eman ta zabal zazu

Universidad del País Vasco
Euskal Herriko Unibertsitatea

# Assisted Interaction for Improving Web Accessibility: An Approach Driven and Tested by Users with Disabilities


*Department of*

Computer Architecture and Technology

*Candidate*

Juan Eduardo Pérez López

*Supervisors*

Myriam Arrue Recondo

Julio Abascal González


Donostia/San Sebastián, 2020

*To my parents.*

# Acknowledgments

A Julio y a Myriam por haber dirigido mi tesis durante todos estos años. Gracias al apoyo y supervisión recibidos este trabajo ha salido adelante.

A Nestor y a Luis por haber estado siempre disponibles en la segunda planta de la facultad para ayudarme cuando lo he necesitado.

Al resto de miembros de EGOKITUZ y demás compañeros de la facultad con los que he compartido laboratorios (302 primero y luego 303) y reuniones: Idoia, Amaia, Ana, Carlos, Xabier G., Aritz, Sandra, Unai M. y Unai B., Raúl, Iñigo, Ainhoa, Javi, Olatz por todos los momentos pasados y haber hecho este viaje más llevadero. En especial a Borja G., gracias a quien tuve la oportunidad de empezar en el grupo de investigación EGOKITUZ, y a Xabier V. con quien más tiempo he tenido el placer de trabajar y compartir laboratorio; sin su colaboración esta tesis no hubiese sido posible.

Esta tesis tampoco hubiese sido posible sin la ayuda de Elkartu y de otras asociaciones de personas con discapacidad. A todos los participantes que han colaborado en los distintos experimentos realizados, sin su participación desinteresada este trabajo no hubiese sido el mismo. Especialmente tengo que agradecer la colaboración en este trabajo de Lourdes de la Universidad Carlos III de Madrid, siempre dispuesta a ayudar.

I would like to express my gratitude to Hironobu Takagi and Chieko Asakawa for hosting me at IBM Research – Tokyo in 2016–17. I would also like to thank the rest the accessibility group for making me feel at home: Daisuke, Masatomo, Masayuki, Kaoru, Mario, Yasunori, Kentaro, Tatsuya and specially to Itoh-san thanks to whom I never got lost in the biggest city in the world.

A strongly appreciate the reviews from Markel Vigo and Yeliz Yesilada. Their comments have contributed to improve this work.

A mis padres, por su paciencia y apoyo incondicional. A mi hermano por haber entendido cuando estaba más ocupado.

To Jennison, Stephanie, Mentxu, José, Sebas and rest of friends who have been supporting me one way or another during this long journey. This has also been thanks to you.

The University of the Basque Country (UPV/EHU) awarded a pre-doctoral fellowship to the author of this PhD dissertation to conduct the research herein presented. The Canon Foundation in Europe awarded a fellowship to support the research done during the period abroad at IBM Research – Tokyo.

# Abstract

The Web has become an essential tool for the daily lives of millions of people. Today, an increasing percentage of the world's population depend on this global information system to work, socialise, and entertain themselves among many other activities. The benefits of the Web are even more crucial for people with disabilities as it allows them to perform some tasks that are restricted for them in the physical word due to the plethora of accessibility barriers. Consequently, their opportunities for social integration as free and equal citizens are hindered. Despite their advantages, most web pages usually ignore the special needs of people with disabilities. Many of them usually include a single design to fit all users, making it more difficult to adapt the web interface to the personal needs. Diverse methods have been proposed to combat this problem. For instance, transcoding systems aim to automatically transform inaccessible Web pages on the fly into accessible ones. In order to improve web accessibility to specific groups of people these methods require information about the most suitable adaptation techniques that should be applied to each one.

This thesis collects a number of in-depth studies about the suitability of adaptation techniques to improve the web navigation for two different groups of people with disabilities: people with motor impairments and people with low vision. Based on literature reviews and observational studies different sets of adaptations have been implemented and evaluated with users both in-situ during single session laboratory experiments and remotely from participants' homes during a longitudinal study. The RemoTest tool was used to assist designing and conducting experimental sessions, as well to gather interaction data from participants. The interaction environments created by the RemoTest were previously evaluated and the results revealed that these were accessible to conduct inclusive experiments both in remote and in-situ contexts.

Various interface adaptations and alternative interaction methods were evaluated in different contexts by means of qualitative and quantitative analysis of participants' performance and satisfaction. Transcoding techniques were evaluated with tablet users with reduced mobility. The results showed that the majority of the participants preferred the adapted interfaces even if some disadvantages were associated to these. Web adaptations techniques were also evaluated with people with low vision using desktop computers. The results showed that the advantages of some adaptations techniques varied depending on the type of assistive technology used by the participants to access the Web. Two cursor enhancements for assisting link selection to people with motor impairments were evaluated with users of different alternative pointing devices.

The results from a preliminary laboratory study showed that the majority of participants with motor impairments improved effectiveness and efficiency on point and click trajectories with one of the two tested cursor enhancements. These results depended on the alternative pointing device used, as well as on subjective preferences. A subsequent longitudinal study showed an improvement in the performance with the cursor enhancements over time.

# Resumen

La Web se ha convertido en una herramienta esencial en el día a día de millones de personas. Actualmente, un porcentaje cada vez mayor de la población mundial depende de este sistema de información global para trabajar, socializar, y divertirse entre otras muchas actividades. Los beneficios de la Web son incluso más cruciales para las personas con necesidades especiales ya que les permite realizar ciertas tareas que en el mundo físico les están restringidas debido a distintas barreras de accesibilidad. En consecuencia, sus oportunidades de integración social como ciudadanos libres e iguales se ven obstaculizadas. A pesar de sus ventajas, la mayoría de páginas web generalmente ignoran las necesidades especiales de las personas con discapacidad. Muchas de estas páginas suelen incluir un diseño único para todos los usuarios, lo que dificulta la adaptación de la interfaz web a las necesidades de cada persona. Se han propuesto diversos métodos para solucionar este problema. Por ejemplo, los sistemas de "transcoding" tienen como objetivo transformar automáticamente páginas web inaccesibles en accesibles. Para mejorar la accesibilidad web a grupos específicos de personas, estos sistema requieren de información sobre las técnicas de adaptación más adecuadas en cada caso.

Esta tesis recoge una serie de estudios en profundidad sobre la idoneidad de aplicar diferentes técnicas de adaptación para mejorar la navegación web a dos grupos de personas con discapacidad: personas con discapacidades motoras y personas con baja visión. En base a una revisión bibliográfica y a estudios de observación, se han implementado diferentes adaptaciones para estos grupos de usuarios con necesidades especiales. Posteriormente, estas adaptaciones se han evaluado con usuarios tanto en experimentos de laboratorio de sesión única, como en un estudio longitudinal realizado en casa de los participantes de forma no supervisada. La herramienta RemoTest se utilizó para ayudar a diseñar y realizar las sesiones experimentales, así como para recopilar los datos de interacción de los participantes. Los entornos de interacción creados por RemoTest se evaluaron previamente y los resultados revelaron que estos eran accesibles para realizar experimentos inclusivos tanto en contextos remotos como in-situ.

Diversas adaptaciones de interfaz y métodos de interacción alternativos fueron evaluados en diferentes contextos mediante análisis cualitativos y cuantitativos del rendimiento y satisfacción de los participantes. Diferentes técnicas de "transcoding" para tabletas táctiles fueron evaluadas con usuarios de movilidad reducida. Los resultados mostraron que la mayoría de los participantes preferían las interfaces adaptadas incluso aunque se detectaron algunas desventajas. Otras técnicas de adaptación web fueron evaluadas en ordenadores de escritorio con

usuarios con baja visión. Los resultados mostraron que los beneficios de algunas técnicas de adaptación variaban según el tipo de tecnología de asistencia utilizada por los participantes para acceder a la Web. Dos mejoras del cursor para facilitar la selección de enlaces a personas con discapacidades motoras fueron evaluadas con usuarios de diferentes dispositivos apuntadores alternativos al ratón. Los resultados de un estudio preliminar mostraron que la mayoría de los participantes con discapacidades motoras mejoraron su efectividad y eficiencia al seleccionar enlaces en la Web con una de las dos mejoras probadas. Estos resultados dependían del dispositivo apuntador alternativo utilizado, así como de las preferencias subjetivas. En un estudio longitudinal posterior se detectó que el rendimiento con las ayudas de cursor mejoraba a lo largo del tiempo.

# Laburpena

Web-a funtsezko tresna bat bilakatu da milioika pertsonen eguneroko bizitzan. Gaur egun, informazio sistema global hau lanerako, gizarteratzeko, ondo pasatzeko eta beste jarduera askotarako erabiltzen duen munduko biztanleriaren ehunekoa gero eta handiagoa da. Web-aren abantailak are garrantzitsuagoak dira behar bereziak dituzten pertsonentzat, izan ere, mundu fisikoan zenbait zeregin egiteko aurkitzen dituzten oztopo eta mugak gainditu ditzakete. Abantailak izan arren, web orrialde gehienek ez dituzte desgaitasuna duten pertsonen behar bereziak kontuan hartzen. Orrialde hauetako askok normalean erabiltzaile guztientzako diseinu bakarra izaten dute, eta, beraz, zaila da web interfazea pertsona bakoitzaren beharretara egokitzea. Arazo hau konpontzeko hainbat metodo proposatu dira. Adibidez, transkodifikazio sistemek irisgarriak ez diren web orrialdeak automatikoki irisgarriak izateko beharrezko aldaketak egitea dute helburu. Sistema hauek, web irisgarritasuna hobetzeko kasu bakoitzean egokienak diren egokitzapen teknikei buruzko informazioa behar dute.

Tesi honetan, desgaitasuna duten pertsona talde batzuentzat (desgaitasun fisikoak dituzten pertsonak eta ikusmen urritasuna duten pertsonak) egokitzapen teknika desberdinak aplikatzeko egokitasunaren inguruko azterketa sakonak biltzen dira. Berrikuspen bibliografiko eta behaketa azterlanetatik abiatuta, behar bereziak dituzten erabiltzaile talde horietarako egokitzapen desberdinak garatu dira. Ondoren, egokitzapen horiek erabiltzaileekin ebaluatu dira, saio bakarreko laborategiko esperimentuetan eta parte-hartzaileen etxean egindako gainbegiratu gabeko luzetarako estudioetan. RemoTest tresna saio esperimentalak diseinatu eta aurrera eramateko erabili zen, baita parte-hartzaileen interfazeekiko elkarrekintza datuak biltzeko ere. RemoTest-ek sortutako elkarrekintza inguruneak aldez aurretik ebaluatu ziren eta emaitzek esperimentu inklusiboak egiteko irisgarriak zirela ondorioztatu zuten, bai urruneko testuinguruan eta baita ingurune lokaletan ere.

Hainbat interfaze-egokitzapen eta elkarrekintza metodo alternatibo ebaluatu ziren testuinguru desberdinetan, partaideen errendimenduaren eta gogobetetzearen analisi kualitatiboak eta kuantitatiboak erabiliz. Transkodetzeko teknika desberdinak ebaluatu ziren tablet gailuen interfazeetan mugikortasun urria duten erabiltzaileekin. Emaitzek erakutsi zuten parte-hartzaile gehienek beraien beharretara egokitutako interfazeak nahiago zituztela nahiz eta zenbait desabantaila antzeman. Web egokitzapen teknika batzuk ordenagailuetan ebaluatu ziren ikusmen urritasuna duten erabiltzaileekin. Emaitzek erakutsi zuten egokitzapen teknika batzuen abantailak webean sartzeko parte-hartzaileek erabilitako laguntza motaren arabera aldatu egiten zirela. Desgaitasun fisikoak

dituztenentzako estekak aukeratzen errazteko bi kurtsore laguntza ebaluatu ziren. Aurretiazko ikerketa baten emaitzek erakutsi dutenez, desgaitasun fisikoa duten parte-hartzaile gehienek beraien eraginkortasuna hobetu zuten estekak aukeratzeko kurtsoreen hobenkutzaren batekin. Emaitza horiek parte-hartzaileek erabilitako sagu gailuaren aukerazko gailuaren araberakoak ziren (joystick, trackball, etab.), baita beraien lehentasun subjektiboen araberakoa ere. Geroko luzetarako azterketa batean, kurtsoreen laguntzekin parte-hartzaileen errendimendua denborarekin hobera egin zela ikusi zen.

# Contents

# 1. Summary

## 1.1. Introduction

Since the Web was first introduced to the public three decades ago [3], this global information system continues to evolve and has become unavoidable in much of our lives. An increasing percentage of the world's population constantly takes advantage of the Internet to perform a wide variety of activities in areas as diverse as education, work, business, entertainment, or public administration, among many others. In addition to the clear advantages that the Web provides to its users, it also promotes the inclusion of some human groups such as people with disabilities. These people frequently find barriers to carrying out several activities in the physical world that hinder or even make impossible their social inclusion and participation. For many of them, the Web provides a valid alternative with which to carry out these activities.

In the early 1990s the challenge was to achieve the technological infrastructure for the Web to work. Nowadays, one of the main concerns is to design web sites that meet the needs of the people who use them [9]. Despite the many guidelines, recommendations and policies [4]–[7][13] that have been developed to promote accessibility to the Web for people with disabilities, this is a challenge that has not yet been overcome [8]. The majority of websites present rigid designs for all users, which cannot be accommodated to the specific needs of the different users with disabilities. Even if a website is fully conformant to strict accessibility guidelines it is likely that not all users with special needs will get the best user experience. This happens when the user interface cannot be adapted to their own needs, preferences, characteristics, and even to their own equipment.

Some interesting solutions that allow the user to personalize certain aspects of a website to improve its accessibility can be found in the literature. However, these are frequently limited and are mainly focused on specific features such as proposed colour schemes or font sizes to meet the needs of some people with visual disabilities. There is, therefore, a shortage of general-purpose proposals that are, at the same time, tailored to the specific needs of each user, covering an ample range of disabilities.

In this sense, an interesting approach is transcoding systems, which allow automated adaptation of inaccessible web content and structure on the fly into accessible pages, without needing the intervention of the authors to modify their pages [2].

Users with limited dexterity in upper limbs may have difficulties to manipulate graphical user interfaces. The largest barrier is frequently the use of pointing devices to select, click, drag, etc., items. There exist in the literature different pieces of software that create virtual devices implementing alternative interaction methods to assist target selection.

This thesis work starts from these two approaches, trying to provide a systematic framework for selecting, adapting, measuring and evaluating their use, in order to improve the user experience on the Web of two groups of people with special needs: people with motor impairments and people with low vision.

#### *1.1.1. Coherence of the thesis work*

This thesis work presents six papers published in diverse international prestigious journals and conferences. All of them form part of a sequential research work with the same objectives and methodology. The common objective is to improve Web accessibility for people with disabilities by means of the use of adaptation techniques and cursor enhancement applications. The general methodology is centred in the study of the behaviour of real users that would later feed user models.

For this purpose, a recursive scheme has been followed in the work presented in this thesis. This scheme always included:

- a literature review for selecting suitable adaptation techniques and cursor enhancement applications
- observation of real users behaviour to detect the problems and barriers they find when accessing the Web
- selection and adjustment of the most suitable aids and
- evaluation of the adaptations and enhancements with users

As a result, this thesis presents an evolving research work that builds a number of proposals for web accessibility advancements supported in evidence obtained from the experimentation with users.

## 1.2. Methodology

Even though the work presented in this thesis has been developed as a sequence of concatenated research activities, the methodology adopted is coherent and common to the whole work. We followed User-Centred Methodology, as described in the next paragraphs:

1. All the activities commenced with a rigorous study of the needs and characteristics of the selected target user population. These studies have taken different forms depending of the type of research. Users with special needs were observed while interacting with selected web pages, as well as interviewed in order to obtain knowledge of their accessibility problems.

2. After that, a literature review was always conducted to seek, select, and evaluate:

    a) relevant guidelines and standards for the selected group of users and

    b) suitable web accessibility supportive techniques for these users.

3. Taking all this information as a starting point, we tackled the development of techniques and aids based on the previous findings, applying User-Centred Design techniques. In all the works, we followed the principles of usability and accessibility to favour their acceptance and to maximize their compatibility with:

    a) the equipment on which the tests are carried out (always allowing users to use their own personal equipment, this being well adapted to their characteristics and needs).

    b) any type of web page in order to foster free navigation in subsequent experimental tasks.

    In most cases this phase involved the design of pieces of software, including interface adaptations and interaction aids, or complete user-testing environments, such as the RemoTest platform, co-designed with other researchers, in order to carry out tests with users and gather interaction data. To this end, the following was required:

a) the definition of the types of data and other parameters such as the sampling frequency, necessary to be able to subsequently evaluate the performance achieved with the adaptations and aids being tested.

b) the design of accessible questionnaires to collect subjective assessments of participants on the evaluated techniques.

4. All the proposed advancements have been formally evaluated with the collaboration of users. For each experiment the following tasks have been performed:

   a) definition of the tasks for the different tests carried out with users (specific activities on the Web to evaluate the use, targeted selection of objectives to evaluate performance, subjective evaluation based on questionnaires).

   b) carrying out qualitative studies that include the analysis of:
      - subjective evaluations of the participants –through usability questionnaires, and their behaviour
      - users' comments obtained by the thinking aloud technique and recording their interaction on video

   c) performing quantitative studies that include the analysis of aspects relating to the performance of the participants, based on different measures of human performance collected from the literature.

   d) conducting supervised single-session studies in order to assess the acceptance (through usability questionnaires) and the achieved performance (by measuring the effectiveness and efficiency of users completing tasks).

   e) making statistical analysis of the data gathered to assess the significance of the results achieved. For instance, in order to compare the effectiveness of different variants tested through relevant statistical tests. On the other hand, the degree of usability of the proposed aids was measured by means of subjective evaluations of the participants answering SUS questionnaires.

The experiments with users have obtained the prior approval of the University Ethics Committee that requires justification of the studies with human beings and ensures the correct treatment of the personal data. The methodology to be used in this research work was also approved by evaluating the type of study (qualitative or quantitative), sample size, types of users (grouped according to type of disability, assistive technology used to access the computer, control users without disabilities), collected data, study variables, and model of analysis of results.

## 1.3. Objectives

The main objective of this thesis is to contribute to the knowledge about the characteristics and needs of users with specific disabilities, people with upper limb motor restrictions and people with low vision, in order to propose suitable methodologies, procedures and software tools to enhance their use of digital applications.

This objective has been structured in a number of sub-objectives:

1. To analyse the needs of users with motor disability and people with specific disabilities (upper limbs motor restrictions and people with low vision).
2. To search, classify, analyse, adapt and test support aids and adaptations, which assist people with specific disabilities to use the web.
3. To develop and test evaluation tools for the accurate measurement of the patterns of use and the validity of the proposed aids.
4. To develop support aids and adaptations which assist people with specific disabilities to use the web.
5. To conduct studies including users with disabilities in order to test and evaluate the quality and validity of the previously selected and developed technical helps.
6. To propose criteria to enhance web accessibility for the selected groups of users.

Objective 1 has been addressed in publications [14] and [10]. In paper [14], a user study with people with motor impairments was conducted to detect their main interaction characteristics by means of different performance measures about rapidity and accuracy of cursor movement. In paper [10], navigation strategies of users with low vision were studied by means of an observational study in order to detect the appropriate web adaptation techniques for them.

Objective 2 has been addressed in publications [15], [10], [11] and [12]. In paper [15], appropriate adaptation techniques and alternative interaction methods to assist web navigation on touch screen tablets for people with reduced mobility in upper limbs were gathered from the literature. In paper [10] a set of adaptation techniques aiming to assist web navigation for users with low vision were collected from the literature. In papers [11] and [12] several cursor enhancements to assist point and click interactions for users with motor impairments were reviewed from the literature.

Objective 3 has been addressed in publications [14] and [1]. In both papers the RemoTest platform to assist experimenters performing user tests was presented. In paper [14], the platform was applied to conduct a formal user test. In paper [1], the accessibility of the environments created by RemoTest (installation process, questionnaires, task description) was evaluated by means of formal in situ user studies with participants with diverse disabilities.

Objective 4 has been addressed in publications [15], [10], [11] and [12]. In paper [15], a set of interface adaptations and alternative interaction methods were developed for assisting people with motor impairments using touch screen tablets. In paper [10], a set of adaptation techniques aiming to assist web navigation for people with low vision were developed. In papers [11] and [12] two different cursor aids were developed for assisting link selection to people with motor impairments.

Objective 5 has been addressed in publications [15], [10], [11] and [12]. In papers [15], [10] and [11] different *in situ* user tests were conducted in order to evaluate the diverse technical assistances proposed for each specific group of people with disabilities. In paper [12], a longitudinal study was conducted with users with motor impairments, in order to study the learning effect on their performance and their satisfaction with the cursor aids being tested. In this study, the users participated from home in order to obtain more naturalistic interaction data from everyday computer use.

Objective 6 has been addressed in publications [15], [10] and [12]. Based on the results of each study, different improvements were proposed for the technical assistances being tested in each case.

## 1.4. Results

1. Unusual patterns on point and click tasks were detected and different user profiles on participants with motor impairments were identified.

   These results were associated with Objective 1, and were published in the paper entitled *"Assisted Interaction data analysis of web-based user studies"* (see Appendix 1).

2. The RemoTest platform proved to be useful to assist with the analysis of the data automatically collected in the experimental sessions, as well as to present accessible environments (questionnaires, task descriptions, installation process) for participants with disabilities. Based on these results, the tool was evaluated as suitable for conducting formal and inclusive experimental sessions both in remote and in situ contexts.

   These results were associated with Objective 3, and were published in two papers: *"Assisted Interaction data analysis of web-based user studies"* (see Appendix 1) and *"Inclusive Web Empirical Studies in Remote and In-Situ Settings: A User Evaluation of the RemoTest Platform"* (see Appendix 2).

3. The tested interface adaptations with users of touch screen tablets turned out to be beneficial for most of the participants with motor impairments, mainly because navigation required less physical effort. In addition, two of the alternative interaction methods tested proved to be helpful for people with low control of finger movement.

   These results were associated with Objective 5, and were published in the paper entitled *"Adapting the web for people with upper body motor impairments using touch screen tablets"* (see Appendix 3).

4. Some improvements were suggested by participants in the experiments with touch screen tablets, including: customization features for the user interface (e.g., scrolling buttons, collapsible menus), and an adaptive system that dynamically selects the most appropriate interaction methods. These proposals were based on user tests results (both from performance and interviews with participants)

   These results were associated with Objective 6, and were published in the paper entitled *"Adapting the web for people with upper body motor impairments using touch screen tablets"* (see Appendix 3).

5. Navigation strategies of people with low vision were identified. They were used to select appropriate adaptation techniques to assist these users in web browsing.

   These results were associated with Objectives 1 and 2, and were published in the paper entitled *"An exploratory study of web adaptation techniques for people with low vision"* (see Appendix 4).

6. It was proved that the advantages of some techniques varied depending on the type of assistive technology used by participants with low vision to access the Web. For example, some of the applied adaptation techniques turned out to be helpful only for users who utilized screen magnifying software, but not for those using the browser zoom feature

   These results were associated with Objective 5, and were published in the paper entitled *"An exploratory study of web adaptation techniques for people with low vision"* (see Appendix 4).

7. Although there was no statistically significant evidence resulting from these experiments, qualitative information was obtained that guided the definition of two new research hypotheses to be validated in future work.

These results were associated with Objective 6, and were published in the paper entitled *"An exploratory study of web adaptation techniques for people with low vision"* (see Appendix 4).

8. The study performed showed that users of alternative pointing devices benefited from point and click facilitators for accessing the Web. In addition, the findings were promising in terms of performance and satisfaction achieved by participants with motor impairments. These results also suggest that the alternative pointing device used was a good indicator of how to provide better cursor assistance. In addition, improvements on performance with the use of cursor aids are predictable.

   These results were associated with Objective 5, and were published in the paper entitled *"Evaluation of two virtual cursors for assisting web access to people with motor impairments"* (see Appendix 5).

9. Significant improvements with both cursor aids compared to the original cursor in six of the seven cursor parameters studied, albeit with performance variations between some participants were supported by the longitudinal study. These results also reported an improvement in performance during the longitudinal study with one cursor aid.

   These results were associated with Objective 5, and were published in the paper entitled *"Longitudinal Study of Two Virtual Cursors for People with Motor Impairments: A Performance and Satisfaction Analysis on Web Navigation"* (see Appendix 6).

10. Suggestions for improving the area cursor, taking into account the influence on performance of distractors (i.e., nearby links to the target) were introduced. Similarly, other improvements were also suggested for the cross cursor based on performance with respect to clicking time and the opinion of participants.

    These results were associated with Objective 6, and published in the paper entitled *"Longitudinal Study of Two Virtual Cursors for People with Motor Impairments: A Performance and Satisfaction Analysis on Web Navigation"* (see Appendix 6).

## 1.5. References

[1] MYRIAM ARRUE, XABIER VALENCIA, J. EDUARDO PÉREZ, LOURDES MORENO, AND JULIO ABASCAL, "Inclusive web empirical studies in remote and in-situ settings: a

user evaluation of the RemoTest platform," *International Journal of Human–Computer Interaction*, vol. 35, no. 7, Apr. 2019, pp. 568–583.

[2] CHIEKO ASAKAWA AND HIRONOBU TAKAGI, "Transcoding," In *Web accessibility: a foundation for research*, S. Harper and Y. Yesilada, Eds., Springer, London, 2008, pp. 231–260.

[3] TIM BERNERS-LEE, ROBERT CAILLIAU, JEAN-FRANÇOIS GROFF, AND BERND POLLERMANN, "World-Wide Web: the information universe," *Electronic Networking*, vol. 2, no. 1, pp. 52–58, Spring 1992.

