On the Assessment of the Interaction Quality of Users with Cerebral Palsy

C. Mauri and T. Granollers
GRIHO Research Group
Lleida University
C/ de Jaume II, 69, Campus Cappont
E-25001 Lleida, Catalonia, Spain
cesar.mauri@urv.cat, tonig@diei.UdL.es

A. Solanas
Dept. of Computer Engineering and Maths,
Universitat Rovira i Virgili,
Av. Països Catalans 26,
E-43007 Tarragona, Catalonia, Spain
agusti.solanas@urv.cat

Abstract

This paper is the continuation of a series of related work about experimentation of alternative ways of interaction with computers for disabled people (concretely with users suffering from cerebral palsy). In order to define an effective methodology to evaluate the usability and/or accessibility levels for this user profile, here, we start studying and evaluating (a) different interactive devices and interaction techniques that enable the use of computers for this people and, (b) existing metric techniques from human-computer interaction field. With this basis the article proposes a new evaluation method that simplifies the existing ones reducing the complexity to the minimum. This proposal is extensively explained and field research will soon be started.

1. Introduction

Human-computer interaction is a well-known discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them [1]. Although most users have standard motion and mental capabilities, some of them have important physical and mental handicaps which make their interaction with a computer a real challenge.

In this article, we follow the research line initiated in [2][3][4][5], and we focus on the interaction between interactive systems (*e.g* computers) and users suffering from cerebral palsy. These handicapped users interact with computers through standard or specialized devices, and the quality of their interaction is mainly evaluated by accessibility professionals in a qualitative manner. Although the qualitative assessment could be useful in some cases, we believe that it is necessary to define a quantitative criterion (*i.e.* a set of measurements) for determining the quality of the interaction. The existence of this criterion would pave the way for a real scientific approach to the problem.

By using a set of quantitative measurements, the professionals will be able to better assess the quality of the interaction and will improve their ability to determine which is the best interactive device for a given user.

Due to the limited interaction capability of users suffering from cerebral palsy, an assessment measurement must take into account the next factors:

- Cognitive capabilities: Depending on the severity of the mental illness, users can suffer from cognitive problems, attention limitations, memory loss, etc.
- Sensing and moving capabilities: Limitations on the sensing and moving capabilities are quite common and avert the use of some measurements. Moreover, these limitations reduce the number of applicable devices.
- Previous knowledge: Some users have patterns of behaviour that can negatively affect their interaction when using different devices. Generally, some users with cerebral palsy have learned how to manage a wheelchair by using a joystick, and this knowledge can be detrimental when using this device for controlling a computer.
- Learning progression: Users will improve their interaction by practising, and the measurements must be able to precisely assess their improvements.

In this paper we give a view of the state of the art in the field of human-computer interaction with users suffering from cerebral palsy. We evaluate the existing methods and propose new ones.

The proposed method differ from the existing ones because it is based on ISO standard metrics and uses novelty HCI metrics from the literature. Also this method is specifically designed for people with cerebral palsy. Special characteristics of this people group are taken into account.

The rest of the article is organized as follows. Section 2 gives a classification of the input methodologies and devices used by very handicapped people when interacting with computers. A view of the state of the art is given in section 3. The proposed method is described in section 4. Section 5 is a conclusion.

2. Interactive devices and interaction techniques

We next give a classification of the input methodologies and devices:

