < .01$. Applying this to a pairwise comparison, as shown in Table 12, the HSD values are .727, .637, 1.189, and 2.018, for distances of zero through three plus, respectively.

Table 11. Analysis of variance distance to the target

	SS	DF	MS	*F*
Blocks	660.09	23	28.70	11.19
Treatments	534.70	11		
Device	223.97	3	74.66	29.10*
Distance	190.48	3	63.49	24.75*
Interaction	120.25	9	13.36	5.21*
Residual	885.12	345	2.57	
Total	2079.91	383		

*$p < .01$.

Table 12. Pairwise comparisons for distance to target. HSD coefficient is 3.741 (* indicates significance)

Number of items	Device	Jump-mouse	Number keys	Touch screen
Zero	Arrow-jump keys	.307	1.422*	.289
One		.732	.444	.961*
Two		.504	.176	1.863*
Three +		1.438	.849	2.193*
Zero	Jump-mouse		1.115*	0.019
One			0.288	1.693*
Two			0.681	2.367*
Three +			2.287*	3.630*
Zero	Number keys			1.133*
One				1.405*
Two				1.687*
Three +				1.344

Table 13. Average number of errors per path (with standard deviations)

Arrow keys	Jump-mouse	Number keys	Touch screen
0.83 (1.03)	.71 (.89)	.79 (1.08)	1.33 (1.11)

Errors. Since the *F* value for the errors was not significant even at the 5% level, a pairwise comparison would be meaningless. Looking at the results, the order is exactly the reverse of the response times. Although not significantly different, there is a clear implication of the speed–accuracy trade-off (Table 13).

2.4.2. *Subjective Evaluation.* The subjective evaluation was calculated from the responses of 46 subjects, 26 male and 20 female. Repeated analysis of variance consistently showed significance due to device. Results of the subjective evaluation showed that the touch screen was rated the best overall device, although not with respect to accuracy. The number keys were rated the most accurate, and the jump-mouse, due to insufficient feedback, was rated the slowest, most uncomfortable, and least accurate and it ranked last. Table 14 shows the main results from the subjective evaluation, and Table 15 shows the pairwise comparison. The following three sections refer to the data presented in these tables.

The analysis of the subjective evaluation did not consider the interaction among the various attributes (e.g., the speed–accuracy trade-off). Instead, the effect of the device was measured for speed, comfort, and accuracy. The results for the above attributes were significant at the 1% level: $F(3,135) = 23.48$, 31.28, and 21.02, respectively. The calculated HSD values were: for speed .975, for comfort 1.010, and for accuracy 1.143.

Speed. The touch screen was rated as the fastest device, and the number keys were rated faster than the arrow-jump keys. Also note that the cursor positioning devices were rated lower than either the touch screen or the number keys.

Table 14. Results of the subjective evaluation (with standard deviations)

Device	Speed	Comfort	Accuracy
Arrow keys	4.80 (1.39)	5.24 (1.29)	4.96 (1.75)
Jump-mouse	3.37 (1.54)	3.04 (1.85)	2.94 (1.66)
Number keys	5.09 (1.59)	5.15 (1.62)	5.78 (1.61)
Touch screen	5.98 (1.29)	6.09 (1.23)	4.83 (1.94)

Rated on a scale 1–7, 7 being the best.

Table 15. Pairwise comparison for subjective evaluation (* indicates significance)

Number of items	Device	Jump-mouse	Number keys	Touch screen
Speed	Arrow-jump keys	1.435*	.283	1.174*
Comfort		2.196*	.087	.848
Accuracy		2.000*	.848	.109
Speed	Jump-mouse		1.717*	2.609*
Comfort			2.109*	3.043*
Accuracy			2.848*	1.891*
Speed	Number keys			.891
Comfort				.935
Accuracy				.957

Performing a RANOVA on the participants' rating of the devices according to speed resulted in an $F(3,135) = 23.48$ for all subjects. Applying the overall within-group variance of 2.30 and the HSD coefficient of 4.36, the Honestly Significant Difference was .975. Four sets of comparisons were significant by the HSD test: the jump-mouse and the touch screen, 2.61, the jump-mouse and the number keys, 1.72, arrow-jump keys and the jump-mouse, 1.44, arrow-jump keys and the touch screen, 1.17. The jump-mouse was significantly slower than the other devices. In retrospect, considering the jump-mouse implementation, this was not surprising. The arrow-jump keys were also significantly slower than the touch screen with respect to speed. The impression left by the discrete movement of the selector box on the participants helps explain why the arrow-jump keys were perceived to be slower.