[4] GIORGIO BRAJNIK, "Barrier Walkthrough: Heuristic evaluation guided by accessibility barriers," Mar. 2009. Accessed on: Apr. 22, 2020. [Online]. Available on: users.dimi.uniud.it/~giorgio.brajnik/projects/bw/bw.html

[5] BEN CALDWELL, MICHAEL COOPER, LORETTA GUARINO REID, AND GREGG VANDERHEIDEN, "Web Content Accessibility Guidelines (WCAG) 2.0," *W3C Recommendation*, Dec. 2008. Accessed on: Apr. 22, 2020. [Online]. Available on: www.w3.org/TR/WCAG20/

[6] WENDY CHISHOLM, GREGG VANDERHEIDEN, AND IAN JACOBS, "Web Content Accessibility Guidelines (WCAG) 1.0," *W3C Recommendation*, May 1999. Accessed on: Apr. 22, 2020. [Online]. Available on: www.w3.org/TR/WAI-WEBCONTENT/

[7] EUROPEAN UNION, "Accessibility of the websites and mobile applications of public sector bodies," *Official Journal of the European Union*, Oct. 2016. Accessed on: Apr. 22, 2020. [Online]. Available on: eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri =CELEX:32016L2102&from=EN

[8] SIMON HARPER AND YELIZ YESILADA, "Web accessibility: a foundation for research," *Springer*, London, UK, Sep. 2008.

[9] JONATHAN LAZAR, "Web Usability: A user-centered design approach," *Addison-Wesley*, Boston, MA, USA, Aug. 2005.

[10] LOURDES MORENO, XABIER VALENCIA, J. EDUARDO PÉREZ, AND MYRIAM ARRUE, "An exploratory study of web adaptation techniques for people with low vision," *Universal Access in the Information Society*, June 2020.

[11] J. EDUARDO PÉREZ, XABIER VALENCIA, MYRIAM ARRUE, AND JULIO ABASCAL, "Evaluation of two virtual cursors for assisting web access to people with motor impairments," *International Journal of Human-Computer Studiers*, vol. 132, pp. 81–98, Dec. 2019.

[12] J. EDUARDO PÉREZ, MYRIAM ARRUE, XABIER VALENCIA, AND JULIO ABASCAL, "Longitudinal Study of Two Virtual Cursors for People with Motor Impairments: A Performance and Satisfaction Analysis on Web Navigation," *IEEE Access*, Jul. 2020.

[13] US CONGRESS, "Section 508 of the Rehabilitation Act of 1973," Accessed on: Apr. 22, 2020. [Online]. Available on: https://www.section508.gov/manage/laws-and-policies

[14] XABIER VALENCIA, J. EDUARDO PÉREZ, UNAI MUÑOZ, MYRIAM ARRUE, AND JULIO ABASCAL, "Assisted Interaction data analysis of web-based user studies," In *15th IFIP TC 13 Conference on Human-Computer Interaction* (INTERACT 2015), Bamberg, Germany, Sep. 14–18, 2015, pp. 1–19.

[15] XABIER VALENCIA, J. EDUARDO PÉREZ, MYRIAM ARRUE, JULIO ABASCAL, CARLOS DUARTE, AND LOURDES MORENO, "Adapting the web for people with upper body motor impairments using touch screen tablets," *Interacting with Computers*, vol. 29, no. 6, pp. 794–812, Nov. 2017.

# 2. Conclusions

The methodology applied in this thesis has proven to be valid for improving web accessibility for different groups of users with special needs. This has been achieved through the design and development of different interface adaptation techniques and technical aids, and their subsequent evaluation with users.

The conducted user-testing has allowed us to:

(a) detect the needs of specific groups of users and compare them with the proposals to improve web accessibility found in the literature;

(b) evaluate the benefits of the developed technical aids based on different measures of performance, satisfaction and usability; and

(c) propose extensions of the technical aids to cover the more specific needs of the participating users

It is worth noting the importance of the final interviews with the participants in the different tests performed. In addition to the multiple measures analysed to evaluate the aids, very valuable information was obtained from the responses and opinions of the participants, fostering the continued improvement of the accessibility techniques by proposing new features to be implemented.

The longitudinal study carried out remotely in the participants' homes proved to be very useful and to provide categories of information that cannot be obtained with other types of studies (for instance, supervised test or single session evaluations). Among the most notable benefits that were detected, the following can be highlighted:

(a) achievement of more authentic and valuable data from unsupervised free interaction;

(b) analysis of aspects, such as the habits of use of the technical aids, in order to find out about significant issues more truthfully, such as the technology acceptance by the participants;

(c) gathering of a greater amount of participant interaction data has allowed us to analyse the significance of the results; and

(d) continuous monitoring over time, allowing us to analyse the effect of learning on the different parameters being studied (for instance, performance and satisfaction).

The technical aids developed in this thesis have proven to be beneficial in improving access to the Web for the specific groups of users with special needs for whom they were intended. The assistive technology used by the participants to access the Web (screen magnifiers, alternative pointing devices, etc.) has been a determining factor in deducing which aids were most appropriate for the user in each case. Among the various improvements proposed for inclusion in the technical aids it is worth noting the addition of different customization options gathered from the experiments with users.

The specific conclusions resulting from each of the research tasks carried out is summarized below:

- The RemoTest platform, described in Appendix 1 and Appendix 2, proved to be an exceptional tool to support experimenters carrying out user-testing. It includes features for assisting with the analysis of interaction data recorded during experiments. A straightforward visualization of each participant's interaction data in an understandable way helps experimenters to discover at a glance important issues which occurred during the experiments, and to save a great deal of time when analysing supplemental video recordings, if available. Additionally, the RemoTest also performs heuristic estimates in order to obtain measures relating to the pointer's trajectory that enable further understanding of the participants' behaviours.
- On the other hand, the RemoTest platform proved to be usable and accessible as a result of some empirical studies in remote and on-site settings (see Appendix 2). These results revealed that all the participants, regardless of their characteristics and the assistive technology they used, were able to install the tool when specific instructions were provided. The stimuli automatically generated by the RemoTest platform proved to be accessible to a wide range of users. In addition, based on comments provided by participants during the interviews, some aspects of the design were improved, such as providing shortcuts, larger text and controls, numbering the questions, using clear and simple language, etc. In addition, the user study presented in

Appendix 2 allowed us to issue a set of recommendations for the design of experiments with users with disability.

- The adaptation techniques created in this thesis were able to be embedded in other accessibility enhancing tools. For instance, some of the adaptation techniques proposed were implemented in a transcoding system (see Appendix 3). This transcoding system is able to adapt websites to touch screen mobile devices used by people with motor impairments. The evaluation showed that most users prefer transcoded pages, although a number of improvements are still required in the user interface and in the interaction methods.
- The results obtained in the exploratory study of web adaptation techniques for people with low vision (see Appendix 4), provided enriched information for selecting techniques beneficial for these users. The application of these accessibility techniques improves their experience by minimizing the number of magnification/demagnification actions needed. In addition, it makes it easier to locate navigation components and the main content of web pages as well as decreasing the number of complex actions such as horizontal scrolling.
- The cross and area cursors proposed in this thesis proved to be beneficial for participants with motor impairments. The results of the evaluation of the two cursor aids designed for assisting web access, included in Appendix 5, showed improvements in the performance of the selection of links on web interfaces. In addition, the experience showed that longer learning periods improved the user performance. These results show that people with motor impairments improve their web navigation experience if they are provided with personalized adaptations in order to assist point and click interactions.
- The quantitative results obtained from the longitudinal study of the two cursor aids proposed showed that the cross and area cursors improved the performance of both groups of participants compared to the original cursor. In addition, several improvements were designed in order to reduce the clicking time with both cursors. These results allowed us to propose other enhancements to improve the performance of the standard area cursor on web environments with closely spaced links, as can be found in Appendix 6.

## 2.1. Contributions

This thesis makes the following significant contributions to the field of web accessibility:

A platform for conducting formal web-based user studies has been developed. The objective was to create a tool to assist researchers to specify and conduct experimental sessions with numerous participants. This tool also collects and analyses user interaction data within the Web. The novelty of this tool is to allow user tests to be conducted both in remote and in situ contexts with the goal of analysing and detecting unusual interaction patterns of users with different impairments. This is achieved by recording the appropriate user data and by applying convenient performance and satisfaction measures. Additionally, our tool proved to be suitable for conducting inclusive experiments by presenting accessible environments for users with impairments (people with physical disabilities, blind people and people with low vision). This work was published in two peer-reviewed papers (see Appendix 1 and Appendix 2).

A new heuristic method for delimiting point and click cursor trajectories for link selection has been defined. This measurement is necessary in naturalistic experimental settings (e.g., free web navigation) in order to identify intended cursor movements as no explicit traces of the cognitive process behind the users' intention are registered. Unlike other similar heuristic methods that perform estimations based on the combined data from all individuals, the proposed one uses an individual approach (e.g., to identify valid pauses of cursor aimed movements for each individual user). This approach allows the high heterogeneity among people with disabilities to be taken into consideration. An evaluation with users with motor impairments showed that the heuristic method proposed was able to distinguish navigation patterns and determine differences between participants, for instance, in the assistive technology being used. This work was published in a peer-reviewed paper (see Appendix 1).

Additional knowledge has been provided about what difficulties are encountered by different groups of users relating to the diverse assistive technologies they use to access the Web, and how to improve accessibility through software enhancements. This knowledge derives from the formal evaluations of different technical assistances. The evaluations usually consisted of user tests with groups of people with specific disabilities. The results of this work were published in four peer-reviewed papers, corresponding to people with motor impairments using touch screen tablets (see Appendix 3), people with low vision using desktop computers (see Appendix 4), and people with physical impairments using desktop computers (see Appendix 5 and Appendix 6).

Finally, this thesis has proved that keyboard-only users with the novel cross cursor significantly improved their performance in link selection over the other options tested, as well as showing the users' preference for this new assistance. The cross cursor was designed and developed to assist web access to keyboard-only users, a group of people with motor impairments for which specific research in this area was barely found. The results also revealed a positive learning effect, with better performances being achieved with frequent use over time. These results were obtained from two experiments (a preliminary single-session test and a longitudinal test) performed to compare the cross cursor with other cursor variants. Users of different alternative pointing devices participated in these experiments. This work was published in two peer-reviewed papers (see Appendix 5 and Appendix 6).

As a result of this thesis work, we were invited to publish part of the knowledge built in a book chapter. The prestigious De Gruyter international publishing house included the following peer-reviewed chapter in the book entitled *"Personalized Human-Computer Interaction"*:

- JULIO ABASCAL, OLATZ ARBELAITZ, XABIER GARDEAZABAL, JAVIER MUGUERZA, J. EDUARDO PÉREZ, XABIER VALENCIA, AND AINHOA YERA, "Personalizing the user interface for people with disabilities," in *Personalized Human-Computer Interaction*, M. Augstein, E. Herder, and W. Wörndl, Eds. Berlin, Germany: De Gruyter, 2019, ch. 10, pp. 253–282.

## 2.2. Future work

With regard to the various software developments resulting from this work, which include the multiple technical aids and the platform for conducting user studies, the next step is to develop downloadable versions of these programs for end users. According to the distributed architecture we have followed to collect and analyse user interaction, a broader amount of participants would become available this way, which in turn would allow us to continue studying behavioural patterns and to improve technical aids.

The target users of this thesis were people with physical impairments and people with visual restrictions. With them, we tested technical aids to allow access to desktop computers and touch screen tablets. For the future, we plan to extend this work, applying the same methodology to other types of users, for example, people with cognitive impairments, and to other devices, for example, smartphones, smartwatches, etc.

In this thesis, various interface adaptations were tested through a transcoding system. This system was supported by an annotation procedure that provides the necessary additional semantic information to the html content. The workload for manual annotation of web pages is a major drawback of this approach. Therefore, we plan to investigate alternatives, such as crowdsourcing or gamification techniques, to speed up the annotation process.

In order to broaden the availability of verified software enhancements for assisting the access to Web interfaces we plan to perform further research on other interactions than point and click for link selection. For instance, opening dropdown menus of navigation bars, filling forms with diverse elements, or interacting with new dynamic content.

# 3. Published Works

## 3.1. Appendix 1

This appendix includes the article entitled *"Assisted Interaction Data Analysis of Web-based User Studies"*, which received the Accessibility Award for the most outstanding contribution on ageing, disability and inclusive design at INTERACT 2015. The 15th IFIP TC13 Conference on Human-Computer Interaction is ranked as a CORE A conference.

In order to evaluate the different adaptations proposed in this thesis work to assist access to the Web for people with disabilities, we have used the RemoTest platform to carry out all user tests. The RemoTest platform was built to assist experimenters in designing and conducting remote and in-situ web-based user studies, as well as to gather different data from participants, both automatically (e.g., user interaction) and by means of questionnaires (e.g., demographics and satisfaction), and to facilitate subsequent analyses. In this article the RemoTest platform was presented and evaluated in a study including 16 users with and without motor impairments. Besides contextualizing our tool with other related works, this paper shows how the RemoTest platform was built and applied to define user studies on the Web, as well as demonstrating its usefulness in assisting researchers to detect interaction issues of different user profiles. To do so, interaction data automatically gathered with RemoTest was studied and contrasted with manual analysis from video recordings from participants as a proof-of-concept of the proposed platform. The results of the observational study were presented at a premier international conference on web accessibility:

- J. EDUARDO PÉREZ, MYRIAM ARRUE, XABIER VALENCIA, AND LOURDES MORENO, "Exploratory study of web navigation strategies for users with physical disabilities," In *Proceedings of the 11th Web for All Conference* (W4A '14), Seoul, South Korea, Apr. 7–9, 2014, article no. 20.

# Assisted Interaction Data Analysis of Web-Based User Studies

Xabier Valencia, J. Eduardo Pérez, Unai Muñoz, Myriam Arrue(✉), and Julio Abascal

EGOKITUZ: Laboratory of HCI for Special Needs,
University of the Basque Country (UPV/EHU), Informatika Fakultatea,
20018 Donostia, Spain
{xabier.valencia, juaneduardo.perez, unai.munoz, myriam.arrue, julio.abascal}@ehu.eus

**Abstract.** User behaviour analysis requires defining experimental sessions with numerous participants. In this context, the specification of experiments is a demanding task, as several issues have to be considered such as the type of experiment, the type and number of tasks, the definition of questionnaires and the user interaction data to be gathered. The analysis of collected data is also complex and often requires repeatedly examining recorded interaction videos. In order to deal with these tasks, we present a platform called RemoTest which assists researchers to specify and conduct experimental sessions as well as to gather and analyse the interaction data. This platform has been applied to define different formal user studies on the web and has assisted researchers in detecting the main interaction characteristics of different user profiles and settings.

**Keywords:** Web accessibility · User testing · User behaviour · Accessibility in use

## 1 Introduction

User behaviour when interacting with the Web has been extensively studied in the last decade. This significant research area requires the conducting of experimental sessions with large and diverse groups of users. Experimental sessions have to be carefully planned in order to obtain meaningful results because a minor fault in the design could lead to an erroneous interpretation of results. Researchers need to clearly define the objectives of the experiment, the type of experiment, the stimuli to be presented to participants, the tasks to be performed and the procedure of the experimental sessions. In addition, sometimes specific questionnaires are required in order to explicitly obtain certain data from the participants. The experiment design process is demanding and involves knowledge from different areas such as human factors, hypertext, web technology, etc.

The designed experiments are intended to gather significant interaction data. This data could be gathered through combining different methods such as the traditional ones of recording sessions with video cameras or using specific software components to conveniently collect and store it.


J. Abascal et al. (Eds.): INTERACT 2015, Part I, LNCS 9296, pp. 1–19, 2015.
DOI: 10.1007/978-3-319-22701-6_1

Interaction data analysis is a tedious task especially when only traditional recording methods have been used. Researchers are required to view the recorded videos repeatedly and annotate every meaningful interaction event and data. Among others, cursor movement events are of vital importance when studying the accessibility-in-use of websites or web navigation strategies. For instance, actions such as pointing to a target (buttons, scroll bars, check boxes and radio buttons, etc.), clicking on a target or providing accurate text entry could be very difficult for people with motor impairments due to their lack of dexterity [23]. Examining the cursor movements on recorded images is a hard task that can be alleviated through the application of software components. These components should be efficient in appropriately storing, presenting and preparing all this interaction data for analysis.

In addition, involving an appropriate number of participants for a specific experiment is also challenging. Frequently, this is due to the location and the rigorous timing of sessions. Nowadays, there is an increasing interest in using software tools for conducting experiments remotely. That means that participants are observed while they perform the tasks in their habitual daily environment. This particularly facilitates the conduction of experiments with disabled people, as their environment is already adapted to their needs. Moreover, this type of experiments gathers real interaction data without any obtrusive observation mechanism [2]. It also makes it possible to involve a larger number of participants, as they do not have to physically get to a specific location.

This paper presents a platform called RemoTest that can be applied in remote and in situ experimental sessions. Its objective is to assist researchers when designing experiments, conducting experimental sessions and analysing data gathered in the sessions. A case study based on a real in situ exploratory study with 16 participants (11 people with physical impairments and 5 able-bodied participants) is described in Sect. 4. The objective of this experiment was to analyse the differing interaction characteristics of users with physical impairments and able-bodied users. The interaction of the 11 participants with physical impairments was manually analysed based on the video recordings and results are presented and discussed in [20]. The RemoTest platform was also applied to analyse the data of all the participants so the results of manual and automated analysis are contrasted in this paper as a proof-of-concept of the proposed platform.

## 2 Related Work

In the last few years, the use of tools for remotely conducting user tests has come to the attention of researchers. These tools can be classified depending upon their architecture: server-side tools, proxy tools and client-side tools. Each one has its drawbacks and advantages. Server-side tools are the most unobtrusive. Participants are not required to install or configure anything on their systems. Among the HTTP requests from the servers, user interaction data could be also gathered by modifying the web pages with code to track the interaction data (for instance, JavaScript). One drawback is that the conducted tests could be only based on websites located on the servers researchers have access to. Contrarily, proxy tools require participants to configure their web browser in order to access to the proxy. This type of tool enables more information to be gathered than via the server logs as well as permitting the tests to be conducted on any website but however, it is not possible to capture all the participants actions, such as the events on the browser. Finally, client-side tools are the most obtrusive ones since they require the installation of some add-on applications to the participant's system. Nevertheless, they are the best option for researchers because they capture all the user interaction data needed in a user behaviour study, such as cursor movements, scroll events, clicking actions, browser events (backward/forward button, print/save page, add bookmark), etc. without modifying the original web pages.

There are some commercial tools such as Google Analytics [11], Loop11 [15] and Morae [17]. Google Analytics is a server-side tool. Its objective is to obtain general connection data about the users of a specific web site but not to conduct formal experiments. It records information such as the network provider or the browser used by users accessing to a specific web site enriched with Google Analytics. Loop11 is a proxy tool developed for conducting user tests. It includes features that facilitate the definition of web tasks or questionnaires. Regarding the data analysis, it provides click stream reports to visualize the paths that the user followed, click heat maps and the option of visualizing the data in real time. There is a version of Loop11 using AccessWorks devoted to performing user tests with disabled people. Nevertheless this tool does not capture browser events. These events can be used as user disorientation indicators (for example, several clicks on the browser backward button in a task may sometimes imply user orientation problems). Morae is a client-side tool which stores user interaction data when interacting with either standalone applications or web sites. In addition, it provides tools to enable the observation and annotation while the user is performing the test. The tool is quite complex to configure due to the amount of available options provided. Sometimes, programmer skills are required in order to gather some interaction events.

There are many examples of tools developed in academic contexts. Webquilt [12] is a proxy-based tool that stores only information obtained from HTTP requests. NIST WebMetricsSuite [5, 22] is a server-side tool that provides methods to assess the usability of web pages by analysing the path followed by the user. These tools only gather events related to clicks. Therefore, no user interaction data is gathered. Other tools such as WET [8] or USAPROXY [3] are more comprehensive as they also enable the detection of problematic web elements by gathering user interaction data such as mouse movements and keyboard events by injecting Javascript code to the original web pages. Nevertheless, these tools do not assist in formal experiment definition, as they do not include features for task definition or the elaboration of questionnaires. WebRemUSINE tool [18] provides researchers with features for defining formal experiments. It is a client-side tool that uses both technologies Java Applets and JavaScript code to capture the interaction events. The tool is based on ConcurTaskTrees (CTT) task model annotation [19] for defining experiments. It detects usability problems analyzing the differences between the path followed by users to perform the defined tasks and the specified task model in CTT. It requires high knowledge levels and expertise to define the task models, making this tool only usable by expert researchers. In addition, this tool does not provide features for defining questionnaires to fill in by participants during the experimental sessions.

On the contrary, Curious Browser [5] presents questionnaires to the user after every new visited page, with the aim of rating its interest. Nevertheless the tool does not provide methods to define formal experiments like defining tasks, pre/post task questionnaires and so on. Uzilla [7] is another comprehensive tool for defining formal experiments. It also provides features for defining questionnaires to be filled in by participants during the session. However, there is not much information about the suitability of the created questionnaires for people with disabilities.

Some other approaches can be found relating to the study of users' interaction performance in the wild. Gajos et al. [9] developed a system devoted to minimize the gap between the results of pointing performance in a laboratory and those in the wild. The tool distinguishes between deliberate and distracted mouse pointer movements. Another similar tool was proposed by Hurst et al. [14]. This tool classifies users' characteristics based on their input events. Both tools provide valuable information about user behaviour or characteristics but do not allow performing formal experiments.

Power et al. list a number of requirements that a remote user tool should meet to be able to conduct experimental sessions with users with disabilities [21]. Some of them are related to participants and other to researchers. The requirements related to participants include the following: provide features to record demographic data (P1), specify the technology used (OS, browser, assistive technology) (P2), select the trials (P3). The ones related to researchers are the following: provide features to test customized and "real" websites (R1), define tasks for a set of users (R2), specify a set of questions to the user, before and/or after the task has been completed (R3), provide instructions and training documents for each trial (R4). Table 1 shows the requirements fulfilled by the academic context tools presented in this section.

**Table 1.** Classification of user testing tools according to the requirements proposed by Power et al. and their location.

| Name | Location | P1 | P2 | P3 | R1 | R2 | R3 | R4 |
|---|---|---|---|---|---|---|---|---|
| NISTWebMetrics Suite | Sever | No | No | No | Partial | No | No | No |
| WET | Server | No | No | Yes | Partial | No | No | No |
| WEBQUILT | Proxy | No | No | No | Yes | No | No | No |
| Curious Browser | Client | No | No | No | Yes | No | Partial | No |
| UsaProxy | Proxy | No | No | No | Yes | No | No | No |
| WebRemUSINE | Server | No | No | Yes | Yes | Yes | No | No |
| Uzilla | Client | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Gajos et al. | Client | No | No | No | Yes | No | No | No |
| Hurst et al. | Client | No | No | No | Yes | No | No | No |

The Remotest platform is a client-side tool that provides all the necessary features for defining formal experiments and fulfils the requirements proposed by Power and colleagues. In addition, the questionnaires created by the tool have proven to be accessible to people with disabilities in several evaluations.

# 3 The Remotest Platform

The RemoTest platform provides the necessary functionalities to assist researchers to define experiments, manage experimental remote/in situ sessions and analyse the gathered interaction data. This platform admits a wide range of experiments. The objective can be, for instance, to study user behaviour when performing a task in different websites, to analyse and compare navigational strategies of different types of participants when interacting with the same website, to evaluate the accessibility-in-use of several websites, to gather significant information through surveys, to measure user satisfaction when using a certain web service and to analyse user performance improvement when interacting with adapted versions of original web pages and so on.

The architecture of the platform has been designed taking all these different types of experiments into consideration. In this case, we opted for a hybrid architecture model that includes some functions in a client-side module and the other ones in some server-side modules. The platform is split into four modules: Experimenter Module (EXm), Participant Module (PAm), Coordinator Module (COm) and Results Viewer Module (RVm).

## 3.1 Experimenter Module

This module provides a set of functionalities for defining all the components of the experiments. It is a server-side module which can be accessed by experimenters from any computer with an Internet connection. All the definition process is performed by the use of a web application and has been divided into five main steps:

Step 1: Specifying the type of experiment. This first step is intended to specify the type of the experiment (survey or navigation tasks on the web) and general characteristics of the experiment such as the stimuli to be presented in experimental sessions, number of questionnaires (demographic data questionnaire, satisfaction questionnaire, etc.)

Step 2: Determining the tasks and stimuli of the experimental sessions. Depending on the type of experiment specified in Step 1 the platform would ask to provide information about the tasks to be performed by participants and the stimuli to present. Two main types of tasks can be defined: "Fill in Questionnaire" and "Web Navigation". There is no limit to the number of tasks per experiment

For each task some details have to be provided. For instance, for the "Fill in Questionnaire" task the questions, the type of questions (open-ended or closed) and possible options for answers (likert scales, ranges) have to be defined. There is some other information regarding title, id, etc., that has to be completed in order to ensure the accessibility of the created questionnaire. The "Web Navigation" task also requires some data. Currently, there are two types of tasks in this category: "searching target" tasks and "free navigation" tasks. The former one refers to tasks where participants are required to find a specific target (such as a specific button, link, form, etc.) whereas the latter one entails navigation without any concrete objective. Both require certain information from the experimenter, for instance, the starting URI, time limit for the task, etc. Searching target tasks also require the specific target URI to be provided. In addition, a title and a description can be added to each task so that, in the experimental sessions, an accessible task explanation web page is presented to participants before starting the task and an alert after performing the task.

Once all the tasks are defined, the dependencies between them have to be explicitly stated. This module provides features to define if any task should be presented just before or after another one. For example, in an experiment to measure satisfaction of users when interacting with a certain website, a "Fill in Questionnaire" task about satisfaction has to be presented just after a "Web Navigation" task.

