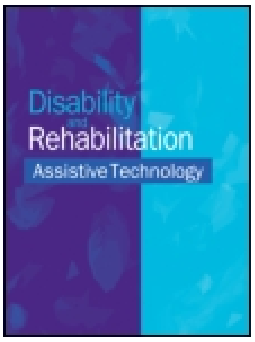

Disability and Rehabilitation: Assistive Technology

ISSN: 1748-3107 (Print) 1748-3115 (Online) Journal homepage: http://www.tandfonline.com/loi/iidt20

Toward automatic adjustment of keyboard settings for people with physical impairments

Heidi Horstmann Koester, Edmund Lopresti & Richard C. Simpson

Disability and Rehabilitation: Assistive Technology, September 2007; 2(5): 261 – 274

informa
healthcare

RESEARCH PAPER

Toward automatic adjustment of keyboard settings for people with physical impairments

HEIDI HORSTMANN KOESTER[1], EDMUND LOPRESTI[1,2] & RICHARD C. SIMPSON[3,4]

[1]*Koester Performance Research, Ann Arbor, MI, USA,* [2]*AT Sciences, Pittsburgh, PA, USA,* [3]*Department of Rehabilitation Science and Technology, University of Pittsburgh, Pittsburgh, PA, USA, and* [4]*Human Engineering Research Labs, VA Pittsburgh Healthcare System, Pittsburgh, PA, USA*

Received 27 November 2006; revised 15 February 2007; accepted 15 February 2007

Abstract

Purpose. We are developing a software system called IDA (Input Device Agent), whose goal is to optimally configure input devices for people with physical impairments. This study assessed IDA's ability to recommend three keyboard parameters in response to measurements of typing performance: repeat rate, repeat delay, and use of StickyKeys.

Method. Twelve typists with physical impairments participated. The study employed a repeated measures design. Each participant typed six sentences in each of four keyboard conditions: default settings, IDA-recommended repeat settings, StickyKeys On, and a repeat of default settings.

Results. Two participants had significant problems with inadvertent key repeats, when using the default repeat settings. For those two participants, use of the IDA-recommended repeat settings reduced the number of repeated characters by 96% and significantly improved text entry rate and typing accuracy. IDA recommended StickyKeys for six participants, each of whom had at least one problem related to modifying keys without StickyKeys. Use of StickyKeys for these individuals eliminated their modifier-related errors and significantly improved typing speed. IDA did not recommend StickyKeys for the six participants who demonstrated no need for it.

Conclusions. The results indicate that IDA can provide useful assistance with repeat settings and StickyKeys.

Keywords: *Accessibility, assistive technology, computer access, evaluation study, keyboard, physical disability, physical impairment, repeat rate, StickyKeys, user-computer interface*

Background

The importance of proper keyboard configuration

An important part of computer access interventions is appropriately choosing and configuring the user's keyboard. There are a variety of keyboards to choose from, ranging from 'standard' desktop keyboards, to keyboards with large keys, to mini keyboards. Once a given keyboard is selected, tuning it to the user's strengths and limitations may yield significant performance and comfort benefits.

Keyboard behavior can be adjusted within the Windows and Macintosh operating systems using the parameters in Table I. These settings are included in the operating system, and there may be others for third party keyboards depending on their design and associated software. The potential consequences of inappropriate settings are many [1]. For example, for someone who types with a mouthstick, not having StickyKeys active makes it cumbersome to type capital letters and impossible to use other key combinations such as shifted punctuation or Ctrl-C. For someone who can target a key reasonably well but has difficulty releasing from the key, the default settings for repeat delay will cause numerous additional characters to appear in documents. In Trewin and Pain's [2] study of keyboard users with physical disabilities, key repeat errors were the most common problem observed, comprising an average of 11% of all keystrokes.

Values for keyboard configuration settings are typically determined in one or more of the following ways. The first, and perhaps most common, is to use the default values. This scenario occurs when the

Correspondence: Heidi Koester, Koester Performance Research, 2408 Antietam, Ann Arbor, MI 48105, USA. Tel: +1 734 663 4295. E-mail: hhk@kpronline.com

ISSN 1748-3107 print/ISSN 1748-3115 online
DOI: 10.1080/17483100701284265

Table I. Keyboard configuration parameters. Parameters in **bold** are addressed in this study.

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.
SlowKeys	How long a key must be held down before it is accepted.
BounceKeys	Tells the operating system to ignore keystrokes that are depressed within x seconds of the previous key release.
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.
ToggleKeys	Gives an auditory signal when locking keys, such as Caps Lock, are depressed.

individual is using a computer without the benefit of any specific intervention relative to accessibility. Moderately inappropriate default values may result in multiple keyboarding errors and/or inability to generate certain key combinations, decreasing user performance and satisfaction. In a more extreme case, the system may be virtually unusable under the default values.

A second method of parameter selection is when the user does his or her own adjustment. This requires that the user know that these settings are available for adjustment and what to do to adjust them. This is a complex task. Performing all adjustments for the keyboard settings within Windows XP requires accessing two separate Control Panel applications and eight different tabbed panels. The terminology can be ambiguous; for example, to invoke BounceKeys, the user chooses to 'ignore repeated keystrokes', while to adjust the repeat settings, the user must select 'ignore quick keystrokes'. Another potential source of confusion is that the repeat settings can be adjusted in two different control panels, with the accessibility settings overriding the keyboard control panel settings.

Even if the user can successfully navigate the options, knowing the most appropriate values for all applicable settings may be even more difficult. Users may not understand how the parameter settings relate to the interface problems they are having, or if they do, the best choice of specific values may be unclear. Finally, interface configuration is secondary to the user's primary computer tasks; even if it can be done effectively, it takes time, physical effort, and cognitive focus away from more central tasks.

The Windows operating system includes an accessory program called the Accessibility Wizard, which does provide some help in reducing the complexity of configuration for keyboards. However, it does not include all available settings (e.g., the repeat settings are not available), nor does it give specific suggestions about how to appropriately set parameter values based on user performance.

A third scenario occurs when a clinician or teacher is available to assist with the configuration process, using clinical observations and knowledge of the possible accommodations as a guide. However, even when clinicians have the skills to do this effectively, configuration in this scenario takes time. Clinician-assisted adjustment may result in more appropriate settings for an individual, but most users with physical disabilities do not have a qualified clinician available to them. For example, Trewin and Pain [3] found that only 35% of 30 computer users with physical disabilities had a 'computer teacher' available to them.

Under each of these three approaches, it may be difficult to define appropriate settings for a user's initial configuration. It is equally if not more difficult to address changes in the user's abilities over time, which may happen over the course of a day, a month, or a year, depending on the nature of the user's disability. Current methods may lead to appropriate keyboard settings in some cases, but it does take special knowledge, additional time, and continued maintenance to do it right [1].

