# Defining Acceptable Interaction for Universal Access

Simeon Keates[(✉)]

University of Greenwich, Medway Campus, Central Avenue, Chatham Maritime,
Kent, ME4 4TB, UK
s.keates@gre.ac.uk

**Abstract.** Many new assistive input systems developed to meet the needs of users with functional impairments fail to make it out of the research laboratory and into regular use by the intended end users. This paper examines some of the reasons for this and focuses particularly on whether the developers of such systems are using the correct metrics for evaluating the functional attributes of the input technologies they are designing. In particular, the paper focuses on the issue of benchmarking new assistive input systems against a baseline measure of useful interaction rate that takes allowance of factors such as input success/recognition rate, error rate, correction effort and input time. By addressing each of these measures, a more complete understanding of whether an input system is practically and functionally acceptable can be obtained.

**Keywords:** Interaction rate · Universal access · HCI · Input technologies · Error rate · Assistive technologies · Acceptability

## 1 Introduction

Much of the research into Universal Access, both past and present, has focused on the development of new and innovative assistive input device and interface design technologies for users with functional impairments. It is widely accepted that the traditional keyboard and mouse input arrangement does not serve those with a range of functional impairments well [1].

For example, a person with severe vision impairment will experience significant difficulties in using a mouse, not least because the feedback on the position of the cursor on the screen is invariably visual only. Similarly, users with motor impairments will typically experience comparable levels of difficulty, because of the challenges presented in generating the quality of limb and digit control usually required to position a mouse, click on its buttons or type on a keyboard [2]. Consequently, many researchers have taken the view that perhaps a new input device/user interface arrangement [e.g. 3] or a re-design of the device/interface [e.g. 4] may alleviate or remedy the difficulties faced by many such users.

However, while the motivation for developing new assistive input and interaction technologies is clear, the success of such devices has been mixed. It is still a common problem that many of the new technologies developed rarely progress beyond the

research laboratory. Of those that do, many end up simply collecting dust on shelves, never really used to the extent anticipated by their developers [5].

There are many reasons why individual assistive input technologies suffer this fate, but there are a few that are reliably useful indicators of the likely success or otherwise of such developments. Jakob Nielsen, for example, has identified that the success of a product depends on it meeting both practical/functional acceptability and social acceptability criteria [6]. He defines practical acceptability as including factors such as cost, reliability, utility/functionality and usability. Social acceptability considers factors such as brand identity, stigma, etc.

There is a large body of work looking at usability theory and overall acceptability of products and systems. This paper focuses on one aspect in particular, that of the challenge of establishing whether the practical acceptability offered by assistive input systems has genuinely been met.

It is accepted that one of the principal reasons for the failure of the uptake of these new solutions is that their development has typically focused on the functional/technical issues, i.e., getting the solution to work, often to the detriment of the softer/social issues, i.e., does it meet the wants, needs and/or aspirations of the users [7].

It is also correct, though, to recognize that a failure to meet the practical acceptability criteria will also translate to a failure of the product or system to succeed in the real world. For assistive input systems, one of the major difficulties has been that the functional aspect of the development often only considers a narrow part of the interaction process as the metric of success. In many cases, this is usually input recognition rate [e.g. 8]. This paper explores the possibility of developing more complete measure of the functionality of new input devices.

## 2 Functional Impairments and Computer Access

There are several approaches to categorizing types of functional impairment. One of the most straightforward was inspired by the work of Card, Moran and Newell on the Model Human Processor [9]. Effectively, they proposed a model of interaction that consists of three elements:

$$\text{Total time} = x\,\tau_{\mathbf{p}} + y\,\tau_{\mathbf{c}} + z\,\tau_{\mathbf{m}} \tag{1}$$

In this equation, $x$, $y$ and $z$ are integers and $\tau_{\mathbf{p}}$, $\tau_{\mathbf{c}}$ and $\tau_{\mathbf{m}}$ correspond to the times for single occurrences of the perceptual, cognitive and motor functions respectively. It is possible to categorize impairments along these lines of functionality.