Step 3: Defining the procedure of the experimental sessions. In this step, the number of groups of participants, the tasks presented to each group and the task sequence for each group has to be defined. Experimenters are asked to provide the number of groups in the sample. Specifying a unique group leads to a within-subject experiment. Therefore, each participant will perform all the defined tasks. The task sequence should be counterbalanced between participants and the experimenter is required to define one method (manual, latin square, rotation or random). In between-groups experiments, the tasks to perform by each group are manually selected by the experimenter. The task sequence for each group can be defined manually or automatically (using random method). In all cases, the EXm considers the dependencies between tasks and any other detected inconsistencies are notified to the experimenter

Step 4: Specifying interaction data to be gathered. The experimenter selects the interaction data to be gathered in the experimental sessions. This data will be automatically gathered by the PAm during the experimental sessions. Currently, the interaction data gathered by the platform are the following:

- Browser related events. Active tab, opening tab, closing tab, changing tab, backward button, forward button, vertical/horizontal scroll movements, screen resolution, window size, mouse context menu, favourites management…
- Cursor/Mouse related events. Clicking left button, clicking right button, using mouse wheel, size of clicked elements, hover events, size of hovered elements, tracking cursor movements.
- Keyboard related events. Key down, key up, key pressed.

Step 5: Selecting the sample of participants. The RemoTest platform provides functionalities to maintain information about participants in a database through the module EXm. This tool includes options such as manual selection of participants in each group, randomly creating groups from the selected participants or establishing some kind of criteria to select the sample, (such as gender, age, assistive technology used, etc.), applying filtering criteria to the query

The information about the experiments is stored in an XML-based language specifically developed for this platform, the Experiment Specification Language.

### 3.2 Coordinator Module

The Coordinator Module is a server-side type module which has been developed as a Web Service. The objectives of this module are the following:

- Storing and managing the experiments defined by different experimenters applying features of EXm.
- Creating stimuli to present during the experimental sessions (questionnaires, task description web page, task completion alert, etc.).
- Defining the personalized experimental sessions specifications.
- Collecting and storing interaction data obtained in experimental sessions.
- Maintaining information about experimenters, experiments, participants in databases.

This module creates all the necessary stimuli for the experiments (surveys, questionnaires, task description web pages, alerts, etc.) according to the definition provided by means of the Experiment Specification Language. In addition, it prepares the participant groups and performs the counterbalancing methods, when necessary, in order to create personalized experimental sessions for each participant or group of participants. This leads to the obtaining of specific personalized experimental session specifications for each participant defined in an XML-based language developed for this purpose, the Experimental Session Controlling Language. The interaction data created in the sessions are managed and stored in a MongoDB database.

### 3.3 Participant Module

The participant module is a client-side type module. Therefore participants have to install this module in their computers. It is currently an add-on for Firefox browser but it can be easily migrated to other platforms, since this module is mostly based on JavaScript and XML. It processes the Experimental Session Controller Language for correctly conducting participants' experimental sessions. In addition, it gathers all the interaction data, as well as the HTTPRequest information during the whole session and asynchronously sends to COm by using AJAX technology.

### 3.4 Results Viewer Module

The RVm is a server-side type module which deals with the presentation of the interaction data gathered in experimental sessions. For this purpose, RVm implements functions for collecting the data from COm, structuring it in understandable blocks of events and presenting them to the experimenter as a web application. Some statistics by pages, tasks or users can be visualized:

- Rapidity measures: Time on page, cursor average speed, cursor acceleration.
- Accuracy measures: Trajectory distance (cursor travel distance), curvature index (CI) relation between optimal path and path followed, distances to the centre of the target and to the last click and ratio of start-end position amplitude to start-target centre amplitude [1].

Figure 1 shows the information presented regarding the performance of a participant in a visited web page.

| Rapidity Measures | |
|---|---|
| Time on page: | 00:00:16 (hh:mm:ss) |
| Average-speed: | 91,92 Pixels/seconds |
| Number of ZA: | 56 |
| Number of acceleration to deceleration phases: | 27 |
| Number of deceleration to a acceleration phases: | 22 |
| **Accuracy Measures** | |
| Trajectory distance: | 1470,78 Pixels |
| Curvature Index: | 1,21 |
| Distance between start to end point: | 1213,66 Pixels |
| Distance between start point to target center: | 1180,15 Pixels |
| Ratio of start-end position amplitude to start-target center amplitude: | 1,03 |

**Fig. 1.** General information about the performance of a participant in a visited web page.

This module includes a tool for comparing the performance of each participant in all the visited pages. Figure 2 show the charts generated for comparing the trajectory distance of a participant in the visited pages.

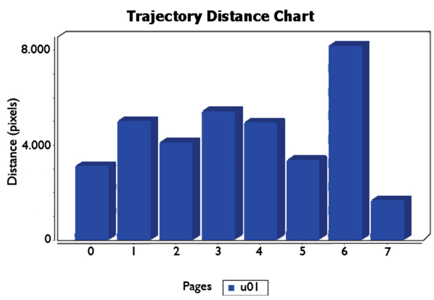


**Fig. 2.** Comparing chart about the trajectory distance parameter of a participant

The general information shown by the RemoTest platform assist researchers to obtain detailed information about the performance of each participant during the experimental session. In addition, this module also provides several graphs for each visited page. One is the "Distance to target chart", where the distance to the target at every moment can be seen. With this chart researchers can easily identify users' problems when aiming at the target, see Fig. 3 Left. As can be seen, once cursor is in the target nearby the user requires several attempts to place the cursor on it.

Furthermore, associated with this first chart the tool also provides another graph that shows the distance to the target but starting from users' intention time to click the

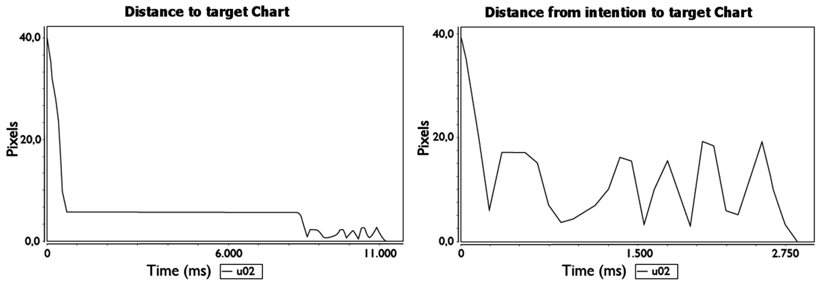


**Fig. 3.** "Distance to target" (Left) and "Distance from intention to target" (Right) charts

target, see Fig. 3 Right. The intention time to click is automatically calculated by the algorithm described in the Sect. 3.4.1.

In addition, graphs of cursor's movement speed and acceleration charts can be found. Through these graphs researchers can appreciate for instance, the different cursor movements taking into account the input device used by the user.

Figure 4 shows one participant (who moves the cursor by pressing the numerical keypad with a head pointer) creates similar speed and acceleration peaks followed by short pauses, the time it takes the user to change direction to press another key with the head pointer.

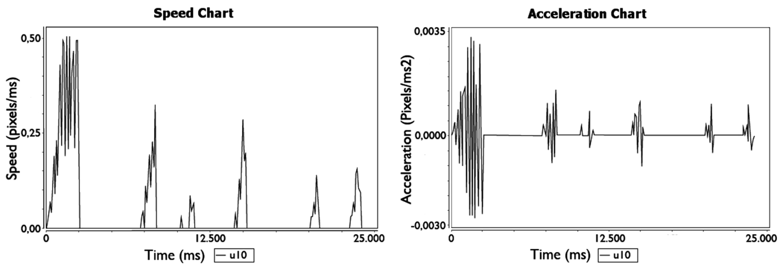


**Fig. 4.** Speed chart (left) and acceleration chart (right) of a participant (U10 in the case study)

Another important feature is the ability of the tool to make comparisons. In first place, researchers can view a comparison between different statistics about all visited pages in a task of a selected user. In the same way, the application also allows to compare these statistics to all participants for a given task. So, RVm is able to display automatically generated bar charts comparing the average or median, of the trajectory distances, curvature index, times to point and click or cursor speeds.

This straightforward visualization may assist experimenters in discovering relevant bits of users' interactions, and in case of also having complementary video recordings, fast locating the corresponding moments and save huge amount of time viewing and analysing video images.

In order to extract aimed movements from all the cursor's kinematic data recorded by the RemoTest platform and study pointing trajectory-related measures, we have processed the data as follows.

#### 3.4.1 Delimiting Pointing Trajectories

A pointing trajectory starts when a user resolves to move the cursor to reach an objective. Controlled laboratory experiments can specify restricted interactions to make the beginning of cursor movement explicit [9]. In contrast to those studies, more naturalistic settings with untagged web interactions do not permit the register of any explicit trace of the cognitive process behind the users' intention. In those cases, as in ours, heuristics are needed to estimate the beginning of the aimed movements [4, 10].

Among the other possible aimed movements within web GUIs (e.g., hovering over an element to see its tooltip), we have decided to focus on pointing trajectories that end with a navigational click and a posterior page load event. In our case we have considered the following bases for delimiting pointing trajectories.

*Beginning of movement.* A pointing to navigational click trajectory may correspond to the complete cursor movement recorded along a page, however behaviours such as moving the cursor while reading the content of the page provoke the need to analyse all the cursor trajectory throughout every page.

We have considered the first cursor move event recorded within every page as the beginning of movement candidates for pointing to navigational click trajectories. Each time a page scroll interaction occurs, actual cursor trajectory (if it exists) is rejected as a page candidate for pointing to navigational click trajectory, restarting a new pointing trajectory when the next cursor move event is triggered.

Additionally, pauses take place along cursor trajectories are useful to segment pointing trajectories. Aimed movements consist of several sub-movements separated by pauses [16], so trying to delimit a pointing trajectory as a cursor movement without pauses is not feasible.

*Calculating Valid User Pauses.* A pause during an aimed movement may correspond to a sub-movement transition, or in case of being long enough (movement stop) to cursor trajectory segmentation (Fig. 5).

Unlike other studies where a unique value for all users serves to distinguish between valid pauses during cursor aimed movements and movement stops [4, 10], we have followed an individual approach for this purpose. As we have observed from the interaction of the physically impaired participants of our study, valid pauses during aimed movements vary among users depending on the computer pointing device used and their ability with it (Fig. 5). For instance, a numeric keypad user needs more time to change cursor trajectory (notable pauses between strokes) than a joystick user. We believe that calculating a movement stop threshold for every user will improve the quality of the segmentation and therefore the pointing trajectory-related measures.

To calculate each user movement stop threshold, we have taken into account all the intervals in which the cursor velocity falls to zero within each every page, registering these time durations by user. From this data, we have calculated the median value of all

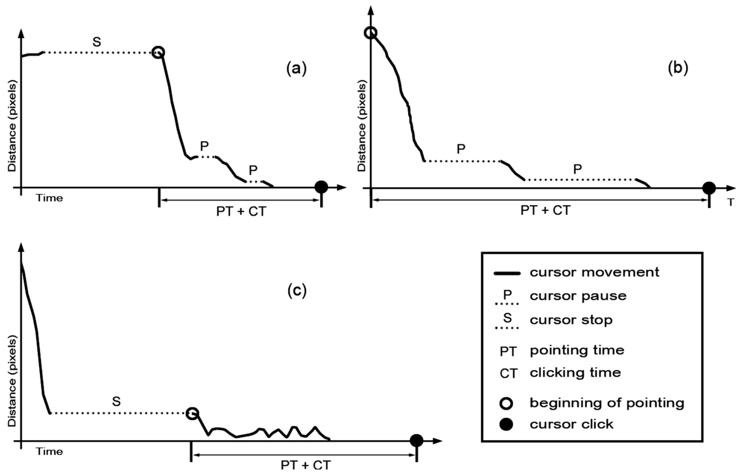


**Fig. 5.** Beginning of pointing trajectories estimations for 3 physically impaired participants using different computer pointing devices: (a) joystick, (b) numeric keypad with a head pointer and (c) oversized trackball.

collected observations by user, and so discarded stops duration that were two quartile deviations or more away from the participant's median. We have used the median and quartile deviation values to reduce the importance of outliers within each distribution. Through an extensive observation of our data we have concluded that values obtained this way were reasonable.

*End of movement.* As mentioned above, we only focus on pointing trajectories that end with a navigational click and a posterior page load event. Thus we have discarded all pages from our data logs without pointing to navigational click trajectories, for instance using keyboard shortcuts or browsing history navigation without moving the cursor.

## 4 Case Study

This section is devoted to describing the benefits of using the RemoTest platform to define experiments, conducting experimental sessions and collecting and analyzing interaction data. This case study is based on a specific experiment carried out by the authors with 16 participants, 11 people with upper-body physical impairments (U01–U11) and 5 able-bodied people (U12–U16). The objective of the experiment was to analyse the different navigation strategies used by the participants. Web expertise varied among participants as well as the assistive technology and input devices used for accessing the Web. Table 2 shows information about the participants.

The experimental sessions were carried out locally in different settings (experimenters moved to the specific location and assisted participants during all the session) and used different computers: 7 experimental sessions were conducted at the Elkartu

**Table 2.** Information about the participants

| User | Expertise | Input devices | Setting |
|---|---|---|---|
| U01 | Medium | Reconfigured mouse and keyboard | Elkartu |
| U02 | High | Oversized trackball and keyboard | Lab |
| U03 | High | Joystick and keyboard with cover | Home |
| U04 | Medium | Joystick and keyboard | Elkartu |
| U05 | Medium | Joystick and on screen keyboard | Elkartu |
| U06 | Low | Joystick and keyboard with handstick | Elkartu |
| U07 | Low | Reconfigured touchpad and keyboard | Elkartu |
| U08 | High | Numeric keypad and keyboard | Home |
| U09 | High | Numeric keypad and keyboard | Home |
| U10 | High | Head pointer and reconfigured keyboard | Elkartu |
| U11 | Medium | Head pointer and reconfigured keyboard | Elkartu |
| U12 | High | Standard mouse and keyboard | Lab |
| U13 | High | Standard mouse and keyboard | Lab |
| U14 | High | Standard mouse and keyboard | Lab |
| U15 | High | Standard mouse and keyboard | Lab |
| U16 | High | Standard mouse and keyboard | Lab |

premises (a local association of people with physical disabilities), 6 participants carried out the experimental session in a laboratory of the Computer Science School at the University of the Basque Country and the remaining experimental sessions (3 of 16) were conducted at the participant's home. The platform used was similar in all cases: Windows operating system (Windows 7 except of U02 who used Windows XP) and Mozilla Firefox browser.

Participants were asked to install the PAm module on the local computer (their usual computer in some cases). To this end, the PAm Firefox add-on was placed on an URL and a web page with clear instructions of how to install it was created. This task was useful for detecting some minor issues in the instructions that needed to be fixed. Even though it was not the object of the study, this preliminary installation task was also useful in order to test the adequacy of using a client-side tool for conducting experimental sessions. Despite only one participant (U02) having previous experience in installing this type of software, all participants were able to install the PAm module including the ones with low-level expertise in using the computer (U06, U07).

### 4.1 Experiment Definition

The experiment consisted of three tasks: filling in a questionnaire about demographic data (Task 1), free navigation task with 5 min duration (Task 2) and searching for a target task with a maximum duration of 10 min (Task 3).

Web navigation tasks (Task 2 and Task 3) were performed on the Discapnet website [www.discapnet.com] which is specialized in providing information to people with disabilities. The order of these tasks was predefined since the free navigation task was intended to familiarize the users with the website.

RemoTest platform was used for defining the experimental sessions. The definition process was as follows:

Step 1: Specifying the type of experiment. This experiment was a web navigation experiment

Step 2: Determining the tasks and stimuli of the experimental sessions

- Task 1: The EXm module guided the researcher in the process of defining the questions and possible options for responses.
- Task 2: This task was a "Web Navigation" type task with free navigation category. The EXm module asked for information regarding this task such as: duration, task description text, URL of the website, task completion text, etc.
- Task 3: This task was "Web Navigation" type task and search target category. The EXm module asked for similar information as in Task 2 and, in addition, the specific target has to be defined. In this case the target is a specific URL in Discapnet.

Step 3: Defining the procedure of the experimental sessions. In this experiment all participants had to do the same tasks and the order was the same in all cases. This information was inserted in the RemoTest platform

Step 4: Specifying interaction data to be gathered. The EXm module presents all the interaction data PAm module is prepared to be gathered so the experimenter could select the most interesting in each case

Step 5: Selecting the sample of participants. The RemoTest platform provides functions to select the participants of an experiment

The experiment definition process leads to obtain an XML file containing the experiment definition based on the Experiment Specification Language.

### 4.2 Creating the Experimental Sessions

The COm module automatically created the necessary stimuli for the experiment based on its definition. In this particular case study, the stimuli to create were the following: the demographic data questionnaire, the free navigation task description web page, the free navigation task completion web page, the target searching task description web page and the target searching task completion web page.

In addition, the information included in the experiment definition XML file is applied to create the personalized experimental sessions. In this specific case study, all the experimental sessions will be similar.

### 4.3 Conducting the Experiment

The PAm module installed on the client-side guided participants throughout the experimental session based on the XML file in Experimental Session Controlling Language. It managed the duration of tasks, the sequence, the presentation of stimuli, the gathering of data provided by each participant in questionnaires and the interaction data. All the information gathered by this module (that explicitly provided by the participant by answering questionnaires, and the interaction data implicitly obtained by the platform) were sent to the COm module in an asynchronous manner without interrupting the participant session. The experimental session completion web page was shown when all tasks were completed.

### 4.4 Interaction Data Analysis

The analysis presented in this section is focused on the cursor movement characterization features included in the RVm module and described in Sect. 3.4. These features were applied to the interaction data gathered in Task 3 (Searching a target task) allowing the researchers to identify navigation patterns and profiles between the participants. In this case, a total number of 323 web pages were visited by participants in this task. Note that U11 had to be excluded from the analysis due to fact that she decided to leave the experimental session before finishing the tasks. Applying the pre-processing algorithm defined in Sect. 3.4.1, the data of 133 web pages were selected to perform the cursor movement characterization analysis (23 web pages were excluded because of their lack of cursor movements as they were the result of repeatedly pushing the browser back button, 167 web pages were removed due to some detected problems in the PAm module when gathering interaction data when the cursor was out of the web browser window, a new web page was loaded, cursor position errors with *iframe* components in the page). Table 3 shows the total number of the visited pages by each participant and the total number of pages to be considered for the cursor movement characterization.

The information in Table 3 allows researchers to select the participants for further analysis. For instance, some of the participants (U04, U05, U06, U07 and U13) have fewer than 5 pages for analysing. Some parameters (curvature index, pointing time, clicking time) may not be accurate enough due to this lack of data. At this point, it is possible to select the participants to be excluded for the cursor movement characterization process. In this case, we excluded those with fewer than 5 pages to be analysed.

Some parameters were calculated based on the interaction data obtained in Task 3 for the rest of participants. Median values of cursor speed were automatically calculated for each participant. Figure 6 shows the resulting values. There is a considerable difference in speed parameter between participants. Able-bodied participants using standard mouse and keyboard (U12,U14,U15,U16), obtained the highest values.

The curvature index parameter allows problematic cursor movement patterns to be detected. It measures the relation between the optimal path and the followed path. The calculated values are shown in Fig. 7. U02 has the highest value meaning that there is a lack of precision in cursor movements. Actually, he was a trackball user with low

**Table 3.** Information about the total number of visited pages (VP) in Task 3 and the analysed pages (AP) for cursor movement characterization.

| User | Visited pages | Analysed pages |
|---|---|---|
| U01 | 8 | 5 |
| U02 | 32 | 10 |
| U03 | 33 | 12 |
| U04 | 14 | 1 |
| U05 | 5 | 3 |
| U06 | 5 | 2 |
| U07 | 6 | 2 |
| U08 | 31 | 15 |
| **User** | **Visited pages** | **Analysed pages** |
| U09 | 15 | 7 |
| U10 | 13 | 6 |
| U12 | 26 | 10 |
| U13 | 9 | 4 |
| U14 | 40 | 24 |
| U15 | 50 | 15 |
| U16 | 36 | 17 |

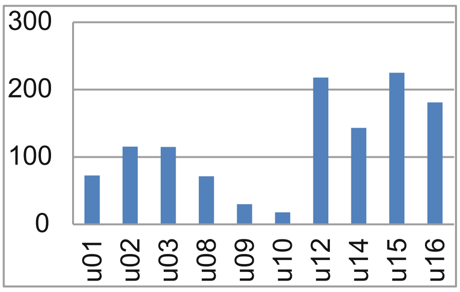


**Fig. 6.** Comparing chart about the speed parameter automatically obtained for each participant.

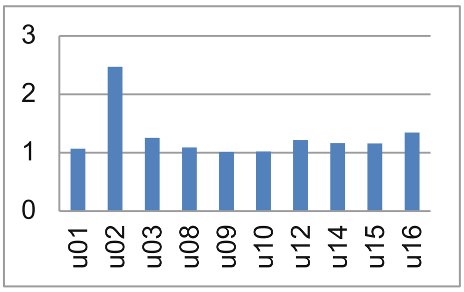


**Fig. 7.** Comparing chart about the curvature index automatically obtained for each participant

upper-body movement precision. Users using numeric keypad (U08,U09,U10) are the ones with the best values (1.09, 1.01 and 1.02 respectively). Actually, the use of keypad for moving the cursor produces linear and more precise movements.

Other significant measures for characterizing cursor movements are the time required for pointing a target and the time required for clicking on a target. Both measures are considered for detecting problematic situations in cursor movements. For instance, a high value of time for clicking on a target may indicate that the user has problems trying to perform the click action due to the size of the target and lack of precision or some distracting content around the target. The algorithm presented in Sect. 3.4.1 is applied for calculating both measures for each participant. Figure 8 shows the means of these measures for each participant automatically obtained by the RemoTest platform. It can be observed that participant U02 has the highest mean value for clicking on targets (3059 ms) whereas U14 is the one with lowest mean value (536 ms). U02 is one of the participants with physical disabilities using a trackball and U14 is an able-bodied participant. Able-bodied participants (U12,U14,U15,U16) obtained the better values for both measures (mean value for clicking is 855.32 ms and mean value for pointing is 1931.41 ms) than participants with physical impairments (mean value for clicking is 1872.53 ms and mean value for pointing is 4351.27 ms).

Recordings of participants with physical disabilities were manually analyzed and the time required for pointing to a target measure was annotated for each visited web page. The mean of automatically obtained values was compared with the mean of those obtained manually. These values can be found in Table 4. It can be observed that the values obtained automatically were generally lower than those manually gathered from the video analysis. Kendall's Concordance Tests was performed to analyse the agreement between rankings provided by both measures. The value obtained by the

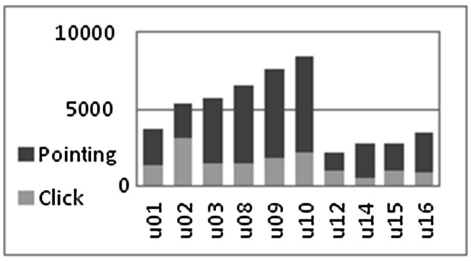


**Fig. 8.** Comparing chart about the time for pointing a target and time for clicking on a target for each participant in the experiment

**Table 4.** Mean values of the time required for pointing to a target automatically obtained by RemoTest platform (APM) and manually obtained based on video analysis (MPM).

| | U01 | U02 | U03 | U08 | U09 | U10 |
|---|---|---|---|---|---|---|
| APM (ms) | 2318 | 2314.6 | 4289.5 | 5107.6 | 5728.7 | 6349.2 |
| **MPM (ms)** | 6440 | 5080 | 3240 | 7190 | 8570 | 20480 |

concordance coefficient was 0.73 (p = 0.055) meaning that there is some correlation between both rankings.

## 5 Discussion

The results automatically obtained by the RemoTest platform proved to be useful for characterizing the cursor movements and detecting different profiles between participants in a formal experiment. In fact, the observation of values obtained in curvature index, cursor speed and the time to clicking on a target assists researchers in detecting problematic situations in experimental sessions due to lack of precision, inappropriate target dimension, features of assistive technology used by participants, etc. It would be possible to detect some problematic situations, even if the experiment were carried out in remote settings. For instance, behaviour of participant U01 who uses a standard mouse differs in speed values for other participants using the same input device. It may be concluded that this participant requires some assistance by means of web interface adaptation mechanisms in order to improve his performance.

Observing the results in Fig. 9 it can be appreciated that participants U08, U09 and U10 obtained considerably higher values for their typical pauses between cursor sub-movements. These values may indicate some difficulties for starting the cursor movements as well as participants' fatigue during experimental sessions. In this case study, U10 is the one with the highest value (955.5 ms). This participant used a head pointer that may cause fatigue and sometimes disorientation as the cursor is out of sight when pressing the key during the cursor movements.

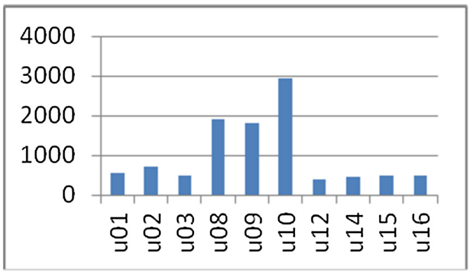


**Fig. 9.** Median values by user for motionless intervals along cursor trajectories

The algorithm for automatically calculating the required time for pointing to a target proved to be useful for ranking purposes. The values obtained are not accurate comparing with the values obtained manually. However, this could be due to lack of data used for automated analysis after the pre-processing. The current version of RemoTest platform will be more accurate in this sense as some of the errors when gathering coordinates of the cursor position have been fixed.

## 6 Conclusions

The RemoTest platform was conceived with a clear objective of assisting experimenters when dealing with the tedious task of carrying out remote/in situ experiments with web systems. The functionalities included in the platform support experimenters throughout the entire process: designing the experiment, conducting experimental sessions, gathering interaction data and analyzing results.

The experiment definition features consider the specification of different kind of studies. For instance, user behaviour studies on concrete web tasks, comparative studies on navigational strategies when using different types of assistive technology, accessibility-in-use evaluations, web surveys for collecting specific information, user satisfaction measuring, user performance improvement analysis when applying adaptation techniques and so on.

