Measuring Keyboard Performance for People with Disabilities

Edmund F. LoPresti, Heidi Horstmann Koester, Koester Performance Research;
Richard C. Simpson, University of Pittsburgh

ABSTRACT

Fourteen individuals with various motor impairments performed a series of text entry exercises. Performance measures were collected which indicate overall performance (speed and error rate) and classification of errors. Such performance metrics could be useful in assessing the computer access skills of clients, and recommending software, hardware, or training interventions.

BACKGROUND

People with disabilities may have difficulty performing text entry tasks using a computer keyboard. The specific types of difficulty will vary depending on the person's intrinsic abilities, the input device being used, and the task. In order to understand and remediate the difficulties faced by a particular individual, it is desirable to quantify performance and classify errors. This aids in comparing performance over time or between input devices; evaluating the success of an intervention; or communicating the client's need for an intervention.

General measures of typing performance include text entry rate and accuracy. A typical measure of text entry rate is the number of words per minute. It is more difficult to measure error rate in a consistent manner, given the variety of ways in which someone can make an error. Soukoreff and MacKenzie have proposed a measure of error rate using the mean string distance (MSD) (1). The MSD between two strings of characters is the minimum number of primitives – character insertions, deletions, or substitutions – necessary to transform one string into the other. This provides a measure of the difference between a string of characters entered by the user and the correct string.

In addition to an overall error rate, it is desirable to measure the rate of different kinds of errors. Trewin and Pain classified text entry errors into seven categories. These categories are shown in Table 1 along with their frequencies for 26 participants typing approximately 100 words each (2). The appropriate intervention for a client will depend on the types of errors which he or she experiences. For example, the types of errors observed may indicate whether someone needs a slower response time, a smaller keyboard, a different keyboard layout, or other hardware or software interventions.

Error type	Description	Rate (Participants with Disabilities)	Rate (Participants without Disabilities)
Long key press errors	A key was pressed long enough to generate repeats.	10.6%	0
Additional key errors	A key near the intended key was activated instead of or in addition to the intended key	1.1%	0.2%
Missing key errors	A movement intended to press a key did not produce a character, because the person either missed the key or did not press hard enough	0.7%	0.1%
Dropping errors	When the person wanted to press two keys simultaneously (e.g. Shift and a letter) he or she only pressed one key	0.2%	0.01%
Bounce errors	A key was unintentionally pressed more than once	0.2%	0
Remote errors	A key that is not near the intended key is pressed by mistake	0.1%	0
Transposition errors	Two keys are pressed in the wrong order (e.g. the letters are transposed)	0.02%	0.03%

Table 1: Error categories identified in [2], with error rates for 20 subjects with no disabilities and 6 subjects without disabilities. Error rates for a type of error are calculated as the number of keystrokes resulting from that type of error divided by the total number of keystrokes.

RESEARCH GOALS

The goals for this study were (1) to collect baseline data on text entry performance, addressing the question of what problems people are having with text entry; and (2) evaluate possible metrics for text entry performance.

METHODS

Fourteen participants with disabilities took part in the study. Each participant took part in one of four group sessions. Each participant used a standard laptop keyboard to type ten sentences using direct selection. Software based on Compass assessment software (3) presented a sentence to be typed, and the participant entered text in a second text field (see Figure 1). Sentences included multiple characters which required the participants to modify a key (e.g. use the Shift key), such as capital letters or certain punctuation (e.g. ?, !, @). Sticky Keys was initially turned off on all computers. Participants could choose whether to correct any errors in their typing; any time spent on correcting was included in the total time for the sentence. After completing the sentence, the participant pressed the "Enter" key to move to the next sentence. If "Enter" was not pressed within a predetermined maximum time, the software automatically presented the next sentence.