Perceptual impairments are those that affect a user's ability to perceive the state of the world around them and are principally focused on the five senses. In the case of computer access, the human senses of most interest are vision and hearing [10]. Indeed, vision impairments have received arguably the lion's share of research effort and also have the most successful assistive technologies to facilitate better interaction, with products such as JAWS achieving strong market positions [11]. Blindness and low vision present challenges with most stages of human-computer interaction, from input actions,

such as text entry and cursor control, to perceiving output, such as reading text on a screen or interpreting a figure or diagram.

Cognitive impairments are those that affect the user's ability to understand or respond to the state of the world around them. Such impairments can include memory loss or reduction, learning and communication difficulties and executive function limitations [12]. It is often argued that cognitive impairments are the most "hidden" ones, since their presence is often more difficult to identify and, once identified, to also diagnose. However, they are beginning to be researched more frequently [e.g. 13] than, say, 10 years ago. Typical solutions can include personalized diaries and reminders for medication and other reminders, assistive word processors for help with typing and dialogue structures, etc. More innovative solutions include emotion and affective state recognition to assist people with Asperger's and forms of autism [14] and also deep question and answer systems, such as IBM Watson [15].

Finally, motor impairments can create difficulties with both text entry and cursor control in a typical computer interaction scenario [1]. Symptoms such as tremor, spasm, restricted range of motion and weakened muscles can make both gross and fine motor control a challenge [1].

Text entry assistance typically focuses on making keyboards more accessible through physical assistance, for example adding keyguards, or using "soft" on-screen keyboards or replacements, such as Dasher [16]. On-screen, soft keyboards are usually activated by a dwell time function (in the case of a cursor control replacement system) or some form of binary switch/scanning combination [5].

Cursor assistance can be in the form of adapted mouse replacement devices, such as tablets or specially design mice/joysticks/trackballs [1]. One area of particular promise is that of haptic assistance, such as through the addition of "gravity" to on-screen targets [17]. Other approaches include adapting or altering the processing of the cursor input stream to make targets more "sticky" by slowing the cursor down over the targets or by fixing a mouse button activation to the location of the button down event, not the button up one [18]. More radical solutions involve changing the input paradigm from the usual windows/icons approach to that of using gestures for the input [19], for example.

As can be seen, there are many forms of functional impairments that can affect human-computer interaction adversely and present specific challenges to particular users. There are also many forms of potential assistance, each of which offer their own particular combination of strengths and weaknesses. As discussed earlier, not all of these assistive solutions are successful in the wild, so the question then becomes whether there are more effective methods for identifying or predicting whether a particular solution has a genuine chance of successful adoption by users in real world circumstances.

As regards determining the social acceptability of a new technology or product, approaches such as focus groups, user evaluations, etc., would usually be used [20]. These methods are generally well understood and widely accepted. However, there is less of a consensus on methods of evaluating the practical acceptability of novel interaction technologies.

## 3   Establishing a Measure of "Acceptable" Interaction

Most research papers addressing the development of novel input systems or interaction paradigms usually focus on only one or two measures of success, principally the rate of successful completion of a specified task, such as clicking on a target or producing a particular gesture that is recognized correctly by the computer. While clearly a very important measure, focusing on this metric only can lead to an exaggerated view of the efficacy of the new input system/interface. There are other important factors to consider, such as the definition of usability used by ISO [21]:

- Efficiency – i.e. the time taken and effort expended to complete a task.
- Effectiveness – i.e. the ability to complete the task.
- Satisfaction – i.e. user contentedness with the interaction.

Using these definitions, satisfaction is typically measured through user surveys, interviews, questionnaires, etc., after completing a series of tasks using the new technology [7]. Efficiency is usually calculated by looking that the task completion rates and times. Measuring effectiveness involves looking at error rates and effort expended to correct for any errors that occur as well as proportion of tasks completed [22]. However, while research papers addressing the development of assistive input systems that include some form of user evaluation with the prototype system usually include a summary of task completion times (i.e. a variant of the efficiency metric above) and task completion rates (i.e. a partial treatment of the effectiveness metric), it is less common to find an exploration of the frequency of errors. It is even less common to find an analysis of the impact of those errors, with some experimental designs not even recognizing the presence of errors.

