0743-4618/96/1203-0155 $3.00/0; Volume 12, September 1996
AAC Augmentative and Alternative Communication

Effect of a Word Prediction Feature on User Performance

Heidi Horstmann Koester and Simon P. Levine
Rehabilitation Engineering Program, Graduate Bioengineering Program, Department of Physical Medicine and Rehabilitation, University of Michigan, Ann Arbor, Michigan, USA

This study examines the influence of a word prediction feature on several measures of user performance during text entry tasks. Fourteen subjects transcribed text with and without a word prediction feature for seven test sessions. Eight subjects were able-bodied and used mouth-stick typing, while six subjects had high-level spinal cord injuries and used their usual method of keyboard access. Use of word prediction significantly decreased text generation rate for the spinal cord injured (SCI) subjects and only modestly enhanced it for the able-bodied (AB) subjects. While fewer selections were required with the word prediction system, each selection took significantly longer to make, which we attribute to the additional cognitive and perceptual processes associated with use of word prediction. Performance was further analyzed by deriving subjects' times for keypress and list search actions during word prediction use. All subjects had slower keypress times during word prediction use as compared to letters-only typing, and SCI subjects had much slower list search times than AB subjects. Under the conditions of this experiment, the cognitive cost of using word prediction largely overwhelmed the benefit provided by keystroke savings. SCI subjects appeared to incur higher cognitive costs than AB subjects, possibly due to their prior expertise in letters-only typing.

KEY WORDS: augmentative communication, evaluation, group designs, rate enhancement, word prediction

Augmentative and alternative communication (AAC) systems have the potential to improve the quality of life of individuals with severe disabilities by providing independent access to spoken communication, written communication, and/or standard computer software. However, the rate of communication achieved with current systems is far from optimal, and attempts to improve performance through system design often do not work as well in practice as they do in theory, due to complexities in the interaction between the user and the system. A better understanding of user-system interaction is needed for all AAC interfaces to improve the design and promote the recommendation of systems that best meet users' needs.

This paper focuses on word prediction in particular, as one example of a technique designed to enhance text generation rate in AAC. Past research shows that word prediction is successful in reducing the motor requirements for use, but usually at the cost of introducing additional cognitive and perceptual loads (Horstmann & Levine, 1991; Soede & Foulds, 1986). The goal of this work is to provide greater insights into how the trade-off between decreased motor and increased cognitive loads affects user performance with a word prediction system.

This goal is being pursued in two complementary ways: through empirical measurement of actual user performance under a range of conditions and by creating and validating quantitative models of user performance with AAC systems. This paper presents results from the empirical aspect of the project only, although the study was designed to address both the empirical and modeling goals. The empirical work adds to the data that have been published to date on user performance with word prediction and attempts to fit this data into a theoretical framework. A subsequent paper will focus on the modeling aspect of the project.

BACKGROUND

A major problem with AAC systems continues to be the slow speed of text generation. For users with severe physical disabilities, reported rates for current systems are well below those achieved by unimpaired typists and speakers, commonly cited at 35 to 40 words/minute (wpm) for typing and 150 to 180 wpm for speaking (Foulds, 1980). For example, well-trained users of two-switch Morse code systems can achieve text generation rates of approximately 12 wpm (Levine, Gauger, Bowers, & Khan, 1986), while novice users may reach only 2 wpm (Beukelman, Yorkston, & Dowden, 1985). Reported rates for letter-based scanning systems range from under 1 wpm to about 7 wpm (Goodenough-Trepagnier, Rosen, & Demsetz, 1982; Horstmann, 1990; Koester and Levine, 1994).

A variety of methods have been devised in the attempt to enhance communication rate. While all of these methods can significantly improve efficiency, as measured by keystroke savings, their effect on overall text generation rate is less clear. One common technique is the use of fixed vocabulary sets of words or syllables, either arranged in hierarchical menus or on one large display (Chopra, Thomason, Kazemi, & Alguindigue, 1988; Goodenough-Trepagnier, Tarry, & Prather, 1982). While estimates of keystroke savings with menu systems are as high as 50% (Pollak & Gallagher, 1989), performance with menu systems has been observed to be only the same as, and sometimes slower than, letter-by-letter spelling (Deroost & Soede, 1989; Goodenough-Trepagnier, Rosen, & Galdieri, 1986; Rowley, 1989).

A second technique is encoding, which involves defining and memorizing abbreviations for a set of words, phrases, and/or sentences, using alphanumeric codes (e.g., Vanderheiden & Kelso, 1987) or graphic symbols (e.g., Baker, 1987). For alphanumeric codes, an estimated 20 to 50% of keystrokes can be saved by abbreviating 200 words (Vanderheiden & Kelso, 1987), but the net effect on rate is unknown, since a comparison of performance with and without alphanumeric codes has not been published, to our knowledge. There has been one comparison of a symbolic encoding system to letter-by-letter spelling, in which the group using symbolic encoding entered text 28% faster than the group using letters only by the conclusion of the 6-week study (Gardner-Bonneau & Schwartz, 1989). While this is a notable improvement, it is considerably less than the more than doubling of speed predicted based on keystroke savings alone (Baker, 1987).

A third technique, called word prediction, takes advantage of the redundancy in the English language to predict a set of words based on the letters entered by the user. Word prediction choices are typically displayed in a short list and are refined as the user inputs additional letters (Bentrup, 1987; Heckathorne, Voda, & Leibowitz, 1987; Swiffin, Arnott, Pickering, & Newell, 1987). Large reductions in inputs (or keystrokes) required per word can be achieved, since many words can be completed by selection from the list rather than by single-letter spelling. However, as with menu and encoding systems, keystroke savings alone does not always yield a significant improvement in rate. In 15 case reports, across several studies and systems, the median keystroke savings was 38% (range 23–58%), while the median improvement in text generation rate was 4% (range −21–162%) (Gibler & Childress, 1982; Koester & Levine, 1994; Newell, Arnott, & Waller, 1992; Newell, Booth, & Beattie, 1991; Scull & Hill, 1988; Venkatagiri, 1993). While some users enjoyed substantial improvement relative to letter-by-letter spelling, others improved only marginally or even decreased in speed.

These studies reveal several limitations in current understanding of the effectiveness of different AAC interface designs. In almost all reported cases, the improvements in text generation rate were lower than those expected based on efficiency improvements. The problem is that improving efficiency (having fewer selections to make) may increase the time it takes to make each selection, due to changes in the cognitive, perceptual, and/or motor processes required to use the system. Therefore, efficiency must be considered in tandem with user selection time when assessing and predicting the effectiveness of an AAC system design.

A simple cost-benefit model has been used to describe the relationship between efficiency, user selection time, and rate enhancement (Koester & Levine, 1994). In this model, cost is defined as the percent decrease in item selection rate, benefit as the percent of keystrokes saved, and net gain as the percent increase in text generation rate.[1] From these definitions, it can be shown that:

$$\text{Net Gain} = \frac{(\text{Benefit} - \text{Cost})}{(100 - \text{Benefit})} * 100.$$

The equation reveals the guideline that use of word prediction (or any rate enhancement technique) will only enhance text generation rate if the benefit in keystroke savings achieved by the user exceeds the cost in item selection rate.

The trade-off between cost and benefit in rate enhancement techniques has been recognized by other researchers as well (Soede & Foulds, 1986; Vanderheiden & Kelso, 1987; Witten, Cleary, Darragh, & Hill, 1982), but very little is known about its impact on overall performance. Some quantitative guidance on the size of the cognitive and perceptual loads associated with word prediction is provided by studies in cognitive psychology and human-computer interaction. Much of the additional load is due to the need to search the word list. The time required for a list search can be estimated from studies of human visual

[1] Text generation rate is measured in characters generated per unit time. Item selection rate is the number of items selected per unit time; for example, in a word prediction system, an item can be a single letter or a word list selection.

search performance (Card, 1982; Card, Moran, & Newell, 1983; Neisser, 1963; Somberg, 1987). For a short list (around six words), ordered either alphabetically or by frequency of use, unpracticed users require approximately 1 second to locate a target word. With practice, search times for fixed lists might be expected to improve to around 0.5 seconds (Horstmann & Levine, 1990; Koester, 1994). Additional time may be required to decide whether or not to search the list and to shift attention between the word list and the text generation task (Horstmann & Levine, 1991). Empirical support for extra selection time within the context of word prediction use was provided in a recent study of able-bodied (AB) users of single-switch scanning, who took at least 0.7 extra seconds to make selections with word prediction as compared to without (Koester & Levine, 1994).