Moreover, the RemoTest platform includes features for assisting with the analysis of interaction data recorded during experiments. A straightforward visualization of each participant interaction data in an understandable mode helps experimenters discovering valuable issues occurred during experiments at a glance, and saving huge amount of time when analysing complementary video recordings if available. Additionally, the tool also performs heuristic estimations in order to obtain pointing trajectory-related measures that enable further understandings within participants' behaviours. Future work will be focused on testing new performance measures to enrich users' characterization, searching for additional parameters that lead to better aimed movements estimations, and studying the application of the RemoTest tool for analysing interaction data of other groups of impaired users.

**Acknowledgements.** The authors gratefully acknowledge the support of the Spanish Ministry of Science and Innovation through Project ModelAccess (TIN2010-15549). The EGOKITUZ Research Laboratory is supported by the Department of Education, Universities and Research of the Basque Government (Grant# IT-395-10) and the University of the Basque Country (Grant# UFI11/45).

## References

1. Almanji, A., Davies, T.C., Stott, N.: Using cursor measures to investigate the effects of impairment severity on cursor control for youths with cerebral palsy. Int. J. Hum. Comput. Stud. **72**(3), 349–357 (2014)
2. Apaolaza, A., Harper, S., Jay, C.: Understanding users in the wild. In: 10th International Cross-Disciplinary Conference on Web Accessibility (W4A 2013), Article 13, p. 4. ACM, New York (2013)
3. Atterer, R., Wnuk, M., Schmidt, A.: Knowing the user's every move: user activity tracking for website usability evaluation and implicit interaction. In: 15th International Conference on World Wide Web, pp. 203–212. ACM, New York (2006)
4. Chapuis, O., Blanch, R., Beaudouin-Lafon, M.: Fitts' law in the wild: a field study of aimed movements. Technical report 1480, LRI, University Paris-Sud, France, p. 11 (2007)

5. Claypool, M., Le, P., Wased, M, Brown, D.: Implicit interest indicators. In: 6th International Conference on Intelligent User Interfaces (IUI 2001), pp. 33–40. ACM, New York (2001)
6. Cugini, J., Scholtz, J.: VISVIP: 3D visualization of paths through web sites. In: 10th International Workshop on Database and Expert Systems Applications, pp. 259–263, IEEE (1999)
7. Edmonds, A.: Uzilla : a new tool for web usability testing. Behav. Res. Meth. Instrum. Comput. **35**(2), 194–201 (2003)
8. Etgen, M., Cantor, J.: What does getting WET (Web Event-logging Tool) mean for web usability? In: 5th Conference on Human Factors and the Web (1999)
9. Gajos, K.Z., Weld, D.S., Wobbrock, J.O.: Automatically generating personalized user interfaces with Supple. Artif. Intell. **174**, 910–950 (2010)
10. Gajos, K., Reinecke, K., Herrmann, C.: Accurate measurements of pointing performance from in situ observations. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2012), pp. 3157–3166. ACM, New York (2012)
11. Google Analytics 2014. http://www.google.es/analytics
12. Hong, J., Heer, J., Waterson, S., Landay, J.A.: WebQuilt: a proxy-based approach to remote web usability testing. ACM Trans. Inf. Syst. **19**(3), 263–285 (2001)
13. HttpFox 2014. https://addons.mozilla.org/es/firefox/addon/httpfox/
14. Hurst, A., Hudson, S.E., Mankoff, J., Trewin, S.: Distinguishing users by pointing performance in laboratory and real-world tasks. ACM Trans. Access. Comput. **5**(2), 27 (2013)
15. Loop11 2014. http://www.loop11.com
16. Meyer, D., Smith, J., Kornblum, S., Abrams, R., Wright, C.: Optimality in human motor performance: ideal control of rapid aimed movements. Psychol. Rev. **95**, 340–370 (1988)
17. Morae 2014. http://www.techsmith.com/morae.html
18. Paganelli, L., Paternò, F.: Intelligent analysis of user interactions with web applications. In: 7th International Conference on Intelligent User Interfaces (IUI 2002), pp. 111–118.3. ACM, New York (2002)
19. Paternò, F., Mancini, C., Meniconi, S.: ConcurTaskTrees: a diagrammatic notation for specifying task models. In: Howard, S., Hammond, J., Lindgaard, G. (eds.) IFIP TC13 Interantional Conference on Human-Computer Interaction, pp. 362–369. Chapman and Hall Ltd., London, UK (1997)
20. Pérez, J.E., Arrue, M., Valencia, X., Moreno, L.: Exploratory study of web navigation strategies for users with physical disabilities. In: Proceedings of the 11th Web for All Conference, Article 20, p. 4. ACM, New York (2014)
21. Power, C., Petrie, H., Mitchell, R.: A framework for remote user evaluation of accessibility and usability of websites. In: Stephanidis, C. (ed.) Universal Access in HCI, Part I, HCII 2009. LNCS, vol. 5614, pp. 594–601. Springer, Heidelberg (2009)
22. Scholtz, J., Laskowski, S., Downey, L.: Developing usability tools and techniques for designing and testing web sites. In: 4th Conference on Human Factors and the Web (1998)
23. Trewin, S.: Physical impairment. In: Harper, S., Yesilada, Y. (eds.) Web Accessibility - A Foundation for Research, pp. 37–46. Springer, London (2008)

## 3.2. Appendix 2

This appendix corresponds to the article entitled *"Inclusive Web Empirical Studies in Remote and In-Situ Settings: A User Evaluation of the RemoTest Platform"* that was published in the International Journal of Human–Computer Interaction (IJHCI) in April 2019. That year the journal reached an impact factor of 1.713 according to Journal Citation Reports, ranking 12$^{th}$ out of 22 journals (Q3) in the "Computer Science, Cybernetics" category.

Although the primary objective of the RemoTest platform is to assist researchers in designing, conducting and analysing data from web-based user studies, it is also necessary to provide accessible environments for the experiment participants (texts, questionnaires, navigation menus…). This issue is even more crucial when user tests are focused on people with disabilities and when they are carried out remotely without the supervision of an experimenter. In this article the RemoTest platform was evaluated based on the ease of the installation process, the accessibility of the automatically generated questionnaires, and user satisfaction. For this purpose we conducted in-situ evaluations with 36 users with different characteristics (physical impairments, blind, low vision and without disabilities). Study results showed the suitability of the RemoTest platform for conducting remote and in-situ user tests with people with different disabilities. Several enhancements, detailed in the article, were also identified to improve its accessibility.

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
2019, VOL. 35, NO. 7, 568–583
https://doi.org/10.1080/10447318.2018.1473941



# Inclusive Web Empirical Studies in Remote and In-Situ Settings: A User Evaluation of the RemoTest Platform

Myriam Arrue[a], Xabier Valencia[a], J. Eduardo Pérez[a], Lourdes Moreno[b], and Julio Abascal[a]

[a]EGOKITUZ, University of the Basque Country (UPV/EHU), Donostia (Gipuzkoa), Spain; [b]Computer Science Department, Universidad Carlos III, Madrid, Spain

**ABSTRACT**
Web accessibility evaluation requires tests to be carried out with real users with disabilities performing real tasks or activities. To recruit an appropriate group of users and to observe their performance in the real world is difficult. For this reason we have developed RemoTest, a platform that assists researchers designing experiments, conducting remote and in-situ experimental sessions and analyzing the data gathered while the users are accessing the Web. Although this tool is oriented to experimenters, it is necessary to check whether the evaluation environments created by RemoTest are accessible or not to the users that participate in the tests. To this end, we conducted formal in-situ evaluations with 36 users with diverse characteristics. For this assessment, the participants were asked to install the platform, to fill in some automatically created questionnaires and to carry out several web navigation tasks. From the data gathered we analyzed the ease of the installation process, the accessibility of the automatically generated questionnaires, and user satisfaction. The results revealed the suitability of the platform for conducting inclusive experiments both in remote and in-situ contexts and provided guidelines on how the experiments should be set out.

## 1. Introduction

The involvement of users in the evaluation of web services is fundamental in order to achieve universal access to the information society. Other methods exist to assess the usability and accessibility of web sites, such as the use of automatic checkers of sets of accessibility guidelines or standards or the hiring of experts to perform manual evaluations (Petrie & Bevan, 2009). However, evaluation by real users is the most valuable technique because it enables the detection of real problems and barriers that users experience while using the web pages. The expertise of each user, the configuration of the system, the assistive technology utilized by the user, are just a few of the variables that can determine whether the user manages to overcome a potential barrier or not.

User evaluation requires conducting experimental sessions with large and diverse groups of users. Researchers need to clearly define the tasks to be performed as well as the specific questionnaires which are required in order to explicitly obtain certain data from the participants such as satisfaction level, socio-demographic data and emotional aspect. Therefore, the experiment design process is demanding and requires experience from different areas: human factors, hypertext, web technology, etc.

In addition, involving an appropriate number of participants for a specific experiment is also challenging. Frequently, this is due to the location and the rigorous timing of sessions. Nowadays, interest in the use of software tools to conduct experiments remotely is increasing because they allow participants to be observed while they perform the tasks in their habitual daily environment. To work with their own resources and devices (which are already adapted to their needs) is particularly important when working with people with disabilities, as it facilitates the conduction of experiments “on the wild.” Moreover, this type of experiment gathers real interaction data without any obtrusive observation mechanism (Apaolaza et al., 2013). It also makes it possible to involve a larger number of participants, as they do not have to physically get to a specific location.

However, remote usability testing also has drawbacks. For example, it may not provide a thorough understanding of the users and their behavior and it is also necessary for all participants to have access to a reliable Internet connection (Albert, Tullis, & Tedesco, 2009). In order to carry out remote experimental sessions the remote tool should meet accessibility requirements to ensure that the tool can be used by a wider range of users. Finally, the set-up, installation and configuration processes must be accessible and user-friendly.

This article presents an evaluation of the accessibility and suitability of the platform RemoTest (Valencia, Pérez, Muñoz, Arrue, & Abascal, 2015) to carry out sessions with people with disabilities. The RemoTest platform objective is to assist researchers to design experiments, conduct

**CONTACT** Lourdes Moreno lmoreno@inf.uc3m.es Computer Science Department, Universidad Carlos III de Madrid, Avda. de la Universidad Nº 30, 28911 Leganés, Spain.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hihc.

experimental sessions and analyze data gathered in the evaluation sessions.

In order to verify whether the evaluation environments created by RemoTest are accessible or not to the users that participate in the test, the accessibility and usability of the test environment created for participants was evaluated by 36 users with different characteristics: 13 people with physical disabilities, 10 blind people, 8 people with low vision and 5 able-bodied people.

An in-situ experimental session was conducted and participants were asked to install the testing tool which is a Firefox add-on and perform different types of tasks such as filling in questionnaires automatically generated by the tool and web navigation tasks. Results revealed the suitability of the platform for conducting inclusive experiments both in remote and in-situ contexts.

## 2. Systems for web testing

Several remote web usability-testing tools have been developed in the last decade. They can be classified as server-side, proxy-based or client-side tools depending upon their architecture. Server-side tools are the most transparent for users, since no installation or configuration is needed (Etgen & Cantor, 1999; Google analytics, 2018; Leiva & Vivó, 2013; Optimizely, 2018; Paganelli & Paternò, 2002; Santana & Baranauskas, 2010; Scholtz, Laskowski, & Downey, 1998). Even though only HTTP requests can be gathered by the tool developed by Scholtz et al. (1998), adding some additional code, usually some JavaScript, to the web pages enables significant user interaction data to be gathered (Claypool, Le, Wased, & Brown, 2001; Etgen & Cantor, 1999; Leiva & Vivó, 2013, 2013; Paganelli & Paternò, 2002). This approach can be considered only when the web pages being evaluated are located in servers to which the researchers have access. On the other hand, proxy-based tools allow the evaluation of un-owned web pages but they require some configuration parameters to be fixed by the users (users have to configure their browser to access via the proxy) (Atterer, Wnuk, & Schmidt, 2006; Hong, Heer, Waterson, & Landay, 2001). Client-side tools (Claypool et al., 2001; Edmonds, 2003; Gajos, Reinecke, & Herrmann, 2012) are the most appropriate for usability testing since researchers can have access to any local interaction data generated in the experimental sessions (browser back, forward, bookmark, print options, mouse contextual menus, vertical/horizontal scrolling actions, etc.). Moreover, this type of architecture facilitates the inclusion of a questionnaire during the experimental session so that explicit data can be gathered from the participants. Other systems, such as USERZOOM (2018) can act as a server-side tool or client-side tool depending on the data to be gathered or the type of web site.

Table 1 presents information regarding the architecture of the most used remote usability-testing tools as well as the implicit interaction data gathered during the experimental sessions (mouse events, keyboard events, window events, browser actions and information in HTTP requests). In addition, other events collected by tools are also specified in Table 1. For instance, the NIST WebMetrics Suite (Scholtz et al., 1998) allows the injection of code to links in order to track the path followed by the user. Almost all the tools gather mouse, keyboard and window events, but there are more differences between them when it comes to browser actions and information in HTTP requests. Only UZILLA (Edmonds, 2003) and MORAE (2018) collect browser actions such as back/forward buttons. WEBQUILT (Hong et al., 2001) only obtains information from the HTTP requests. The last five tools indicated in Table 1 are commercial whereas the others were developed in an academic environment.

The system developed by Gajos et al. (2012) is devoted to gathering interaction data for data mining purposes. The USAPROXY tool (Atterer et al., 2006) injects tracking code automatically via proxy. Neither tool includes features for analyzing or visualizing the gathered interaction data. The rest of the tools have some functionality in order to facilitate the visualization and analysis of collected interaction data in remote experimental sessions.

Regarding the different types of experiments that can be executed by the analyzed tools, two main kinds of tasks are found: target searching and free navigation tasks. Target searching tasks require some features to be included in the testing tool: defining the target of the tasks, determining their duration, giving instructions to the participants and informing them when the target has been reached or when they are out of time. UZILLA, USERZOOM, Loop11 (LOOP11, 2018) and Morae are tools that include all these features. The other tools are devoted to conducting experiments based on free navigation tasks.

Table 1. Web usability-testing tools classification.

| | | Implicit interaction data | | | | | |
|---|---|---|---|---|---|---|---|
| Tool name | Architecture | Mouse | Keyboard | Window | Browser | HTTP requests | Other |
| NIST WebMetrics Suite (Scholtz et al., 1998) | Server-side | No | No | No | No | No | Path |
| WET (Etgen & Cantor, 1999) | Server-side | Yes | Yes | Yes | No | No | No |
| SMT2 (Leiva & Vivó, 2013) | Server-side | Yes | Yes | Yes | No | No | No |
| WELFIT | Server-side | Yes | Yes | Yes | No | No | Customized events |
| WebRemUSINE (Paganelli & Paternò, 2002) | Server-side | Yes | Yes | Yes | No | No | No |
| Curious Browser (Claypool et al., 2001) | Client-side | Yes | Yes | No | No | No | No |
| USAPROXY (Atterer et al., 2006) | Proxy-based | Yes | Yes | Yes | No | No | No |
| WEBQUILT (Hong et al., 2001) | Proxy-based | No | No | No | No | Yes | No |
| UZILLA (Edmonds, 2003) | Client-side | Yes | Yes | Yes | Yes | No | No |
| Gajos et al. (2012) | Client-side | Yes | Yes | Yes | No | No | No |
| Optimizely (2018) | Server-side | Yes | Yes | Yes | No | No | Customized events |
| MORAE (2018) | Client-side | Yes | Yes | Yes | Yes | No | Customized events |
| Google analytics (2018) | Server-side | Yes | Yes | Yes | No | No | Customized events |
| LOOP11 (2018) | Proxy-based | Yes | Yes | Yes | No | No | No |
| USERZOOM (2018) | Client/Proxy | Yes | Yes | Yes | No | No | No |

When it comes to collecting explicit data from participants by means of questionnaires (for measuring satisfaction, emotions, etc.), Morae, USERZOOM, UZILLA and LOOP11 include features for presenting and getting information through questionnaires before and after completing the tasks. Curious Browser (Claypool et al., 2001) presents questionnaires after every visited new page in order to study the relation between the events gathered during the session with users' interest in the page being evaluated.

Many of those systems, such as Loop11, Morae, Usaproxy, WELFIT (Santana & Baranauskas, 2010) or the one presented by Gajos et al. have been or are being used by people with disabilities. But only Morae and Loop11 can be used to perform guided user testing. The other tools are more focused on free navigation tasks. No accessibility evaluations could be found about the use or installation of Morae or Loop11. Morae is a powerful tool with which to perform user behavior studies but it is quite difficult to use or to be installed by people with disabilities. On the contrary, Loop11 can be easily used due to its proxy-based architecture. It does not require any installation as user testing starts when accessing a predetermined URL. One drawback of this tool is that it does not gather any information about the browser events occurring during the experimental sessions since the system acts as a proxy within the user and the evaluated web pages.

The RemoTest platform (Valencia et al., 2015) is a web testing tool which gathers most of the implicit interaction data presented in Table 1: mouse, keyboard, window and browser events and HTTP Requests. In addition, it includes features for defining different types of experiments and tasks as well as questionnaires for gathering explicit data from participants. It was developed as an inclusive testing platform which takes accessibility into account throughout the process. The following section describes the general architecture of the platform.

## 3. Remotest, platform for inclusive web experiments

The RemoTest platform provides evaluators with functionalities to facilitate the definition of experiments, manage experimental remote/in-situ sessions, describe questionnaires/surveys to be displayed to participants and to gather interaction data produced during the sessions and analyze this interaction data. This platform admits a wide range of experiments with a variety of objectives, for instance, to study user behavior when performing a task on different websites, to analyze and compare the navigational strategies of different types of participants when interacting with the same website, to evaluate the accessibility-in-use of several websites, to gather significant information through surveys, to measure user satisfaction when using certain web services, to analyze user performance improvement when interacting with adapted versions of original web pages and so on.

The architecture of the platform has been designed taking all these different types of experiments into consideration. In this case, we opted for a hybrid architecture model that includes some functionalities from a client-side module and other ones from some server-side modules. The platform is split into four modules: experimenter module (EXm), participant module (PAm), coordinator module (COm) and results viewer module (RVm). Figure 1 shows the architecture and interactions between these modules.

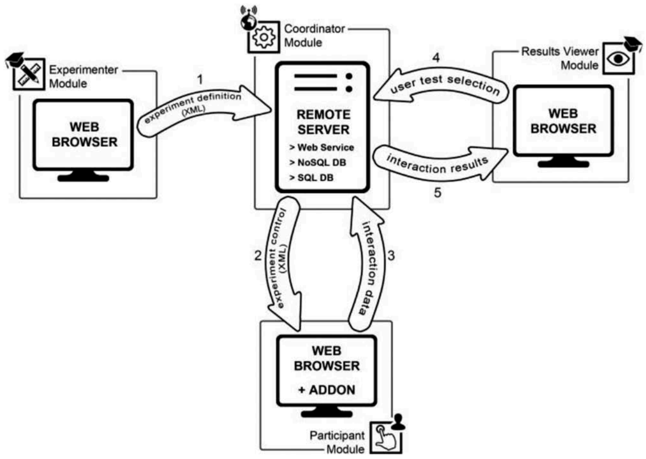


Figure 1. The RemoTest platform general architecture and interactions between modules.

Each module has specific functions and uses different technologies. The EXm is responsible for assisting researchers during the experiment definition process. The experiment definition is stored in an XML file based on specific vocabulary created for specifying experiments. This vocabulary is comprehensive enough to define the tasks, objectives, stimuli to be presented, task time limits, questionnaires to be filled in by participants and so on. The COm exploits the information in this XML file (Step 1 in Figure 1) in order to create personalized experimental sessions for each participant. These personalized sessions are transferred to the corresponding PAm (Step 2 in Figure 1). The PAm guides participants during the experimental sessions, presents the stimuli to participants and gathers the interaction data created during the experimental sessions. The interface of the PAm has been designed with accessibility aspects taken into account so the initial login screen, task description screens and the presented questionnaires conform to WCAG 2.0 accessibility guidelines (W3C, 2008). This module is developed as an add-on for Firefox and has to be locally installed in the participants' computer. The interaction data are centrally stored in a remote server for future analysis (Step 3 in Figure 1). The RVm organizes and presents the abundant interaction data gathered in the experiments (Step 4 and 5 in Figure 1).

## 4. Evaluation methods for assessing web-based tools

The evaluation of web-based tools entails significant challenges, as different aspects have to be considered. This work focuses on the evaluation of the user-testing tool installation process (PAm of the RemoTest platform), accessibility and usability of the interfaces automatically created and displayed by the tool and users' overall satisfaction and acceptance. This system has been developed to be used by people with different skills and ways of access, including people with and without disabilities as well as users employing different system configurations.

Several methods have been considered for carrying out the evaluation of the RemoTest platform: rating pragmatic quality (PQ) attributes of the installation process and emotional aspects during the installation, user testing and expert-based evaluations for detecting accessibility barriers, observational methods and inquiry methods for assessing the overall satisfaction and acceptance of the tool.

### 4.1. User experience (UX) evaluation

UX can be defined as

> the entire set of affects that is elicited by the interaction between a user and a product, including the degree to which all his or her senses are gratified (aesthetic experience), the meanings we attach to the product (experience of meanings) and the feelings and emotions that are induced (emotional experience). Law, Roto, Hassenzahl, Vermeeren, & Kort, 2009)

UX covers different aspects such as aesthetics, emotions, usability/pragmatic attributes, hedonic attributes, cognitive load, interactivity, social responses, persuasion and acceptability (Brajnik & Giachin, 2014). Each of these aspects is evaluated with different metrics or established methods. Pragmatic attributes and emotions were especially considered for evaluating the installation process of the RemoTest platform.

Pragmatic attributes are connected to the users' need to achieve behavioral goals (Hassenzahl, 2004). "A product may be perceived as pragmatic because it provides effective and efficient means to manipulate the environment" (Hassenzahl, 2005). Thus, PQ can be understood as perceived usability. Hassenzahl's Attrakdiff questionnaire[1] was selected as the method to gather the perceived PQ of the installation process. The set of seven word pairs reflecting opposite adjectives that can be rated on a 7-point scale to measure users' perceptions of the PQ attributes was translated to Spanish and introduced to the RemoTest platform so that a questionnaire was automatically generated and displayed to the participants after installing the tool.

Gathering information about emotions of participants during the installation process enables the appraisal of situations from an affective point of view (i.e., assigning arousal and valence value). In this work, emotional feedback was collected from participants in terms of the self-assessment manikin (SAM) scale to measure dimensions of valence (pleasantness of the emotion) and arousal (strength of the emotion) (Bradley & Lang, 1994).

### 4.2. Accessibility evaluation

Effective evaluation of websites for accessibility remains problematic. Automated evaluation tools still require significant manual testing and human judgment. Furthermore, the evaluation methodologies such as the one proposed by the web accessibility initiative presupposes that the evaluator has considerable knowledge about accessibility and assistive technology. There are other methods such as the Barrier Walkthrough Method[2] that can be performed more easily and is reliable and efficient in terms of the time required for carrying out the evaluation. This evaluation method in combination with conformance to WCAG 2.0 guidelines was applied during the RemoTest platform development and experiment preparation.

However, accessibility barriers that can only be discovered by user testing may be overlooked. In this work, user testing has been conducted in order to analyze the interaction of participants with the interfaces which are automatically created and displayed by the RemoTest platform. Participants were observed during the experiment and their interaction was video recorded so any accessibility barrier was detected during the analysis.

### 4.3. Interaction data gathering and inquiry methods

Participants' interaction data, such as the time required to complete the given navigational tasks or the time required for filling in the questionnaires, were automatically collected by the RemoTest platform and were analyzed in order to detect any barrier. In addition, participants were interviewed to record their accessibility/usability perceptions. Some questions were directly related to rating the ease of filling in the automatically generated questionnaires and to comment upon any difficulty presented by this process.

Inquiry methods were also used to measure participants' overall satisfaction and their acceptance of the RemoTest platform.

## 5. Experimental study

### 5.1. Research goals

The aim of the experimental study was to evaluate the RemoTest tool from the participants' perspective. In this case, an in-situ setting was chosen for the experimental sessions in order to obtain first-hand direct feedback from users and to be able to help them in any problem occurring during the interaction. The main objectives were evaluating the suitability and accessibility of the tool for carrying out experimental sessions with different groups of users. The study included the evaluation of the installation process as well as performing different tasks managed by the tool and obtaining participants' information through automatically generated questionnaires. Any accessibility barrier encountered during the process was immediately communicated to and annotated by experimenters. Participants' perceptions and opinions about the accessibility, usability and usefulness of the tool were gathered through semi-structured interviews.

The experimental study was designed to explore the following research questions:

- Q1: The RemoTest installation process is accessible and usable regardless of the participants' characteristics and the assistive technology used.

- Q2: The questionnaires automatically generated by RemoTest are accessible regardless of the participants' characteristics and the assistive technology used.
- Q3: Participants are satisfied with the tool performance, consider that RemoTest is easy to use and would use it in future experimental sessions even in remote settings.

### 5.2. Participants

The evaluation required participants of different groups of users. A call for participation was disseminated through several organizations of people with disabilities, social networks and email distribution lists. A total of 36 users were recruited. As required by the study all of them had some experience in using computers and Internet browsing.

Table 2 shows the description of the 36 participants who took part in the study. These participants have been grouped together into four user groups: physical disability (13 participants, 36.1%), blind (10 participants, 27.8%), low vision (8 participants, 22.2%) and participants without disabilities (5 participants, 13.9%). Figure 2 shows the frequency distribution bar chart of each user group.

The assistive technology used by participants is also included in Table 2. With regard to the participants with physical disabilities, four users employed joysticks, four adapted mouse, two head pointers, two users did not use any specific assistive technology but some specific configuration (such as switching right and left button functions in the mouse) and one interacted using the touchpad. Regarding the blind users, nine users employed a screen reader (one of them used it jointly with a braille display), all of them used JAWS screen reader. In the low vision user group, four users employed the ZoomText screen magnifier software, three users applied browser zoom functionalities and one user configured system settings to obtain high contrast interfaces.

Of the 36 participants in the study 17 were female and 19 were male. Mean age was 44.06 (SD = 9.9), see Figure 3 for the frequency distribution bar chart.