- Discrete order interaction: This kind of interaction is mainly achieved by means of a switch-based device:
  - Mechanical switch: This device is very simple. It consist of a simple switch that closes and electrical circuit when activated.
  - Blow-based switch: This device has been designed for being used by people with very important disabilities but with some capability for controlling their breathing. The device consist of a tube through which the user can blow or suck to activate the switch.
  - Inclination-based switch: This switch is activated when an inclination change is produced. Generally, this kind of devices are placed on the arm or head of the user and are mainly based on a mercury measurer.
  - Sound-based switch: In this case the switch is activated by an arbitrary or predefined sound that the user generates.
  - Artificial vision-based switch: These systems use a camera connected to a computer. The images captured by the camera are analysed in order to detect movements (*i.e.* changes in the images). The detection of any relevant change activates the switch.
- Multiple Discrete order interaction: This interaction is achieved by means of using a number of discrete devices. Using a greater number of discrete devices let the user to obtain a better interaction. The Multiple discrete order devices can be also used as a simple discrete order device, by using only one of their switches. Next we give a list of this kind of devices:
  - Keyboard: This device is without a doubt the most widely used since the end of the 19th century. Its most common form is the QWERTY, which is an inheritance of old typewriters.
  - Digital joystick: It is mainly based on 4 micro-switches placed at the bottom of the device. This micro-switches are activated when the joystick is moved, and it can handle 8 movement directions.
  - Array of switches: It is basically a set of switches distributed over a platform. Each switch has a special function (*e.g.* controlling the cursor movement).
- Force-based interaction: This kind of interaction is mainly achieved by exerting a force over a surface or an object. Generally, these devices are used to move a pointer in a screen. Next we give a couple of examples of this kind of device:
  - Analogical joystick: This device has a similar shape to the digital joystick. However, it uses potentiometers instead of switches. The value returned by the potentiometers is proportional to the exerted force.
  - Pointing stick: It is also known as isometric joystick. The pointer is moved by applying force to the stick. The stick itself does not move, or moves very little. The most common input-output mapping is known as ”velocity-control”, whereby the applied force controls the velocity of the pointer. In computing technology, isometric joysticks are usually associated with notebook computers [6].
- Movement-based interaction: This kind of interaction is based on the movement of some parts of the body of the user. Devices implementing this kind of interaction are mainly used for moving a pointer in a screen. Some common examples of this kind of device are mouses, trackballs, touch pads, head tracking systems, tablets and eye tracking systems.

Actually, it is possible to achieve interaction with almost all the devices described (*e.g.* it is possible to look for a word on the Internet by using a simple switch). Thus, the election of one device or another will mostly depend on the motion and cognitive capabilities of the user.

3. Previous works on HCI metrics

Collecting quantitative (performance) measures from human-computer interaction sessions is not a new idea. First studies were carried out in the earlies of the twenty century. For instance [7] describes an error analysis when using a QWERTY typewriter. When the mouse was invented during the sixties, various usability tests were conducted [8] and variables like the error rate or task time were studied.

More recently and, in the scope of special needs, several human-computer interaction performance analysis works were proposed. Many of them are conceived as tools for helping professionals during the assistive technology assessment process. These systems propose several interactive tasks to the user that must be performed as best and as quick as possible. During the interactive session the computer collects several parameters like time or accuracy. These data provide quantitative information to the professional to help in the assessment process. Although there are many proposed systems (see [9] for a survey), we want to highlight the Single Switch Performance Test [10], the Compass project [11] and the ERGOLAB system [12, 13].

The Single Switch Performance Test is a piece of software to facilitate the measurement of the ability to activate a single switch. SSPT software measures the average time required to activate or release the switch and also can measure the speed of repetitive activations. SSPT provides simple performance data. It is available on the Internet for free.

The Compass project is a software that measures the performance in skills needed for computer interaction. It includes eight skill tests, covering text entry, pointing device input, and switch use. Compass collects speed and accuracy data during the performance test and reports the results. Compass is available as shareware.

ERGOLAB is a software that also measures human-computer interaction performance variables. ERGOLAB includes three progressive difficulty steps with several exercises on each. The first step analyses switch based interaction for simple activation, maintaining button activation and scanning selection. The second step analyses the use of a pointing device for pointing and dragging. The third step consist of exercises based in the manipulation of hidden data. ERGOLAB was specifically conceived for people with cerebral palsy, thus part of our proposal is based on it.

The ISO 9241 [14] is an standard covering a number of aspects for people working with computers. The standard is divided in 17 parts. Part 4, entitled "Keyboard requirements" defines the physical characteristics of the keyboard and qualitative performance measures for text input. Part 9, entitled "Requirements for non-keyboard input devices" also defines procedures and measures to test efficiency and effectiveness for pointing devices. None of the previous works described above uses the ISO measures.