Previous work on keyboard configuration agents

To address the challenges of manual configuration of keyboards and other input devices, several groups have been working toward configuration agents that would support this process [3–8]. A configuration agent models a user's strengths and limitations, and based on the model, helps configure the user's input devices appropriately. In general, a configuration agent can operate in one of four modes [9], as shown in Table II. A given implementation of an agent supports one or more of these modes. The choice of the most appropriate mode depends on the technical feasibility of increased agent responsibility as well as the desirability of retaining user control.

Of the previous work on configuration agents, only Trewin and colleagues have focused on keyboard settings, with the others focusing on pointing devices and switch use. Trewin's agent for keyboard settings focuses on agent-initiated continuous monitoring of user performance [3,9–12]. The agent creates a user model based on free typing and determines settings for a range of parameters such as StickyKeys, Repeat Delay, and BounceKeys (see Table I). They have empirically evaluated the agent's recommendations with 20 keyboard users who have physical disabilities. For StickyKeys, the agent's recommendation

Table II. Matrix of operating modes for a configuration agent.

	Who initiates the change?	
	User	Agent
Who controls the change?		
User	The user initiates the configuration process with an explicit action, and the agent suggests a configuration. The user decides what if any changes should be made.	The agent continuously monitors user performance, suggesting configuration changes as needed. The user decides what if any changes should be made.
Agent	The user initiates the configuration process. The agent determines and automatically implements any configuration changes.	The agent continuously monitors user performance. The agent determines and automatically implements any configuration changes.

correlated significantly with users' opinions on how useful StickyKeys would be for them. However, the discrimination of the agent was imperfect, as nine users felt that StickyKeys was useful for them, even though the agent did not recommend it for them. For repeat delay, use of the agent-recommended settings significantly reduced but did not eliminate key repeat errors (from a total of 2610 to 151 errors) [10]. The agent accurately recommended use of BounceKeys for five of seven subjects who made bounce errors. Effects on productivity measures, such as typing speed, were not measured.

One of the challenges in Trewin's approach is that it makes inferences based on unconstrained typing tasks. The difficulty of this is shown in the fact that the agent accurately detected only 55% of inadvertent keypress errors [3]. The use of unconstrained typing tasks allows for continuous monitoring, which is less obtrusive to the user, but it may compromise the success of the agent's suggestions. An evaluation version of this keyboard agent, called the Keyboard Optimizer, is available from IBM at http://www.alphaworks.ibm.com/tech/keyboardoptimizer. It uses unconstrained typing tasks and has both user-initiated and continuous monitoring modes.

Research goals

The approaches reviewed above demonstrate that agents to assist in configuring keyboard settings are feasible and have real potential to improve computer access outcomes for users with physical impairments. However, the problem has not yet been completely solved. A complementary (and somewhat simpler) approach to that of Trewin and colleagues may yield better results, by using a specific, known typing task on which to base recommendations. This approach provides the recommendation system with better, more targeted information on which to base configuration decisions, since the system defines the measurement task.

Because we have already developed software (called Compass[1]) that presents a typing task and measures speed and accuracy during performance of that task, we are in a good position to implement such an approach [13,14]. Our overall goal is to use the performance data recorded by Compass to make appropriate keyboard configuration recommendations. This is one of three input device agents that comprise the current IDA system, with the other two in the domains of single switch scanning and pointing device use [15,16].

Our focus for this initial work is on the key repeat delay, repeat rate, and the use of StickyKeys (see Table I for definitions). These three settings were chosen because inappropriate values for them are a major cause of keyboarding errors [2] and because we believe there is a reasonably straightforward way to build a configuration agent for these settings. Additional keyboard settings will be addressed in future studies.

Hypotheses

In general terms, we want to determine whether IDA is sensitive enough to suggest changes to settings when such changes would assist the user, and specific enough to avoid suggesting changes when they are not necessary. Further, when IDA does recommend an adjustment, that adjustment should yield improvements in keyboard performance. Those general criteria for assessing IDA's appropriateness can be expressed in the following specific hypotheses:

1) For participants who experience repeating keystrokes under the default setting:
   a. IDA will recommend a longer repeat delay and a slower repeat rate compared to the default.
   b. The repeat settings recommended by IDA will result in fewer repeat episodes and fewer repeated characters.
   c. The repeat settings recommended by IDA will result in fewer typing errors and faster typing speed (due to less time correcting errors).

   d. Participants will rate their computer as easier to use with the IDA recommendation for repeat settings.
2) IDA will not recommend a longer repeat delay for participants who do not have repeating keystrokes under the default setting.
3) For participants who have difficulty using modifier keys under the default setting:
   a. IDA will recommend Sticky Keys.
   b. The use of StickyKeys will result in fewer modifier errors.
   c. The use of StickyKeys will result in fewer typing errors and faster typing speed.
   d. Participants will rate their computer as easier to use with StickyKeys.
4) IDA will not recommend StickyKeys for participants who can perform necessary modifier functions under the default setting.

Methods

Agent algorithms

We built a keyboard agent for recommending repeat settings and StickyKeys, using Trewin's work as a foundation and examining various approaches using keyboard data from a preliminary study [17]. The algorithm IDA uses for repeat delay is almost identical to Trewin and Pain's [3]. The duration of a user's key presses during typing are recorded, excluding the backspace key, arrow keys, and modifier keys. IDA calculates the mean and standard deviation (SD) of these key press length measurements. It then computes the 'raw' repeat delay (in milliseconds) as the maximum of (mean + 3*SD) and (2*mean + 50). The raw value is converted to the nearest Windows setting, rounding up. The raw repeat rate, in characters per second, is simply the reciprocal of the raw repeat delay. This is then mapped to the closest available Windows setting (rounding to the slower setting).

IDA recommends activation of Sticky Keys based on the three indices of difficulty: (1) the user's reliance on Caps Lock for capitalizing letters; (2) the number of characters left unmodified that should have been; and (3) the number of times the Shift key was pressed without modifying anything. Activation of StickyKeys is recommended if these indices exceed a particular proportion of the number of characters that should have been modified. If none of the 'StickyKeys ON' conditions are met, it recommends that StickyKeys not be activated.

This approach is similar to Trewin's approach of accumulating evidence of difficulty by observing patterns of keystrokes, but our method relies on knowing exactly what should have been typed. This allows us to know, for example, when a period in the typed text should have actually been an exclamation point.

Protocol overview

The performance of this agent was then assessed in a study with 12 keyboardists with physical impairments. The study employed a repeated measures design, in which repeat settings and StickyKeys were each within-subjects factors. Each participant typed six sentences in each of the following keyboard conditions:

1. Default_1 – a baseline condition in which all keyboard settings were at the default value for Windows XP. IDA recommendations for repeat delay, repeat rate, and StickyKeys are made at the conclusion of this condition.
2. KeyDelay – an experimental condition in which the repeat delay and repeat rate were set to the values recommended in the Default_1 condition. StickyKeys was Off.
3. SK_On – an experimental condition in which StickyKeys was On. All other keyboard settings were set to the default values.
4. Default_2 – a baseline replication in which all keyboard settings were at the default values.