Comfort. The touch screen again was rated as the most comfortable device. The keyboard tactics were similarly judged, the arrow-jump keys rated slightly higher than the number keys. The jump-mouse with a score of 3.04 (1.85) was significantly slower than the other devices.

The RANOVA on the comfort rating of the devices resulted in a larger F value than either speed or accuracy, $F(3,135) = 31.28$. With a coefficient of 4.36 and within group variance of 2.47, the HSD is 1.01. There was a significant difference in the comparison between the jump-mouse and the touch screen 3.04, the jump-mouse and the arrow-jump keys 2.20, and jump-mouse and number keys 2.11: the jump-mouse was rated less comfortable (Table 14).

Accuracy. Participants were asked to rate the devices according to accuracy by responding to the phrase "least likely to cause errors." The number keys were clearly rated the most accurate device followed by the arrow keys and touch screen. The standard deviation of the rating of the touch screen was the highest of all devices, probably a reflection of a mixed attitude toward the touch screen. Having rated it fast and comfortable, people may have carried over that positive disposition to the question of accuracy. The jump-mouse again was judged the poorest with a 2.94 rating for accuracy, even though the data show it to be relatively the most accurate. Again, confusion in using the jump-mouse may have caused personal bias against it.

The HSD for accuracy was 1.14, given the coefficient of 4.36 and the within-group variance of 3.16. The calculated F value for all participants was $F(3,135) = 21.02$. The pairwise comparisons again showed the jump-mouse significantly slower, at the 5% level, than the other three devices: with numbers 2.85, arrow-jump keys 2.00, and touch screen 1.89. It is interesting to note that the difference between the number keys and the touch screen, .96, and between the number keys and the arrow-jump keys, .85, were almost significant.

Ranking. The participants were given four blanks on which the devices were to be ranked (Tables 16–18). A low ranking indicated preference for that device. The touch screen received an overall ranking of 1.80, whereas the arrow-jump keys edged out the number keys with a ranking of 2.33 to 2.37. With an $F(3,135) = 19.86$ and an HSD of .698, figured from a within-group variance of 1.18 and an HSD coefficient 4.36, again the jump-mouse's ranking was significantly poorer ($p < .05$) from that of the other devices. The largest difference was that between the jump-mouse and the touch screen 1.70, followed by the difference between the jump-mouse and the arrow-jump keys 1.17 and number keys 1.13.

Most people found the arrow-jump and number keys similarly appealing. People with more computer usage favored the keyboard selection mechanisms. Infrequent computer users tended to be awed by the technology of both the touch screen and jump-mouse.

Participants' comments. The written comments provided at the end of the evaluation form well describe the participants' attitude toward the four selection tactics. The touch screen was seen as comfortable and fast but sometimes unreliable: "Touch was very convenient," but another wrote that it was "unstable at screen extremities." "Touch screen is the easiest, but can be frustrating when it does not work." "I found the touch system extremely easy to follow, especially for a person who never uses a computer."

The jump-mouse was resoundingly deplored, but a sympathetic note occasionally appeared. "I had frustration when I was using [the] mouse, because I couldn't see the cursor movement." "The mouse was awkward and difficult to control." "The mouse definitely

Table 16. Ranking of devices on a scale of 1–4, where 1 is the "best" (with standard deviations)

Arrow keys	Jump-mouse	Number keys	Touch screen
2.33 (0.84)	3.50 (.83)	2.37 (1.07)	1.80 (.97)

Table 17. Analysis of variance for ranking of devices

	SS	DF	MS	F
Total	230.00	183		
Device	70.43	3	23.48	19.86
Subject	.00	45		
Residual	159.57	135	1.18	

Table 18. Pairwise comparisons ranking of the devices. HSD coefficient is 4.36. HSD value is .698 (* indicates significance)

Device	Jump-mouse	Number keys	Touch screen
Arrow	1.174*	.043	.522
Mouse		1.130*	1.696*
Number			.565

moved too quickly . . ." "If the mouse can be improved in a few ways it would be as adequate as the touch." "Mouse key is cute but not easy to control."

The arrow-jump keys and the number keys were often compared, and, as the performance and preference data show, the arrow keys were regarded as faster but less accurate than the number keys. "The [arrow] Jump keys and Number keys are so similar it's hard to objectively differentiate their qualities; however, the numbers made it easier to spot and reduce chance for error." "The number [keys] took too much time." The arrow-jump keys had "buttons too close together—hit wrong one often."

The number keys were pointed out to have positive and negative traits, and suggestions were given: "I thought the number keys were most effective . . . By using the number keys there was less chance of error." "Numbers clutter up the article." "If the number keys for the NEXT PAGE and other functions [items] that appear with frequency had a constant symbol, this method would be more comfortable."