The Internet usage experience (1–3 years, 4–6 years, more than 7 years), the Internet expertise level (beginner, intermediate or advanced) and the Internet use frequency (daily, weekly, monthly) varied among participants. Generally, most of them claimed to have a usage experience of more than 7 years (80.56% of participants), have an intermediate expertise level (58.33% of participants) and to use the Internet daily (80.33% of participants). Figure 4 shows the frequency distribution bar chart of Internet usage experience, Internet expertise level and Internet use frequency. No significant variations are noticed in those values among user groups.

Participants were also asked about their experience with Mozilla Firefox browser as the RemoTest has been developed as an extension of this browser. The data obtained indicated that it is not the favorite browser among participants, eight participants (P2, P7, P9, P14, P23, P25, P26 and P30) said they have never used it. Only seven participants (19.4% of participants) stated that it is

**Table 2.** Description of the participants in the study.

| Id | Sex | Age | Disability \| nondisabled | Assistive technology (AT) | Expertise level | Setting |
|---|---|---|---|---|---|---|
| P1 | Male | 45 | Physical disability | Joystick | Intermediate | Elkartu |
| P2 | Female | 39 | Physical disability | Joystick | Intermediate | Elkartu |
| P3 | Male | 53 | Physical disability | Adapted mouse | Beginner | Elkartu |
| P4 | Male | 40 | Physical disability | Joystick | Advanced | Home |
| P5 | Female | 59 | Physical disability | Nothing | Beginner | Elkartu |
| P6 | Male | 50 | Physical disability | Nothing | Beginner | Elkartu |
| P7 | Female | 54 | Physical disability | Touchpad | Beginner | Elkartu |
| P8 | Female | 42 | Physical disability | Head pointer | Intermediate | Elkartu |
| P9 | Male | 41 | Physical disability | Head pointer | Advanced | Elkartu |
| P10 | Female | 76 | Physical disability | Joystick | Intermediate | Elkartu |
| P11 | Male | 50 | Blind | Screen reader | Intermediate | LabUC3M |
| P12 | Male | 44 | Low vision | Screen magnifier | Intermediate | LabUC3M |
| P13 | Female | 54 | Low vision | Browser zoom | Intermediate | LabUC3M |
| P14 | Female | 52 | Low vision | Screen magnifier | Intermediate | LabUC3M |
| P15 | Female | 44 | Low vision | High Contrast | Intermediate | Servimedia |
| P16 | Female | 39 | Low vision | Screen magnifier | Intermediate | Servimedia |
| P17 | Male | 47 | Blind | Screen reader | Intermediate | LabUC3M |
| P18 | Male | 32 | Blind | Screen reader | Intermediate | LabUC3M |
| P19 | Female | 54 | Low vision | Screen magnifier | Intermediate | LabUC3M |
| P20 | Female | 34 | Blind | Screen reader | Intermediate | LabUC3M |
| P21 | Male | 45 | Low vision | Screen magnifier | Intermediate | Servimedia |
| P22 | Male | 45 | Blind | Braille display screen reader | Advanced | Servimedia |
| P23 | Female | 41 | Low vision | Browser zoom | Intermediate | Servimedia |
| P24 | Male | 36 | Blind | Screen reader | Intermediate | Servimedia |
| P25 | Male | 23 | Blind | Screen reader | Intermediate | Servimedia |
| P26 | Female | 30 | Blind | Screen reader | Beginner | Servimedia |
| P27 | Female | 31 | Blind | Screen reader | Advanced | LabEHU |
| P28 | Male | 40 | Blind | Screen reader | Advanced | LabEHU |
| P29 | Male | 43 | Physical disability | Adapted mouse | Advanced | LabEHU |
| P30 | Female | 57 | Physical disability | Adapted mouse | Advanced | Home |
| P31 | Female | 52 | Physical disability | Adapted mouse | Intermediate | Home |
| P32 | Female | 32 | Nondisabled | Nothing | Advanced | LabEHU |
| P33 | Male | 41 | Nondisabled | Nothing | Intermediate | LabEHU |
| P34 | Male | 34 | Nondisabled | Nothing | Advanced | LabEHU |
| P35 | Male | 45 | Nondisabled | Nothing | Advanced | LabEHU |
| P36 | Male | 42 | Nondisabled | Nothing | Intermediate | LabEHU |

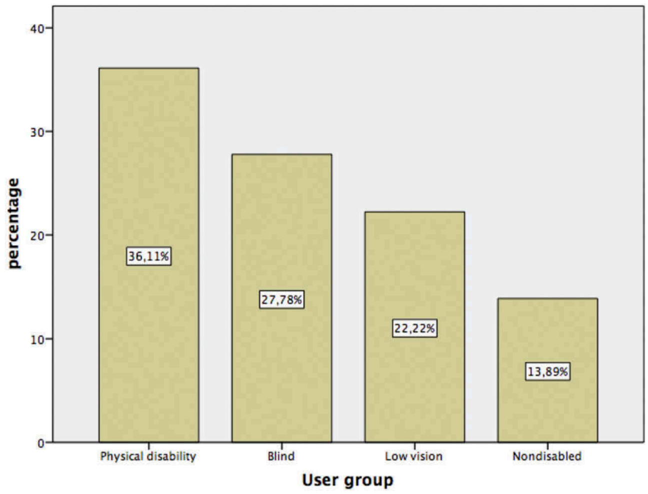


**Figure 2.** Frequency distribution bar chart of user group.

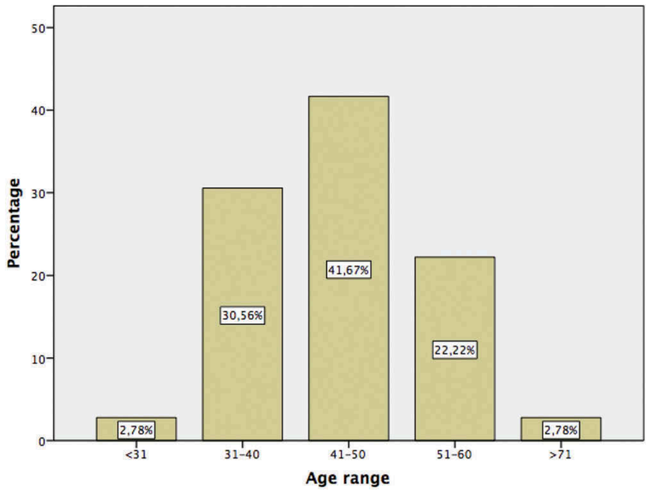


**Figure 3.** Frequency distribution bar chart of age.

their habitual browser and that they always use it for connecting to the Internet, 38.9% of participants used it sporadically and 19.4% of participants usually used Firefox. These frequencies do not have many variations segmented by user group.

### 5.3. Setting and equipment

The experimental sessions were carried out in different settings. There were four main different settings: the computer room at Elkartu (an association of people with physical disabilities), a laboratory located in the Carlos III University of Madrid (LabUC3M), a laboratory of the Computer Science School at the University of the Basque Country (LabEHU) and Servimedia (a news agency that employs people with visual disabilities). Nine experimental sessions were conducted on the Elkartu premises and eight sessions were conducted in each location: LabUC3M, LabEHU and Servimedia. In addition, the experimental sessions of three participants (P4, P30, P31) were conducted at the participant's home.

Participants were encouraged to use their own laptop and assistive technology whenever possible. The objective was to evaluate the RemoTest on different platforms and settings adapted to the participants. However, a desktop PC or laptop was configured for the sessions in Elkartu, LabUC3M and LabEHU. The desktop PC in Elkartu and the laptop in LabUC3M run Microsoft Windows 7 and Mozilla Firefox 25.0. The desktop PC in Elkartu was utilized by seven participants with their own assistive technology (joystick, head pointer, etc.), the other two participants (P3, P7) used their own laptop with the same configuration (Microsoft Windows 7 and Mozilla Firefox 25.0). The laptop in LabUC3M was utilized in 5 sessions in which a ZoomText magnifier and a JAWS 15 screen reader were also installed. The other three participants (P12, P13 and P17) used their own laptop with different configurations: Microsoft Windows XP and Mozilla Firefox 25 (P12), Windows Vista and Mozilla Firefox 25 (P13) and Microsoft Windows 7 and Mozilla Firefox 9.0.1 (P17). The desktop PC in LabEHU runs Microsoft Windows XP and Mozilla Firefox 22.0. All participants except for P27 and P28 used it. The laptops of these participants did not differ on the Mozilla Firefox version from the one installed in the PC but the laptop of P28 ran Microsoft Windows 7. Participant P29 used his own trackball to interact with the PC. Different configurations were found in Servimedia as the experimental sessions were conducted in the participant workplace. All computers run Microsoft Windows XP but differ in the version of Mozilla Firefox (we found 8.0.1, 19.0, 21.0, 23.0, 25.0 and 26.0 versions). Two of the participants conducting the session at home (P30, P31) had the same configuration: Microsoft Windows 7 and Mozilla Firefox 22.0. Finally, participant P4 used his own desktop PC running Microsoft Windows 7 and Mozilla Firefox 25.0.

### 5.4. Tasks and stimuli

Users were asked to perform different types of tasks. The first task (Task 1) was to install the RemoTest tool based on the instructions provided on a web page. Then, participants were asked to log in and fill in a set of questionnaires and complete some web navigation tasks. The tasks proposed to participants by the RemoTest were the following: filling in a questionnaire to provide some feedback about their perception about the tool installation process (Task 2), a free navigation task on a website (Task 3), a target searching task on a website (Task 4) and filling in a questionnaire to provide their socio-demographic data (Task 5).

Some of the stimuli presented by the RemoTest in the experimental sessions were manually designed web pages. For instance, a web page was developed for giving instructions for the installation process of the RemoTest and was displayed in Task 1. Other stimuli were automatically generated by the RemoTest such as the questionnaires and task description pages displayed in the rest of the tasks. Table 3 shows the description of the stimuli presented in the experimental sessions. It describes each stimuli by indicating the task to which it is related, the type of stimuli (informational web page, task description web page, questionnaire to be fulfilled, task completion indication), the description

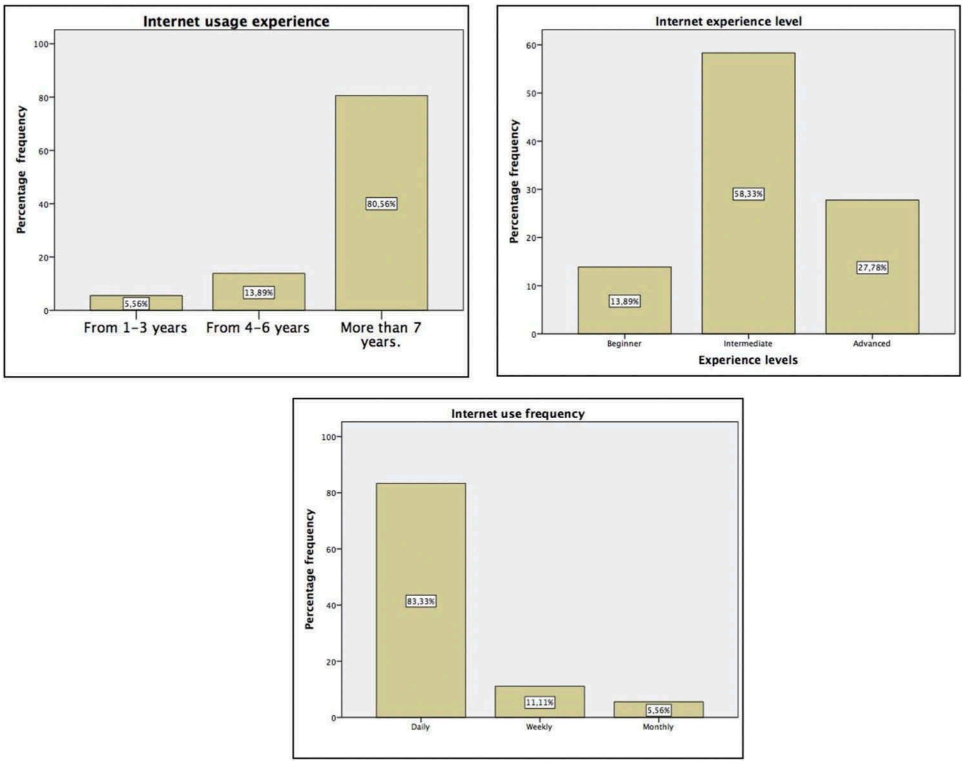


**Figure 4.** Frequency distribution bar chart of internet usage experience, internet expertise level and internet use frequency.

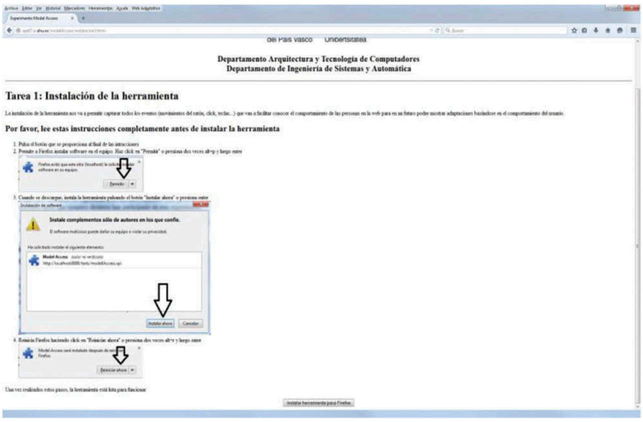

**Figure 5.** Manually developed web page containing the description of the RemoTest installation process (S1).

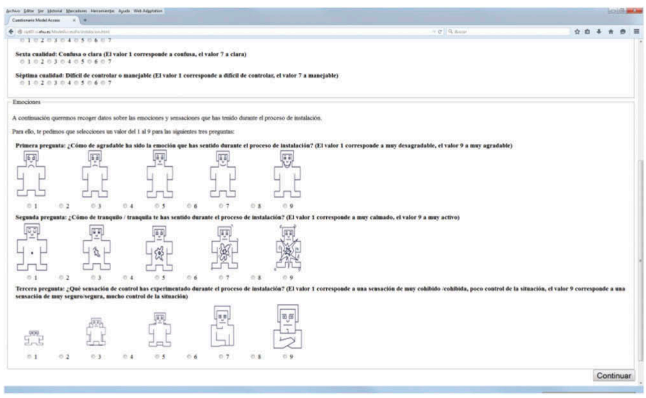

**Figure 6.** Automatically created questionnaire for gathering participants' perceived pragmatic aspects of the installation process and their emotional state during Task 1 (S2).

of the stimuli and whether it had been manually developed for the experiment or automatically generated by the RemoTest tool.

The Figures 5, 6, 7 and 8 show some screenshots of the stimuli displayed during the experimental session (S1, S2, S3 and S7). This set of screenshots includes all the different types of web pages displayed by RemoTest. The remaining stimuli (S4, S5, S6 and S8) are the ones created automatically by the tool for describing a task or for indicating the completion of the task and they all are similar to S3.

As stated above, S1 was manually generated with all the information needed by participants about the installation process. The different popup windows that would be presented by Mozilla Firefox were displayed and explained in order to install the add-on. Once they started the process (clicking on the button at the bottom of the web page) all the installation process was guided by Mozilla Firefox as in the installation of any other add-on.

The stimuli automatically generated by the RemoTest (S2–S8) were created first in XUL but an expert evaluation carried out by two blind people showed some critical accessibility issues (Valencia, Arrue, Rojas-Valduciel, & Moreno, 2014). Therefore, the stimuli generation process was updated to create the stimuli in HTML to avoid accessibility barriers. A Barrier Walkthrough Method was carried out by two of the authors in order to detect any accessibility barrier in these stimuli in HTML for blind people, people with physical

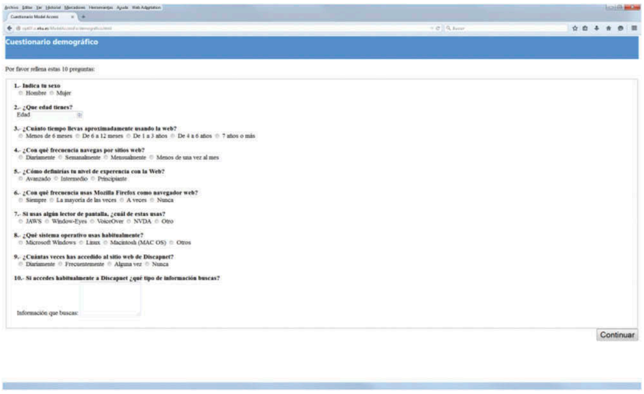


Figure 7. Automatically created questionnaire for gathering participants' socio-demographic data (S7).

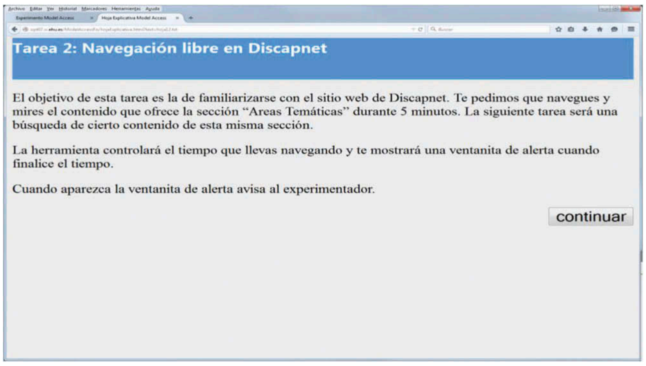


Figure 8. Automatically created web page with the description of the Task 3 (S3).

disabilities or low vision users. There were no significant accessibility barriers detected though there were some minor issues which will be improved upon in future versions of the tool such as including shortcuts for activating the button in the web pages (the button with the text "Continuar") and skipping links to directly access specific questions in the questionnaires (S2 and S7).

The free navigation and target searching tasks (Tasks 3 and 4, respectively) were carried out on the Discapnet website [www.discapnet.com]. This website focuses on providing information to people with disabilities. It officially conforms to the AA level defined in WCAG 1.0 accessibility guidelines.

## 5.5. Procedure

The sessions with participants were conducted one at a time. The whole test was conducted in the participants' mother tongue, Spanish. Each session started by providing information about the objectives of the study. Participants were told that their contribution to the scientific experiment was voluntarily, and that they could withdraw from the study at any point. All participants followed the same sequence of tasks in the experimental session. Then, all of them started installing the RemoTest tool and carried out the questionnaire completion tasks and navigation tasks. Finally, they were briefly interviewed. All the interactions with RemoTest platform were video recorded and the interviews were audio recorded.

## 5.6. Data Collection

The following methods were used for data collection:

- Interaction data: Every user interaction with the Discapnet website and the web pages automatically generated by RemoTest was monitored and stored in XML files. These files contain information such as the time at which each task was started, web pages visited, cursor movements and browser events.
- Video recordings: User interactions were recorded with a video camera. These recordings provided us with information about the users' interaction with the interfaces displayed by RemoTest.
- Observations: Interaction-specific aspects that drew the attention of the experimenters were noted (for instance, problems that occurred during the interaction or installation of the tool).
- Semi-structured interview: Two short post-interaction interviews were carried out and were audio recorded. Both interviews focused on getting information about users' satisfaction levels and opinions on the RemoTest tool, displayed interfaces, difficulties encountered when accomplishing tasks, etc. The objective was to gain direct feedback from participants.

Table 4 presents the data collected during each task in the experimental sessions.

This section is devoted to the analysis of the data gathered in the experimental session regarding the problems detected in the RemoTest tool installation process and in the automatically generated questionnaire completion tasks. Thus, data collected in Task 1, Task 2, Task 5 and the interviews are analyzed in this section. The analysis of the user interaction data automatically gathered by RemoTest (data collected in Task 3 and Task 4) is beyond the scope of this article and was carried out in other previously published research papers (Pérez, Arrue, Valencia, & Moreno, 2014; Valencia et al., 2015).

### 5.6.1. Remotest installation process

The installation process was evaluated based on the data collected in Task 1, Task 2 and the first short interview. Task 1 was completed by all of the participants even though some of them, by means of the responses given in the questionnaire of Task 2 and the comments in the short interview, reported several issues which could be improved upon and minor accessibility barriers they were faced with.

**Analysis of the data gathered through the questionnaire in Task 2:**

Task 2 consisted in filling in a questionnaire about the installation which was used to measure the perceived usability of the installation process and the emotions felt by participants during the installation. This questionnaire consisted of two parts: the first one was devoted to gathering users'

**Table 3.** Information about the stimuli presented by RemoTest during the experimental sessions.

| Id | Task | Type | Description | Manually/ Automatically |
|---|---|---|---|---|
| S1 | Task 1 | Task description | Web page containing the description of the RemoTest installation process. | Manually |
| S2 | Task 2 | Questionnaire | Participants were asked to fulfill a questionnaire about perceived pragmatic aspects of the installation process and their emotional state during the installation task (Task 1). | Automatically |
| S3 | Task 3 | Task description | Web page containing the description of Task 3 (free navigation on a website). | Automatically |
| S4 | Task 3 | Task completion | Web page indicating the completion of Task 3. | Automatically |
| S5 | Task 4 | Task description | Web page containing the description of Task 4 (searching a target link on a website). | Automatically |
| S6 | Task 4 | Task completion | Web page indicating the completion of Task 4. | Automatically |
| S7 | Task 5 | Questionnaire | Participants were asked to fulfill a questionnaire about their socio-demographic data. | Automatically |
| S8 | Task 5 | Task completion | Web page indicating that the provided data has been correctly stored and informing about the end of the experimental session. | Automatically |

**Table 4.** Information about the data collected in each task of the experimental session.

| Task | Description | Data collected |
|---|---|---|
| Task 1 | Installing the RemoTest tool based on the instructions provided on a web page (stimuli S1) | • -Annotations of problems occurred during the installation process based on direct observation and video recordings. |
| Task 2 | Filling in a questionnaire automatically generated by the RemoTest tool (stimuli S2) for gathering participants' perception about the installation task (Task 1). | • -Time for completing the task.<br>• -Perceived UX usability and pragmatic attributes based on the Hassenzahl's model.<br>• -Participants' emotions during installation task based on the SAM scale. |
| Short/brief Interview | Interview with questions about the previous tasks (Task 1, Task 2) and user satisfaction. | Semi-structured interview focused on gathering data about:<br>• -Accessibility barriers<br>• -User satisfaction |
| Task 3 | Free navigation on the Discapnet website for 5 minutes. | Interaction data collected by the RemoTest tool:<br>• -Visited web pages<br>• -Time in each web pages<br>• -Cursor movements, etc. |
| Task 4 | Target searching task on the Discapnet website for a maximum of 10 minutes. | • -Completing the task<br>• -Task completion time<br>• -Interaction data collected by the RemoTest tool |
| Task 5 | Filling in a questionnaire automatically generated by the RemoTest (stimuli S7) for collecting socio-demographic data. | • -Time for completing the task.<br>• -Participants' socio-demographic data. |
| Short/brief interview | Semi-structured interview with questions about the overall user satisfaction with the tool. | Semi-structured interview focused on gathering data about:<br>• -Accessibility barriers<br>• -User satisfaction |

perceptions about the PQ of the installation process and the second one for collecting participants' emotions during the installation process.

The first part of the questionnaire was based on the Attrakdiff questionnaire and consisted of a set of seven word pairs reflecting opposite adjectives to be rated on a 7-point Likert scale: technical-human, complicated-simple, impractical-practical, cumbersome-straightforward, unpredictable-predictable, confusing-clearly structured and unruly-manageable.

The second part of the questionnaire consisted of three questions to be rated on a 9-point scale based on the SAM: pleasure, arousal and dominance. The original SAM scale is a non-verbal pictorial assessment technique and so alternative texts (in Spanish) were added in order to make it accessible for the participants with visual disabilities.

The reliability of the used scales in this questionnaire was analyzed based on the Cronbach coefficient. For the PQ attributes all the Cronbach coefficients were greater than 0.70 indicating moderate-to-good reliability. On the contrary,

**Table 5.** Cronbach for the different scales that were used.

| | PQ attributes | Emotions (SAM) |
|---|---|---|
| N items | 7 | 3 |
| Cronbach $\alpha$ | 0.864 | 0.233 |

no reliable data could be gathered from the emotions scale as can be appreciated in Table 5. We found that some users did not understand the semantics of the images of the SAM scale or the alternative text provided. No further quantitative analysis was made with the values gathered with this scale. However, more information about the feelings during the installation was gathered in the interviews.

Each PQ attribute was analyzed separately and the results are presented in Table 6. It can be observed that the Corrected Item-Total Correlation for each item is favorable or positive as all the values are above 0.4 except one: the technical-human item (0.387). This result confirms that this item had less consistency compared with the other attributes. However, the general Cronbach's Alpha value increment is low when removing this item (from 0.864 to 0.883). These results suggest that the item has moderate reliability. During the interview, which was carried out after the installation task, some users commented that they had difficulties understanding the Technical-Human adjectives pair.

Participants tended to give high values to the PQ attributes. The median, mean confidence interval and standard deviation for each item are presented in Table 7. The mean of six of the attributes was higher than 5 (5.22–5.92). The technical-human question has the lowest value (4.25). This could be due to its moderate reliability, previously calculated with the Cronbach coefficient. The same tendency can be observed for the median and the standard deviation values.

The obtained values reveal a moderately good perception of the PQ of the installation process of the RemoTest tool.

*5.6.1.1. Insights from the short interview.* Once the participants installed the RemoTest tool and filled in the questionnaire a short interview was carried out to evaluate the process. They were asked to quantify the ease of the installation process based on a Likert 7-point scale. Only one participant said that the installation process was very complex. The 58.33% of the participants rated the installation process as simple or very simple (values 6 and 7). The mean value of the rates given by participants was 5.48. Figure 9 shows the values given by participants.