The order of conditions was fixed for all subjects. Order effects were mitigated by including baseline conditions at the beginning and end of the protocol. This design also allows for single case analysis, as it includes two separate ABA designs for examining the repeat settings (Default_1, KeyDelay, Default_2) and StickyKeys (Default_1, SK_On, Default_2). The protocol was approved by the University of Pittsburgh Institutional Review Board.

Participants

Twelve keyboardists with physical impairments participated. The criteria for inclusion in the study were: ability to access a 'typical' desktop or laptop keyboard, presence of a physical impairment which affects keyboard use, and the ability to see and interpret the test stimuli. There were no requirements for prior computer or typing experience; however, ten of 12 participants reported using their computer 21 or more hours per week. Participants were recruited from the local Center for Independent Living and the United Cerebral Palsy center, and written informed consent was obtained. Table III provides some basic characteristics of the participant group, and Table IV summarizes their typing behaviors.

Pre-test interview

Participants completed a short interview with a researcher, covering basic demographic information, as well as information about their current computer use, especially keyboard use: how much they use the computer, current keyboard, current repeat rate/delay settings (if available), current Sticky Keys status (if available), and satisfaction with current setup.

Test procedures

Participants typed on a laptop computer with Windows XP and IDA installed. Participants could choose whether to use the laptop keyboard or a Dell 102-key desktop keyboard, depending on which one was closest to the type of keyboard they usually used. Prior to beginning any test conditions, a researcher oriented the participant to the typing task, to ensure participant comfort, understanding, and appropriate positioning. Participants were asked to 'Try to type the way you normally do, just as quickly and accurately as you can'.

Table III. Basic characteristics of the participant group.

Subject	Sex	Age	Diagnosis	Education
20	M	48	SCI (C5)	Grad degree
18	M	42	TBI	Some college
28	F	24	Freidrich's Ataxia	College degree
10	F	28	SCI (C7)	Grad degree
25	M	29	CP	High school
31	M	40	SCI (C5/6)	College degree
16	F	44	CP	Assoc degree
15	M	24	MD	High school
7	M	39	SCI (cervical)	Some college
30	M	24	SCI (C4/C5)	College degree
5	M	36	CP	Grad degree
4	F	69	MS	High school

SCI, spinal cord injury; TBI, traumatic brain injury; CP, cerebral palsy; MD, muscular dystrophy; MS, multiple sclerosis.

Each of the four typing conditions followed the same basic procedure. The keyboard settings were first set to those defined by the condition. Software based on Compass assessment software [13,14] presented a sentence to be typed, and the participant entered matching text in a second text field (see Figure 1). 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.

For each condition, participants first typed a single practice sentence, as a warm-up. The SK_On condition also included a brief orientation and practice with StickyKeys, to be sure participants knew how to use it properly. They then typed six unique sentences, presented one at a time. Sentence sets were developed specifically for this study, to include multiple characters requiring modifier keys, such as capital letters or certain punctuation (e.g., ? or !). A different set of six sentences was used for each condition, but sentence sets were equivalent in average word length, sentence length, and reading level. Between conditions, subjects rested for the few minutes it took for the experimenter to set the appropriate keyboard settings and prepare for the next sentence set.

Performance variables

During performance of the typing test, the Compass-IDA software recorded each key pressed, the time at which it was pressed, and how long it was held down. From these basic keystroke data, the software

Table IV. Typing characteristics of participants, as reported in pre-test interview.

Subject	Typing method	Use of Shift key (without StickyKeys)	Uses StickyKeys?	Repeat settings?
20	R index	L thumb Shift, R index character	Yes	Unsure
18	R index	L index Shift, R index character	No	Default
28	Bilateral; mostly L index and R middle fingers	Bilateral	No	Off
10	L hand	L pinky Shift, L finger character; or bilateral	No	Unsure
25	R index	L index Shift, R index character	Yes	Unsure
31	Typing splint near R index	L index Shift, R index character	Yes	Unsure
16	Bilateral index fingers; favors L hand	L thumb or pinky Shift, L index character	Yes	Unsure
15	Bilateral, multiple fingers	Bilateral	Yes	Unsure
7	R index	L thumb Shift, R index character	Yes	Off
30	Bilateral index fingers; favors R hand	L index Shift, R index character	Yes	Unsure
5	Mouthstick	Gets question mark by precise mouthstick use; otherwise cannot use Shift	Yes	Default
4	R index	Cannot use Shift without StickyKeys	No	Default

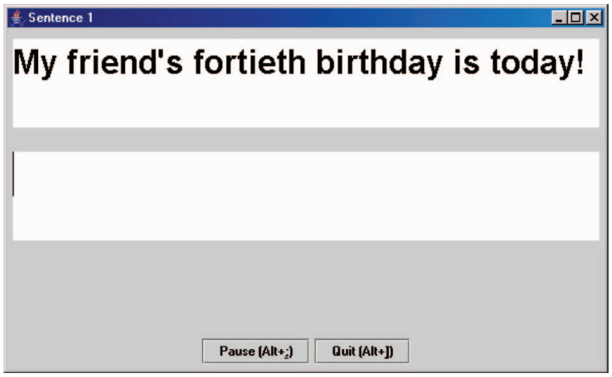

Figure 1. Screenshot of the IDA Sentence test. The text to be entered is displayed in the top text field. The participant types the sentence into the lower text field.

calculated and reported the following metrics related to overall speed and accuracy:

- TypSpeed – correct words per minute (correct characters per minute divided by five characters per word)
- TotErr – total errors made during the test, as a percentage of keystrokes entered
- NetErr – net error rate, % (includes only errors remaining at completion of trial)

It also measured and reported the following variables related specifically to key press behavior and repeated keys:

- KeyPressLength – Length of time (in milliseconds) a key is held down during a keystroke
- RepEvents – Number of episodes in which a key was held down long enough to repeat at least once
- RepChars – Number of extra characters generated across the repeat events

The following variables relate specifically to modifier key behavior:

- ShiftUsed – Number of times the Shift key was used correctly to modify a character
- CapsLockUsed – Number of times that Caps Lock was used to correctly modify a character
- CapsLockExtras – Number of extra capital letters produced, due to Caps Lock being left on
- DropLetters – Number of times a letter was not capitalized
- DropPunct – Number of times the '?' or '!' character was not shifted

The first four of these metrics were counted by the software, as they were used to make decisions about StickyKeys. They were also counted manually, along with DropPunct, since there are a few known situations where the software miscounts. The manual counts are reported here.

Post-test survey

Participants answered two subjective questions following completion of each condition. The first question asked participants to rate the ease of accurate typing during the test, on a scale from 1 to 5, with 1 being 'Difficult' and 5 being 'Easy'. The second question asked how desirable it would be to use the keyboard with these settings on a daily basis, again on a scale from 1 to 5, with 1 being 'Undesirable' and 5 being 'Desirable'.

After completing all four conditions, participants were asked for additional information about the keyboard settings they typically use, including how often the settings are adjusted, whether they know how to set them independently, and whether their typing ability changes with time. Finally, they were asked, 'How useful would it be if your computer helped adjust your keyboard settings at times to better match your abilities?' Response choices ranged from 1 to 7, with 1 = 'Not at all useful' and 7 = 'Extremely useful'.