Even though most participants felt that the experiment as a whole was "good and interesting," one participant's comment strikes to the heart of the matter by suggesting that "perhaps a timed experiment is not the situation under which most people would want to use such an encyclopedia."

2.5 *Discussion*

2.5.1 *Subjective evaluation.* The discrete movement likely explains the user's perception that the jump-mouse is slower. The user's perception may explain why the touch screen was not significantly different from the number keys. Both tactics are used without an intervening selector box and so did not have the problems associated with its manipulation. The surprising high rating for the number keys, although not significantly different from the other devices, can be explained by the feeling of having mentally grasped the command, which gives the user a sense of control.

2.5.2 *Location of target as given by type of screen.* Since the location of the item imparted semantic meaning, the time for selecting fixed place items when turning a page or accessing an article should be lower than for selecting a definition. For article access, however, the faster time of the preparation stage should be offset by the longer system execution time.

As can be seen from the data in Table 2, the conjecture that system execution time was a significant factor in total time cannot be true, since the article access had the lowest times for three of the devices. This may be explained by the fact that the cursor would always default to the article access, or be at most one item away. That the time taken for selection was the major factor in total time is further proven by the longer times for selecting a definition where the location was not constant and thus not readily apprehended.

The reason for the significant difference between the touch screen and number keys and between the arrow-jump keys and number keys is that the command strategy used has no benefit of a default cursor or direct pointing. For the arrow-jump and number keys, the difference is not based on location of target, but on distance to the target. In contrast to all the other devices that had, relative to themselves, the lowest response times for article access, the touch screen had the highest times. Here, the location is an explanation, for the article accesses were at the bottom of the screen, the most likely place for touching errors. Accordingly, there was no significant difference between the touch screen and the cursor pointing devices in contrast to the other target locations that showed significant differences between the touch and all the other devices.

2.5.3 *Distance to target.* The ability to go directly to an item explains why it was significantly different from other devices. The exception was when the selector box defaulted to the target item, in which case the arrow-jump keys and the jump-mouse required no cursor movement. Likewise, the speed of use for number keys was invariant with target distance, and significantly slower than the cursor movement devices when the distance to the target was zero.

2.5.4 *Touch screen.* The nature of the touch tactic shows why it would be the fastest and least accurate and is a good example of the trade-off between speed and accuracy. In this touch screen tactic, the latter was sacrificed for the former. The "directness" or "naturalness" of touching with a finger means that there is minimal switching of domains of operation. There is no locating a cursor or device. There is no input stage or confirmation; the only action is the pointing of the finger. In this respect, the benefit for novice users is clear, but the lack of feedback indicating the actual touch point does result in errors. Also, the selector box during the input is a visual clue to the item under selection; the touch screen version did not have this clue.

Arrow-jump keys and number keys. As Shinar *et al.* [35] showed, the user response times for the arrow-jump keys were expected to be fast because of the relatively few items per screen and general familiarity with step keys. The notion of one-to-one movement of cursor to keystrokes is not foreign to many, nor is the keyboard a strange device. The input stage for arrow-jump keys is likely to be slightly shorter than for the number keys since there are fewer arrow-jump keys. The user's hand could remain positioned on the arrow-jump keys so as to make homing time negligible, just as some participants kept the left hand on the number keys. From observation, most had to search for the number keys more often than for the arrow-jump keys.

Examining the difference between the two keyboard tactics, the cursor positioning task involves no recall, only recognition and reaction, whereas use of the number keys requires remembering the associated number after spotting the target, then entering. The answer seems to be that number keys exact a heavier cognitive load.

The speed–accuracy trade-off also comes into the picture. The number keys were rated more accurate and less error-prone than the arrow-jump keys, but the difference was not significant. Still, though the two tactics entail similar physical motions, they invoke different cognitive functions. The result is an edge in speed for the arrow-jump keys and likewise in accuracy for the number keys. Bear in mind, however, that the speed of the arrow-jump keys and likewise in accuracy for the number keys is dependent upon distance to the target, unlike the number keys.

Jump-mouse. The statistical tests show that the jump-mouse was the slowest device and the poorest rated. Although it was the most accurate, people were unaccustomed to the jump-mouse and therefore were more careful in manipulating it. One would then expect a corresponding increase in the response time.

The "blind-jump" mouse tactic illustrates a poor interface design because of the inability to know where the actual cursor is. There was uncertainty about the effect a move would have, so the participants proceeded slowly. The jump-mouse has the drawback of being an unfamiliar device, despite the prevalence of video games and assorted uses of other positioning devices.