Even in the comparatively rare instances where such analyses exist, it can be argued that the final piece of the jigsaw is still missing – i.e. a comparison with an accepted baseline measure. Fundamentally, even where the developers do such analysis, they often fail to reflect on whether the assistive input system that they have developed meets an acceptable level of interaction. It is all well and good to say that it takes $x$ seconds to complete a task, with an error rate of $y$ %, but the real question is whether those task completion and error rates are acceptable to the intended end users [23]. Basically, the question that really needs to be asked is:

- *Does this new assistive input system equal or outperform the other systems available to the end users?*

If the answer to this question is negative then that immediately casts doubt upon the likely successful adoption of the system being developed by users outside of the research laboratory. Fundamentally, if users can obtain better interaction rates using an existing, and most likely proven, assistive input system then they are less likely to wish to switch to a new or different one.

Even where the answer to the question above is positive, there is still a further question to be asked:

- *Does this new assistive input system meet the full needs, wants and aspirations of the end users?*

A simpler, more direct formulation of this question is:

- *Is this new assistive input system good enough?*

The reason for asking this second question is that for users with more severe impairments there may not be a suitable or practical input system readily available. However, in all but the most extreme cases, some form of input is usually possible through the use of simple binary, i.e. on/off, switches and a scanning on-screen keyboard. Consequently, it can be argued that the very minimum target for user acceptance of a new assistive input system is that it should at least outperform the scanning/binary switch input approach.

### 3.1   Measuring Text Input

Text input is typically reported in terms of words per minute [e.g. 24]. It may also be reported as characters per minute, if that is a more meaningful metric, such as when typing rates are unusually slow or where a more detailed analysis is required [25].

However, defining a "word" is not straightforward. Many approaches simply assume that a word is 5 characters in length, with a following space implicitly (5 characters) or explicitly (5 + 1 characters) associated with it. In many modern systems, the impact of word prediction systems needs to be considered. It is not clear how often users need to actually enter all 5 characters to make a word when a predictive system is also being used, thus raising a question over the calculations made using the 5 or 5 + 1 assumptions.

There is a choice to be made over how to handle errors. Some researchers simply choose to ignore that errors may exist, e.g. by not supporting or allowing error correction in the design of the experiment. Others remove words with errors in them from the data analysis. Neither of these are ideal solutions when looking at users with motor impairments where errors will most often carry a significant correction penalty, i.e. the amount of effort required to correct any errors will be non-trivial, and also where the frequency of errors can be expected to be significant.

Where errors are identified, they are typically reported through metrics that capture deviations from the expected minimum, error-free input, such as Mean String Distance (MSD) or Keystrokes per Character (KSPC) [26]:

$$MSD \approx \frac{INF}{C + INF} * 100\% \tag{2}$$

$$MSD \approx \frac{C + INF + IF + F}{C + INF} \tag{3}$$

where INF = Incorrect and Not Fixed character entries, IF = Incorrect but Fixed, F = Fixing non-character entries (e.g. a backspace or other edit function) and C = Correct character entries. Other measures are possible [26], but are not used as often as MSD and KSPC.

## 3.2  Measuring Cursor Input

The most common approach to measuring cursor input is to use a Fitts' Law type experiment. Fitts' Law has undergone a number of modifications since first proposed and one of the more common formulations is the Shannon one [e.g. 27]:

$$Movement\_Time = a + b * ID \tag{4}$$

where $a$ and $b$ are constants and the Index of Difficulty (ID) is:

$$ID = \log_2 \left( \frac{A}{W} + 1 \right) \tag{5}$$

in which $A$ is the distance travelled towards the target and $W$ is the width of the target along the direction of travel.

Although experiments have confirmed that Fitts' Law can be applied to users with motor impairments, there is again little explicit handling of errors. A more sophisticated set of cursor measures has been developed to look at the detail of the quality of cursor control [28] and these measures have been applied successfully to examine the quality of cursor control for users with severe motor impairments [29]. Again, though, while these measures can tell a lot about what is happening to the cursor input, they do not necessarily help researchers and designers determine if the quality of the input is sufficiently good by themselves. There is a clear need for a baseline measure to compare against.