Critical questions still remain regarding the balance between cost and benefit with word prediction as well as other rate enhancement systems, but a comprehensive approach to finding the answers remains elusive. A necessary component of a comprehensive approach is to gather additional performance data, since the available data is quite sparse in terms of the number of subjects and conditions that have been studied. Only a few case studies have employed users with disabilities, and only a subset of these have reported quantitative performance (Mathy-Laikko, West, & Jones, 1993; Newell et al., 1991, 1992; Scull & Hill, 1988). The few group studies that have measured performance have generally involved only AB subjects (Gardner-Bonneau & Schwartz, 1989; Koester & Levine, 1994). Additionally, since previous studies have measured performance only for specific user-system combinations, the performance effects of system configuration changes or different user characteristics remain unknown. More empirical performance data must be acquired experimentally across a range of conditions in order to gain a better understanding of the factors that determine user performance. This paper presents the results of an experiment intended to contribute to that understanding.

HYPOTHESES

The general hypothesis is that the additional cognitive and perceptual loads imposed by word prediction have a measurable effect on user performance as compared to typing without the use of word prediction. This will be assessed through multiple indicators, including several measures of performance time, an indicator of mental workload, and users' perception of the difficulty of word prediction use. Both AB and physically disabled subjects are employed across a multisession protocol. The specific hypotheses to be tested are listed below:

1. Text generation rate with word prediction will show a positive relationship with keystroke savings, but text generation rate will be slower than that expected based solely on consideration of keystroke savings.
2. Item selection rate will be markedly slower with word prediction than without, reflecting the additional time needed to make each selection.
3. The time required to search the word prediction list will improve with practice to roughly 0.5 seconds. Search times will be independent of a user's physical disability.
4. Users will perceive word prediction to be somewhat more difficult to use than letters-only typing.
5. Performance with word prediction can be manipulated by instructing users to employ a particular strategy. A strategy involving fewer list searches will impose lower cognitive/perceptual loads on users, resulting in faster item selection rates than a more search-intensive strategy.
6. Performance with word prediction will improve with practice, and the effect of cognitive and perceptual loads will decrease over time but not disappear.
7. The cognitive and perceptual loads incurred during use of word prediction will be similar for users with or without physical disabilities.

METHODS

To test these hypotheses, user performance was measured with and without a word prediction feature across a multisession protocol under a range of conditions. These conditions were defined to include different subject characteristics, task characteristics, and strategies of use. Subjects with physical disabilities were included to represent at least a subset of the actual user population. Because the differences in these subjects' abilities were expected to lead to higher between-subject variance, an alternating treatments design was used, allowing subjects to serve as their own controls (Girden, 1992). AB subjects were also employed in order to address potential problems with between-subject variance and to assess the extent to which AB subjects' performance is similar to that of users with disabilities. Subjects were taught a particular strategy with which to use the word prediction list in order to provide some insight into the influence of strategy on user performance.

Subjects

Fourteen subjects participated. All shared the following characteristics: at least some college-level education; high familiarity with the standard keyboard; no significant prior experience with word prediction; and no reported cognitive, perceptual, or linguistic impairments. All received $75 at the completion of the protocol. Eight of the subjects were able-bodied—four men and four women, recruited from the university community. The remaining six subjects had spinal cord injuries (SCIs) at levels ranging from C4 to C6. All were men recruited from the university community and from the client population seen at our center. These subjects incurred their SCI at least 3 years before the start of the study, with the average time since injury being 9.3 years.

User Interfaces

Two interfaces were employed: a "Letters-only" system involving letter-by-letter spelling on a standard computer keyboard and a "Letters+WP" system using single letter entry augmented by a word prediction feature. Predictions were based on overall frequency of use in English (Beukelman, Yorkston, Poblete, & Naranjo, 1984; Jones & Wepman, 1966; Kucera & Francis, 1967) and presented in a fixed order in a six-word vertical list at the top left corner of the screen, as shown in Figure 1. Software implementing the interfaces was developed specifically for this project to provide sufficient control over the system configuration as well as the means of data collection. AB subjects used mouthstick typing, while subjects with SCIs used their usual method of keyboard access, which was mouthstick typing for two of the subjects and hand splint typing for the other four.

Experimental Design

An alternating treatments design was used, in which subject performance with and without word prediction was recorded in each of seven test sessions. This design was used for three major reasons: to allow a direct comparison of performance with and without word prediction at multiple points across time, to control for learning effects with both interfaces, and to control for variability between subjects. The keystroke savings provided by word prediction was fixed across Sessions 1 to 4 and varied in Sessions 5, 6, and 7 (as

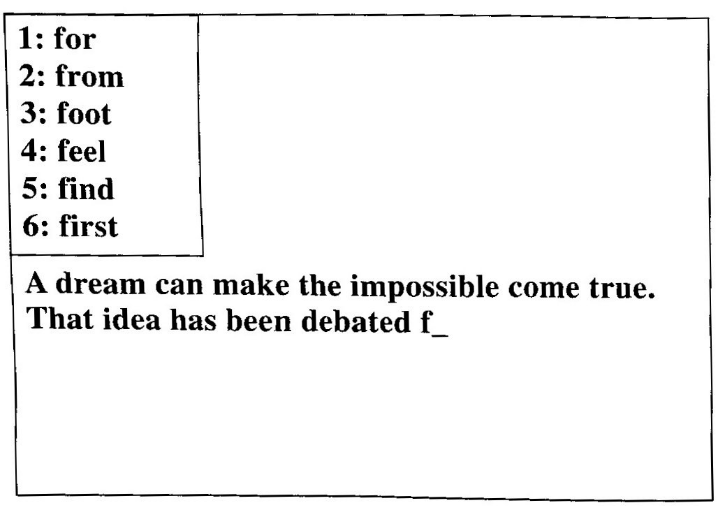

Figure 1. Schematic representation of the Letters+WP system display. The six-word list is fixed in the upper left corner, with the transcribed text displayed below the list. A sample of actual text transcribed by subjects is shown, and the list contents are those that follow selection of the letter "f."

discussed in more detail below). Each subject was randomly assigned one of two strategies with which to use the word prediction feature. Labels for the four subject groups are shown in Table 1.

Procedures

Subjects were tested in individual sessions, conducted in laboratory space for 11 subjects and at subjects' homes for three SCI subjects who had difficulty arranging travel to the university. The protocol incorporated training and testing and spanned an average of 21 days.

In the first part of training, subjects practiced using the Letters-only system. Each AB subject was provided with a 17" anodized aluminum mouthstick for the duration of the study.[2] Proper use of the mouthstick was demonstrated for these subjects, and the keyboard was tilted at an angle of 45 degrees relative to the desk surface. The SCI subjects used their own mouthstick or typing splints. The keyboard was placed to match their normal set-up; all used a flat keyboard. Subjects were instructed to type as quickly as possible, while keeping mistakes to a minimum. They then practiced for 24 sentences, divided evenly into six blocks of text.

The second part of training introduced subjects to the Letters+WP system and their assigned strategy for its use. The rules for the two strategies were defined as follows:

1. *Always-search.* Search the list before every selection.
2. *Two-then-search.* Choose the first two letters of a word without searching the list, then search the list before each subsequent selection.

A search was not required in either strategy when the list was empty. These strategies were chosen to be realistic enough to represent at least a subset of actual user approaches and simple enough to be learned in a single training session. Subjects were asked to follow the rules as closely as possible, even if they thought of other strategies to use. Subjects practiced using their strategy for 16 sentences, which

[2] AdLib Incorporated, 5142 Bolsa Avenue, Suite 106, Huntington Beach, CA 92649.

TABLE 1: **The Four Subject Groups**

Strategy	*SCI No*	*SCI Yes*
Always-search	AB1 (n = 4)	SCI1 (n = 3)
Two-then-search	AB2 (n = 4)	SCI2 (n = 3)

AB = able-bodied; SCI = spinal cord injured.

proved sufficient for each to learn to use the strategy correctly without prompting.