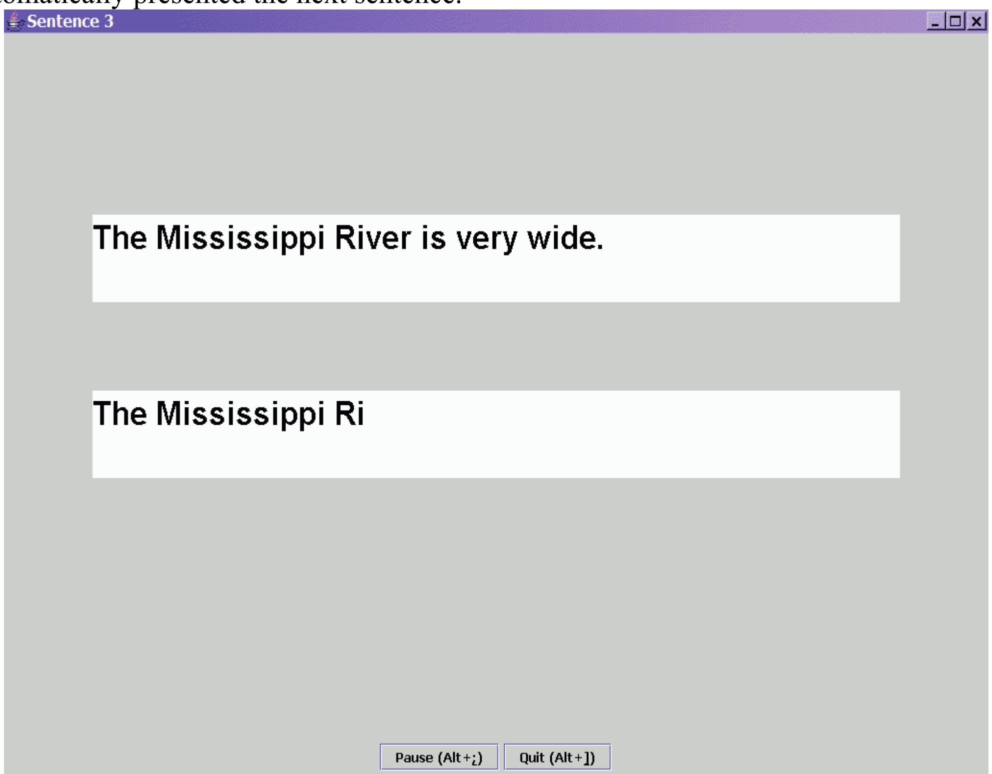

Figure 1: One trial of the text entry task. The top text field presents the sample text to the participant, and the participant enters text in the lower field.

Twelve participants performed a second trial with a different keyboard. Six participants used an ergonomic keyboard (Chicony, www.chicony.com.tw), 5 used a WinMini one-hand keyboard (Tash, www.tashinc.com/), and one used a FingerWorks mini-keyboard (www.fingerworks.com/). Data for the WinMini and FingerWorks are combined as "reduced-size keyboards" below.

The software recorded performance metrics for each participant and each hardware device, including the total time for each sentence, the total number of errors, and the number of uncorrected errors. In addition, the software recorded the list of characters typed and the time between keystrokes; and these data were used to identify and classify specific errors. Errors were classified using the categories presented in Table 1, and were identified as described in Table 2.

Trewin and Pain used the category of "dropping errors" to record instances of people having difficulty with combined key presses (in particular, simultaneously pressing the Shift key and a character key). In this study, five metrics were used to address this issue, as described in Table 3. For purposes of comparing these data to the other error rates, "failures to use a modifier" and "incorrect uses of a modifier" were combined into a single measure of "dropping errors".

Error type	Method of Identification
Long key press errors	Repeated characters with long keystroke durations (e.g. greater than the repeat delay); value indicates the number of repeating characters generated, not the number of keystrokes (e.g. holding down the 'd' key could generate 5 'd' characters with just 1 keystroke)
Bounce errors	Repeated characters with shorter keystroke durations (indicating a series of separate keystrokes)
Additional key errors	A character which did not belong in the sentence, and for which a neighboring character in the entered text was also a neighboring key in the physical keyboard layout; e.g., the key erroneously pressed by the user was adjacent to a key that was correctly pressed by the user.
Missing key errors	Characters which were missing from the entered text
Remote errors	A character which did not belong in the sentence, and which did not qualify as an "additional key error"
Dropping errors	Sum of "failures to use a modifier" and "incorrect uses of a modifier" (see Table 3)
Transposition errors	Neighboring characters for which the order is reversed.