Analyses for repeat settings

To assess IDA's ability to appropriately recommend repeat delay and repeat rate, we first identified those participants who had at least one repeat event in the Default_1 condition. Hypotheses 1 and 2 state that, for that subgroup, we expect to see an IDA-recommended delay longer than the 500-ms default value, and we expect the KeyDelay condition to provide meaningful improvement in performance and satisfaction relative to the Default_1 and Default_2 conditions. Conversely, for individuals who did **not** have any repeat events in the Default_1 condition, IDA should not increase the repeat delay, nor should the KeyDelay condition show any meaningful improvement relative to the Default_1 and Default_2 conditions. Performance variables were RepEvents, RepChars, TypSpeed, and TotErr, as well as the satisfaction responses from the post-test survey.

A 'meaningful improvement' was defined in two ways. First, if the subgroup is large enough, a paired *t*-test between the KeyDelay and pooled baseline conditions could be performed. This should demonstrate statistically significant differences for participants with repeat events, and no significant differences for participants without repeat problems. Second, if there are not enough cases for a *t*-test, each participant could be analyzed in a single case fashion. If the Default_1 and Default_2 conditions are each at least 15% different than the

KeyDelay condition, for a given performance variable, we considered that to be a meaningful improvement.

Analyses for StickyKeys

To assess IDA's ability to appropriately recommend StickyKeys, we identified those participants who had any modifier-related errors in the Default_1 or Default_2 baseline conditions. A modifier-related error was defined as any non-zero value for CapsLockExtras, DropLetters, or DropPunct. Hypothesis 3 states that, for that subgroup, we expect IDA to recommend StickyKeys, and we expect the SK_On condition to provide meaningful improvement in performance and satisfaction relative to the Default_1 and Default_2 conditions. Conversely, for individuals who did *not* have any modifier-related errors in the baseline conditions, IDA should not recommend StickyKeys, nor should the SK_On condition show any meaningful improvement relative to the Default_1 and Default_2 conditions. Performance variables are CapsLockExtras, DropLetters, DropPunct, TypSpeed, and TotErr, as well as the satisfaction responses from the post-test survey. A 'meaningful improvement' was defined as above for the repeat settings analyses.

We also compared IDA's StickyKeys recommendations to an alternative benchmark recommendation. We defined functional criteria to determine whether a participant needed StickyKeys, and compared the decisions resulting from the 'functional standard' to IDA's. The criteria for the 'functional standard' were as follows:

1. Does not need StickyKeys, if no modifier errors were made in the default condition;
2. Does need StickyKeys, if any modifier errors occurred in the default condition, and the use of StickyKeys decreased modifier errors.

Applying these criteria to both the Default_1 and Default_2 conditions for the 12 participants yields 24 'functional standard' classifications, to compare to the 24 recommendations made by IDA using Default_1 and Default_2 performance data. The degree of agreement between IDA and the functional criteria were determined using Fisher's exact test and by measuring the Kappa agreement value.

Results

Typing speed and total error rate

Tables V and VI show the typing speed and total error rate for each participant across all four study conditions. These provide the basis for the baseline performance analysis and specific comparisons across conditions that are presented in subsequent sections.

Table V. Typing speed across the four study conditions.

	Typing Speed (wpm)			
Subject	Default_1	KeyDelay	SK On	Default_2
20	11.37	13.04	14.62	14.55
18	6.05	5.76	5.36	5.07
28	4.95	5.73	4.45	4.54
10	33.87	35.53	32.38	35.22
25	2.97	2.73	2.87	2.54
31	16.18	17.58	22.50	22.59
16	8.94	8.88	11.16	9.10
15	3.03	3.63	4.01	4.22
7	4.90	5.30	6.65	4.98
30	8.61	9.32	10.98	9.75
5	7.06	6.52	7.35	7.00
4	2.56	3.17	1.98	1.80
Average	9.21	9.77	10.36	10.11
SD	8.71	9.19	9.05	9.82
Minimum	2.56	2.73	1.98	1.80
Maximum	33.87	35.53	32.38	35.22

SD, standard deviation.

Table VI. Total error rate across the four study conditions.

	Total Error Rate (%)			
Subject	Default_1	KeyDelay	SK On	Default_2
20	2.42	0.86	1.64	0.45
18	2.52	1.89	2.52	2.76
28	19.64	4.21	33.19	20.72
10	2.40	2.41	1.78	0.81
25	5.59	2.27	1.97	5.19
31	6.42	3.58	3.47	4.30
16	2.37	2.56	0.38	1.20
15	40.76	33.87	33.08	25.85
7	2.85	3.59	0.76	0.81
30	8.75	3.65	2.85	2.77
5	4.40	5.70	3.97	8.25
4	50.30	27.71	65.76	68.63
Average	12.37	7.69	12.61	11.81
SD	16.35	10.94	20.59	19.69
Minimum	2.37	0.86	0.38	0.45
Maximum	50.30	33.87	65.76	68.63

SD, standard deviation.

Baseline typing performance

Table VII shows the average typing speed and error rates across the two baseline conditions (Default_1 and Default_2) for each participant. Typing speed and total error rate varied widely across participants. While the average baseline typing speed was 9.7 wpm, the median typing speed was 6.3 wpm, and only three participants typed faster than

10 wpm. For errors committed during typing, three participants had significant difficulty, with total error rates over 20%, while the remaining nine had total error rates of approximately 5% or lower. All participants were conscientious about fixing their errors, as the net error rate averaged less than 1% and did not exceed 2.3% for any individual. (Net errors were well below 5% for every subject and every condition in the study.)

Key repeat behavior across conditions

As shown in Table VIII, three participants had at least one key repeat event in the Default_1 condition. Two in particular, P4 and P28 had notable difficulty with the default repeat delay setting, resulting in dozens of unwanted repeated characters. For both of these individuals, IDA recommended a repeat delay longer than the 500 ms default. This reduced repeat events by 100% and 63.2% for P28 and P4, respectively, with a corresponding reduction in repeated characters of 100% and 92.7%. IDA did not recommend an increased repeat delay for any other participants, and no other participant had any repeat events in the KeyDelay condition.

Table IX provides more insight into the subjects' key press behavior by showing the actual length of time each subject held a key down, during the Default_1 condition. As expected, P28 and P4, who had the most difficulty with inadvertent key repeats, had the longest key press lengths. Both averaged over 300 ms, and P4 exhibited considerable variability in press length as well. No other participant averaged over 200 ms, and two were under 100 ms.

Table VII. Baseline typing performance using default keyboard settings.

Subject	TypSpeed (wpm)	TotErr (%)	NetErr (%)
20	12.96	1.44	0.23
18	5.56	2.64	1.43
28	4.75	20.18	0.23
10	34.55	1.61	0.00
25	2.76	5.39	2.24
31	19.39	5.36	0.39
16	9.02	1.79	0.21
15	3.63	33.31	0.30
7	4.94	1.83	0.00
30	9.18	5.76	2.26
5	7.03	6.33	0.41
4	2.18	59.47	1.03
Average	9.66	12.09	0.73
SD	9.22	17.70	0.82
Minimum	2.18	1.44	0
Maximum	34.55	59.47	2.26

Data are averaged across the Default_1 and Default_2 conditions. SD, standard deviation.