Each of the seven test sessions involved four sentences of warm-up using word prediction, an eight-sentence test with word prediction, then a two-sentence typing test without word prediction. Unique text blocks were developed for each session based on published typing tests, matched with respect to syllable intensity, average word length, and percent of words that occur with high frequency (Lessenberry, 1975). These texts were revised to provide the levels of keystroke savings shown in Figure 2. The keystroke savings in Sessions 1 to 4 was intended to approximate the level that might be expected with an "average" word prediction system (Higginbotham, 1992). Session 5 reflects a higher level of keystroke savings, with Sessions 6 and 7 providing successively poorer-than-average levels.

Sentences were presented singly on index cards, which remained in view for reference throughout transcription. Subjects had 20 seconds to read the sentence before an audio cue signalled them to begin transcription. Errors could be corrected by selecting the "backspace" key as well as a special key for correcting word list selections.

Data Collection

All items selected by subjects were timed and stored by the software in real time. Entries were also encoded to store various information such as the type of selection made (i.e., a single letter or a word list selection) and the number of words in the list when the item was selected. The raw data were used to produce an entry log in which each line shows the selected item, its various characteristics, and the time at which it was selected.

An experimenter was present throughout each session to record observations of subject behavior as well as any comments made by subjects. After each block of text, subjects were asked to rate the difficulty of the task by making a mark on a continuous 7" scale ranging from "very easy" to "very difficult."

All sessions were videotaped with the camera focused close enough on the subject's face to determine the direction of eye gaze. Item selections were recorded on the video using a mirror to reflect a view of the keyboard into the camera and a speech synthesizer to echo the selected item onto the audio track (without being audible to the subject). The camera's clock was synchronized with the computer, so the times on the videotape matched those on the entry log. A diagram of the experimental set-up is shown in Figure 3.

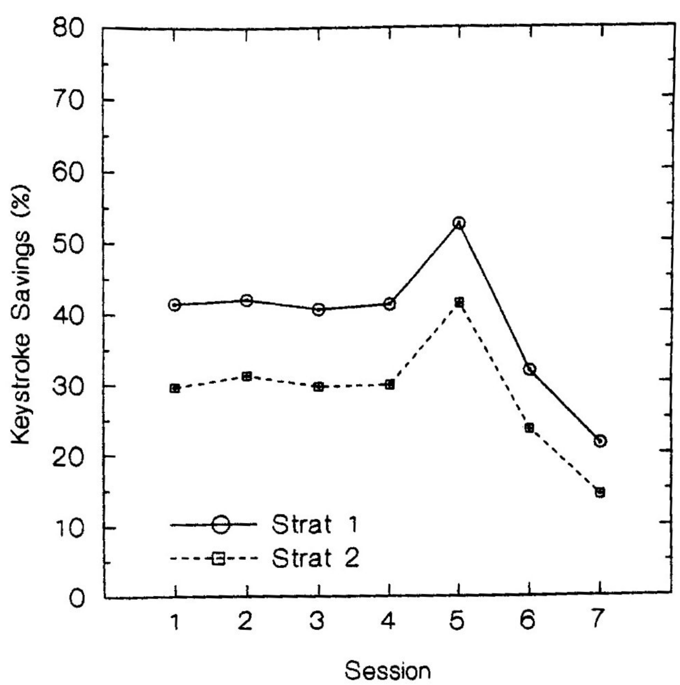

Figure 2. Keystroke savings provided by the Letters+WP system at each session for Strategy 1 and Strategy 2. Note that the text transcribed was identical for users of both strategies, so the difference in keystroke savings between groups was due solely to differences in the strategies.

Data Filtering

The raw data were filtered to remove items judged to be in any of the following three categories. The first category was errors and error corrections, including typographical and transcription errors. The second included words that were not entered in a manner consistent with the assigned strategy. Items in these two categories were identified by comparing a subject's generated text to an error-free template. The final category consisted of "card reads," when a subject referred back to the text card during transcription, as identified using the videotape record. Carriage returns and periods were filtered out as well.

The rationale for filtering arises from the planned use of the data for both empirical measurement and

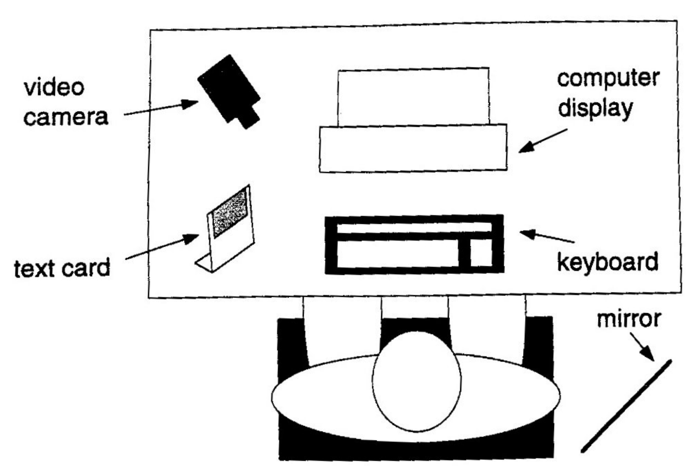

Figure 3. Overhead view of the experimental set-up.

model validation. Empirically, the primary interest is in error-free performance with the assigned strategy, and filtering the "card reads" removes the influence of a task artifact on performance. Additionally, valid derivation of list search and keypress times (discussed below) requires the use of filtered data. From a modeling standpoint, it is reasonable to initially validate the model against error-free data to measure accuracy under best case conditions. The amount of filtered data was analyzed to guard against misleading results, as presented below.

The first author and a trained assistant served as raters for coding events to be filtered. Inter-rater reliability was measured at 99.4%, based on a sample of four sessions analyzed by both raters. A point-to-point reliability measure was used in which the total number of agreements between raters was divided by the total number of agreements and disagreements (McReynolds & Kearns, 1983).

Dependent Measures and Statistical Analyses

Text generation rate and item selection rate were measured in each session for both the Letters+WP and Letters-only systems. Text generation rate was defined as the number of characters generated during the test divided by the total time required to generate those characters. Item selection rate was defined as the number of items selected (including letter and word selections) during the test divided by the total time. The percent change in each of these measures with Letters+WP relative to Letters-only was also calculated. Filtered items were not counted in the number of characters generated or in the total time.

In addition to these overall performance measurements, durations for the component actions of list search and keypress while using Letters+WP were derived from the filtered data for each subject. The technique used for this followed the subtractive methods of Card et al. (1983) and Olson and Nilsen (1988). Based on the strategy used with Letters+WP, each selection made was labelled according to whether it involved a keypress preceded by a list search or a keypress with no list search. For example, when using the "two-then-search" strategy, the first two letters of every word involve no list searches, so they are labelled as a keypress only. The keypress time during use of Letters+WP was then calculated by averaging the times for all keypress-only selections in the session. The list search time was derived by subtracting one keypress time from the time recorded for each selection preceded by a list search, then averaging the remaining times. Keypress time during use of Letters-only was simply the average time for all selections made in the session.

Subjects' subjective ratings of difficulty were measured as the distance in inches from the "very easy" mark to the subject's rating mark. For example, a "very easy" rating score would be 0, "average" would be 3.5, and "very difficult" equal to 7.

Statistical differences in dependent measures across subject groups, systems, and test sessions were determined using a repeated measures analysis of variance (ANOVA) technique. There were four experimental factors, with strategy and presence/absence of SCI as the between-subject factors, and system and session as the repeated measures factors. Statistical significance for each effect was judged at a family-wise p-value of .05, using the Bonferroni procedure to divide by the number of effects examined within the test (Girden, 1992). For example, a test analyzing all four experimental factors examines 14 different effects (4 main effects and 10 interactions), so the critical p-value used for any one of these 14 would be .05/14 = .003. The corresponding critical values for three- and two-factor tests were .007 and .017, respectively. Additionally, all p-values examined and reported are those adjusted based on the Greenhouse-Geisser epsilon as an additional precaution against Type I errors (Girden, 1992). None of the third or fourth order interaction terms in any test were statistically significant, so significant main effects and two-factor interactions could be interpreted without concern for confounding factors.