**Table 6.** Cronbach for the items of perceived usability/ PQ attributes.

| Pragmatic quality (PQ) | Corrected item-total correlation | Cronbach's alpha if item deleted |
|---|---|---|
| Technical – human | .387 | .883 |
| Complicated – simple | .649 | .844 |
| Impractical – practical | .725 | .831 |
| Cumbersome – direct | .681 | .838 |
| Unpredictable – predictable | .774 | .827 |
| Confusing – clear | .641 | .843 |
| Unruly – manageable | .652 | .842 |

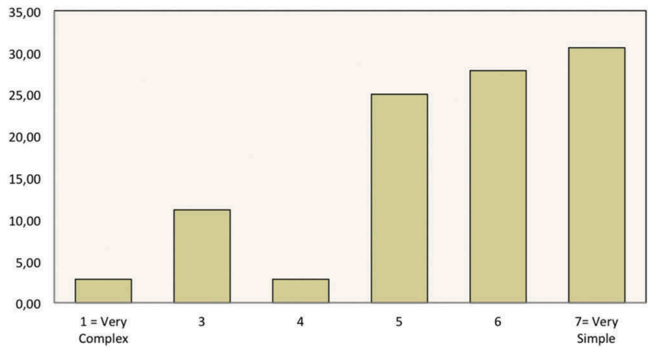


**Figure 9.** Graph showing the frequency of the answers to the question regarding the ease of the installation process.

In order to detect any cognitive barriers with the provided installation process instructions, participants were asked if they were easy to follow and understand. The results were positive since 24 of 36 of the users told us that the instructions were simple. Nevertheless, they did make some suggestions for improvement. The same suggestions from participants within the same user group were not obtained in all cases:

- Some suggestions were about the aesthetics and were not regarding the comprehension of the content. Two participants (P4 and P2) with physical disability commented on the text style. P4 stated that the text was too close together and could be quite confusing and P2 said that the font size was small. Two participants with low vision (P12 and P24) said that the font used was not very accessible due to the use of a font with serif. P12 suggested not using the whole screen for the text since going across from left to right tires people with low vision. Another participant also asked for more colorful instructions.
- One suggestion made by some participants was related to the difficulty they had to read all the instructions at once, since this required them to remember all the steps needed to install the tool. The blind participant P11 and the participant with low vision P12 asked for a step-by-step installation. P18, a blind participant, instead, asked for a shorter installation with fewer steps and simpler instructions. Participant P10, a participant with physical disability, had difficulties to understand the instructions and recommended using clearer and simpler language.

Participants were asked about accessibility barriers they found in the installation process and the responses given by those participants who rated the process as very complex or complex were thoroughly analyzed. Some of the barriers reported by participants are related to the Mozilla Firefox browser. In the case of accessibility barriers, coincidences in the answers of the participants corresponding to the same user group were obtained. These are the main barriers classified by user groups detected through the interview:

- Most of the screen reader users (P11, P18, P22, P25 and P26) did not notice the popup alert window opened by the browser agent in order to initiate the installation process.

**Table 7.** Mean, median and standard deviation for each of the PQ attributes.

| | Technical – human | Complicated – simple | Impractical – practical | Cumbersome – direct | Unpredictable–predictable | Confusing – clear | Unruly–manageable |
|---|---|---|---|---|---|---|---|
| Mean (μ) | 4.25 | 5.92 | 5.22 | 5.56 | 5.39 | 5.58 | 5.28 |
| Lower limit | 3.60 | 5.43 | 4.65 | 4.99 | 4.87 | 5.03 | 4.65 |
| Upper limit | 4.90 | 6.41 | 5.80 | 6.12 | 5.91 | 6.13 | 5.90 |
| Median | 4.00 | 7.00 | 6.00 | 6.00 | 6.00 | 6.00 | 6.00 |
| Standard deviation (σ) | 1.991 | 1.500 | 1.758 | 1.731 | 1.591 | 1.680 | 1.907 |

- Problems with popup alerts were also detected by low vision users using magnification software. The installation alert was positioned by Firefox in the upper left corner, which most of the time was out of their field of vision (P8 and P16). On the other hand, participants P12 and P14 saw the window but they did need more time to find it. Moreover, P12 had to change his strategy using the magnifier to decrease the zoom in order to access the alert window more easily. P23 said that the popup window size was too small and that she would prefer a bigger alert window in a centered position.
- One participant with physical disability (P29) proposed an improvement for stimuli S1 (installation process description). He suggested breaking the web page into smaller ones to avoid the use of scrolling.

Despite the accessibility barriers encountered by participants during the installation process, all of them were able to overcome them and install the tool without significant difficulties. The comments gathered in the interview are valuable for improving future versions of the RemoTest tool.

Chi-square and Fisher's exact tests were carried out in order to determine whether there were significant associations between variables of characteristics of users (user group, with or without disability, assistive technology used, age, Internet usage experience, Internet expertise level, Internet use frequency, Mozilla Firefox browser experience, etc.) and the opinions of the participants about the complexity-simplicity of installation process (Task 1). No statistical evidence of associations in the results was found.

### 5.6.2. Questionnaire completion tasks

Task 2 and Task 5 consisted of filling in some questionnaires. The questionnaires (stimuli S2 and S7) were automatically generated by the RemoTest tool based on the parameters specified by researchers when defining the experiment. The objective was to generate accessible questionnaires. All the participants were able to complete these tasks despite using different assistive technologies. This section explores the data gathered during these tasks such as the time required for filling in the questionnaires and any barrier detected by participants.

#### *5.6.2.1. Time required for filling in the questionnaires.*

Table 8 shows the time required by participants to fill in each of the questionnaires presented to users by Remotest during the experimental session.

Due to some technical problems the demographic questionnaire could not be presented to participant P32. These data were gathered through specific questions in the short interview.

**Table 8.** Time required to complete the questionnaires.

| Participant | Task 2 | Task 5 | Total |
|---|---|---|---|
| P1 | 582 | 321 | 903 |
| P2 | 638 | 269 | 907 |
| P3 | 381 | 257 | 638 |
| P4 | 230 | 208 | 438 |
| P5 | 342 | 142 | 484 |
| P6 | 280 | 92 | 372 |
| P7 | 797 | 371 | 1168 |
| P8 | 1784 | 543 | 2327 |
| P9 | 993 | 377 | 1370 |
| P10 | 756 | 485 | 1241 |
| P11 | 262 | 183 | 445 |
| P12 | 375 | 313 | 688 |
| P13 | 207 | 143 | 350 |
| P14 | 430 | 279 | 709 |
| P15 | 648 | 235 | 883 |
| P16 | 456 | 170 | 626 |
| P17 | 299 | 212 | 511 |
| P18 | 243 | 289 | 532 |
| P19 | 182 | 90 | 272 |
| P20 | 455 | 374 | 829 |
| P21 | 329 | 161 | 490 |
| P22 | 411 | 222 | 633 |
| P23 | 270 | 80 | 350 |
| P24 | 175 | 164 | 339 |
| P25 | 316 | 192 | 508 |
| P26 | 224 | 328 | 552 |
| P27 | 229 | 158 | 387 |
| P28 | 881 | 828 | 1709 |
| P29 | 328 | 176 | 504 |
| P30 | 184 | 263 | 447 |
| P31 | 645 | 314 | 959 |
| P32 | 143 | – | – |
| P33 | 275 | 76 | 351 |
| P34 | 148 | 39 | 187 |
| P35 | 103 | 46 | 149 |
| P36 | 272 | 58 | 330 |

Regarding the time required to fill in the questionnaires, P8, P28 and P9 needed appreciably more total time to complete both questionnaires: 2327, 1709, and 1370 seconds respectively. Participants P8 and P9 used a head pointer which takes considerably more time to point and click on the answers. In addition, video recordings of P8 showed that she needed longer to read the questions and response options than other participants. She also had some difficulties when answering the questions related to the PQ attributes in Task 2. Moreover, she clicked on the radio buttons in order to select her answers even though the clickable area was wider (in fact it encompassed all of the text of the answer as well as the radio button). Participant P28 was a blind user. He experienced some problems with the JAWS screen reader. He did not get any advice on the option selected in a question and sometimes the screen reader cursor returned to the beginning of the form and he had to navigate through all the questions he had already answered.

**Table 9.** Mean and standard deviation for the time required to complete the questionnaires by user group.

| User Group | Task 2 | | Task 5 | |
|---|---|---|---|---|
| | Mean ($\mu$) | Standard deviation ($\sigma$) | Mean ($\mu$) | Standard deviation ($\sigma$) |
| Blind | 349.50 | 205.69 | 295.00 | 200.76 |
| Low vision | 362.13 | 151.84 | 183.88 | 84.79 |
| Physical disability | 610.77 | 430.36 | 293.69 | 129.29 |
| Nondisabled | 188.20 | 79.80 | 43.80 | 28.21 |

Video recordings of the participants using a screen reader were observed in order to detect any general barrier experienced by them. Results showed that participant P20 also experienced some inconvenience with the use of the screen reader, as it did not read out all the questions as apparently it skipped some of them. The participant became aware of this problem when an alert advising that all questions had to be filled in appeared when the "Continue" button was clicked. Even though these issues did not prevent participants from completing Task 2 and Task 5 they have to be analyzed and fixed for the next version of the tool.

Analyzing all of the questionnaires and all the participants it was seen that users required a mean time of 334 seconds to fill in a questionnaire. Table 9 presents the mean and standard deviation for the time required for each questionnaire by the different user groups.

As can be appreciated in the table, the time needed to complete the Task 2 questionnaire is higher than the Task 5 questionnaire for all the user groups. The user group of participants with physical disabilities obtained the slower mean value of 610.77 seconds for filling in the questionnaire in Task 2, but the standard deviation is also a very high value (430.36). Thus, the high time required by participants with physical disability is not data which gives us a central trend. The data scatter is due to certain specific participants such as P8 and P9 who obtained very high values due to using a head pointer and some problems they had to understand the PQ attributes.

The difference of the time distribution between user groups was able to be confirmed by running the Kruskal–Wallis one-way test. A rejection in the null hypothesis of independence (Task 2 ($\chi^2(3) = 8.6$, $p = .035$) and Task 5 ($\chi^2(3) = 13.9$, $p = .03$)) was obtained. The ranking average by user group was 7.5, 15.3, 17.8 and 23.5 for nondisabled people, people with blindness, with low vision and with motor impairments respectively in Task 2. While in Task 5 2.5, 14.3, 21.0, 22.8 ranking averages were obtained by nondisabled people, people with low vision, with blindness and with motor impairments respectively.

*5.6.2.2. Insights from the short interviews.* An interview about the participants' feelings about the questionnaire completion task was carried out just after filling in each questionnaire. This semi-structured interview consisted of some questions relating to the complexity of the completion process, previous experience with such kind of questionnaires and about barriers detected when filling them in. One participant (P15) was not interviewed as she performed the experimental session at her workplace and she was interrupted by a phone call. Nevertheless, she was able to complete both questionnaires without any noticeable problem. In total, 35 participants were interviewed about the questionnaire completion tasks.

We asked participants to rate the ease of filling in the questionnaire on a 7-point Likert scale. Figures 10 and 11 show the result of this question for the questionnaire in Task 2 and Task 5 respectively. The mean value for the questionnaire in Task 2 was 5.2 CI [4.61, 5.78]. It was rated as very simple (value 6–7) by 19 participants with adjusted-Wald 95% binomial confidence range (BC) of [38.18, 69.54%]. Only 6 from 35 participants gave a value lower than 4 BC [7.72, 33.06%]. Regarding the questionnaire in Task 5, the mean

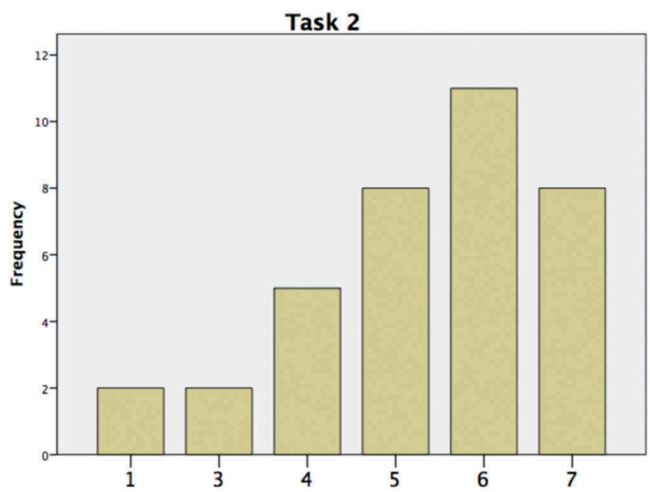


**Figure 10.** Ease score frequency for completing the questionnaire in Task 2.

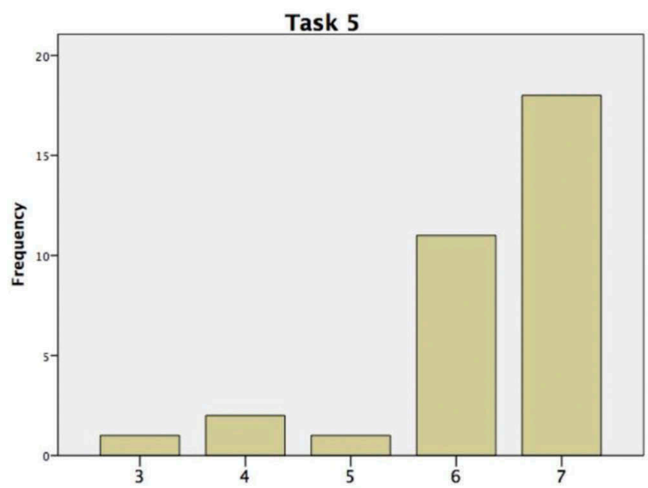


**Figure 11.** Ease score frequency for completing the questionnaire in Task 5.

value was 6 CI [5.61, 6.39]. Most of the participants, 29 out of 35, rated it as very easy (value 6–7) BC [66.94, 92.28%] . Only two participants rated it with a value lower than 4 BC [6.2, 19.57%].

All of them were able to access to all the content of the questionnaire in Task 2 but participant P14 with low vision had some problems with the figures due to her using a screen magnifier and losing the context of the images. 10 BC [16.19, 45.20%] participants had problems understanding some of the pairs of words of the PQ attributes, mainly the technical-human word pair. 6 BC [7.72, 33.06%] participants, on the other hand, stated that it was very simple and clear. It is worth mentioning that near the half of them (15 participants out of 35 BC [27.97, 59.16%]) had never filled in this kind of questionnaires before.

Regarding the questionnaire in Task 5, one participant commented on the screen reader cursor problem. Another participant reported some doubts about a question. In contrast to the previous questionnaire, most of the users had previously filled in similar questionnaires (29 participants from 35 BC [66.94, 92.28%]).

As in Task 1, Chi-square and Fisher's exact tests were carried out in order to determine whether there were significant associations between variables of characteristics of users and the opinions of the participants about the complexity-simplicity of Tasks 2 and 5. As a result, statistical evidence of associations were not found for either.

#### 5.6.3. User satisfaction

The final interview of the experimental session was also intended to obtain participants' satisfaction and acceptance of the RemoTest tool. It was a semi-structured interview and participants were asked for their opinion about the tool and whether they would use it in remote settings. Participants were asked three questions to in order to obtain their opinion:

- Question 1: "What do you think about having a system for conducting inclusive experiments remotely?
- Question 2: "Would you participate in remote experiments? Would you encourage friends to participate?"
- Question 3: "Would you feel more comfortable doing the experiment remotely (e.g. from home, office, etc.)?"

All participants except P35 thought it would be interesting to have such a system to conduct experiments remotely. P35 expressed concerns about security. He did not intend to install the add-on on his computer as he thought such a system could get personal data from the system. However, he would try if the experimenters were people whom he trusted.

Most participants responded that they would participate in remote experiments and would encourage friends to do so:

- P28: "Sure, I have done some tests in remote before and everything was OK"
- P10: "It would be very interesting to carry out experiments from home and I would participate"
- P11: "I think it is amazing to have such a system to conduct experiments remotely. There is a lot of work to do and people often have a great sense of helplessness"

However, some participants indicated they would participate only if it did not take a long time:

- P21: "Yes, I would participate if I had enough time. I think that this is to help others and it is easier if you can do it from anywhere".
- P22: "Yes I would, depending on the time it would take"
- P5: "Yes I would if it is not difficult and it is not everyday"
- P6: "I don't know if I would be available to do tests at home"

Finally, other participants revealed some concerns about doing experiments in remote settings:

- P9: "I think that I could have some problems when installing and it would be necessary to provide good instructions"
- P34: "Yes I would participate if it is not too difficult and the experiments are helpful"
- P31: "Yes I would participate. If I had any problem I would email you"

Regarding Question 3, 16 out of 35 participants replied that they would be more comfortable performing the experiment at home or in the office:

- P2: "Yes I would be more comfortable without cameras and a tape recorder and I think I would be more efficient completing the tasks at home"
- P13: "I would prefer to do it at home because I have a huge screen and I see everything much better there"
- P14: "I would be more comfortable at home because I have all my tools there"
- P23: "The best place for me is the office because I have everything adapted"

Three participants (P11, P18 and P28) would be more comfortable in remote settings as long as they were provided with a chat system to resolve any problem encountered during the experimental sessions.

Another three participants (P7, P10 and P30) would prefer to carry out experiments in local settings. They claimed that they were more comfortable when the experimenters were on hand so they could ask about any doubt concerning the tasks.

The remaining participants responded that they would be comfortable in both a local and a remote setting.

### 5.7. Discussion

The results gathered served to explore the research questions defined for the experimental study. In relation to research question Q1, the results obtained revealed that the installation process of the RemoTest tool proved to be successful irrespective of the disability and the assistive technology used by participants. All participants were able to install the tool even if some screen reader users and low vision users had to cope with problems when managing the popup windows which appeared during the installation process by the Mozilla Firefox browser. It is worth mentioning that Mozilla Firefox was not their usual web browser. This issue ought to be thoroughly tested for future experiments so that the screen reader used by each participant is proven to be compatible with the Mozilla Firefox version used.

The usability of the installation process was tested according to the PQ attributes. All participants gave positive scores to the seven pairs of adjectives displayed in the questionnaire of Task 2. Their feelings about the installation process were also positive. 83.33% BC [66.73, 92.51%] the participants rated the ease of the installation process with a value greater than 4 in a 7-point Likert scale. Statistical associations between variables of characteristics of users and their perception of the complexity of Task 1 were not found. This supports the assumption that the tool is accessible, simple and usable irrespective of the participants' characteristics and the assistive technology used.

However, some potential improvements on the presentation of the instructions for the installation were discerned during the interviews. There were some similar comments from participants within the same user group such as those from two participants with low vision (P12 and P24) who questioned the choice of text font. We also found opposing comments from participants from the same user group such as two blind participants (P11 and P18) one of whom requested a step-by-step installation process whereas the other asked for a shorter installation with fewer steps and simpler instructions. There were also some coincidences between participants within different user groups. These comments showed us the importance of including some mechanism of personalization so the installation and the presented stimuli can be adapted according to participants' preferences.

As regards research question Q2, the stimuli automatically generated by RemoTest tool proved to be accessible to all participants. All of them were able to fill in the questionnaires displayed by the tool and they rated them positively. According to the time required for completing the questionnaires, results revealed variations between user groups and among users within the same group. Participants in the group of physical disability required more time on average than others to fill in questionnaires but they showed high deviations in their results. These results led us to consider defining parameters in the configuration of tasks in future versions of the tool so that enough time is provided to each participant independent of their user group (this is related to Guideline 2.2 of WCAG 2.0). There were some participants using screen readers who reported some minor problems. They were able to fill in all the questions but the compatibility of the questionnaires with different versions of screen readers should be thoroughly analyzed and improved.

Participants gave a mean value of 5.2 out of 7 to the ease of filling in the questionnaire of Task 2 and a mean value of 6 to the questionnaire in Task 5. Moreover, there was no significant evidence found to contradict that these automatically generated questionnaires were accessible and simple to fill in regardless of the participants' characteristics.

The main concerns commented on by participants in the interviews were related to the semantics of the questions. This highlights the importance of using a clear and easy language, even more so when the studies are specifically focused on people with disabilities. In addition, some aspects of the design could be improved according to comments gathered in the interviews and these will be considered in future versions of the tool.

The lack of reliability of the SAM questionnaire about the feelings of participants during the installation might be due to the fact that some participants had problems understanding the semantic of the pictures or the alternative texts provided. The SAM questionnaire was selected because of its simplicity, however due to the results obtained the suitability of alternatives should be explored, such as, for example, the hedonic quality measure of Hassenzahl (2001).

Regarding research question Q3, the last interview revealed interesting data about user satisfaction with and acceptance of the RemoTest tool. All but one participant found this kind of tool interesting for conducting experimental sessions in local or remote settings. However, some of them showed some reluctance to perform remote experiments. Their main concerns were the length of time experiments would take and the problem solving mechanism during the sessions. They would appreciate some kind of support such as chat systems to ask for help if they are locked in any step or need some clarification about tasks. Nevertheless, most of them showed a positive attitude toward participating in other experiments even if they were to be remotely performed.

## 6. Conclusions

The need for remote web usability-testing tools in order to test web services in real contexts is growing. Several tools have been developed in the last decade and the most used ones are discussed in this paper. One of the common drawbacks of such tools is their lack of inclusiveness for people with disabilities. They have not been formally evaluated with users from different user groups.

We presented the RemoTest platform for designing, conducting and analyzing the data gathered in experimental sessions. It has been evaluated from the perspective of the experiment participants in a formal empirical study including 36 participants with different characteristics. Results revealed that all the participants, irrespective of their characteristics and the assistive technology used, were able to install the tool when provided with specific instructions. However, it was seen that the installation process could be improved by applying some of the suggestions made by participants. Future work will be focused on improving this process so it will be more personalized to tailor for the specific characteristics and the assistive technology used by participants.

Regarding the stimuli automatically generated by the RemoTest platform, results from the empirical study showed that they were accessible for a wide range of users. All participants were able to complete the questionnaires presented by the tool in a reasonable amount of time and most of the problems or barriers detected by participants and pointed out in the interviews were more related to the complexity of the questions rather than difficulties encountered operating with the form controls. However, there were some compatibility issues between the assistive technology employed by users and the web browser version that should be considered in future versions of the platform. Moreover, based on feedback provided by participants during the interviews, some design aspects could also be improved such as providing shortcuts, bigger text and controls, numbering the questions, use of clear and simple language, etc. Future versions of the platform will incorporate these suggestions.

All in all, participants expressed their satisfaction with the platform and in general they were confident enough to take part in remote experimental sessions. One aspect to be taken into consideration in designing future versions of the platform is the possibility of dividing the experimental session into shorter sessions to encourage more participants to take part. However, the implications of such a feature at experiment design time would need to be thoroughly analyzed. Another potential feature suggested by participants was a mechanism, such as a chat system, for problem solving during experiment time. This will be also considered for future work.

Summarizing, in order to carry out an inclusive remote usability study with RemoTest or other remote usability tools, the following recommendations should be followed.

- Provide clear instructions about the installation or configuration of the user-testing tool including explanatory images when required
- Task descriptions, questionnaires, alarms etc. should be set up based on standards and be accessible in order to ensure their compatibility with the assistive technologies
- Texts must be short and clear and technical language should not be used
- Long experimental sessions should be divided into shorter sessions to avoid tiring the user and to encourage a greater number of participants to get involved.
- Due to the diversity of users, personalization features should be included, for instance allowing the preferred contrast, color, text size etc. to be set or allowing images, shortcuts etc. to be removed.

## Notes

1. http://attrakdiff.de/index-en.html.
2. https://users.dimi.uniud.it/~giorgio.brajnik/projects/bw/bw.html.

## Funding

This research work was developed within the project eGovernability, funded by the Spanish Government, Ministry of Economy, Industry and Competitiveness, and the European Regional Development Fund, under grant (TIN2014-52665-C2-1-R MINECO/FEDER). J.E.P. holds a PhD Scholarship from the University of the Basque Country (UPV/EHU). Some of the authors are members of the EGOKITUZ/ADIAN research team, supported by the Basque Government, Department of Education, Universities and Research under grant (IT980-16).

## References

Albert, W., Tullis, T., & Tedesco, D. (2009). *Beyond the usability lab: Conducting large-scale online user experience studies*. Morgan Kaufmann, 1–314. ISBN: 9780080953854,

Apaolaza, A., Harper, S., & Jay, C. (2013). Understanding users in the wild. In *Proceedings of the 10th international cross-disciplinary conference on web accessibility*, 1–13.

Atterer, R., Wnuk, M., & Schmidt, A. (2006). Knowing the user's every move: User activity tracking for website usability evaluation and implicit interaction. In *Proceedings of the 15th international conference on World Wide Web*, 203–212.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49–59. doi:10.1016/0005-7916(94)90063-9

Brajnik, G., & Giachin, C. (2014). Using sketches and storyboards to assess impact of age difference in user experience. *International Journal of Human-Computer Studies*, *72*(6), 552–566. doi:10.1016/j.ijhcs.2013.12.005

Claypool, M., Le, P., Wased, M., & Brown, D. (2001). Implicit interest indicators. In *Proceedings of the 6th international conference on intelligent user interfaces*, 33–40.

Edmonds, A. (2003). Uzilla: A new tool for web usability testing. *Behavior Research Methods, Instruments, & Computers*, *35*(2), 194–201. doi:10.3758/BF03202542

Etgen, M., & Cantor, J. (1999). What does getting WET (web event-logging tool) mean for web usability. In *Proceedings of fifth human factors and the web conference*.

Gajos, K. Z., Reinecke, K., & Herrmann, C. (2012). Accurate measurements of pointing performance from in situ observations. *Proceedings of the SIGCHI conference on human factors in computing systems*, 3157–3166.

Google analytics. 2018. Retrieved from http://www.google.com/analytics/

Hassenzahl, M. (2001). The effect of perceived hedonic quality on product appealingness. *International Journal of Human-Computer Interaction*, *13*(4), 481–499. doi:10.1207/S15327590IJHC1304_07

Hassenzahl, M. (2004). The interplay of beauty, goodness, and usability in interactive products. *Human-Computer Interaction*, *19*(4), 319–349. doi:10.1207/s15327051hci1904_2

Hassenzahl, M. (2005). The thing and I: Understanding the relationship between user and product. In Blythe, M. A., Overbeeke, K., Monk, A. F., & Wright, P. C., (Eds.), *Book funology: From usability to enjoyment* (pp. 31–42). ISSN 1571-5035, ISBN 978-1-4020-2966-0.