4. Proposed evaluation method

The proposed evaluation method simplifies the existing ones reducing the complexity to the minimum. We focus on simple interactive tasks, in the sense that these are parts of more complex ones. There are good reasons to choose this kind of tasks. First, simple tasks are more appropriate for users suffering from cerebral palsy because they are easy to learn and reproduce even when severe impairments are present. Second, metrics are simple and clear and the tasks are easily reproducible, so they are useful to compare interactive devices. Finally, these measures establish the basis for developing more complex ones.

This section describes the five proposed interactive tasks that conform the test. Performance data will be collected during the interactive sessions with real users. This information will be stored in a database for further analysis.

Each task could be configured to run in demo or in production mode. The demo mode is used to instruct the user how to perform the task and no data is collected. In production mode the user should perform the task by itself without assistance and performance data is collected and stored in the database.

The only limitations that should prevent the user from executing a task are the cognitive impairments or educational stage (illiteracy for instance). Physical impairments can be overcome if the appropriate assistive technology is used.

4.1. Task 1. Simple response to an stimulus

A visual and/or auditive stimulus is presented to the user. The stimulus can be specific (relative's photo or voice recording, toy drawing, bell ring, music, etc.) or abstract (geometric shapes, colour changes, synthetic sounds, etc.). The decision of the appropriate input modality (visual and/or auditive) and the type of stimulus (specific or abstract) depends on factors like visual, auditive and cognitive capabilities and motivation. For users with deeper affectations it should be interesting to use black and white images because colour can distract them [15]. A discrete input interactive device is required.

Once the stimulus is presented, the user should respond as fast as possible, typically pressing a switch. Optionally, after the answer, the system can present a reinforcement to the user in the form of a new rewarding stimulus. This process will be repeated a finite number of times. The assistant should configure the number of iterations accordingly to the user. The time elapsed between the user response and the presentation of new stimuli is chosen randomly from an interval of possible values.

The measures to be collected and stored are:

- Response time: It is the time elapsed between the stimulus presentation and the user response.
- Number of iterations.
- Total time to complete the task.
- Error count: We consider errors the responses (switch activations) without previous presentation of stimulus.

The response time and the total time are common measures used in all other systems. The error count came from Compass system.

4.2. Task 2. Maintaining a discrete response

This task is similar to the previous one, but with the difference that now it's necessary to maintain the response (pressing the switch) until the stimulus disappears. The time that the stimulus lasts varies randomly inside a configurable range. Optionally the stimulus itself can point out the remaining time somehow like a progress bar or sound scale. This task can be useful for more complex ones like Morse code based interaction (the user must be able to make short and long responses) or GUI drag & drop function.

The measures to be collected and stored are:

- Initial response time. It is the response time of the task 1. In fact the results should be the same than in task 1 (although this should be experimentally proved).
- Final response time. It is the time elapsed between stimulus extinction and response extinction (switch released). Normally this value should be very similar to the previous one.
- Activation time. It is the duration of a non erroneous response.
- Initial error count. The error count of task 1.
- Final error count. Number of responses prematurely interrupted before stimulus has disappeared.
- Total time to complete the task.

These measures are similar or can be derived from these found in ERGOLAB and Compass.

4.3. Task 3. Selection of an element by scanning

Selection by scanning is a technique that enables, by means of one or more discreet orders, people to select elements from a set (options from a menu, letter on a virtual keyboard or the box in a communication device, for example). The various options change at a determined speed called scanning speed. At the moment when the correct option is highlighted, the user must send an order to choose it. Although, in essence, it is a discreet response to a stimulus with a greater cognitive load, we believe that due to its importance it needs to be included in the assessment method proposed.

The scanning speed is a critical parameter that must be carefully regulated in such a way that it is the greatest possible without involving an unacceptable rate of errors. The regulation of this speed is usually done in an experimental way with the help of the assessor, although techniques have also been proposed to regulate the speed automatically as the user interacts with the system [16]. Whatever, before starting the exercise, it is necessary to set a determined scan speed so that the results are not altered by the variations that an automatic regulation system would introduce and so that it is possible to make comparisons between different devices.