RESULTS

Filtering

The percentage of data removed from analysis was 16.3% of all Letters+WP selections and 7.3% of all Letters-only selections, averaged across all subjects and sessions.[3] The amount of data filtered was independent of SCI ($p = .31$), strategy used ($p = .70$), or session ($p = .38$). Significantly more data were filtered from Letters+WP selections than from Letters-only selections ($p < .001$). Part of the reason for this is that Letters+WP had the additional category of strategy compliance. The other factor is that subjects referred to the text card over twice as often with Letters+WP, which was a statistically significant difference ($p < .001$). Errors comprised less than 5% of all selections, with no difference in the error rate for the two systems ($p = .66$).

The performance results reported below are based on filtered data. Graphic analysis of the unfiltered data was performed, and it was found that filtering did not change the pattern or apparent significance of the results. Statistical tests were also performed for the unfiltered data, yielding the same pattern of results as for the filtered data.

[3] For all dependent measures, average values across test sessions are reported, regardless of whether session has a statistically significant effect. This is simply to provide an overall picture of the results.

Text Generation Rate

Figure 4a shows the average text generation rate for the Letters-only system for the four subject groups. Subjects with SCIs were significantly faster at Letters-only typing than the AB subjects (p = .005). SCI subjects averaged 116 characters/minute (cpm) (20 wpm) with a standard deviation of 34 cpm, as compared to the AB subjects, who averaged 70 cpm (12 wpm) with a standard deviation of 6 cpm. The finding that SCI subjects were faster is not surprising given that they had much more prior experience with their keyboard access method than the AB subjects had with mouthstick typing. The large magnitude of the difference was surprising, however, and indicates the high skill of the SCI subjects at Letters-only typing.

The average text generation rate with Letters-only typing improved across the seven sessions by 11% and 8% for the AB and SCI subjects, respectively. However, this modest improvement was not statistically significant, with p = .023 for the main effect of session, which is greater than the Bonferroni critical value of .007 for this test.

Subjects' text generation rates with the Letters+WP system were strikingly similar for both groups, as shown in Figure 4b, averaging 71 cpm (12 wpm). There were no statistically significant differences between any of the groups due to strategy (p = .68) or SCI (p = .52). Session was found to have a significant effect (p < .001) across all seven sessions, largely because the higher keystroke savings provided in Session 5 increased Letters+WP performance for all subjects, while the lower keystroke savings in Session 7 decreased it. Over the first four sessions, in which keystroke savings was fixed, there was also a moderate (13%), but statistically significant (p < .001), increase in text generation rate with Letters+WP. As with Letters-only typing, variability between subjects was larger for those with SCIs, with a standard deviation of 19 cpm as compared to 2 cpm for AB subjects.

The net change in text generation rate using Letters+WP relative to Letters-only for each subject group is shown in Figure 5. There was a strong main effect of SCI (p < .001), so data from AB and SCI subjects were analyzed separately. For SCI subjects, use of Letters+WP decreased text generation rate by an average of 41%, as compared to Letters-only typing. For the AB subjects, text generation rate was not significantly affected by the use of word prediction, except during Session 5, which had the highest level of keystroke savings and improved rate by 32%, and during Session 7, which had the lowest keystroke savings and inhibited rate by 14%. Strategy of using Letters+WP had no effect on rate improvement for the AB subjects, while SCI subjects who used the "two-then-search" strategy had a statistically significant advantage over those who used "always-search" (p = .014). Across all subjects, the average text generation rate improvement increased slightly across the first four sessions, in which keystroke savings was fixed, but the effect was not statistically significant (p = .049 vs. criterion of .007).

The between-subject variability for text generation rate improvements was quite consistent for the different subject groups, illustrating the effectiveness of the

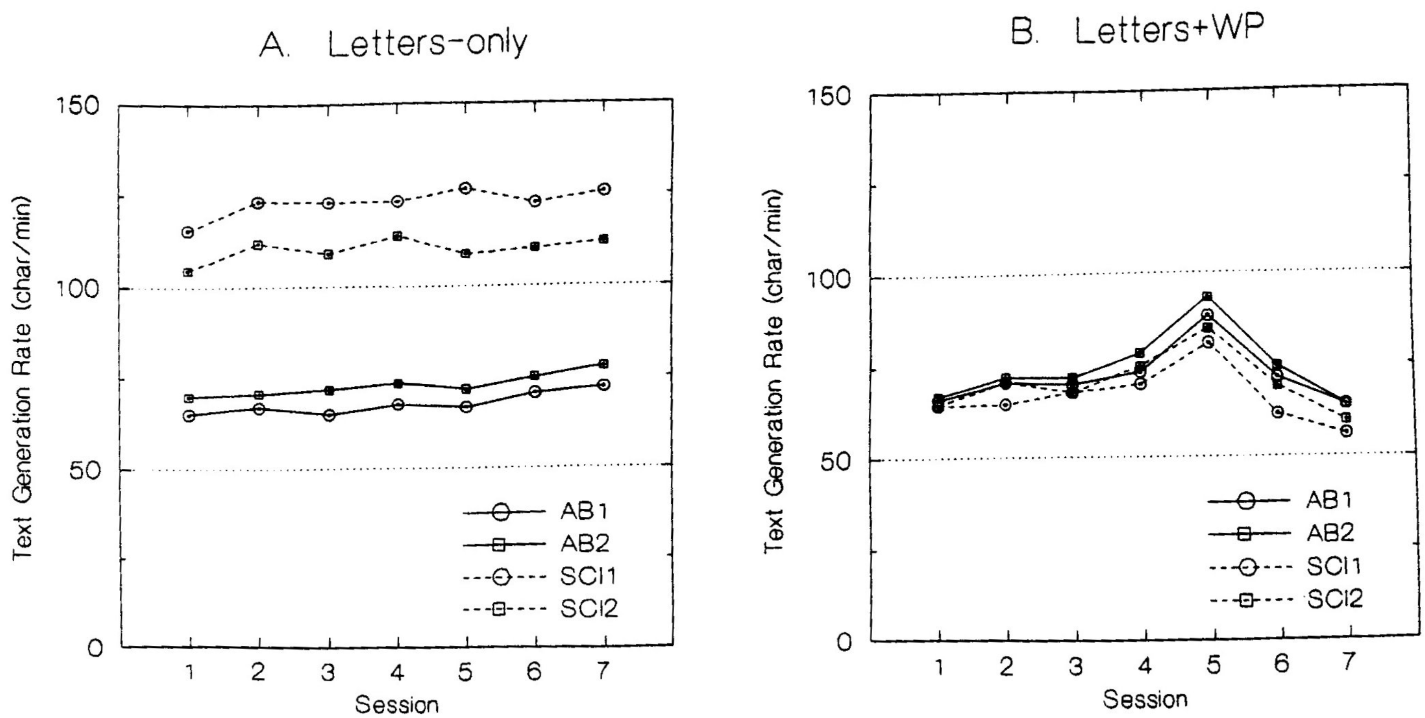

Figure 4. Average text generation rates achieved by each subject group for each session.

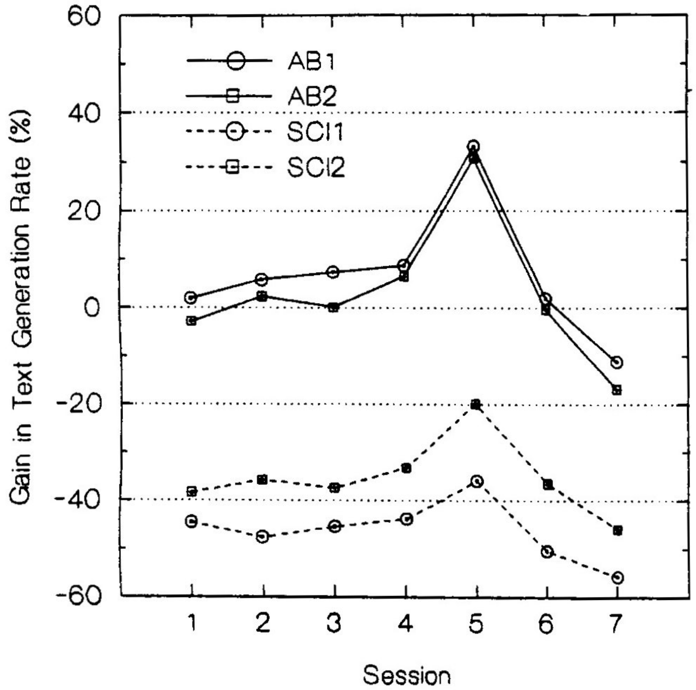

Figure 5. Average improvement in text generation rate during use of the Letters+WP system, relative to Letters-only rate.