Hong, J. I., Heer, J., Waterson, S., & Landay, J. A. (2001). WebQuilt: A proxy-based approach to remote web usability testing. *ACM Transactions on Information Systems*, *19*(3), 263–285. doi:10.1145/502115.502118

Law, E. L. C., Roto, V., Hassenzahl, M., Vermeeren, A. P., & Kort, J. (2009). Understanding, scoping and defining user experience: A survey approach. In *Proceedings of the SIGCHI conference on human factors in computing systems*, 719–728.

Leiva, L. A., & Vivó, R. (2013). Web browsing behavior analysis and interactive hypervideo. *ACM Transactions on the Web (TWEB)*, *7*(4), 1–28. doi:10.1145/2540635

LOOP11. (2018). Retrieved from http://www.loop11.com/

MORAE. (2018). Retrieved from http://www.techsmith.com/morae.html

Optimizely. (2018). Retrieved from https://www.optimizely.com/

Paganelli, L., & Paternò, F. (2002). Intelligent analysis of user interactions with web applications. *Proceedings of the 7th international conference on intelligent user interfaces*, 111–118.

Pérez, J. E., Arrue, M., Valencia, X., & Moreno, L. (2014). Exploratory study of web navigation strategies for users with physical disabilities. In *Proceedings of the 11th web for all conference*, 1–4.

Petrie, H., & Bevan, N. (2009). The evaluation of accessibility, usability and user experience. In Stephanidis, C. (Ed.), *The universal access handbook* (pp. 10–20). Boca Raton, FL: CRC Press Reference. ISBN 9780805862805

Santana, V. F., & Baranauskas, M. C. C. (2010). Bringing users of a digital divide context to website evaluation using WELFIT. In *Proceedings of the IX symposium on human factors in computing systems*, 31–40.

Scholtz, J., Laskowski, S., & Downey, L. (1998). Developing usability tools and techniques for designing and testing web sites. *Proceedings HFWeb*, *98*, 1–10.

USERZOOM. (2018). Retrieved from https://www.userzoom.com

Valencia, X., Arrue, M., Rojas-Valduciel, H., & Moreno, L. (2014). Interdependent components for the development of accessible XUL applications for screen reader users. *Webist*, *2*, 65–73.

Valencia, X., Pérez, J. E., Muñoz, U., Arrue, M., & Abascal, J. (2015, July). Assisted interaction data analysis of web-based user studies. *Human-Computer Interaction – INTERACT 2015* (Vol. 9296; pp. 1–19), Vancouver, BC. Springer International Publishing.

W3C, WAI. (2008, December 11). *Web content accessibility guidelines (WCAG) 2.0*. W3C Recommendation. Retrieved from http://www.w3.org/TR/WCAG20/

## About the Authors

**Myriam Arrue** is an Assistant Professor at the University of the Basque Country/Euskal Herriko Unibertsitatea (UPV/EHU) and a member of the Egokituz Laboratory of HCI for Special Needs at UPV/EHU. Her research focuses on web accessibility evaluation, web navigation behavior analysis, and the application of transcoding techniques for improving UX in web interactions.

**Xabier Valencia** is a researcher at the Egokituz Laboratory of HCI for Special Needs (UPV/EHU). His main research interest is focused on providing higher accessibility to the Web to people with diverse capabilities, principally applying transcoding techniques.

**J. Eduardo Pérez** is a PhD student at the University of the Basque Country/Euskal Herriko Unibertsitatea (UPV/EHU) and a member of the Egokituz Laboratory of HCI for Special Needs at UPV/EHU. His research focuses on studying web navigation behaviors within diverse groups of users with special needs and testing different GUI enhancements in order to improve their web browsing ability.

**Lourdes Moreno** is Visiting Associate Professors in the Computer Science Department at Carlos III University of Madrid, Spain (UC3M), and researcher at the LaBDA Research Group in UC3M. Her research is focused on the design, development, and validation of accessible technology targeted to users with disabilities in different application domains.

**Julio Abascal**, BSD in Physics (U. de Navarra, 1978) and PhD in Informatics (University of the Basque Country-Euskal Herriko Unibertsitatea, 1987), is a Professor of the Computer Architecture and Technology Department of the University of the Basque Country, where he has worked since 1981. In 1985 he co-founded Egokituz Laboratory of Human–Computer Interaction for Special Needs which has participated in several R&D projects at national and international level.

## 3.3. Appendix 3

This appendix includes the article entitled *"Adapting the Web for People With Upper Body Motor Impairments Using Touch Screen Tablets"* that was published in the international journal Interacting with Computers (IwC) in November 2017. That year the journal reached an impact factor of 0.809 according to Journal Citation Reports, ranking 18th out of 22 journals (Q4) in the "Computer Science, Cybernetics" category.

People with lack of dexterity in upper limbs often encounter aggravated accessibility barriers when using mobile devices to access the Web. In order to alleviate the problems faced by this group when using mobile devices, we extended a previously developed transcoding-based system that adapts non-accessible web pages to the needs of specific users in order to enhance their accessibility. The objective of this work was to evaluate the suitability (by means of performance and satisfaction measures) of different interface adaptation techniques and alternative interaction methods for touch screen tablets with eight users with motor impairments. Additional work related to the research included in this paper was presented in several international conferences: including the description and validation of a conceptual model to provide a personalized navigation experience, a comprehensive description of the transcoding system for adapting websites according to user characteristics, and the identification of barriers and adaptations for users with motor impairments within the development of a text-to-speech tool following a user centred approach:

- J. EDUARDO PÉREZ, XABIER VALENCIA, MYRIAM ARRUE, AND JULIO ABASCAL, "Elaborating a web interface personalization process," In *Proceedings of the XVI International Conference on Human Computer Interaction* (Interacción '15), Vilanova i la Geltrú, Spain, Sep. 7–9, 2015, article no. 23.

- XABIER VALENCIA, MYRIAM ARRUE, J. EDUARDO PÉREZ, AND JULIO ABASCAL, "User individuality management in websites based on WAI-ARIA annotations and ontologies," In *Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility* (W4A '13), Rio de Janeiro, Brazil, May 13–15, 2013, article no. 29.

- J. EDUARDO PÉREZ, MYRIAM ARRUE, AND JULIO ABASCAL, "Mintzatek, Text-to-Speech Conversion Tool Adapted to Users with Motor Impairments," In *Proceedings of the 16th International Conference on Enterprise Information Systems* (ICEIS'14), Lisbon, Portugal, Apr. 27–30, 2014, pp. 112-119.

Advance Access publication on 9 August 2017 doi:10.1093/iwc/iwx013

# Adapting the Web for People With Upper Body Motor Impairments Using Touch Screen Tablets

Xabier Valencia[1,*], J. Eduardo Pérez[1], Myriam Arrue[1], Julio Abascal[1], Carlos Duarte[2] and Lourdes Moreno[3]

[1]*EGOKITUZ: Laboratory of HCI for Special Needs, University of the Basque Country (UPV/EHU), Informatika Fakultatea, Manuel Lardizabal pasealekua 1, E-20018 Donostia (Gipuzkoa), Spain*
[2]*Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa Edifício C6, Campo Grande, 1749-016 Lisboa, Portugal*
[3]*Computer Science Department, Universidad Carlos III de Madrid, Avda. de la Universidad Nº 30, Edificio Sabatini, 28911 Leganés, Spain*
**Corresponding author: xabier.valencia@ehu.eus*

**People with disabilities frequently use the Internet to perform a variety of common activities; however, they may often encounter aggravated accessibility barriers when using mobile devices to access the Web. In order to alleviate the problems faced by this group when using mobile devices, we have extended a previously developed transcoding-based system that adapts non-accessible web pages to the needs of specific users in order to enhance their accessibility. In this version, we included new adaptation techniques gathered from the literature in order to apply transcoding techniques to mobile devices. The enhanced system was evaluated with eight users with reduced mobility using tablets. The exploratory study suggests that alternative interaction methods such as the ones named 'end tap' and 'steady tap' are beneficial for some participants with reduced mobility, dexterity or strength in the upper limbs. Other results show that six of the eight users preferred the adapted version with enlarged interaction elements which required less physical effort, even if this adaptation increases the size of the page with the disadvantages associated with such a change.**

## RESEARCH HIGHLIGHTS

- Introduction and evaluation of a web transcoding system that adapts web pages for people with motor impairments who use touch screen devices.
- Tagging the target websites with an extension of the WAI-ARIA mark-up language enabling web transcoding.
- From the evaluation, transcoded pages were revealed to be the preferred option for most of the participants, as they require less physical effort.
- Participants were classified in function of the type and size of finger movement for target selection on touch screen devices, in order to evaluate the performance of diverse alternative interaction methods or gestures.
- Alternative interaction methods were tested by motor impaired users: 'end tap' and 'steady tap' proved to be helpful for people with less finger movement control.

*Keywords: adaptation and personalization; web-based interaction paradigm*

*Editorial Board Member: Dr. Fabio Paternò*

*Received 9 September 2016; Revised 8 May 2017; Editorial Decision 14 July 2017; Accepted 21 July 2017*

## 1. INTRODUCTION

The Internet is an increasingly valuable tool for anyone but especially for people with physical, sensory or cognitive disabilities because it allows them to perform numerous activities relating to labour, leisure, learning, etc. that would be difficult or impossible for them in a physical environment. For this reason, it is vital that digitally provided services are accessible to as many people as possible.

It is known that people with disabilities experience difficulties when they access the web from desktop computers that are similar to the problems experienced by people without disabilities using mobile devices (Yesilada *et al.*, 2011). Evidently, people with motor disabilities are faced with aggravated accessibility barriers when accessing the web from mobile devices. For example, they find the icons are too small or they are not given adequate feedback from their actions. Physical buttons are frequently substituted by gestures (such as 'swipe', 'double tap' and 'pinch'), which can be difficult for people with certain types of disability (Guerreiro *et al.*, 2010; Nicolau *et al.*, 2014; Trewin *et al.*, 2013).

Developers should create pages that are accessible for all in order to alleviate or eliminate these barriers. Unfortunately, although efforts to make the web more accessible are rapidly increasing, the number of non-accessible pages is growing even faster. In addition, universally accessible pages can be problematic because a single design may not work for everyone due to the different characteristics and needs of each person.

An alternative method to enhance Web accessibility is to make existing and currently inaccessible web content accessible. Transcoding is one of the existing approaches for converting non-accessible pages into accessible ones by automatically modifying their code. The tool presented in this paper is framed in transcoding methods. In Valencia *et al.* (2013), we described how our system, based on an extension of the WAI-ARIA (2016) annotation language, adapts web pages for people with disabilities using desktop devices.

In this paper, we present the elements added to that system to enable the adaptation of web pages to touch devices, with the aim of helping people with motor restrictions in their upper limbs to access the Internet via mobile devices. For this purpose, we gathered a number of adaptation techniques from the technical literature intended to enhance the user experience of people with disabilities. The adaptation techniques found are very diverse. For instance, they propose 'to increase the size of the interaction elements' or 'to enable the possibility of performing actions (e.g. scrolling) by tap gestures'. Nevertheless, the interaction methods considered may be insufficient for some users. For this reason, in this paper we propose new interaction methods such as 'end tap', 'steady tap' or 'augment tap', to replace the traditional 'tap' gesture.

An evaluation of this system with eight users with reduced mobility allowed us to examine in detail how users select targets (e.g. finger movements or distances from where the fingers landed in or lifted from to the target). We also measured the usefulness of diverse interaction techniques in different settings.

Advantages and disadvantages of the final design after the application of the different adaptation techniques were also evaluated. Assessment was based on both quantitative metrics (such as time, number of finished tasks) and subjective metrics (such as preferences or estimation of mental workload using the NASA TLX questionnaire) (Hart and Staveland, 1988).

The following sections reviews published works devoted to the creation of personalized user interfaces. In addition, we describe some alternative interaction methods for people with reduced mobility which are used in the desktop domain because these may serve as inspiration for creating alternative interaction methods for touch screen devices.

Subsequently, the implemented adaptation techniques are described, grouping them according to the WCAG 2.0 criteria (WCAG 2.0, 2016) (see Section 3). The proposed transcoding system and the improvements made to adapt it to allow access to the Web for people with upper body impairments are detailed (see Section 4). Finally, the adaptations made and the alternative interaction methods included in the tool are evaluated (Section 5), analysed (Section 6) and discussed (Section 7). Some conclusions are presented in Section 8.

## 2. RELATED WORK

User interfaces can be adapted to the needs of users and to the devices' characteristics in design-time or in run-time adaptation. Languages such as UIML, MARIA XML or UsiXML allow the definition of abstract interfaces in design time. Final user interfaces are generated in runtime, creating a user and device-adapted interface (Abram's *et al.*, 1999; Limbourg *et al.*, 2004; Paterno *et al.*, 2009). These languages can also be used to generate user interfaces in runtime. For instance Ghiani *et al.* (2014) present a system that transforms web pages to MARIA by means of machine learning techniques and subsequently it tailors them to the user.

An example of design-time adaptation is SUPPLE (Gajos and Weld, 2004). The developer creates a declarative description of the interface, device and user model in the design process. Following this, the system creates the final user interface based on functions that take into account the restrictions imposed by the input device, the user and the interface specifications. Although languages of this type are very promising they are not widespread because designers are often reluctant to use them due to the extra requirements of expertise and time.

Our system is a run-time tool that uses adaptation techniques (that is, models or templates) to automatically generate tailored user interfaces. It is applied to enriched web pages, typically present in the semantic Web. Previous annotation of web pages is only required if the target pages are not provided with semantic tags. In the future, when the semantic web prevails, previous manual annotations will not be necessary.

### 2.1. Transcoding techniques

The *transcoding* techniques can be framed in the latter approach. Transcoding is a method that alters the code in run-time in order to adapt web pages to the user or to the device (Asakawa and Takagi, 2008). The *transcoder* application may be located in the web server, in a proxy, or in the client (the user device). If the *transcoder* is located in the server, the user does not need to install or configure it, but it will only be valid for those pages managed by the specific server. Conversely, when the *transcoder* is located in a proxy or a client, it can adapt any page. Transcoders implemented in proxy systems do not require installation, but they may require some configuration. Installation is required when the transcoder is on the client side. This makes it more obtrusive, but it has some advantages, such as providing better control over the final result of the adaptations (Richards and Hanson, 2004). In addition, it can interact with websites under 'secure connection', unlike the proxy version which can have problems with such connections.

One of the first transcoding systems to improve accessibility was developed by IBM Research in Tokyo. Among other features, it numbered the links and serialized the content in order to make Web pages more accessible to screen readers (Asakawa, 2005).

However, this kind of adaptation was rather limited. In order to be able to produce more thorough adaptations it is necessary to know the semantics of the page elements. Semantics can be added by means of annotations. Aurora (Takagi and Asakawa, 2000) was one of the first systems to use semantic annotation to adapt the web. This system characterized user goals by means of a transaction model. 'Adapters' were responsible for adapting the elements involved in user goals and eliminating the remaining elements. The annotation task in Aurora required the creation of a transaction model and a set of rules that applied to each web element, which turned out to be rather time consuming.

Later, Takagi *et al.* (2002) presented a system to improve Web navigation for blind people, using annotation that was made, element by element, through XPATH (2016). This system was able to propose annotations based on the similarity of previously annotated pages on the site, to alleviate some of the burden of the annotation process.

The Sadie system, also created to aid navigation for the blind, proposed a new annotation system consisting in labelling the elements as a type of menu (main menu, submenu, concertina) or assigning a priority to each of them (Harper and Bechhofer, 2007). The identification of the elements in the annotation was made through the CSS (id, class) of the site. This procedure eased the annotation process as CSS elements are used throughout the site.

Our system uses XPATH in addition to CSS when CSS is not sufficient. That is, there is no consistent semantic meaning across the site or CSS is not present. Moreover, as the annotations we proposed are based on the WAI-ARIA standard (WAI-ARIA, 2016), our system can perform a large number of adaptations in the pages that include this standard, even if they have not been manually annotated.

GAPforAPE (Mirri *et al.*, 2012) is a scripting system based on Greasmonkey (2016) that utilizes a user profile to store the preferences. The user profile is stored locally and it follows the XML-based IMS ACCLIP (2016) standard. Among other adaptations, such as CSS transformations or DOM manipulations, it also adds or modifies the scripts of a web page to improve its accessibility, for example, to avoid automatic refreshing of the page. Every time the user requests a page, the system checks if any specific script for the requested page exists, if not it applies a general script. Even if general scripts can be created, they are usually tailored to a specific web page. The application of the user profile enables the personalization of the content, but the profile is locally stored adaptations. Therefore, these preferences are lost when the user accesses the Internet from a different device. Conversely, our system stores the preferences and the user model in a server. In this way the user preferences can be used across different devices.

Akpınar and Yeşilada (2015) presented an eye-tracking experiential transcoding system that sets the role of the visual elements and detects the most common eye path in the visited page in order to transcode it. This is a highly interesting approach, although somewhat limited by the requirement of eye-tracking data which hinders its use on numerous websites.

On the other hand, references to transcoding systems specifically devoted to adapting the web to people with restricted mobility are not frequent. Among the few available, Ivory *et al.* (2003) proposed various adaptations, including the addition of navigation buttons (skip to links, back, forward), and 'making evident the focus'. Although these adaptation techniques appear useful, they do not provide any evaluation of the resulting systems.

To summarize, the transcoding tools found in the literature have complex annotation models or time-consuming annotation processes. Our system, by contrast, uses an extension of the WAI-ARIA language, which is not complex and can be efficiently applied. A large number of adaptations can be applied to pages that are previously annotated with WAI-ARIA language without requiring any further annotation. On the other hand, the annotation of CSS elements, such as ids or classes, is valid for all the web pages that share these specific CSSs. In addition, the use of XPATH expressions allows the annotation of any web page lacking CSS and WAI-ARIA.

Besides, transcoding systems found in the literature, propose a limited number of adaptations targeted to specific groups of users. On the other hand, our system can be used for diverse types of users and different devices. What is more, the granularity of the adaptation techniques and the use of an ontology to decide which adaptation should be applied, enables the easy creation of different adapted web pages, without modifying the tool's code.

### 2.2. New interaction methods

New methods of interaction for people with restricted dexterity to assist them in selecting targets have also been proposed for desktops devices. They include 'steady click', 'bubble cursor', 'angle mouse' and 'adaptive click and cross', among others.

The 'steady click' (Trewin *et al.*, 2006) method allows users to move the cursor away from the target to a certain distance after having clicked on it. Since the adaptation of this technique to touch screen mobile devices appears to be very useful we decided to implement the 'steady tap' version proposed by Trewin *et al.* (2013).

The 'bubble cursor' (Grossman and Balakrishnan, 2005) enlarges or reduces the size of the cursor activation area depending on the proximity of potential targets in order to allow the selection of only one target. We also included a version of this interaction method, which we call 'augmented tap'.

The 'angle mouse' changes the C-D gain depending on the angles between the samples of mouse movements (Wobbrock *et al.*, 2009). When the angle of the trajectory of the mouse does not change the C-D gain is maintained however, if the angle changes the C-D gain decreases, thus smoothing the movement of the mouse. Obviously, this technique does not work on mobile device touch screens because, unlike mouse interactions, there is no cursor path.

The 'adaptive click and cross' technique modifies both the interface and the interaction (Li and Gajos, 2014). When the links are small and close to one another, the user clicks in the target location and then they cross the target element in a circle that appears with all the possible targets. This procedure can be combined with enlarging those elements that are regularly accessed. Despite this technique appearing to be helpful for mobile devices, it can be troublesome for users who have difficulties with the 'slide' gesture.

## 3. ADAPTATION TECHNIQUES

Transcoding techniques convert non-accessible web pages into accessible ones by means of adaptations. In order to select the most adequate adaptation techniques for each case we searched the literature to find what problems were experienced by people with reduced mobility when interacting with touch-input mobile devices and what proposals were put forward to fix them. In addition, we complemented the set of techniques found with a number of generic guidelines to improve mobile accessibility issued by W3C/WAI (Mobile Accessibility, 2016; Mobile Web Best Practices, 2016).

Conventional gestures, such as tapping, for selecting elements or directional gestures required for scrolling or zooming can be troublesome or even impossible for some people with motor impairments (Guerreiro *et al.*, 2010; Nicolau *et al.*, 2014; Trewin *et al.*, 2013). In addition, inadequate element size or position can make the target selection even harder (Guerreiro *et al.*, 2010). Text entry is also a challenging task due to the small size of screen keyboards or the lack of edges between keys (Belatar and Poirier, 2008; Wobbrock *et al.*, 2003). Moreover, some people with motor impairments can also have other associated conditions, such as cerebral palsy, that may include vision problems.

Only adaptation techniques devoted to improving navigation (such as target selection, readability or scrolling) were undertaken for the first version of the tool, a. The implemented adaptation techniques, grouped according to the WCAG 2.0 criteria (WCAG 2.0, 2016) are discussed below.

### 3.1. Perceivable—information and user interface components must be presentable to users in ways they can perceive

Kane *et al.* (2009) found that some users might have problems with the contrast or the font size. These barriers are often due to the restrictions of the users but they can be also caused by environmental conditions. To ensure good contrast for the *main* content, a cream yellow background, black text, and blue links are recommended (which provides a contrast ratio of, at least, 9.41:1). In other sections of the page (*navigation, content-info, complementary, banner*) white background, and black and blue letters are recommended (which ensure a contrast ratio of 9.65:1).

With respect to the font size, in the study carried out by Trewin *et al.* (2013), participants preferred font sizes ranging from 20 pt to 56 pt. For testing purposes we established a 24 pt font size, but in future versions the users themselves will be able to choose the font size that best suits their needs.

### 3.2. Operable: user interface components and navigation must be operable

Guerreiro *et al.* (2010) state that objectives with a 12 mm (or larger) diameter provide good ratio size/error. For this reason, we selected 12 mm as the minimum size of the interaction elements (such as links or buttons). On the other hand, the document W3C Mobile Best Practices (2016) recommends a minimum separation, or inactive space, for small interaction elements. We added a 20 px (4.46 mm) space between links and buttons.

Regarding gestures, Trewin, Swart and Pettick (2013) found that the 'tap' gesture was easily performed by 10 users out 14. Three users had some level of difficulty and only one, encountered serious difficulties. Yet they found that only in 48% of the interactions did the finger movement begin and end in the same point (or near), which is a necessary condition for a valid 'tap'.

They also found an average distance of 17.5 mm between the target and the starting or ending point of the tapping for 28% of users. At the same time, they noticed that actions such as 'slide' and 'pinch' were difficult for a large number of users. Similarly, Nicolau *et al.* (2014) verified that directional gestures were difficult, and that the most effective interaction technique was 'tap' followed by 'crossing'.

As a solution, Trewin *et al.* (2013) proposed new interaction techniques such as 'steady tap' or 'end tap'. 'Steady tap' allows the user to select an item even if the finger moves away from the target within an established threshold. While, 'end tap' allows the activation of an element when lifting the finger from it. Both methods allow the selection of an item even if there are uncontrolled finger movements during the process.

Since the 'slide' gesture may be difficult or even impossible for some users, as attested by Trewin *et al.* (2013) and Nicolau *et al.* (2014), we introduced buttons for scrolling in order to avoid forcing people to use the 'slide' gesture.

Bearing in mind that lack of precision can also be a problem, we decided to increase the activation area around the position where the finger landed: 'augmented tap'. This decreases the precision requirements for selecting the target item (Findlater *et al.*, 2010; Grossman and Balakrishnan, 2005). However, this can be problematic when there are other interactive elements close to the target. For these cases, a disambiguation list was added. The list is ordered by the distance from the finger to the targets, and from bottom to top, as shown in Fig. 1.

### 3.3. Understandable: information and the operation of user interface must be understandable.

The layout for the adapted website created by our system is based on the WAI-ARIA landmarks (*banner, navigation, main, content-info and complementary*), common elements being grouped as recommended by the W3C standard. The *banner* is located at the top, *navigation* elements on the left, *main* content in the middle and *content-info* at the bottom of the page. If there is any *complementary* content it is placed on the right. In addition, *breadcrumbs* were inserted into the top of the *main* content, 'provideConsitentNavigation'.

Finally, a technique that eliminates non-essential page elements (such as advertising or unnecessary images) was also applied in order to make the interface clearer. Figure 2 shows the page before adaptation and Fig. 3 shows the same page after applying all the adaptations techniques.


**Figure 1.** Disambiguation user interface for the 'augmented tap' method of interaction.

## 4. ADAPTATION SYSTEM

### 4.1. Introduction

The transcoding system we presented previously (Valencia *et al.*, 2013) has substantially evolved. In order to adapt the system to the mobile environment, new adaptation techniques were added such as the ones presented in the previous section. The system makes use of an ontology (Gruber, 1993) to model adaptation techniques, the user or the web page, etc. The ontology defined in OWL (2016) was modified in order to adapt the system to the needs identified in the experiments conducted previously (Pérez *et al.*, 2014, 2015; Valencia *et al.*, 2015) and for it to work within mobile environments.

### 4.2. System architecture

Transcoding systems are usually classified as client, proxy or server tools depending the location in which they are placed.


**Figure 2.** Discapnet website.


**Figure 3.** Discapnet website after applying the suitable adaptation techniques.

The designed system has a hybrid architecture since one module is located in the client and others are on a server:

- The Presentation Module runs on the user device (PC, mobile, tablet).
- The Adaptation and Coordinator Modules and the Knowledge Base run on a server.

The adaptation process is roughly as follows. A non-accessible page is cached by the Presentation Module and sent to the Coordinator Module. The Adaptation Module carries out the pertinent adaptations following the information collected from the Knowledge Base. Subsequently, the Presentation Module presents the modified page to the user. Figure 4 shows the process for adapting a previously annotated page and another page with the WAI-ARIA (2016) annotations already integrated.