Effect of IDA repeat settings

Use of the IDA repeat settings yielded a meaningful improvement in typing speed and error rate for the two subjects who had notable difficulty with the default repeat settings. While a test of statistical significance cannot be performed with two subjects, viewing each subject as an ABA single case design shows a consistent enhancement with the IDA repeat settings (see Figures 2 and 3). Typing speed improved by an average of 50% for P4 and 21% for P28. The effect on total error rate was even more pronounced, with an average reduction in errors of 52% for P4 and 79% for P28. Both variables met the criterion of at least a 15% improvement in the KeyDelay condition relative to each baseline condition.

Conversely, none of the remaining 10 participants exhibited this reversal pattern for either typing speed

Table VIII. Key repeat behavior for each participant in the Default_1 and KeyDelay conditions.

	Default_1 Condition			KeyDelay Condition		
Subject	Repeat delay (ms)	RepEvents (N)	RepChars (N)	Repeat delay (ms)	RepEvents (N)	RepChars (N)
20	500	0	0	500	0	0
18	500	0	0	500	0	0
28	500	13	26	1000	0	0
10	500	0	0	250	0	0
25	500	0	0	500	0	0
31	500	0	0	250	0	0
16	500	0	0	500	0	0
15	500	2	14	500	0	0
7	500	0	0	500	0	0
30	500	0	0	500	0	0
5	500	0	0	500	0	0
4	500	19	109	750	7	8

Default_1 used the Windows default value for key repeat delay, and KeyDelay used the IDA-recommended value.

Table IX. Statistics on key press length in the Default_1 condition for each participant.

	KeyPressLength		
Subject	Average (ms)	SD (ms)	CV (%)
20	151.7	41.5	27.4
18	125.1	20.1	16.0
28	398.9	82.7	20.7
10	95.1	23.4	24.6
25	172.4	23.2	13.5
31	75.4	14.6	19.4
16	179.6	28.9	16.1
15	176.1	42.8	24.3
7	185.2	39.8	21.5
30	130.8	37.7	28.8
5	135.9	50.7	37.3
4	316.3	143.8	45.5

Key press length is the length of time a key is held down during a keystroke. SD, standard deviation; CV, coefficient of variation.

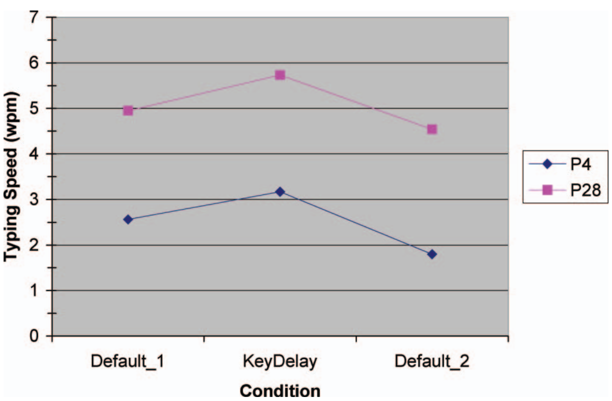

Figure 2. Effect of IDA repeat settings on Typing Speed for Participants P4 and P28.

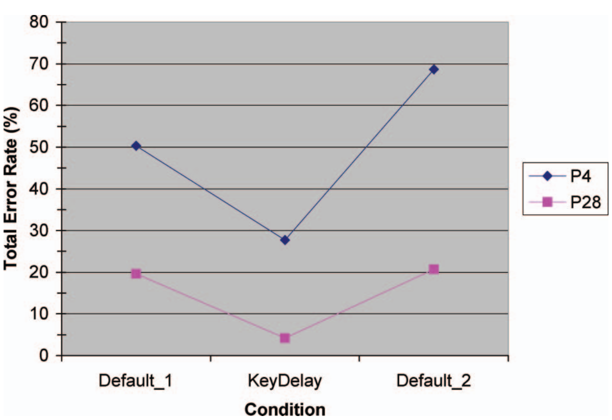

Figure 3. Effect of IDA repeat settings on Total Error Rate for Participants P4 and P28.

or total errors (see Table VI). And there were no significant differences in typing speed or total error rate in the baseline versus KeyDelay conditions. The average baseline typing speed was 10.90 wpm for this group, as compared to 10.83 in the KeyDelay condition. Similarly, total error rate averaged 6.54% in baseline, compared to 6.04% in the KeyDelay condition.

Subjective ratings generally concurred with the performance results. Since IDA did not increase the repeat delay for 10 participants, we would not expect to see a pattern of change between the baseline and KeyDelay conditions for these participants. And indeed, the average ratings for accuracy and desirability were almost identical across conditions for this group. The average accuracy ratings were 3.75 and 3.80 for baseline and KeyDelay conditions, respectively. For desirability, the average ratings were 3.60 for both baseline and KeyDelay conditions. For the two participants who did have an increased IDA repeat delay, the KeyDelay condition generally received a more positive rating, but this was only pronounced for participant P28 (see Figure 4).

Modifier key behavior without StickyKeys

Table X shows participant's modifier key behavior across both baseline conditions, in which StickyKeys was off. Six of the participants (in the first six rows of the table) had little or no difficulty using the Shift key to modify letters and punctuation. For all of these participants, IDA recommended that StickyKeys remain Off.

The remaining six participants displayed varying degrees of difficulty with modifying keys. Three of these (P16, P15, P7) used the Shift key to modify characters most of the time, but also used Caps Lock part of the time. This approach generally accomplished the task, but with a few errors, as there were a total of three dropped letters and three extra capitalizations among this group. This behavior provided mixed signals to the IDA decision algorithm, and IDA recommended the use of StickyKeys after at least one, but not both, of the baseline conditions.

Three participants (P30, P5, and P4) relied heavily on Caps Lock. P5 and P4 had particular trouble typing the question mark and exclamation point, which cannot be produced with Caps Lock, and they also generated extra capitalizations by forgetting to unlock Caps Lock. For this subgroup, with more consistent difficulty with modifier characters, IDA unequivocally recommended the use of StickyKeys.