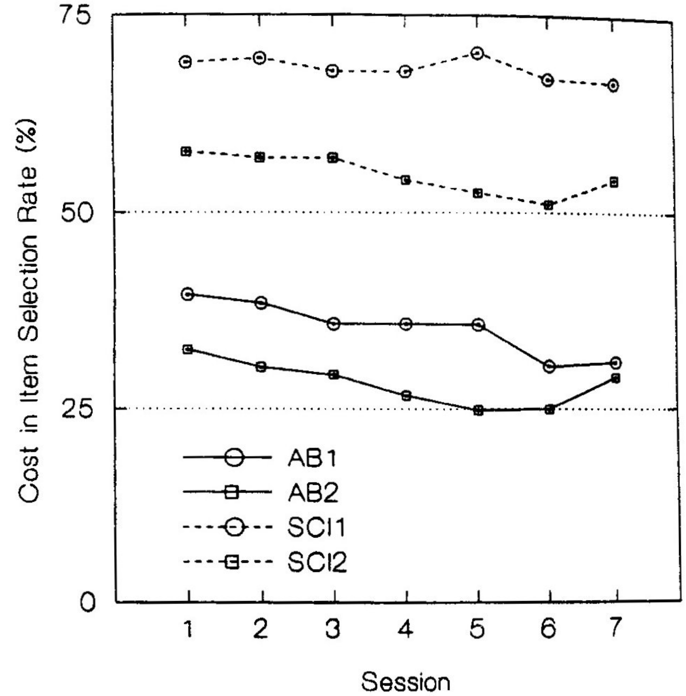

Figure 6. Percent decrease in item selection rate with Letters+WP relative to Letters-only.

alternating treatments design in controlling for variation between subjects. The standard deviation of subjects' average improvements in text generation rate was 7% for SCI subjects and 8% for AB subjects.

Item Selection Rate

For letters-only typing, each selected item generates a single character, so the item selection rate results for the Letters-only system are identical to the text generation rate results shown in Figure 4a above. For word prediction, each item generates an average of more than one character, so item selection rate is not the same as text generation rate. In this study, item selection rate with Letters+WP averaged 46 items per minute (ipm) across all subjects, with a standard deviation of 10 ipm, as compared to the text generation rate average of 71 cpm. There was no significant effect of strategy ($p = .10$) or SCI ($p = .48$) on item selection rate with word prediction. Session was a statistically significant effect, as the average improved from 42 to 50 ipm over the seven sessions.

For all subject groups, item selection rate was significantly slower for the Letters+WP system than for Letters-only ($p < .001$). Figure 6 illustrates this decrease as a relative percentage of the item selection rate with Letters-only. This percent change is the cost of word prediction use discussed above in the Background section. The effect of word prediction on item selection rate was larger for the SCI subjects ($p < .001$). Their item selection rate averaged 62% slower with Letters+WP as compared to Letters-only, while for AB subjects, item selection rate was 32% slower with Letters+WP. The strategy with which Letters+WP was used also influenced the cost in item selection rate ($p < .003$). For both AB and SCI subjects, the cost in item selection rate with word prediction was less for those who used the "two-then-search" strategy, which involved fewer list searches. Across all subjects, practice was a significant effect ($p = .001$), as the cost in item selection rate decreased across sessions. However, the average cost at the end of seven sessions was still 30% for AB subjects and 60% for SCI subjects. The standard deviations were consistent across the four subject groups, averaging 4% with a maximum of 7% for Group AB2.

The decrease seen in item selection rate means that the time required to make each Letters+WP selection was longer, on average, than each Letters-only selection. For SCI subjects, each selection made with the Letters+WP system took an average of 0.91 seconds longer than each Letters-only selection. For AB subjects, this extra selection time was much lower, at 0.41 seconds, but still appreciable.

List Search Time

Figure 7 shows the average list search times at each session for each of the four subject groups. A three-factor ANOVA analysis of list search times showed that strategy of use did not significantly affect list search time ($p = .058$). SCI, however, did have a significant effect ($p < .001$); the average list search times of subjects with SCI were an average of 96% (560 msec) slower than the AB subjects. Across all sessions, the average for SCI subjects was 1.14 sec-

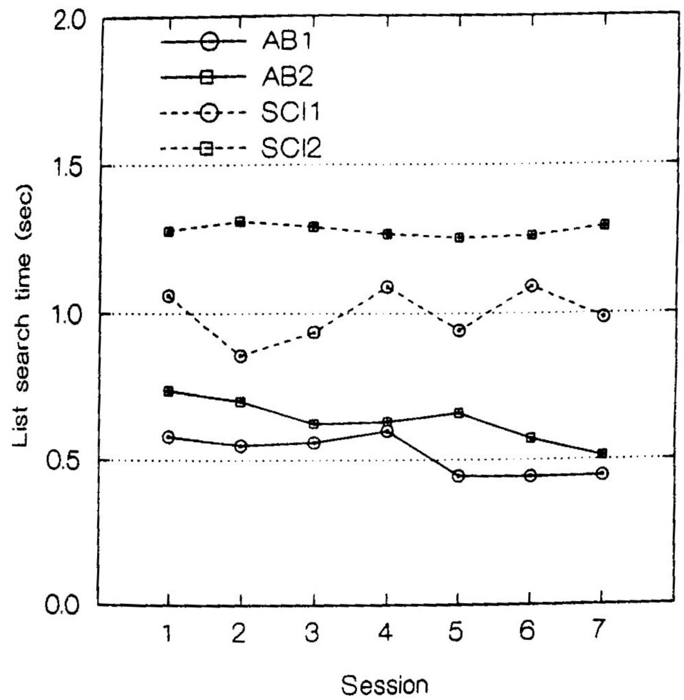

Figure 7. Average list search times during use of Letters+WP for each subject group. SCI subjects had significantly slower search times.

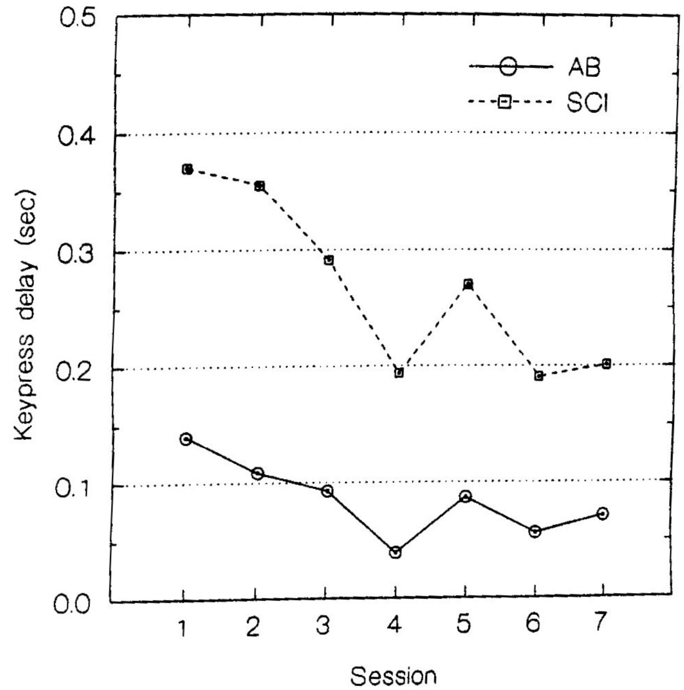

Figure 8. Average keypress delays for each subject group. Keypress delay represents the extent to which keypress time was slower during use of Letters+WP relative to Letters-only.

onds with a standard deviation of 0.26 seconds, while for AB subjects, the average was 0.58 seconds, with a standard deviation of 0.13 seconds.

Because of the large difference between AB and SCI subjects, within-subject effects for these two groups were examined separately, using two-factor ANOVA tests. For AB subjects, session had a significant main effect ($p < .001$), with list search time improving by an average of 27% (180 msec) from Session 1 to Session 7. For SCI subjects, however, average list search time improved only 3% over these sessions, which was not significant ($p = .39$). So, in addition to having slower list search times overall, subjects with SCIs did not significantly improve their search time with practice.