Table 2: Methods of identifying errors from the recorded data.

Performance Metric	Method of Identification
Correct uses of modifier	Number of instances when a key that should have been modified was modified with the Shift key
Uses of other methods to generate capital letters	The number of instances in which a character which should have been modified was modified, but not using the Shift key (e.g. by using Caps Lock for a single character); use of Caps Lock for a single character was taken as evidence that the user was having difficulty with combined keypresses
Failures to use a modifier	Number of keys that should have been modified but were not
Incorrect uses of a modifier	The number of instances in which a key was modified when it should not have been modified (this is often due to leaving Caps Lock on rather than using the Shift key)
Extra modifiers	Number of instances when Shift was pressed without typing a character (possible evidence that the user is having trouble holding the Shift key down long enough to press another key)

Table 3: Errors related to combined keystrokes.

RESULTS

Average text entry rates for each keyboard type are shown in Table 4. Overall error rates for each keyboard type are shown in Table 5. In Table 5, "total errors" refers to all errors made by the user, and was calculated by comparing the string of all characters typed with the intended text. "Net errors" does not count any errors which the participant fixed, and was calculated by comparing the participant's completed sentence with the intended text.

	Mean Text Entry Rate	Standard Deviation	Minimum	Maximum
Laptop Keyboard	6.0 wpm	3.8	1.1 wpm	13.4 wpm
Expanded Keyboard	5.7 wpm	4.1	1.2 wpm	12.4 wpm
Reduced Size Keyboards	3.1 wpm	2.2	0.3 wpm	5.2 wpm

Table 4: Text entry times for each keyboard, in words per minute (wpm). Statistics are across 14 subjects for the laptop keyboard and across six subjects each for the expanded keyboard and reduced size keyboards.

	Mean Error Rate	Standard Deviation	Minimum	Maximum
Laptop Keyboard - total errors	32.3%	29.6%	2.6%	85.8%
Laptop Keyboard - net errors	27.0%	27.6%	0.6%	74.8%
Expanded Keyboard - total errors	30.1%	31.1%	1.6%	83.7%
Expanded Keyboard - net errors	25.9%	30.7%	1.2%	56.0%
Reduced Size Keyboards - total errors	54.4%	32.6%	19.4%	96.5%
Reduced Size Keyboards - net errors	39.4%	30.2%	2.7%	71.6%

Table 5: Error rates for each keyboard. Statistics are across 14 subjects for the laptop keyboard and across six subjects each for the expanded keyboard and reduced size keyboards.

Table 6 provides the rates of specific error types. The last row of Table 6 gives the rate of "missed characters" which resulted from participants not completing sentences before the allowed time elapsed. These were counted toward the error rates in Table 5. However, they were not counted as "missed errors" in Table 6 because they may have a different cause than missed characters within a sentence. Table 7 provides further performance measures related to use of the modifiers, as defined above.

Error type	Rate (laptop keyboard)	Rate (expanded keyboard)	Rate (reduced size keyboard)
Long key press errors	2.5%	1.8%	3.5%
Additional key errors	1.3%	0.7%	0.1%
Missing key errors	0.8%	0.5%	2.0%
Dropping errors	2.0%	2.0%	2.7%
Bounce errors	0.4%	4.5%	7.7%
Remote errors	3.2%	0.5%	5.8%
Transposition errors	0.1%	0	0
Unfinished sentence errors	24.0%	22.9%	23.5%

Table 6: Error rates for different error types.