2.6 *Suggested revisions.*

In light of the results from both studies, the arrow-jump keys were faster than and preferred over the jump-mouse, although the study by Card *et al.* [8] showed the opposite. The limited number of targets and the nature of the cursor movement with the jump-mouse may be reasons. Nevertheless, there are improvements to the selection tactics presented here. In the first study (Ewing *et al.* [13]), the person using the jump-mouse could move the cursor throughout the screen, but only when the cursor was on the target did the selector box appear. In the second study, the selector box appeared on the item closest to the cursor, but the actual cursor location was invisible. A suggested improvement, which has already been implemented, is to combine these two approaches: allow the selector box to always be on the screen, highlighting that item closest to the cursor as well as displaying the cursor. One of the problems in this revised jump-mouse tactic is the emulation of the cursor on the screen that is not connected to the display memory. The software has to grab the screen area, save it, display the cursor, and then replace it with the original contents. This process was too slow for the fluid motion of the cursor. Improvements in hardware and a faster implementation could easily dispel these drawbacks. The advantage of this new tactic is that the user is never outside the domain of selectable items, whereas the cursor motion is continuous and visible.

The infrared touch screen, although it was the preferred device, had several drawbacks, including detecting touches at the edge of the screen and inability to restart the selection process anew. A suggested improvement, which was also implemented, is the repeat-confirm touch. When the user touches at item, the selector box appears in yellow. If the same item is touched again, then the selector box turns red and the selection is effected. If another item were touched instead, the selector box would move from the previous item and appear in yellow around the item that is currently being touched. The clear drawback is the extra time to touch twice, but considering that touching is fairly rapid, the ability to catch errors before they are made is desirable. Another approach is to enable users to place their fingers on the screen and to give feedback on placement by showing a cursor. Users can then move the cursor around until it rests on the desired target. Removing the finger triggers the selection. This strategy results in a substantial reduction in errors (Potter *et al.* [30]).

Acknowledgements—The authors wish to thank the College of Library and Information Services at the University of Maryland, Dr. Janis Morariu of CIDE, Dr. Kent Norman of the Department of Psychology, and Wan-Lee Ngiam who assisted in the running of the experiment.

REFERENCES

1. Albert, A.E. The effect of graphic input devices on performance in a cursor positioning task. Proc. of Human Factors Society—26th Annual Meeting, 54-58; 1982.
2. Beaumont, J.G. System requirements for interactive testing. International Journal of Man Machine Studies, 17: 311-320; 1982.
3. Beaumont, J.G. Speed of response using keyboard and screen-based micro computer response media. International Journal Man-Machine Studies, 23: 61-70; 1985.
4. Brown, C.M. Human-computer interface guidelines. Norwood, NJ: Ablex Publishing Co.; 1987.
5. Brown, E.; Buxton, W.; Murtagh, K. Windows on tablets as a means of achieving virtual input devices. In ACM CHI '85 Tutorial Notes; 1985. Boston.
6. Buxton, W. There's more to interaction than meets the eye: Some issues in manual input. In ACM CHI '85 Tutorial Notes; 1985. Boston.
7. Buxton, W.; Baecker, R. Topics in manual input: Technologies, techniques, and tools. In ACM CHI '85 Tutorial Notes; 1985. Boston.
8. Card, S.K.; English, W.K.; Burr, B.J. Evaluation of mouse, rate-controlled isometric joystick, and text keys for text selection on a CRT. Ergonomics, 21(8): 601-613.
9. Card, S.K.; Moran, T.P. The keystroke-level model for user performance time with interactive systems. Communications of the ACM. 23(7): 396-409; July 1980.
10. Card, S.K.; Moran, T.P.; Newell, A. The Psychology of Human-Computer Interaction. Hillsdale, NJ: Lawrence Erlbaum Associates; 1983.
11. Embley, D.W.; Lan, N.T.; Leinbaugh, D.W.; Nagy, G.A. Procedure for predicting program editor performance, from the user's point of view. International Journal Man-Machine Studies. 10: 639-650; 1978.
12. English, W.K.; Engelbert, D.C.; Berman, M.L. Display-selection techniques for text manipulation. IEEE Transactions on Human Factors in Electronics. HFE-8(1): 21-31; 1967.
13. Ewing, J.; Mehrabanzad, S.; Scheck, S.; Ostroff, D.; Sheiderman, B. An experimental comparison of a mouse and arrow keys for an interactive encyclopedia. International Journal of Man-Machine Studies. 24(1); January 1986.
14. Goodwin, N.C. Cursor positioning on an electronic display using light-pen, lightgun, or keyboard for three basic tasks. Human Factors. 17(3): 289-295; June 1975.
15. Grudin, J.; Maclean, A. Adapting a psychophysical method to measure performance and preference tradeoffs in human-computer interaction. In: Shackel, B., ed., Proceedings of INTERACT '84, IFIP Conference on Human-Computer Interaction. pp. 338-342; 1984. London: IEE.
16. Haller, R.; Mutschler, H.; Voss, M. Comparison of input devices for correction of typing errors in office systems. INTERACT '84: 218-223; 1984. London.
17. Hutchins, E.; Holland, J.; Norman, D.A. Direct manipulation interfaces. In: Norman, D.A. and Draper, S.W., eds. User Centered System Design: New Perspectives on Human-Computer Interaction, pp. 87-124. Hillsdale, NJ: Lawrence Erlbaum Associates; 1986.
18. Karat, J.; McDonald, J.; Anderson, J. A comparison of selection techniques: Touch panel, mouse and keyboard. INTERACT '84: 149-153; 1984. London.
19. Koved, L.; Shneiderman, B. Embedded menus: Menu selection in context. Communications of the ACM; 29(4): 312-318; 1986.
20. Landauer, T.K.; Nachbar, D.W. Selection from alphabetic and numeric menu trees using a touch screen: Breadth, depth, and width. CHI Proceedings: 73-78; 1985.
21. Maclean, A.; Barnard, P.J.; Wilson, M.D. Evaluating the human interface of a data entry system: User choice and performance measures yield different tradeoff functions. In: Johnson, P., ed., People and Computers: Designing the Interface. Cambridge, England: Cambridge University Press; 1985.
22. Marteniuk, R.G. Information Processing in Motor Skills. New York: Holt, Rinehart and Winston; 1976.
23. Mehr, M.H.; Mehr, E. Manual digital processing in 2 axis: A comparison of joystick and trackball controls. Proceeding of the Human Factors Society—16th Annual Meeting; 1972. Santa Monica, CA.
24. Meyer, D.; Smith, J.E.K.; Wright, C.E. Models for the speed and accuracy of aimed movements. Psychology Review, 89(5): 449-479; 1982.