## 3.3  Measuring Overall Interaction Rate

As can be seen from the discussion above, there are many ways of examining the details of human-computer interaction. However, while those methods may make good research tools, they do not typically answer the question raised earlier – specifically: *is the input system good enough*?

To answer this question succinctly, a simple metric needs to be considered, one that can help a developer or researcher know immediately if the new system is operating in the correct ballpark. A likely candidate for such a measure is the bit rate of useful information transfer between the user and the computer utilizing the assistive input system.

An example of how such a calculation can be made is illustrated by a gesture recognition system [30]. In that experiment, users were able to generate a range of possible gestures (the vocabulary). Rather than using a simple recognition rate, a scoring system was implemented where correctly recognized gestures were scored as +1, non-recognized gestures were scored as a 0 or null return and misrecognized gestures were scored as −1 to reflect that a corrective action would be needed to fix the error. The overall input samples gathered from each user were then normalized and scaled to a range of −100 to +100 to remove any data collection inconsistencies.

That score was then combined with the vocabulary size and the time taken to produce and recognize each gesture into a single measure, the bit rate of useful information transfer between the user and the system:

$$Bit\_rate = \frac{\log_2 (Vocabulary\_size) * \frac{Score}{100}}{Time\_taken} \qquad (6)$$

It can be seen from the formulation of Eq. (6) that a system scoring 0 or less will not generate any useful bit rate since the user will be permanently trying to correct incorrect inputs, which is intuitively correct.

### 3.4   Benchmarking the Interaction Rate

If the notion of the bit rate of useful information transfer is taken as the most appropriate measure for benchmarking the practical acceptability of an assistive input system, then it is further possible to establish a baseline to compare the bit rate against.

As discussed earlier, the most basic working input system for almost all users with severe motor impairments is the simple binary switch used in conjunction with a scanning on-screen keyboard. Each successful binary switch input will generate 1 bit of information by definition. It is known from the work on the Model Human Processor [9] that for an able-bodied user the typical response time to a stimulus is $\approx 250$ ms, where the perceptual response time ($\tau_p$) $\approx 100$ ms, cognitive cycle time ($\tau_c$) $\approx 70$ ms and motor response time ($\tau_m$) $\approx 70$ ms. Thus, if we assume no prediction, the idealized input interaction for an able-bodied user would look something like:

$$Time\_per\_input = \tau_p \text{ [see the choice]} + 2\tau_c \text{ [identify each of the options]}$$
$$+ \tau_c \text{ [decide on which option]} + \tau_m \text{ [operate the switch]} + f(t) \qquad (7)$$

where f(**t**) is the mean time for the scanning input to land on the option to be selected. In the limiting case, and without the ability to predict ahead, the fastest scanning speed possible is anticipated to be 250 ms per target. If standard able-bodied performance parameters are used in Eq. (7), the mean idealized time per bit of useful information using such a scanning keyboard is approximately 100 ms + 140 ms + 70 ms + 70 ms + 250 ms = 630 ms, giving a useful information transfer bandwidth of (1/0.63) = 1.59 bits/s. For comparison, the bits rates seen for the gesture recognition system used in [30] ranged from 0.56 bits/s to 0.77 bits/s.

Of course, the values used in (7) above were derived for able-bodied users. The comparable values for motor impaired users have also been determined empirically [31]. Typical values for each of the Model Human Processor parameters were found to be: perceptual response time ($\tau_p$) $\approx 100$ ms, cognitive cycle time ($\tau_c$) $\approx 110$ ms and motor response time ($\tau_c$) $\approx 110, 210$ or 310 ms, depending on the severity of the impairment. From these values, it can be seen that a baseline idealized interaction time for the binary switch/scanning input is approximately 100 ms + 220 ms + 110 ms + 110|210| 310 ms + f(t). Note that f(t) may have to be varied to allow for the range of reaction times, i.e. 320 ms, 420 ms or 520 ms depending on the severity of the impairment and thus also the associated motor function time.