The exercise will present the user with a series of options along a dimension, from which he or she must choose one (the correct option must be shown by means of a visual and/or acoustic stimulus), these options must be alternated consecutively, creating a visual and/or acoustic stimulus each time. At the appropriate moment, the user will send a response. The correct or erroneous choice may optionally be reinforced with a visual and/or acoustic stimulus. Whether the person succeeds or fails, the system will wait for a random time, which can be configured, for the new test to be presented.

To the variables registered for exercise 1, we will need to add the scan speed used and the errors committed by late responses.

4.4. Task 4. Positioning the pointer and choosing

The most common assessment measurements for pointer devices are speed and exactness; to the extent that there is an ISO standard for assessing this kind of device. The complete standard is the ISO 9241 "Ergonomic requirements for office work with visual display terminals (VDTs)." Part 9 of this standard "Requirements for non-keyboard input devices" [14] proposes a single measurement: performance. The performance units are bytes per second and are obtained based on speed and precision.

$$Performance = \frac{ID_e}{MT} \quad (1)$$

In which, the term ID_e indicates the difficulty of the task (in bytes) and MT the time taken to carry out the movement. The difficulty is calculated as:

$$ID_e = log_2\left(\frac{D}{W_e} + 1\right) \quad (2)$$

In which the term D refers to the distance to the target and W_e to the effective size of the target; unlike Fitt's law [17], it takes into account the effective size instead of the real size of the object W. The effective size is obtained from the distribution of the coordinates chosen during various tests, and therefore reflects the precision with which the test has been carried out. In addition, the effective size can help us determine the size that the elements of the graphic interface should have. The formula for obtaining the effective size is the following:

$$W_e = 4.133 \times SD_x \quad (3)$$

In which SD_x is the standard deviation of the coordinates chosen during the exercise. For more detailed information about these measurements, please see [18].

In addition to the standard measurements, some studies [19, 20] propose a series of additional measurements that enable us to finish characterising the task of choosing with the pointer. Of these, we are going to use the target re-entry rate (TRE) and the deviation with regard to the ideal trajectory (Movement Offset [MO]) for our study. The ideal trajectory is considered to be a straight line between the origin and the destination. The deviation from this is calculated as the average of the distance between the different samples of the real trajectory and the ideal trajectory.

We are going to define the exercise based on the standard but with some modifications to adapt it to the kind of users with whom we are dealing. This will consist of a window, the centre of which will be the starting point for all the exercise tests. This point may optionally be indicated with a visual stimulus. The target stimuli will appear around this point in the form of images that may be accompanied by sound. The distance and the size of the stimuli will be configured; the angle will be made to vary at random. Additional visual and acoustic stimuli may optionally be used to reinforce correct or incorrect execution.

Once the task has been started, the pointer will be placed in the central point of this window and will be maintained in this position until the instant that the target stimulus appears. This is done to facilitate the task in the case of using devices in which it is more difficult to maintain the position fixed (such as eye or head following systems). The user must move the pointer and click on the target as centrally as possible, for this the stimulus can be designed so that it encourages signalling the central point (for example in the shape of targets or spirals). The test will be repeated a finite number of times at random intervals within a configurable margin.

The measures to be collected and stored are:

- Reaction time: This is the time that passes between the target stimulus being presented to the user and the pointer starting to move. It should be taken into account that this value may not be valid, depending on the kind of device used.
- Execution time: This is the time that passes between the target stimulus being presented to the user and it being chosen. The value of the execution time is what is used to calculate the performance.
- Performance.
- Number of re-entries.
- Target re-entry (TRE).
- Movement Offset (MO).
- Effective size of the target (calculated for all the exercise).
- Number of errors: Clicking outside the target is considered to be an error.
- Total time to complete the exercise.

It may be interesting to register all the events sent by the pointer device to be able to visualise the trajectories made on the screen at a later date.