25. Norman, D.A. Design principles for human-computer interfaces. Proceedings of CHI 1983: Human Factors in Computer Systems, ACM: 1-10; 1983.
26. Norman, D.A. Design rules based on analyses of human error, Communications of the ACM, 26(4): 254-258; 1983.
27. Norman, D.A. Four stages of user activity. INTERACT '84: 81-85; 1984. London.
28. Pearson, G.; Weiser, M. Of moles and men. ACM CHI: 333-339; 1986. Boston.
29. Pfauth, M.; Priest, J. Person-computer interface using touch screen devices. Proceedings of the Human Factors Society—25th Annual Meeting: 500-504; 1981.
30. Potter, R.M.; Weldon, L.; Shneiderman, B. Increasing the accuracy of touchscreens: An experimental evaluation of three strategies. Proceedings of ACM CHI 88 Conference: 27-32; 1988.
31. Poulton, E.C. Tracking Skill and Manual Control. New York: Academic Press; 1974.
32. Reinhart, W.; Marken, R. Control systems analysis of computer pointing devices. Proceedings of the Human Factors Society—29th Annual Meeting: 119-121; 1985.
33. Rubenstein, R.; Hersh, H. The Human Factor: Designing Computer Systems for People. Burlington, MA: Digital Press; 1984.
34. Runyon, R.P.; Haber, A. Fundamentals of Behavioral Statistics, Fifth Edition. Reading, MA: Addison-Wesley Publishing; 1984.
35. Shinar, D.; Stern, H.; Bubis, G.; Ingram, D. The relative effectiveness of alternative selection strategies in menu driven computer programs. Proceedings of the Human Factors Society—29th Annual Meeting: 645-649; 1985.
36. Shneiderman, F. Designing menu selection systems. Journal of the American Society for Information Science, 37(2): 57-70; 1986.
37. Shneiderman, B. Direct manipulation: A step beyond programming languages. IEEE Computer, 16(8): 57-69; August 1983.
38. Shneiderman, B. Software Psychology: Human Factors in Computer and Information Systems, Boston, MA: Little, Brown and Co.; 1980.
39. Tyler, M. Touch screen: Big deal or no deal? DATAMATION: 146-154; January 1984.
40. Wells, M. Representing the user's model of an interactive system. INTERACT '84: 145-149; 1984. London.
41. Whitfield, D.; Ball, R.G.; Bird, J.M. Some comparisons of on-display and off-display touch input devices for interaction with computer generated displays. Ergonomics, 26: 1033-1053; 1983.