Effect of StickyKeys use

Use of StickyKeys eliminated modifier errors and homogenized participants' modifier key behaviors. Every participant successfully used the Shift key for all modifiers in the SK_On condition, and no dropped letters or punctuation occurred for any participant. Relative to the baseline conditions,

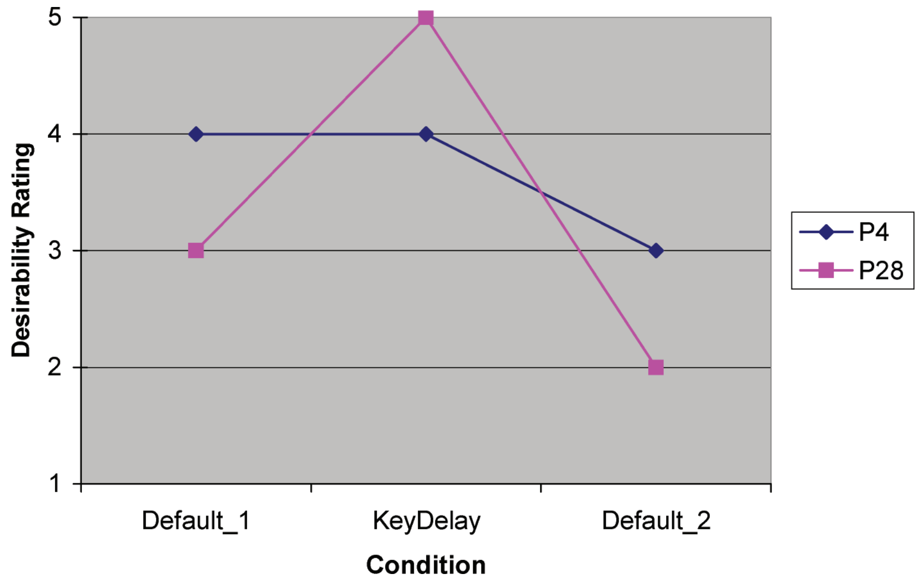

Figure 4. Effect of IDA repeat settings on Desirability Rating for Participants P4 and P28. The rating reflects how desirable it would be to use that setting on a daily basis, with 1 = Undesirable and 5 = Desirable.

Table X. Modifier key statistics, pooled across both baseline conditions, Default_1 and Default_2.

Subject	ShiftUsed (%)	CapsLockUsed (%)	CapsLockExtras (N)	DropLetters (%)	DropPunct (%)	SKeys rec'n	
						Def1	Def2
20	100	0	0	0	0	Off	Off
18	100	0	0	0	0	Off	Off
28	100	0	0	0	0	Off	Off
10	100	0	0	0	0	Off	Off
25	100	0	0	0	0	Off	Off
31	100	0	0	5.0	0	Off	Off
16	89	11	0	5.6	0	Off	**On**
15	78	20	0	3.2	0	**On**	Off
7	53	47	3	0	0	Off	**On**
30	36	64	0	7.1	0	**On**	**On**
5	10	85	14	50.0	33	**On**	**On**
4	0	94	13	0	100	**On**	**On**

IDA's recommendations for StickyKeys in each condition are also presented, as 'Skeys rec'n.'

behavior was unaffected for those who had full use of Shift without StickyKeys. For those who had relied on Caps Lock at least part of the time, StickyKeys enabled the consistent use of Shift. And for those who had had notable errors or inability to generate certain characters at baseline, none of those problems occurred during use of StickyKeys.

The subjects in the last six rows of Table X represent those for whom IDA recommended the use of StickyKeys at least once. While StickyKeys reduced modifier-related errors, it did not have a significant effect on total error rate for this subgroup, which averaged 18.1% without and 17.8% with StickyKeys ($P = 0.845$). (This is because modifier keys represented only a small portion of the total keystrokes in the typing task.) However, the use of StickyKeys significantly enhanced typing speed for this subgroup, improving it from an average of 6.0 wpm in the pooled baseline condition to 7.0 wpm with StickyKeys ($P = 0.049$). On a single case basis, only four of the six subgroup members enjoyed a consistent improvement for *each* baseline condition (see Figure 5). Two of these exceeded the 15% 'meaningful improvement' threshold for each baseline condition (P16 and P7). (P30 was close to this, but had only a 13% improvement in SK_On relative to Default_2.)

For the other six participants, for whom IDA did not recommend StickyKeys, the use of StickyKeys made no difference in typing speed or error rate. Typing speed for these six averaged 13.3 wpm in the pooled baseline condition, as compared to 13.7 wpm

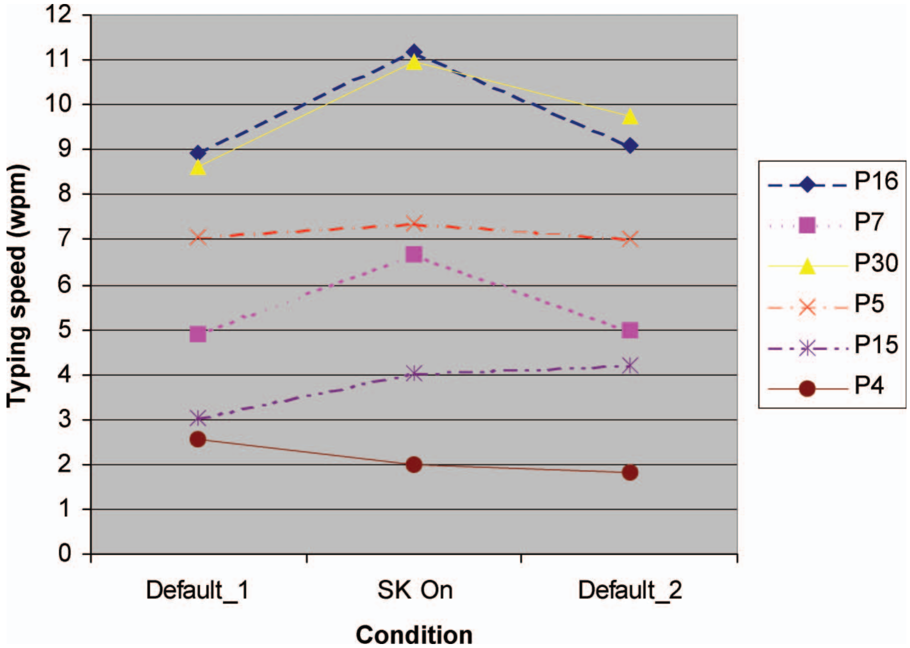

Figure 5. Typing speeds across conditions for six participants. This is the subgroup for whom IDA recommended the use of StickyKeys in at least one of two baseline conditions.

with StickyKeys, while total error rate was 6.1% without and 7.3% with StickyKeys.

Participants' subjective ratings of StickyKeys also support IDA's ability to discriminate those who need StickyKeys from those who do not. Looking at subjects for whom StickyKeys was not recommended, the ease of accurate typing was given the same rating in the SK_On condition as in the baseline conditions (mean difference = 0.0). Similarly, for this subgroup, the desirability of StickyKeys for daily use was rated no different than no StickyKeys (mean difference = 0.08). In contrast, examining subjects for whom IDA did recommend StickyKeys, the ease of accurate typing with StickyKeys was rated 4.5 on average, 1.17 points higher than the baseline rating. This subgroup also rated the desirability of StickyKeys at 4.67, significantly higher than not using it (mean difference = 1.42). Both of these differences were significant at $P < 0.05$ using Wilcoxon signed rank tests to compare the ordinal rating values.