4.5. Task 5. Production of characters and text

To analyse the production of characters and text the metrics defined in the ISO 9241-11 standard [21] will be used, although with the relevant nuances to be used with people who may have serious cognitive alterations or those who are still in the process of learning to read and write. For this, the user must transcribe pre-set texts, the length of which may go from a single character to various lines of text. Two modes of operation may be configured, one without edition (the characters introduced incorrectly will be rejected) and one with edition. The device used to produce text need not necessarily be a conventional keyboard.

The system shows a window (with a preferably dark background with white characters of a suitable size). In the window there will be two areas, one which will display the text to be transcribed and the other into which the user must copy the text. The sample text is reproduced integrally. As an option the next character to be introduced can be reinforced by means of a visual stimulus. The introduction of each character or of the entire phrase, whether correct or incorrect, may be accompanied by a visual and/or acoustic reinforcement. The exercise will consist of various tests in which the words on a list will appear in a random order.

The measures to be collected and stored are:

- Character per second (CPS).

$$CPS = \frac{Characters\ introduced}{Time\ required} \quad (4)$$

- Character Error Rate (CER).

$$CER = \frac{errors}{n} \times 100 \quad (5)$$

  In the case of doing the exercise with the edition mode errors=$s+d+i$ where s=number of substitutions, d=number of deletes, i=number of insertions and n=number of characters. In the mode without edition, errors are considered to be the characters that are introduced incorrectly.

- Length of the sample text.
- Time used to write the text.
- Total time to complete the task.

5. Conclusions and future work

It is essential to have a tool for quantitatively assessing the human-computer interaction. Such a tool gives the possibility to experts of monitoring the interaction achieved by users with special needs and of improving the devices they use.

We have proposed an interaction assessment method that accomplish with the next goals:

- Limited cognitive capabilities: The difficulty of the proposed exercises gradually increases. They have been designed tacking into account the user limitations. Thus, simple stimuli and positive and negative reinforcements have been used.

- Limited motion and sensing capabilities: The utilized stimuli are generally presented in to forms (*i.e.* visual and acoustic). Moreover, it is possible to add different kind of stimuli, namely haptic or smell devices.

- Device interferences: Once the user becomes an expert in using a given device, this expertise could be a shortcoming when learning how to use a new device. The proposed method could be use to successively analyse the adaptation of the user to a new device.

- User interaction evolution: The evolution of the user can be softly monitored because the given measures are based on precise values of time or even on events generated at a high frequency. The planning of a number of control sessions will help the experts to analyse the user evolution with accuracy.

Although our method is based on tested assessment methods and renown standards, its value has not been tested yet. In the near future we are planning to test our method on disabled people to evaluate its accuracy by applying statistical analysis as ANOVA or Tukey and Scheffe tests.

References

[1] T. Hewett, R. Baecker, S. Card, T. Carey, J. Gasen, M. Mantei, G. Perlman, G. Strong, and W. Verplank, “Curricula for human-computer interaction.” ACM SIGCHI, 1996, g. Perlman eds. [Online]. Available: http://sigchi.org/cdg/

[2] M. Garcia and C. Mauri, “Experiencia de interacción persona-ordenador a través de webcam con usuarios con discapacidad motriz grave y moderadamente afectados [HCI experience using webcam with serious and moderately affected motor disabled users],” in *Tecnoneet*, Murcia, 2004.

[3] C. Mauri, “Interacción persona-ordenador mediante cámaras webcam [HCI using web cameras],” in *Interacción*, Lleida, 2004.

[4] M. Garcia and C. Mauri, “Parálisis cerebral y nuevas tecnologías: Ayudas técnicas basadas en visión artificial [Cerebral palsy and new technologies: Assistive technology based on computer vision].” in *IV Jornadas Onubenses Sobre Parálisis Cerebral*, Huelva, 2005.

[5] C. Mauri, T. Granollers, J. Lores, and M. García, “Computer vision interaction for people with severe movement restrictions,” *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, vol. 2, no. 1, pp. 38–54, April 2006. [Online]. Available: http://www.humantechnology.jyu.fi/articles/volume2/2006/mauri-granollers-lores-garcia.pdf

[6] M. Silfverberg, I. S. MacKenzie, and T. Kauppinen, "An isometric joystick as a pointing device for handheld information terminals," in *Proceedings of Graphics Interface*. Toronto, Canada: Canadian Information Processing Society, 2001, pp. 119–126.