Keypress-only Time

The average keypress times during use of Letters+WP averaged 0.90 seconds across all subjects with a standard deviation of 0.21 seconds. Keypress time averaged 0.83 seconds (SD = 0.31) for SCI subjects and 0.96 seconds (SD = 0.06) for AB subjects. There were no statistically significant differences between the groups, either on the basis of Letters+WP strategy or SCI. The one significant difference that did emerge was a main effect of session ($p = .001$), as keypress time improved an average of 18% from Session 1 to Session 7.

Keypress times during use of Letters+WP were significantly slower than during Letters-only typing ($p < .001$), by an average of 23%. The average keypress delay across sessions is shown in Figure 8 for AB and SCI subjects. (The times are collapsed across strategy of Letters+WP use for greater clarity, since strategy had no significant effect [$p = .63$].) Keypress times with word prediction relative to letters only were an average of 46% (267 msec) slower for SCI subjects and 10% (85 msec) slower for AB subjects. The amount of delay also appeared to decrease with practice, particularly across Sessions 1 to 4, but this effect was not statistically significant ($p = .006$ vs. criterion of .003).

Difficulty Ratings

On a scale of 0 (very easy) to 7 (very difficult), the average rating for Letters-only across all subjects was 1.9, and the average for Letters+WP was 2.5. This suggests that subjects found the transcription task to be relatively easy with both systems. Letters+WP was rated significantly more difficult overall ($p < .001$), but the 0.60 average difference is relatively small, with a standard deviation across subjects of 0.62. SCI subjects in particular found Letters+WP to be more difficult than Letters-only ($p < .001$); the average difference for these subjects was 1.13, while the average difference for AB subjects was only 0.23. The data also suggest that subjects who used the "always-search" strategy perceived Letters+WP to be more difficult relative to Letters-only than did subjects who used the "two-then-search" strategy, with average differences of 0.93 and 0.42, respectively. However, this effect was not statistically significant ($p = .035$).

Ratings across the first four sessions suggest that the task got easier for subjects with practice, and the gap between Letters+WP and Letters-only also decreased over these sessions, although neither of these effects was statistically significant. The effect of practice cannot be assessed at Sessions 5, 6, and 7 since the keystroke savings provided by Letters+WP was manipulated during these sessions.

DISCUSSION

In this discussion, the overall effect of Letters+WP on subjects' performance is assessed to examine the general question of how the increased cognitive and perceptual loads associated with this word prediction feature affect performance. The main effects of user strategy, practice, and presence of SCI on user performance are then discussed. Alternative interpretations of these results are presented, followed by a discussion of the limitations of this study and suggestions for future work.

Overall Effect of Word Prediction

The cluster of results from this study supports the hypothesis that word prediction can place additional cognitive and perceptual loads on users, and that this increase has a major impact on user performance with word prediction. Every dependent measure assessed in the study is consistent with this hypothesis. For all subjects, any improvements in text generation rate with the Letters+WP system relative to Letters-only were much less than would be expected based on keystroke savings alone. AB subjects did enjoy a statistically significant improvement in text generation rate with word prediction in the test session that provided a higher-than-average keystroke savings (above 50%). However, in all other sessions, AB performance with word prediction was not significantly faster than without. For SCI subjects, text generation rate performance with Letters+WP was significantly slower than for Letters-only typing, by over 40%. Thus, the benefit in keystroke savings was generally offset or even exceeded by the cost of making each selection.

The item selection rate results provide a direct measure of this additional cost. AB subjects selected items between 25 to 40% more slowly with Letters+WP compared to Letters-only. For SCI subjects, this cost ranged from 50 to 70%. Recalling from the cost-benefit model that keystroke savings must be at least as high as cost in order for word prediction to yield a net gain in text generation rate, it is easy to see why word prediction did not lead to large improvements in text generation rate in this study.

Derivation of subjects' list search and keypress times during use of word prediction gives insight into how the extra time per selection was spent. List search represented a large portion of the item selection time with the Letters+WP system. By the end of the study, the average search time for AB subjects was 0.47 seconds, which is very close to the expected 0.5 seconds (Horstmann & Levine, 1990). For subjects with SCIs, however, search times were almost twice as long as for AB subjects, which contributed to the higher cost they incurred.

Extra time per selection was observed even for Letters+WP items that did not involve a list search. These include the first two letters of a word entered with the "two-then-search" strategy and letters entered after the prediction list was known to be empty. Each keypress-only selection with Letters+WP took an average of 0.17 seconds (or 23%) longer than keypress selections during Letters-only typing. As with list search, this extra time was observed for all subjects, but was especially pronounced for SCI subjects. The keypress delay suggests that there may be some general cognitive overhead involved in using Letters+WP that is not required during Letters-only typing. This overhead may be used for executive processing, such as determining whether or not to search the list for any given selection or switching one's attention from the prediction list to the keyboard (Card et al., 1983).

The other indicators of cognitive and perceptual loads—subjects' card reads and difficulty ratings—are also consistent with a picture of increased loads during use of Letters+WP. Subjects referred to the text card significantly more often during word prediction use, and often volunteered the comment that searching the word lists interfered with their memory for the transcription sentence. The Letters+WP system was also rated more difficult than Letters-only typing, although subjects found both systems relatively easy to use.

Effect of Search Strategy

Search strategy had the expected effect on cognitive and perceptual loads during use of Letters+WP, as subjects who used the "two-then-search" strategy had a lower cost in item selection rate than those who used the "always-search" strategy. Additionally, "two-then-search" users found the Letters+WP system slightly easier to use than did the "always-search" users. For text generation rate performance, however, strategy had only a limited effect; fewer searches led to better overall performance for the SCI subjects, but not for the AB subjects. One reason for this is that the effect of strategy depends on the user's characteristics (e.g., their keypress and list search times). A quantitative treatment of this relationship requires the application of user performance models (Koester & Levine, 1995) and is beyond the scope of this paper. However, qualitative reasoning suggests that a strategy that requires fewer searches and more keypresses will favor individuals whose search times are relatively slow and whose keypress times are relatively fast. The SCI subjects in general had these characteristics, so they benefitted more from the "two-then-search" strategy than the AB subjects.

Across all subjects, strategy was somewhat effective at mitigating the effect of cognitive and perceptual loads, but the effect was not large. There may be, of course, other prescribed strategies that would be more successful at this. It is also possible that allowing users to follow their own "natural" strategy in searching the list would yield different results; this is being investigated in a follow-up study.

Effect of Practice

As expected, practice across test sessions was a factor in improving performance both with and without word prediction. Text generation rates for both systems improved at a similar pace over the first four sessions, so the net difference between the systems did not change significantly with practice. The cognitive and perceptual loads associated with word prediction did decrease with practice but were still significant factors by the end of the study, again as expected. Over the seven test sessions, cost in item selection rate decreased significantly (see Fig. 6). However, the average decrease was only 3 to 6 percentage points. List search times also improved with practice, although for SCI subjects the improvement was only 3%.

Additional practice would likely have resulted in further decrements in cognitive load. Visual search skills might improve with increasing familiarity with the prediction lists, and a user may also improve in the ability to anticipate the list contents (i.e, in deciding when to search the list). In this study, the strategy rules limited the amount of anticipation that could be employed by subjects, so these data do not directly address that issue. Relative to visual search, human information processing studies suggest that performance on a visual search task of the sort required in word prediction is unlikely to become truly automatic, even with extended practice, because the words in the list serve as both targets and distractors (Schneider & Fisk, 1982). This makes it unlikely that cognitive and perceptual loads with word prediction would become negligible in skilled users, although more empirical work is certainly necessary to address the complex and critical question of long-term practice effects.

Effect of SCI

An interesting and unexpected result is the large difference in word prediction's effect on SCI as compared to AB subjects. On all dependent measures, Letters+WP exacted greater cognitive and perceptual loads on subjects with SCIs. Two phenomena converged to produce this result. First, subjects with SCIs were faster at letters-only typing than AB subjects. For the same search time, then, these SCI subjects would incur a higher cost than the AB subjects. Second, subjects with SCIs in fact had much slower search times and longer keypress delays than the AB subjects, which further adds to the increased cost.