Suitability of IDA's StickyKeys classification

Table XI compares the functional classifications for StickyKeys to the recommendations made by IDA. IDA matched the functional 'gold standard' 21 of 24 times. There were two instances in which IDA did not recommend StickyKeys but the functional standard did. Both of these were for P31, who experienced one DropLetter in each of Default_1 and Default_2 conditions. A single instance by itself is generally not enough for IDA to recommend an intervention, as compared to the relatively strict criteria for the functional standard. There was also one instance in which IDA recommended StickyKeys, but the functional standard did not. This was for P30, who performed the Default_1 modifier tasks without errors, but relied on Caps Lock frequently enough to trigger IDA's suggestion of StickyKeys use.

Table XI. Comparison of 'functional standard' classification of StickyKeys to the recommendations made by IDA.

	IDA Recommendation	
Functional standard	Off	On
Off	13	1
On	2	8

Classifications were made for the Default_1 and Default_2 conditions for the 12 participants.

Statistical analysis of the classification table confirms high agreement between IDA and the functional standard. A Fisher's exact test showed significant association between the two classification schemes ($P < 0.001$). The Kappa value of 0.739 (on a scale of 0–1) demonstrates a high level of agreement.

Survey responses

Participants generally reported high satisfaction with their current keyboard, averaging 5.4 on a scale of 1–7. Before participating in this study, nine of 12 participants knew that their keyboard settings could be changed, and all of those knew how to make the changes. All nine determined their own settings, although it is not clear if they had any initial guidance from anyone else. These responses seem to apply to use of StickyKeys, since the majority of participants did not know what their repeat settings were. All except one participant said that their 'ability to type accurately' changes from time to time, and nine people mentioned fatigue when describing that change.

After completion of the study, participants were asked to rate the usefulness of an agent that would 'help adjust their keyboard settings at times'. The average rating was 5.4, where 7 = 'Extremely useful', with a 95% confidence interval of [4.3, 6.6]. seven of 12 participants rated the usefulness as a six or seven. Paradoxically, the lowest usefulness rating (with a value of 2) came from participant P4, who was the subject who benefited the most from the agent's recommendations. It is unclear why this subject's usefulness rating was lower than expected.

Discussion

Effectiveness of IDA's recommendations

Overall, the results suggest that IDA is sensitive enough to suggest needed changes to keyboard settings, and specific enough to avoid suggesting unnecessary changes. For repeat settings, IDA lengthened the repeat delay for the two participants who had significant problems with inadvertent repeats (supporting Hypothesis 1a), and introduced no problems for the remaining 10 participants (supporting Hypothesis 2). For StickyKeys, IDA consistently recommended its use for the three subjects who relied heavily or exclusively on Caps Lock and had notable modifier-related errors (supporting Hypothesis 3a). It avoided recommending StickyKeys for the six subjects who used the Shift key exclusively and made at most one modifier error in the baseline conditions (supporting Hypothesis 4). The remaining three subjects, who used mixed strategies for modifying characters and made few if any modifier errors, did not appear to have a strong need for StickyKeys, and IDA gave mixed recommendations in those cases.

Additionally, the changes suggested by IDA generally led to an improvement in keyboard performance. Use of the IDA repeat settings reduced unintended repeat events by an average of 81%, and repeated characters by an average of 96%, for the two participants who experienced notable difficulty with the default settings (supporting Hypothesis 1b). This yielded a large and consistent reduction in total error rate, averaging 66% fewer errors with the IDA settings, as well as a 36% improvement in typing speed (supporting Hypothesis 1c).

While significant repeat problems had a relatively low incidence in this experiment, the impact of those problems for those who experienced them was high. For participant P4, the effect was particularly striking. With the default repeat settings, which she used in daily computer use, she often found herself in a vicious cycle, first making repeat errors, then deleting correct characters as well as the errors due to extra repeats of the backspace key. With a typing speed of 2 – 3 wpm, and a tendency to fatigue due to her multiple sclerosis, this extra work had considerable consequences, and was almost completely preventable by using appropriate key repeat settings.

Problems with modifier keys affected more participants, as six individuals relied on Caps Lock at least part of the time and had at least one modifier-related error. For this group, use of StickyKeys eliminated modifier errors, and led to a modest but significant 14% improvement in typing speed (supporting Hypotheses 3b and 3c). The largest improvements in typing speed were for two participants who had the ability to use the Shift key without StickyKeys, at least part of the time. Typing speed with StickyKeys averaged 35 and 24% faster for P7 and P16, respectively, as compared to the baseline conditions. For these two participants, using Shift without StickyKeys may be physically possible in many cases, but more difficult and time-consuming than with StickyKeys. Part of the benefit may also be due to a reduction in cognitive load. Since StickyKeys offers one consistent method for modifying characters, it eliminates the need to choose whether to use Shift or Caps Lock for a given character, a decision that can take time [18].

Interestingly, neither IDA's recommendations nor the functional classification (see Table XI) agreed very strongly with participant's self-reported usage of StickyKeys. Comparing IDA's recommendations to self-reported usage yields a low Kappa of 0.154 and a non-significant association (Fisher's exact test $P = 0.657$). IDA recommended StickyKeys for one participant who did not use it, P4. She had a clear need for StickyKeys, as she could not modify any keys without relying on Caps Lock. The only reason she did not already use it is that she did not know about the feature. Of 24 cases, there were nine instances where IDA did not recommend StickyKeys for participants who were already StickyKeys users. Six of these instances were for three individuals (P20, P25, P31) who were able to use the Shift key appropriately and made almost no modifier errors in the baseline conditions. While they were certainly able to type functionally without StickyKeys, it may be that they preferred to use it, or had particular tasks or contexts in which it was especially valuable. Or it may simply be that they did not really need the feature, but had it on anyway.

Limitations of the study

Significant problems with repeat settings were relatively rare in this participant group, affecting only two of 12 participants. This eliminated the possibility of statistically analyzing the effect of IDA's repeat settings across the entire group. However, using the ABA design for each individual did allow us to validly measure the impact on an individual basis. Repeat problems, while low incidence, had a high impact on typing performance for these two participants, and IDA's settings clearly improved the situation. Replicating this performance over several additional users would build additional confidence in this result.

It is interesting to note that in Trewin's study [2], the number of repeat errors was much higher than observed here; in fact, repeat errors were the most frequently observed error type and affected 16 of 20 participants. However, Trewin's counts of repeat events were based on a default repeat delay of 267 ms, considerably shorter than the 500 ms default in this study. If the shortest Windows delay of 250 ms had been used here, eight participants would have experienced inadvertent key repeats, based on their recorded keypress times for each character typed.

Note that both the repeat delay and repeat rate were adjusted by IDA in this study. The reduction observed in repeat events is associated with the increase in repeat delay, while the reduction in repeated characters is related to the combined effect of increasing the repeat delay and slowing the repeat rate. A repeat rate that is too slow could affect the efficiency and ease of using keys that are often intentionally held down, such as the cursor keys. The minor editing and navigational requirements of the typing tasks studied here did not provide much opportunity to investigate this issue, however.