[7] D. Lessenberry, *Analysis of Errors.* Syracuse, NY: L.C. Smith and Corona Typewriters., 1928.

[8] W. K. English, D. C. Engelbart, and M. L. Berman, "Display selection techniques for text manipulation," *IEEE Transactions on Human Factors in Electronics*, vol. HFE-8, pp. 5–15, March 1967.

[9] H. Koester and E. LoPresti, "Tools and methods for assessment in computer access." in *RESNA 2006 Annual Conference*. Atlanta, GA. Arlington, VA: RESNA Press, 2006. [Online]. Available: http://www.kpronline.com/files/toolshandoutresna2006.pdf

[10] B. Liffick, B. Romich, and K. Hill, "A single-switch performance evaluation tool." in *11th International Conference on Human-Computer Interaction.*, Las Vegas, NV., 2005.

[11] H. Koester, E. LoPresti, G. Ashlock, W. McMillan, P. Moore, and R. Simpson, "Compass: Software for computer skills assessment," in *CSUN 2003 International Conference on Technology and Persons with Disabilities*, Los Angeles, CA, March 2003. [Online]. Available: http://www.kpronline.com/files/compasscsun2003.pdf

[12] M. Noirhomme-Fraiture, C. Charriere, J. M. Vanderdonckt, and C. Bernard, "Ergolab: a screen usability evaluation tool for children with cerebral palsy," in *CHI '93: INTERACT '93 and CHI '93 conference companion on Human factors in computing systems*. New York, NY, USA: ACM Press, 1993, pp. 23–24.

[13] M. Noirhomme-Fraiture, C. Charriere, and J. Vanderdonckt, "A laboratory of ergonomic analyses for children suffering form cerebral palsy," in *Tools for Working with Guidelines.* TFWWG, 2000, pp. 35–49. [Online]. Available: citeseer.ist.psu.edu/noirhomme-fraiture00laboratory.html

[14] ISO, *Ergonomics of human-system interaction—Guidance on accessibility for human-computer interface*. ISO/TS 16071, 2003.

[15] F. Van-Nes, J. F. Juola, and R. J. A. M. Moonen, "Attraction and distraction by text colours on displays," in *INTERACT 87 - 2nd IFIP International Conference on Human-Computer Interaction*, B. S. Hans-Jorg Bullinger, Ed. North-Holland, 1987, pp. 625–630.

[16] G. W. Lesher, J. D. Higginbotham, and B. J. Moulton, "Techniques for automatically updating scanning delays," in *Proceedings of the RESNA 2000*. Resna Press, VA, 2000, pp. 85–87. [Online]. Available: http://citeseer.ist.psu.edu/490181.html

[17] I. S. MacKenzie, "Fitts' law as a research and design tool in human-computer interaction." in *Human-Computer Interaction, 7*, 1992, pp. 91–139.

[18] S. A. Douglas, A. E. Kirkpatrick, and I. S. MacKenzie, "Testing pointing device performance and user assessment with the ISO 9241- part 9 standard," in *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA: ACM Press, 1999, pp. 215–222.

[19] S. I. Mackenzie, T. Kauppinen, and M. Silfverberg, "Accuracy measures for evaluating computer pointing devices," in *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA: ACM Press, 2001, pp. 9–16. [Online]. Available: http://dx.doi.org/10.1145/365024.365028

[20] S. Keates, F. Hwang, P. Langdon, P. J. Clarkson, and P. Robinson, "Cursor measures for motion-impaired computer users," in *Assets '02: Proceedings of the fifth international ACM conference on Assistive technologies*. New York, NY, USA: ACM Press, 2002, pp. 135–142.

[21] ISO, "ISO 9241 international standard. ergonomic requirements for officework with visual display terminals (vdts) - part 9: Requirements for non-keyboard input devices." International Organization for Standardization,, Genève, Switzerland., Tech. Rep., 2000.