A major source of the disparity may be the difference in the groups' expertise in the Letters-only condition. All subjects had roughly equal familiarity with the keyboard layout. However, the AB subjects had no prior experience in mouthstick typing, while the SCI subjects had extensive experience with their particular method of keyboard access and had highly developed skills in letter-by-letter typing using that method. When the subjects with SCI were asked to use word prediction, additional cognitive effort may have been required to inhibit their habit of typing the word as they normally would, above and beyond the loads associated with word prediction itself. The AB subjects would not incur these additional loads, since they did not have highly developed motor patterns for letter-by-letter mouthstick typing. Indirect support for this hypothesis is provided by a few occasions in which SCI subjects typed an entire word letter by letter, forgetting to attend to the list.

One explanation for how prior skill with letters-only typing could adversely affect performance with word prediction is known as "negative transfer," whereby the skills learned for one system interfere with developing new skills necessary for a second system. This has been observed when people trained on one form of text editor are forced to switch to another one (Singley & Anderson, 1988). However, the effect in that study was relatively mild and decreased rapidly with practice, while in this study the effect was large and durable (for the experimental time course), at least for the SCI subjects.

The strong negative transfer observed in the SCI subjects may have been due to a qualitative shift in their general mode of information processing in moving from the highly practiced skill of typing to the new task of using word prediction. It is possible that for at least some of these subjects, single-letter typing was largely an "automatic process," requiring a minimum of cognitive effort, not unlike 10-finger touch typing. In contrast, use of word prediction may have required a mode known as "control processing," which is slower and more effortful than automatic processing (Schneider & Fisk, 1982). AB subjects would not have experienced this shift, since the relative novelty of both mouthstick typing and word prediction would require control processing with both systems.

In assessing the effect of word prediction, we have focused on *relative* performance as compared to text generation without word prediction, and it is in this relative sense that word prediction exacted a larger cost from subjects with SCIs. In *absolute* terms, SCI subjects were able to achieve text generation rates with word prediction that were only 5 to 10% slower than those of AB subjects, despite having much slower list search times. The reason for this is that their keypress times during use of word prediction were an average of 13% faster than those of AB subjects.

Although the difference in keypress times was not as dramatic as that for search times, it succeeds in counteracting the effect of slower search times because keypress time is a stronger factor in determining text generation rate than is search time. A quantitative understanding of the relative strength of the factors that determine text generation rate is possible but requires the use of analytical models. This is beyond the scope of this paper but has been discussed in previous work (Koester, 1994; Koester & Levine, 1995) and will continue to be a focus of our work.

Alternative Interpretations

In attributing the additional selection time in word prediction to cognitive and perceptual loads, the implicit assumption is that the motor components of the task are the same with and without word prediction. It is recognized, however, that part of the extra selection time may be due to differences in motor activities. For example, searching the word prediction list requires the user to shift eye gaze and occasionally head position. The time for these activities may form a portion of the extra selection time and may also be a reason for the higher difficulty perceived during use of Letters+WP. Anecdotal observations of subjects, however, suggest that users rarely moved their head to search the list and they generally shifted their eye gaze while completing the previous keypress, which reduces the amount of extra time involved in these processes.

An additional motor issue is that use of Letters+WP required the use of the numerical as well as the alphabetical keys. Depending on a user's range of motion and other motor abilities, the need to reach the numerical keys may influence average item selection times. The extent to which this was a factor in this study was not explicitly analyzed, but based on observations of the ability and skill of these subjects, there was no obvious reason to believe that the difficulty or speed of pressing the keys was different for numerical and alphabetical keys. (Note that the Letters+WP system used the regular number keys, which are closer to the alphabetical keys and presumably faster to access than either the function keys or the numerical keypad, which are sometimes used in commercial word prediction systems.)

Limitations of This Study

The focus here has been primarily on text generation rate and other time-based measures of user performance. It should be noted that there are other factors that also contribute to a user's satisfaction and success with a system. For example, for some users, word prediction's facilitation of spelling may be more important than sheer speed. Improving motor efficiency may reduce fatigue for some users, allowing them to work longer or more comfortably. These are all issues that deserve further empirical investigation. However, time-specific measures of performance remain important factors for many users and therefore are worthy of careful attention.

While this study was carefully designed to provide performance information across a relatively broad range of usage conditions, its design also has features that limit the general application of this result. Subjects were constrained in the strategy they were to use with Letters+WP, the text they were to generate, and the number of sessions in which they used the systems. The limited time course is perhaps the most important of these factors, since none of the subjects could be considered true word prediction experts by the end of the experiment. Additionally, the SCI subjects represent only one subgroup of the actual user population, which includes individuals with more variable motor skills as well as those with cognitive impairments.

The implementation of the Letters+WP system is another potential limitation. While these results may not necessarily generalize to commercial systems, the quality of the Letters+WP system used was comparable in many ways to existing systems. The system dictionary contained a high percentage of the words to be entered, and the keystroke savings offered by the system across the first four sessions was similar to that reported for some other systems (discussed in Background above) (Higginbotham, 1992). The largest difference between the Letters+WP system used in this study and many commercial systems is that an adaptive prediction algorithm was not used, so the word lists seen by subjects were fixed throughout the study. Adaptive lists incorporate a cost-benefit trade-off of their own: they may enhance the keystroke savings provided by the system, but the dynamic lists may reduce subjects' ability to learn the list contents, which may increase search time. Further empirical and modeling work is necessary to address the effect of adaptive lists.

These considerations highlight the need for continued empirical work, focusing on the performance of users with different abilities and levels of expertise than the subjects studied here as well as other implementations of word prediction, to either corroborate these results or reveal conditions under which word prediction provides a significant improvement in rate. Additionally, these empirical limitations underscore the motivation for developing modeling techniques that can be used to understand the many situations that cannot be empirically examined.

CONCLUSIONS

Subject performance under the conditions of this experiment is consistent with the hypothesis that word prediction can impose cognitive and perceptual loads on its users. These loads are associated with a time cost that can offset and even overwhelm the benefit provided by keystroke savings. SCI subjects appeared to incur higher cognitive costs than AB subjects, which we attribute to their prior expertise in letters-only typing. The strategy with which word prediction is used can affect the cognitive load imposed on users, but since it also influences keystroke savings, strategy may not result in a net change in the text generation rate achieved. Finally, the cognitive and perceptual loads associated with word prediction decreased with practice but were still significant factors by the end of the study. Further work is encouraged to assess the effects of system properties, user characteristics, and usage conditions in greater depth.

ACKNOWLEDGMENTS

Many thanks to Aram Sherman for his assistance in performing this study and to the subjects for their generous contributions of time and energy. This work was supported by the National Science Foundation, the University of Michigan Rackham School of Graduate Studies, and the U-M Rehabilitation Engineering Program. This paper was presented in part at the 18th Annual RESNA Conference, Nashville, TN, June 1994 and is based on a doctoral dissertation entitled "User Performance with Augmentative Communication Systems: Measurements and Models," written by the first author in partial fulfillment of the requirements for a Ph.D. degree at the University of Michigan. Simon P. Levine was the chair of the dissertation committee.

Address reprint requests to: Heidi Horstmann Koester, 1C335 University Hospital, Ann Arbor, MI 48109-0032, USA; e-mail: hhk@umich.edu.

REFERENCES

Baker, B. (1987). Semantic compaction for sub-sentence vocabulary units compared to other encoding and prediction systems. In *Proceedings of the 10th Annual RESNA Conference* (pp. 118–120). Washington, DC: RESNA.

Bentrup, J.A. (1987). Exploiting word frequencies and their sequential dependencies. In *Proceedings of the 10th Annual RESNA Conference* (pp. 121–123). Washington, DC: RESNA.

Beukelman, D. R., Yorkston, K. M., & Dowden, P. A. (1985). *Communication augmentation: A casebook of clinical management.* San Diego: College Hill Press.

Beukelman, D. R., Yorkston, K. M., Poblete, M., & Naranjo, C. (1984). Frequency of word occurrence in communication samples produced by adult communication aid users. *Journal of Speech and Hearing Disorders, 49*, 360–367.