Another limitation is that the functional 'gold standard' used as a benchmark for IDA's StickyKeys recommendations is not itself a perfect classification. However, it does provide a reasonable independent classification against which to compare IDA's recommendations. By itself, this comparison would not be sufficient evidence about IDA's suitability, but combining it with the direct evidence of speed and accuracy effects, as well as subjective ratings, provides consistent and cumulative evidence.

This study took place in a single experimental session and involved about 30 sentences of typing. Thus, we could not evaluate the performance of IDA's settings over a long-term timeframe, such as several hours of typing. This study was designed to demonstrate IDA's adequacy in making recommendations for initial settings, and we will need to perform additional work to see how long the initial settings remain suitable, and what the best approach is for revising those initial settings over time.

Limitations of IDA's approach

While in general, the IDA algorithms made helpful decisions about keyboard settings, the results did indicate some possible revisions. An interesting and unexpected result from the StickyKeys recommendations is that some subjects changed their modifier key behavior over time, leading to inconsistent IDA recommendations. For example, participants P16 and P7 definitely benefited from StickyKeys, but IDA recommended it for them only 1 of 2 times. It may be useful to loosen the threshold for StickyKeys recommendation so that it recommends its use under a broader range of behavior.

For the key repeat settings, P4's setting did not completely eliminate her repeat errors. A main reason for this is that her typing ability decreased fairly quickly with fatigue, so that her key presses got longer as her fatigue increased. But even in the case where key presses are highly stable, an occasional repeat error will occur in a statistically based algorithm like IDA's, where the repeat delay is based on the 98th percentile of the keypress length. Because many users want to use intentional repeats for things like cursor keys, IDA tries to keep the repeat delay as short as possible while still preventing the vast majority of inadvertent key repeats. In the future, it may be useful to allow users to set their own preference of how the balance between sensitivity and specificity should be set.

A major distinction between the current IDA system and Trewin's Keyboard Optimizer is that IDA requires the user to perform a short, specific typing task before it makes any recommendations. The Keyboard Optimizer has an option that allows the user to simply do their normal typing work while it 'observes.' IDA's separate-task approach works well in an evaluation or initial set-up situation, and when performance and behavior are relatively stable. In cases where typing ability changes significantly with time, the user would need to initiate interaction with IDA to get a new settings recommendation. We need to learn more about when users prefer the separate-task approach, and when they do not.

Implications of typing performance

One somewhat surprising result is the relatively slow baseline typing speed of some of these participants. While the average typing speed was 9.7 wpm, only three of 12 participants typed faster than 10 wpm.

Five of 12 typed slower than 5 wpm. Not coincidentally, three of these five were the only participants whose total error rate was above 10% in the default conditions, and all three had error rates above 20%. While speed and accuracy are only two indicators of the suitability of a computer access system, these results suggest that several participants might benefit from a more extensive access intervention than simply adjusting the keyboard settings. It may be that a different type of keyboard would be beneficial, or that a text entry method other than the keyboard would be more suitable. A full treatment of this issue is beyond the scope of this paper, but it does suggest that individuals ought to have a simple way to monitor how well their access system is working for them and to understand when improvements might be necessary.

Future plans

The results to date indicate that IDA can provide useful assistance with repeat settings and StickyKeys. We still need to consider the other adjustable keyboard settings, such as BounceKeys and SlowKeys, to provide a comprehensive solution.

The current system implements the tasks and decision algorithm, but we have not yet designed a user interface to facilitate use of the system. Determining the method or methods of operation that users prefer is an important next step. We look forward to continuing this work toward a scenario in which the computer is able to gracefully adapt itself to the user's abilities.

Acknowledgments

This work was supported by the National Institutes of Health, under SBIR grant #1R43HD045015-01, and by National Science Foundation grant #0133395. We thank the participants for their generous contributions of time and effort.

Note

1. Compass version 1.0, Koester Performance Research, www.kpronline.com

References

1. Anson D. Alternative computer access: a guide to selection. Philadelphia: F.A. Davis Company; 1997.
2. Trewin S, Pain H. Keyboard and mouse errors due to motor disabilities. International Journal of Human-Computer Studies 1999;50:109 – 144.
3. Trewin S, Pain H. A model of keyboard configuration requirements. Behavior and Information Technology 1999; 18(1):27 – 35.
4. McGill R. A blackboard knowledge-based approach towards implementing an adaptive force joystick computer input device for persons with tremor disability. Proceedings of the RESNA 1990 Annual Conference. Washington DC: RESNA Press; 1990. pp 431 – 432.
5. Tracey M, Winters J. Neuro-Fuzzy Advisor for Mouse Setting in Microsoft Windows. Proceedings of the BMES-EMBS 1st Joint Conference; 1999. p 664.
6. Simpson R, Koester H. Adaptive one-switch row-column scanning. IEEE Transactions on Rehabilitation Engineering 1999;7(4):464 – 473.
7. Lesher G, Higginbotham J, Moulton B. Techniques for automatically updating scanning delays. Proceedings of the RESNA Annual Conference (RESNA 2000). Arlington, VA: RESNA Press; 2000. pp 85 – 87.
8. LoPresti E, Brienza D. Automatic customization software for computer head controls. Proceedings of RESNA 2002. Arlington, VA: RESNA Press; 2002. pp 109 – 111.
9. Trewin S. Configuration agents, control and privacy. Proceedings of the Conference on Universal Usability 2000. New York: ACM; 2000. pp 9 – 16.
10. Trewin S, Pain H. Dynamic modelling of keyboard skills: supporting users with motor disabilities. Proceedings of the 6th International Conference on User Modeling. Wien, New York: Springer; 1997. pp 135 – 146.
11. Trewin S. An invisible keyguard. Proceedings of ASSETS 2002. New York: ACM; 2002. pp 143 – 149.
12. Trewin S. Automating accessibility: the dynamic keyguard. Proceedings of ASSETS 2004. New York: ACM; 2004. pp 71 – 78.
13. Ashlock G, Koester HH, LoPresti E, McMillan W, Simpson R. User-centered design of software for assessing computer usage skills. Proceedings of RESNA 2003 Annual Conference, Atlanta, GA. Arlington, VA: RESNA Press; 2003.
14. Koester HH, LoPresti E, Ashlock G, McMillan W, Moore P, Simpson. Compass: Software for computer skills assessment. Proceedings of CSUN 2003 International Conference on Technology and Persons with Disabilities. Los Angeles, CA; 2003.
15. Simpson R, Koester H, LoPresti E. Evaluation of an adaptive row/column scanning system. Technology and Disability 2006;18:1 – 12.
16. Koester HH, LoPresti E, Simpson R. Factors influencing user performance with pointing devices. Proceedings of RESNA 2006 Annual Conference, Atlanta, GA. Arlington, VA: RESNA Press; 2006.
17. LoPresti EF, Koester HH, Simpson RC. Measuring keyboard performance for people with disabilities. Proceedings of RESNA 2006 Annual Conference, Atlanta, GA. Arlington, VA: RESNA Press; 2006.
18. Olson JR, Nilsen E. Analysis of the cognition involved in spreadsheet software interaction. Human Computer Interaction 1988;3(4):309 – 350.