Participant	Keyboard	Correct uses of modifier	Other methods (Caps Lock)	Failures to use modifier	Incorrect uses of modifier	Extra modifiers
A	Laptop	28	0	2	9	4
	Expanded	24	0	1	0	4
B	Laptop	16	0	0	0	0
	Reduced	16	0	2	1	8
C	Laptop	27	0	3	0	15
	Expanded	28	0	1	0	1
D	Laptop	27	3	0	0	1
	Reduced	2	26	4	0	6
E	Laptop	4	0	9	0	2
	Extended	0	0	13	0	0
F	Laptop	0	2	14	0	0
	Reduced	0	0	11	0	0
G	Laptop	0	0	14	0	0
	Expanded	0	0	21	0	5
H	Laptop	8	0	4	0	0
	Reduced	0	0	12	0	0
I	Laptop	3	28	0	0	0
	Reduced	0	27	2	0	0
J	Laptop	4	0	29	0	1
	Reduced	0	0	26	0	0
K	Laptop	24	0	7	0	6
	Expanded	17	0	2	1	3
L	Laptop	32	0	1	0	6
	Expanded	32	0	2	1	10
M	Laptop	0	7	6	43	0
N	Laptop	29	1	3	0	2

Table 7: Measures of performance related to using modifier keys.

DISCUSSION

The performance measures recorded in this study reflect multiple levels of analysis. Speed and error rate provide an overall measure, with low speed and/or high error rate indicating the user's level of difficulty. Considering both total errors and net errors allows for analysis of the individual's ability to recognize and correct errors, as well as ability to avoid errors in the first place. Participants in this study varied widely in terms of error rates, but all participants had fairly slow performance.

Classification of errors allows for a more detailed understanding of the difficulties which a person is experiencing. These classifications may suggest possible solutions. Some errors can be reduced through changes to software settings, while others can be reduced through use of a different keyboard; for example a larger keyboard, smaller keyboard, or a keyboard with different spacing between keys. Some software configuration settings available in Windows are shown in Table 8. Increasing repeat delay can reduce the number of long press errors, activating bounce keys can reduce the number of bounce errors, and increasing the acceptance delay could reduce the number of additional key errors.

Parameter	Description
Repeat Delay	How long a key must be held down before it begins to repeat.
Repeat Rate	Once the keyboard begins to repeat a character, the rate at which it repeats.
Bounce Keys	Tells the operating system to ignore keystrokes that are depressed within *x* seconds of the previous key release.
Acceptance Delay	How long a key must be held down before it is accepted
StickyKeys	When StickyKeys are activated, the typist can enter key combinations (e.g., Shift-A to type a capital A) by pressing the modifier key (e.g., Shift) and other keys (e.g., "A") in series, rather than holding down multiple keys simultaneously.

Table 8. Keyboard configuration parameters in Windows.

Further metrics can indicate difficulties not reflected in the number of errors. For example, difficulties with combined keystrokes (reflected by "dropping errors" in Table 6) can be indicated by a variety of behaviors, as shown in Table 7; including not using a modifier at all (e.g. subjects E, F, G, H, and J), overuse of the Caps Lock key (reflected by the "Other methods" column for subjects D and I), or repeatedly pressing the Shift key without pressing another modifier (as may be indicated by the "extra modifiers" column for subject C). The latter two difficulties would be invisible when looking only at the text output, but may still indicate a need for a tool such as Sticky Keys.

Performance measures such as these can be used by clinicians to interpret a client's performance. They could also contribute to the development of software which automatically adjusts to difficulties which a person is experiencing with his or her computer (2,4).

REFERENCES

1. Soukoreff RW, MacKenzie IS. (2003). Metrics for text entry research: An evaluation of MSD and KSPC, and a new unified error metric. *Proc. CHI 2003*.
2. Trewin, S., Pain, H. (1999). Keyboard and mouse errors due to motor disabilities. *International Journal of Human-Computer Studies*. 50:109-144.
3. Ashlock G, Koester HH, LoPresti EF, McMillan WW, Simpson RC. (2003). User-centered design of Software for Assessing Computer Usage Skills. *Proceedings of the RESNA 2003 Annual Conference.*
4. Koester HH, LoPresti EF, Simpson RC. (2005). Toward Goldilocks' Pointing Device: Determining a "Just Right" Gain Setting for Users with Physical Impairments. *Proceedings of Assets 2005*. 84-89.

ACKNOWLEDGMENTS

This work was funded by the National Institutes of Health, SBIR grant #1R43HD045015-01.

Edmund LoPresti, Ph.D.
Koester Performance Research
Pittsburgh, PA 15213
edlopresti@kpronline.com