Card, S. K. (1982). User perceptual mechanisms in the search of computer command menus. In *Human factors in computer systems proceedings* (pp. 190–196). Gaithersburg, MD: National Bureau of Standards.

Card, S., Moran, T., & Newell, A. (1983). *The psychology of human-computer interaction.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Chopra, P. S., Thomason, V. A., Kazemi, H., & Alguindigue, I. E. (1988). Modular single-switch wordprocessor and calculator for classroom applications. In *Proceedings of ICAART 88* (pp. 410–411). Washington, DC: RESNA.

Deroost, G.,& Soede, M. (1989). A text entry system based on frequent letter sequences. In J. Presperin (Ed.), *Proceedings of the 12th Annual RESNA Conference* (pp. 280–281). Washington, DC: RESNA.

Foulds, R. A. (1980). Communication rates for nonspeech expression as a function of manual tasks and linguistic constraints. In *Proceedings of the International Conference on Rehabilitation Engineering* (pp. 83–87). Washington, DC: RESNA.

Gardner-Bonneau, D. J. & Schwartz, P. J. (1989). A comparison of Words Strategy and traditional orthography. In J. Presperin (Ed.), *Proceedings of the 12th Annual RESNA Conference* (pp. 286–287). Washington, DC: RESNA.

Gibler, C.D. & Childress, D.S. (1982). Language anticipation with a computer-based scanning communication aid. In *Proceedings of the IEEE Computer Society Workshop on Computing to Aid the Handicapped* (pp. 11–15). Charlottesville, VA: IEEE.

Girden, E. R. (1992). *ANOVA: repeated measures.* Sage University Paper Series on Quantitative Applications in the Social Sciences, 07–084. Newbury Park, CA: Sage.

Goodenough-Trepagnier, C., Rosen, M. J., & Demsetz, L. (1982). Determinants of rate in communication aids for the non-vocal motor handicapped. In *Proceedings of the Human Factors Society—26th Annual Meeting* (pp. 172–175).

Goodenough-Trepagnier, C., Rosen, M. J., & Galdieri, B. (1986). Word menu reduces communication rate. In *Proceedings of the 9th Annual RESNA Conference* (pp. 354–356). Washington, DC: RESNA.

Goodenough-Trepagnier, C., Tarry, E., & Prather, P. (1982). Derivation of an efficient nonvocal communication system. *Human Factors, 24*, 163–172.

Heckathorne, C. W., Voda, J. A., & Leibowitz, L. J. (1987). Design rationale and evaluation of the Portable Anticipatory Communication Aid - PACA. *Augmentative and Alternative Communication, 3*, 170–180.

Higginbotham, D. J. (1992). Evaluation of keystroke savings across five assistive communication technologies. *Augmentative and Alternative Communication, 8*, 258–272.

Horstmann, H. M. (1990). Quantitative modeling in augmentative communication—a case study. In J. Presperin (Ed.), *Proceedings of the 13th Annual RESNA Conference* (pp. 9–10). Washington, DC: RESNA.

Horstmann, H. M. & Levine, S. P. (1990). Modeling of user performance with computer access and augmentative communication systems for handicapped people. *Augmentative and Alternative Communication, 6*, 231–241.

Horstmann, H. M. & Levine, S. P. (1991). The effectiveness of word prediction. In J. Presperin (Ed.), *Proceedings of the 14th Annual RESNA Conference* (pp. 100–102). Washington, DC: RESNA.

Jones, L., & Wepman, J. (1966). *A spoken word count.* Chicago: Language Research Associates.

Koester, H.H. (1994). *User performance with augmentative communication systems: Measurements and models.* Ph.D. Dissertation, University of Michigan, Ann Arbor, MI.

Koester, H. H., & Levine, S. P. (1994). Learning and performance of AB individuals using scanning systems with and without word prediction. *Assistive Technology, 6(1)*, 42–53.

Koester, H. H., & Levine, S. P. (1995). Simulations of user performance with word prediction. In A. Langton (Ed.), *Proceedings of the RESNA '95 Annual Conference* (pp. 714–716). Washington, DC: RESNA.

Kucera, J., & Francis, W. N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Lessenberry, D. D. (1975). *College typewriting.* Cincinnati: South-western Publishing.

Levine, S. P., Gauger, J. R. D., Bowers, L. D., & Khan, K.J. (1986). A comparison of mouthstick and Morse code text inputs. *Augmentative and Alternative Communication, 2*, 51–55.

Mathy-Laikko, P., West, C., & Jones, R. (1993). Development and assessment of a rate acceleration keyboard for direct-selection

augmentative and alternative communication users. *Technology and Disability, 2(3)*, 57–67.
McReynolds, L.V., & Kearns, K.P. (1983). *Single-subject experimental designs in communicative disorders.* Baltimore: University Park Press.
Neisser, U. (1963). Decision time without reaction time: Experiments in visual scanning. *American Journal of Psychology, 76*, 376–385.
Newell, A. F., Arnott, J., & Waller, A. (1992). On the validity of user-modeling in AAC: comments on Horstmann and Levine (1990). *Augmentative and Alternative Communication, 8*, 89–92.
Newell, A. F., Booth, L., & Beattie, W. (1991). Predictive text entry with PAL and children with learning difficulties. *British Journal of Education Technology, 22*(1), 23–40.
Olson, J. R., & Nilsen, E. (1988). Analysis of the cognition involved in spreadsheet software interaction. *Human Computer Interaction, 3*, 309–350.
Pollak, I. V., & Gallagher, B. (1989). A fast communication aid for non-verbal subjects with severe motor handicaps. *Journal of Medical Engineering and Technology, 13*(1/2), 23–27.
Rowley, B. A. (1989). Evaluation of text processing using a V-KIM window. In J. Presperin (Ed.), *Proceedings of the 12th Annual RESNA Conference* (pp. 252–253). Washington, DC: RESNA.
Schneider, W., & Fisk, A. D. (1982). Degree of consistent training: Improvements in search performance and automatic process development. *Perception and Psychophysics, 31*, 160–168.
Scull, J., & Hill, L. (1988). A computerized communication message preparation program that "learns" the user's vocabulary. *Augmentative and Alternative Communication, 4*, 40–44.
Singley, M.K., & Anderson, J.R. (1988). A keystroke analysis of learning and transfer in text editing. *Human Computer Interaction, 3*, 223–274.
Soede, M., & Foulds, R. A. (1986). Dilemma of prediction in communication aids. In *Proceedings of the 9th Annual RESNA Conference* (pp. 357–359). Washington, DC: RESNA.
Somberg, B. L. (1987). A comparison of rule-based and positionally constant arrangements of computer menu items. In *Proceedings of CHI'87* (pp. 255–260). New York: ACM.
Swiffin, A. L., Arnott, J. L., Pickering, J. A., & Newell, A.F. (1987). Adaptive and predictive techniques in a communication prosthesis. *Augmentative and Alternative Communication, 3*, 181–191.
Vanderheiden, G. C., & Kelso, D. P. (1987). Comparative analysis of fixed-vocabulary communication acceleration techniques. *Augmentative and Alternative Communication, 3*, 196–206.
Venkatagiri, H. S. (1993). Efficiency of lexical prediction as a communication acceleration technique. *Augmentative and Alternative Communication, 9*, 161–167.
Witten, I. H., Cleary, J. G., Darragh, J. J., & Hill, D. R. (1982). Reducing keystroke counts with a predictive computer interface. In *Proceedings of the IEEE Computer Society Workshop on Computing to Aid the Handicapped* (pp. 3–10). Charlottesville, VA: IEEE.

The Craft of Research

Recently, I read the book *The Craft of Research* (1995), written by Wayne Booth, Gregory Colomb, and Joseph Williams for student researchers. According to the authors, the book was written to introduce beginning researchers to the "nature, uses and objectives of research," to guide researchers through the planning, organizing, and drafting of their work, and to show writers how to diagnose passages that will give their readers difficulty and to revise these passages quickly and efficiently.

For those *AAC* subscribers who are responsible for training beginning researchers, *The Craft of Research* will be a useful tool. To introduce this book more completely, the Table of Contents is listed below. (*AAC* Editor)

Table of Contents

Booth, W., Colomb, G., & Williams, J. (1995). *The craft of research.* Chicago: The University of Chicago Press.