Consequently, using these assumptions, the best-case interaction rate for a user with a motor impairment is (1/0.86) = 1.16 bits/s (based on $\tau_m$ = 110 ms). For users with

severe motor impairments, that rate decreases to $(1/1.05) = 0.95$ bits/s. It can be seen from these calculations that the binary switch and scanning input outperforms the gesture input system described in [30].

## 4    Conclusions

To improve the success of assistive input systems outside of the research laboratory, it is necessary for researchers and developers to take a more sophisticated view of how well the systems that they are developing genuinely meet the needs of the users. While methods for assessing the social acceptability of such systems are widely understood, although not necessarily undertaken, there is much more variability over the approaches to measure the practical acceptability of such systems.

This paper has discussed the notion of focusing on a single measure, *the bit rate of useful information transfer*, as a possible more sophisticated metric than measures such as recognition rate. It has also introduced a method for establishing a straightforward baseline for such a measure to be compared with.

The measure and baseline can be further improved – this paper proposes them as a work in progress and not as a definitive set of baselines. For example, the scoring system used in [30] could be modified to penalize incorrect recognitions further to better reflect the effort required to correct an error. Equally, the approximations for f(**t**) in Eq. (7) should be determined empirically and an error rate could be introduced. In practice, it would be rare for a user to be faced with an unexpected choice for the scanning input. An element of prediction and anticipation would usually be expected, where f(**t**) could perhaps tend to significant reductions in the times used above.

Overall, though, the use of such a metric would help designers and researchers understand the likely success or otherwise of a new assistive input system more clearly than the metrics that currently prevail.

## References

1. Keates, S.: Motor Impaired Users and Universal Access. Universal Access Handbook, pp. 5.1–5.14. Taylor and Francis, London (2009)
2. Paradise, J., Trewin, S., Keates, S.: Using pointing devices: difficulties encountered and strategies employed. In: Proceedings of the 3rd International Conference on Universal Access in Human-Computer Interaction, Las Vegas, NV, July 2005, pp. 22–27 (2005)
3. Keates, S., Trewin, S.: Effects of age and Parkinson's disease on cursor positioning using a mouse. In: Proceedings of 7th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2005), Baltimore, MD, October 2005, pp. 68–75 (2005)
4. Hanson, V., Brezin, J.P., Crayne, S., Keates, S., Kjeldsen, R., Richards, J.T., Swart, C., Trewin, S.: Improving web accessibility through an enhanced open-source browser. IEEE/IBM Syst. J. (Spec. Ed. Accessibility) **44**(3), 573–588 (2005). IEEE
5. Keates, S., Potter, R., Perricos, C., Robinson, P.: Gesture recognition - research and clinical perspectives. In: Proceedings of RESNA 1997, Pittsburgh, PA, pp. 333–335 (1997)
6. Nielsen, J.: Usability Engineering. Morgan Kaufman, San Francisco (1993)

7. Keates, S.: Designing for Accessibility: A Business Guide to Countering Design Exclusion. CRC Press, Boca Raton (2007)

8. Tsourakis, N.: Using hand gestures to control mobile spoken dialogue systems. Int. J. Univ. Access Inf. Soc. **13**(3), 257–275 (2014). Springer

9. Card, S.K., Moran, T.P., Newell, A.: The Psychology of Human-Computer Interaction. Lawrence Erlbaum Associates, Hillsdale (1983)

10. Keates, S., Clarkson, P.J.: Countering design exclusion: bridging the gap between usability and accessibility. Int. J. Univ. Access Inf. Soc. **2**(3), 215–225 (2003). Springer

11. Freedom Scientific: JAWS for windows: screen reading software (2014). http://www.freedomscientific.com/Products/Blindness/JAWS

12. Keates, S., Adams, R., Bodine, C., Czaja, S., Gordon, W., Gregor, P., Hacker, E., Hanson, V., Kemp, J., Laff, M., Lewis, C., Pieper, M., Richards, J., Rose, D., Savidis, A., Schultz, G., Snayd, P., Trewin, S., Varker, P.: Cognitive and learning difficulties and how they affect access to IT systems. Int. J. Univ. Access Inf. Soc. **5**(4), 329–339 (2007). Springer

13. Keates, S., Kozloski, J., Varker, P.: Cognitive impairments, HCI and daily living. In: Stephanidis, C. (ed.) Universal Access in HCI, Part I, HCII 2009. LNCS, vol. 5614, pp. 366–374. Springer, Heidelberg (2009)

14. El Kaliouby, R., Robinson, P., Keates, S.: Temporal context and the recognition of emotion from facial expression. In: Proceedings of HCI International 2003, Crete, Greece, pp. 631–635 (2003)

15. Keates, S., Varker, P., Spowart, F.: Human-machine design considerations in advanced machine-learning systems. IEEE/IBM J. Res. Dev. **55**(5), 4:1–4:10 (2011). IEEE

16. Welton, T., Brown, D.J., Evett, L., Sherkat, N.: A brain-computer interface for the Dasher alternative text entry system. Int. J. Univ. Access Inf. Soc. Springer, (Springer online first) (2015)

17. Hwang, F., Keates, S., Langdon, P., Clarkson, P.J.: Movement time for motion-impaired users assisted by force-feedback: effects of movement amplitude, target width, and gravity well width. Int. J. Univ. Access Inf. Soc. **4**(2), 85–95 (2005). Springer

18. Trewin, S., Keates, S., Moffatt, K.: Individual responses to the steady clicks cursor assistance technique. Disabil. Rehabil. Assist. Technol. Informa Healthcare **3**(1&2), 2–21 (2008)

19. Keates, S., Perricos, C.: Gesture as a means of computer access. Commun. Matters **10**(1), 17–19 (1996)

20. Keates, S.: A pedagogical example of teaching universal access. Int. J. Univ. Access Inf. Soc. **14**(1), 97–110 (2015). Springer

21. International Standard Organization (ISO) ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs), Part 11: Guidance on Usability Specification and Measures. Technical report, ISO, Geneva (1998)

22. Frokjaer, E., Hertzum, M., Hornbaek, K.: Measuring usability: are effectiveness, efficiency and satisfaction really correlated? In: Proceedings of CHI 2000, The Hague, NL, pp. 345–352 (2000)

23. Keates, S., Clarkson, P.J.: Countering Design Exclusion: An Introduction to Inclusive Design. Springer, London (2003)

24. Rodrigues, E., Carreira, M., Goncalves, D.: Enhancing typing performance of older adults on tablets. Int. J. Univ. Access Inf. Soc. Springer (Springer Online First) (2014). doi:10.1007/s10209-014-0394-8

25. MacKenzie, I.S., Soukeroff, R.W.: A character level error analysis technique for evaluating text entry methods. In: Proceedings of the Second Nordic Conference on Human-Computer Interaction (NordiCHI 2002), pp. 243–246. ACM (2002)

26. Soukoreff, R.W., MacKenzie, I.S.: Metrics for text entry research: an evaluation of MSD and KSPC, and a new unified error metric. In: Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2003), Fort Lauderdale, FL, pp. 113–120 (2003)
27. Hwang, F., Keates, S., Langdon, P., Clarkson, P.J.: A submovement analysis of cursor trajectories. Behav. Inf. Technol. (BIT) **24**(3), 205–217 (2005). Taylor & Francis
28. MacKenzie, I.S., Kauppinen, T., Silfverberg, M.: Accuracy measures for evaluating computer pointing devices. In: Proceedings of CHI 2001, Seattle, WA, pp. 9–15 (2001)
29. Keates, S., Hwang, F., Langdon, P., Clarkson, P.J., Robinson, P.: The use of cursor measures for motion-impaired computer users. Int. J. Univ. Access Inf. Soc. (UAIS) **2**(1), 18–29 (2002). Springer
30. Keates, S., Robinson, P.: Gestures and multimodal input. Behav. Inf. Technol. **18**(1), 36–44 (1999). Taylor and Francis Ltd.
31. Keates, S., Langdon, P., Clarkson, P.J., Robinson, P.: User models and user physical capability. User Model. User-Adap. Inter. (UMUAI) **12**(2–3), 139–169 (2002). Wolters Kluwer Publishers