Even if the Adaptation Module performs the adaptations, the logic of the adaptations is in the Knowledge Base. The Knowledge Base decides which adaptations are applied to specific elements, according to defined rules. This architecture enables easy creation or a set of adaptations without changing the code. For instance, it is possible to add new types of users or devices by simply updating the Knowledge Base as the Adaptation Module is agnostic with respect to both the device and the user. The different modules are explained in detail below.

#### 4.2.1. *Presentation module*

The current implementation of the Presentation Module is an add-on for the Firefox web browser and runs on PCs, smartphones and tablets. Even if it is running as an add-on, the architecture facilitates migration to other platforms (Chrome add-on, proxy, etc.) whenever the new Presentation Module satisfies the following requirements:

- Identify the user
- Catch the web page
- Send the page to the Coordinator Module
- Get and present the modified page

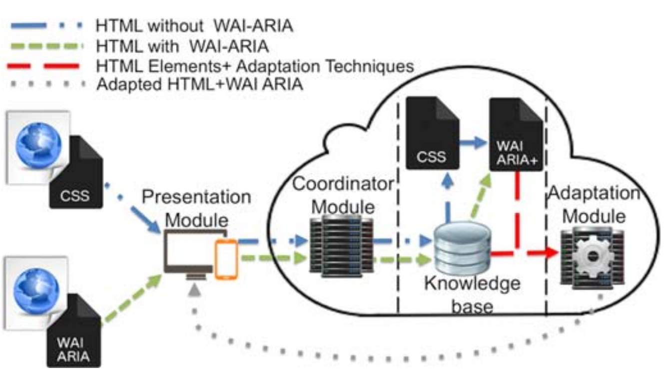


**Figure 4.** General architecture and workflow.

In addition to these requirements, the Presentation Module manages preferences and collects user-generated events. These events can be used to feed specific data mining programs and to detect changes in user skills, such as fatigue, deterioration, etc.

A preference manager for proposing possible preferences to the user has been implemented as a component of the Presentation Module. The changes made in the preferences are stored in the knowledge base and shown in real-time, so the user can immediately notice the consequences of the chosen option. Figure 5 depicts the preferences selection menu for enabling or disabling the scroll buttons.

#### 4.2.2. *Coordinator module*

The Coordinator Module has an instrumental role. It was implemented as a Web service and is responsible for mediating between the various existing modules (Presentation Module, Knowledge Base, Adaptation Module), performing the following tasks:

- Establish communication with the Presentation Module
- Ask for the necessary adaptations to the Knowledge Base
- Update the Knowledge Base
- Communicate with the Adaptation Module

The Coordinator Module (see Fig. 6) first receives the page to be adapted from the Presentation Module along with the user credentials (username, password, device). Afterwards it obtains the adaptation techniques from the Knowledge Base which are suitable for the specific user, device and website.

Using this list the Adaptation Module performs the adaptations and returns the recoded page to the Presentation Module.

Each time the user preferences are modified, a request also arrives to the Coordinator Module, which then updates them in the Knowledge Base. Following this, the aforementioned process is repeated to adapt the page according to the new preferences.

#### 4.2.3. *Knowledge base*

The system is based on a Knowledge Base, implemented in the OWL language (OWL, 2016) created with the ontology editor tool Protégé (2016), which defines the user models, the


**Figure 5.** Bidasoa Tourism website with the preference manager user interface.

adaptation techniques, the devices, the assistive technologies and the annotation model of the web pages. Let us describe the structure (Fig. 7) and content of the ontology.

*Annotation model.* We extended the WAI-ARIA (2016) annotation model in order to be able to perform further adaptations. Using this model the annotator can describe the role of the interaction elements to allow the system to match the most adequate adaptations. Our system can automatically adapt pages with the original WAI-ARIA annotations but it cannot take advantage of the whole set of system features.

Among the new roles added there are 'helping roles', such as '*ContextInfo*', '*FAQ*' or '*Tutorial*', that can be used to provide help to complete a task or to clarify the operation of the element. Another new role is '*SiteMap*', to create a site map of the website when it is not available or to identify an element with such a role, when one is present. '*Caption*' indicates where the video captions are located and '*GeoMap*' provides written directions, instead of a visual map.

In addition, new properties were added: dimming, hide, stretch, remove and priority. '*Dimming*', '*hide*' and '*stretch*' can be used to hide part of the content, such as leaving only the news headings in the starting page to make it simpler and smaller. The '*priority*' property can be used to mark the elements as being necessary or otherwise for the purposes of the task or the point of view of the user, so the page can be reordered with this property taken into account. Finally '*remove*' property is used to remove those elements that can be harmful for the user. For instance, a flashing element is tagged as *remove = 'flashing'* so it would be removed when the user has photosensitivity.

Every annotated element present in the website is stored in the knowledge base in this way: firstly a reference to the website (for example *discapnetsite)*, is stored in the knowledge base as a *website* class. Then, all the website annotated elements are included as *htmlElement* class and linked with the created *website* element with a property assertion (e.g. 'discapnetsite *hasHMTLElement* discapnetFooter', 'discapnetsite *hasHTMLElement* discapnetAdvert'). After that, a role or property is assigned to the html element (e.g. 'discapnetFooter *hasRole* content-info', 'discapnetAdvert hasProperty distractor'). Finally, the html element is identified with the CSS id, the CSS class or the XPATH, with a property assertion (e.g. '*id = foot*' can be a property assertion for a page footer or 'class = publicidadGoogle' for announcement elements).

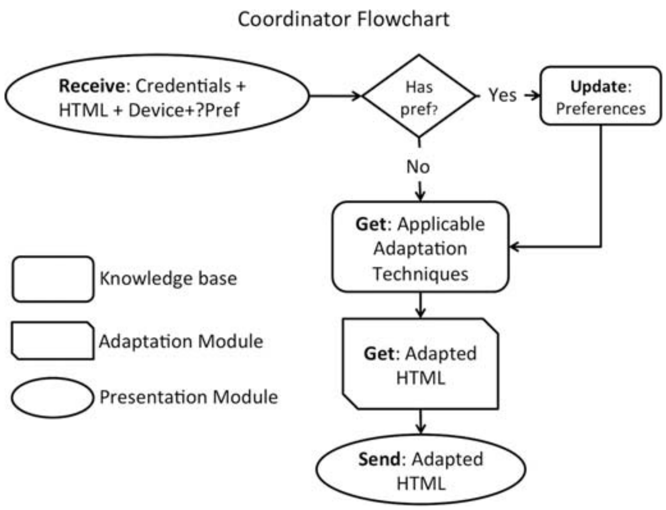


**Figure 6.** Coordinator module flowchart.

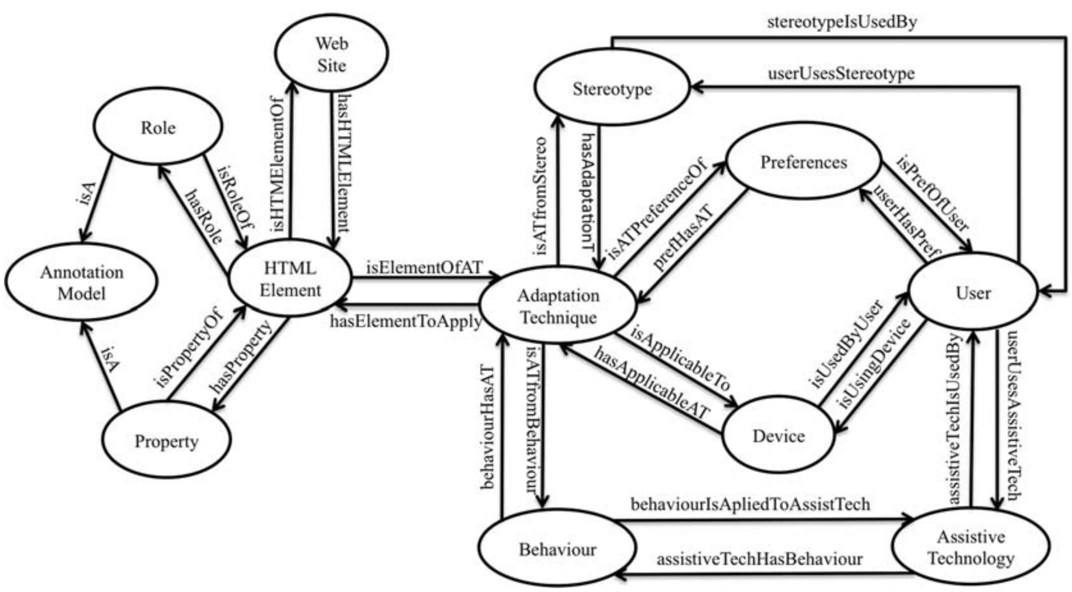


**Figure 7.** Simplified ontology structure and property assertions.

*Adaptation techniques.* The Knowledge Base models the adaption techniques but it does not implement them. In this way it decides which adaptations are applicable. Adaptations are first classified into three main groups: content, presentation and navigation adaptations (Knutov *et al.*, 2009). The parameters of adaptation techniques are defined, when necessary, by identifying the roles and properties of the annotation model. For instance, the technique 'provideConsistentNavigation' can be classified as a navigation adaptation taking the roles *banner*, *navigation*, *main*, *content-info*, *complementary* and *breadcrumb* as parameters.

To date, more than 50 adaptation techniques have been implemented encompassing font style changes, the provision of site maps, the removal of elements, etc. The approach followed enables the addition of new adaptation techniques, thus enabling system development to become a continuous process.

*User model.* Three general user interaction factors were included in order to build the user model *t*: cognitive (C), physical (S) and sensory (S), Table 1. A more detailed user model can be created, adding when necessary new user groups or subgroups.

*Stereotypes and rules.* When inferring adaptation techniques from the characteristics of a given user, inconsistencies between the adaptation techniques can appear. Stereotypes help to avoid this situation. Stereotypes are sets of predefined clusters of adaptation techniques well suited for specific user groups. To date, stereotypes for people with low vision, blind people and for people with motor impairments have been defined.

Stereotypes were created using property assertions:

'Stereotype *motImpaired* hasAdaptationTechnique *removeDistractor*'
'Stereotype *motImpaired* hasAdaptationTechnique *provideEndTap*'

Reasoning rules are used to match users with suitable stereotypes, to connect annotated web elements with adequate adaptation techniques, and to determine which ones are applicable. Rules were coded using the Semantic Web Rule Language (SWRL, 2016) for OWL. Next, some of the rules related with the *removeDistractor* and *provideEndTap* adaptation techniques are explained.

**Table 1.** User model classified by user interaction factors for each subgroup.

| General group | Subgroup | |
|---|---|---|
| C | C0.1 Decline in maintaining attention | |
| | C0.2 Learning disabilities | |
| | C0.3 Language disabilities | |
| | C0.4 Reduced memory capacity | |
| P | P0.1 Limited movement | |
| | P0.2 Inability to use mouse | |
| S | S.S Sight | S.S0.1 low vision |
| | | S.S0.2 blindness |
| | | S.S0.3 colour blindness |
| | | S.S0.4 photosensitivity |
| | | S.S0.5 eye strain |
| | S.H Hearing | S.H0.1 hearing loss |
| | | S.H0.2 deafness |

Firstly, the elements that are part of the adaptation technique are defined with the rule:

'HtmlElement(?el), hasRemoveProperty(?el, *distractor*) -> elementIsPartOfTechnique(?el, *removeDistractors*)'

This step is not required for adaptations techniques without parameters as for example the technique *provideEndTap*.

Subsequently, the target devices for the adaptation technique are set by means of rules such as the following ones:

'Device(?d), deviceHasInput(?d, *touch*) -> isApplicable (*provideEndTap*, ?d)'
'Device(?d) -> isApplicable(*removeDistractor*, ?d)'

The former rule sets the *provideEndTap* as only applicable to devices with touch screens, such as tablets or mobiles. The later rule sets the *removeDistractor* as applicable to all the devices.

As a result, the coordinator module can use the ontology defined in this manner to gather the specific adaptation techniques and their related elements, which are suitable to adapt a concrete website.

*Ontology enhancement.* In the experiments we carried out previously (Pérez *et al.*, 2014, 2015; Valencia *et al.*, 2015), we found that the variability of the characteristics within a user group can be enormous. For this reason, we included the following items in the Knowledge Base: 'Assistive Technology', 'Behaviour', 'Preferences' and 'Device'.

'Behaviour' includes the adaptive techniques that are beneficial to all users who use a particular assistive technology. However, since users of the same assistive technology can also have different experience, strategies, etc, user preferences are also defined in the Knowledge Base. So users can choose which adaptation techniques they want to be applied. Moreover, they can choose added elements such as the navigation bar (up, down, etc.), font colour, sizes, etc.

In addition to preferences or assistive technology, the 'Device' was added to the ontology to define the input/output methods, the operative system of the device, the screen size, etc. The device can be gathered from the Participant Module since it can be obtained from the web browser properties. Assistive Technology instead must be set manually by the user before using the system. This also required the creation of new application rules to determine which adaptation techniques are applicable.

For instance, the next rule determines that a 'behaviour adaptation' technique is applicable to the current web page if the user is using a specific device and a concrete assistive technology.

'behaviourAdaptationHasAdaptationTechnique(?b, ?adt), isApplicable(?adt, ?d), userIsUsingDevice(?u, ?d), usesAssistiveTechnology(?u,?at), assistiveTechHasBehaviour(?at, ?b) -> adaptationTechniquesAreAppliedToUser(?adt, ?u)’

Finally, the adaptation techniques described in Section 3 were also added.

Part of the information contained in this ontology could be obtained from other ontologies, such as the Needs and Preferences part and the ICT Solutions part of the ontology presented by Koutkias *et al.* (2016). Even if we have not discarded this possibility, for this version of the system, we keep a simple set of user data and we combined all the diverse parameters that our system uses to create an adapted web page.

In order to use other ontologies, a thesaurus should be created to match the relevant elements (such as user, preferences, device model, etc.). After harvesting this information, new rules to select suitable adaptation techniques for each case may be required.

#### *4.2.4. Adaptation module*

This module is responsible for performing the necessary adaptations. Its inputs are the web page and the list with the applicable adaptation techniques.

Adaptation techniques can have parameters, such as the web page elements identified by the annotation model. Some techniques do not require parameters. For example, ‘*provideEndTap*’ requires no parameters because it is an interaction aid with no further configuration. Other techniques such as ‘*removeDistractors*’ have those as parameter elements considered to be distractors.

After the application of all the suitable adaptation techniques the result is the adapted web page which is presented to the user.

In order to include new adaptation techniques, its Java code is stored in the adaptation module, and its definition and applicability conditions are inserted in the knowledge base, as mentioned above. Figures 8–10 show the application of a simple adaptation case to remove distractors.

When the conditions for this specific adaptation are met, the original web page (or of the associated CSS) is modified by the adaptation technique, removing inadequate code and/or adding suitable JavaScript or HTML code. For instance, the ‘remove’ adaptation technique receives as parameters the elements to be deleted, as it can be seen in Fig. 8. In this case, the elements ‘class = publicidadGoogle’ tagged as ‘distractors’ in Fig. 9, were removed, resulting the HMTL code in Fig. 10.

Finally, it should be noted that the adaptation techniques are based on the roles and properties, therefore, it is possible to apply them to any annotated website. New adaptation techniques may be required to solve new sources of problems, such as the adaptation of multimedia content, or difficult texts.

```
public void RemoveAdaptation(Vector<HtmlElement> element)
{
        if(trf!=null)
        if(!elements.isEmpty())
        {
                for(HtmlElement el:elements)
                {
                        NodeList nodeList=trf.getELements(el.getIdentification().getType(),
                        el.getIdentification().getValue());
                        if(nodeList!=null)
                        for(int i=0;i<nodeList.getLength();i++)
                        {
                                Element node=(Element)nodeList.item(i);
                                node.getParentNode().removeChild(node);
                        }
                }
        }
}
```

**Figure 8.** Remove adaptation technique code.

```
<script type="text/JavaScript">⊖</script>
<script type="text/javascript">⊖</script>
<!—Inicio Contenedor //—>
<div class="publicidadGoogle">
  <div id="publicidad" class="publicidad_superior">⊖</div>
</div>
<div id="contenedor">⊖</div>
<!—Fin contenedor—>
<div>⊖</div>
<script type="text/javascript">⊖</script>
```

**Figure 9.** Discapnet HTML code.

```
<script type="text/JavaScript">⊖</script>
<script type="text/javascript">⊖</script>
<!—Inicio Contenedor //—>
<div id="contenedor">⊖</div>
<!—Fin contenedor—>
<div>⊖</div>
<script type="text/javascript">⊖</script>
```

**Figure 10.** Discapnet after removing distractors.

## 5. EVALUATION

We conducted a formal evaluation, with three main objectives:

- To collect general characteristics of the users
- To test different interaction techniques
- To measure the results of applying various adaptation techniques

With this purpose in mind we divided the evaluation into two different parts: navigation tasks (Phase 1) and target acquisition tasks (Phase 2). Navigation tasks were used to compare the user interface generated by the results of the applied adaptation techniques with the original un-adapted version. Target acquisition tasks, on the other hand, were used to collect general knowledge about how participants select targets and to evaluate the different interaction techniques.

### 5.1. Users

Eight users with motor impairments in their upper limbs took part in the experiment. Half of them did, in fact, own tablets

**Table 2.** Demographic data.

| User | Age | Gender | Disability | Used Hand | Owns touch device? | Wheelchair |
|---|---|---|---|---|---|---|
| User1 | 41 | Female | Cerebral Palsy | Right | No | Yes |
| User2 | 43 | Female | Cerebral Palsy | Left | Yes (not widely used) | Yes |
| User3 | 44 | Male | Cerebral Palsy | Head pointer | No | Yes |
| User4 | 55 | Male | Cerebral Palsy | Left | Yes | Yes |
| User5 | 55 | Female | Glutaric Aciduria Type I | Left | No | Yes |
| User6 | 55 | Male | Lack of Sensibility | Right | Yes | Yes |
| User7 | 47 | Male | Cerebral Palsy | Right | No | Yes |
| User8 | 60 | Female | Glutaric aciduria Type I | Left | Yes (not widely used) | Yes |

or smartphones, but only two of them claimed to use them regularly. Demographic data can be found in Table 2.

A Firefox Web browser with the RemoTest add-on (Valencia *et al.*, 2015), in charge of presenting stimuli and gathering interaction data from participants for posterior analyses, was installed to carry out the experimental session. In order to avoid possible side effects caused by browser characteristics, participants were encouraged to use only the website elements (not web browser menus).

### 5.2. Tasks and materials

*5.2.1. Phase 1, navigation tasks*

Users were asked to perform a set of navigation tasks in two websites: Bidasoa Turismo (2016) and Discapnet (2016). Discapnet is a website specialized in providing information to people with disabilities, organizations or relatives of people with disabilities. They provide news, documentary collections, information about the rights of people with disabilities, etc. While Bidasoa Turismo provides information related to tourism, such as locations of interest, events, tourism facilities, etc.

Participant's performance under a condition (original or adapted) can be influenced by the experience acquired in the tasks performed under the previous condition. In order to avoid this learning effect, all tasks have two equivalent versions so each one could be assigned to each condition (adapted and original) indifferently. The original 'Bidasoa Turismo' website has a 9-category 'toggle' menu with more than 50 selectable items, Fig. 11, up. In the adapted version, all navigation items were displayed, increasing the page size, Fig. 11, down.

In order to contrast the adequacy of the adapted version against the toggle menu, a number of item selections were set up. Selection of the tasks was based on the scrolling requirements to reach the target (Table 3): NAV1 did not require any scroll, NAV2 required some scroll, and NAV3 was the task requiring the largest scroll, due to the target being located at the bottom of the web page.

In the original 'Discapnet' web page the size of the elements is, in general, quite small and the selectable elements


**Figure 11.** Bidasoa Tourism original (up) and adapted (down) with the navigation menu highlighted.

are often surrounded by other elements. By contrast, the adapted version contains larger elements separated by larger spaces. As a result, this also increased the page size, thus requiring larger scrolls.

In ACT task, Fig. 12, users had to select three elements while in CA, Fig. 13, participants were required to tap on a small element that was close to other selectable elements (Table 4). In the Adapted version, these elements were larger and were more widely separated. Consequently, the page became larger in the adapted version and larger scrolling was required.

On the other hand, SD1, SD2, SD3, SD4 and SB1, SB2, SB3, SB4 tasks were search tasks that allowed participants to use the adapted and original versions more naturally (Table 5). All the search tasks in Discapnet (SD1, SD2, SD3 and SD4) were three clicks away from the homepage. In the Bidasoa Tourism website (SB1, SB2, SB3, SB4), they were two clicks away from the homepage.

*5.2.2. Phase 2, target acquisition tasks*

Target acquisition tasks were carried out to evaluate the implemented new interaction methods and to gather data about the selection of the targets (Table 6). Three web pages

**Table 3.** Tasks to contrast scroll in the adapted version against the toggle menu in the original.

| Task | Description | Website |
|---|---|---|
| NAV1 and NAV1′ | Target selection with no scroll in the adapted page and toggle menu in the original | Bidasoa Turism |
| NAV2 and NAV2′ | Target selection with medium scroll in the adapted page and toggle menu in the original | Bidasoa Turism |
| NAV3 and NAV3′ | Target selection with large scroll in the adapted page and toggle menu in the original | Bidasoa Turism |

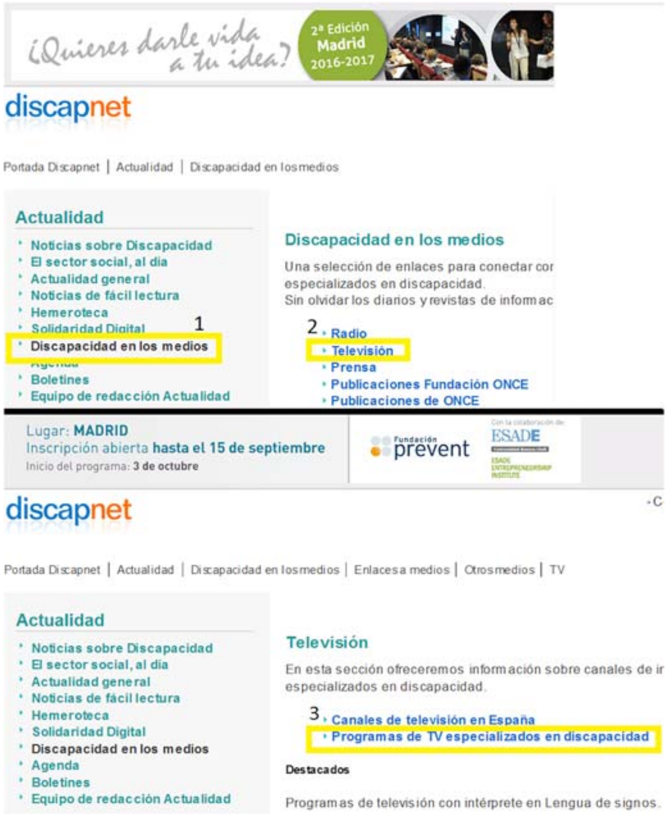


**Figure 12.** ACT task with the three links highlighted.

were created, each one containing nine links to be selected, with the distribution shown in Fig. 14 (TA2). The target was highlighted with a yellow background. Once the highlighted link was selected, the next link was highlighted.

In the task TA1, the link to be selected had no other links around it. In the tasks TA2 the link was bordered by two links with a standard separation (Fig. 14). The task TA3 was similar to TA2 but with a separation of 20 px between links. Four interaction methods were tested: ‘standard’, ‘end tap’, ‘steady tap’ and ‘augmented area’.

### 5.2.3. *Subjective measures*

In order to collect subjective data from participants, a NASA TLX questionnaire (Hart and Staveland, 1988) was used. The objective of the NASA TLX questionnaire was to analyse the adapted and original versions for each questionnaire dimension.

We also conducted a short final interview in order to ascertain which condition was their favourite and which one was

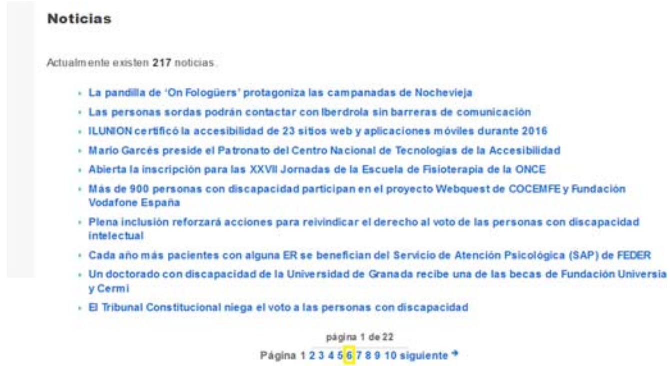


**Figure 13.** CA task with the target link highlighted.

more comfortable for reading and selecting elements. In addition they were also asked whether they preferred the toggle menu or the open menu.

## 5.3. Procedure

Four participants carried out the study session in the installations of their organization, two at their home and the remaining two in the university lab. Each session lasted between one and two hours. Sessions began with the experimenter explaining the system and the experimental session.

The experimenter asked the participant where the tablet should be placed for maximum user-comfort. Different strategies were used to enable the use of the tablet depending on each participant’s needs. Two subjects used a device mount to fix the tablet to their wheelchairs (see Fig. 18). Four subjects placed the tablet on a table: one used a holder (Fig. 17), two fixed it with Velcro (Fig. 15), and a third did not use any additional help. The last two participants placed the device on a lectern (Fig. 16).

After placing the tablet, a training session was carried out to detect if additional adaptations were needed to interact with the tablet. One subject used a glove with an attached touch pen. Another user used a glove with a finger cut out and the last one used a head pointer with a touch pen.

Once the participants were ready, they were asked whether or not they wanted specific buttons to scroll the page in the interface in order to avoid the ‘swipe’ gesture. The decisions made by participants were stored as a preference in the adaptation system. Three users decided to perform direct slide gestures without assistance, while another three preferred the buttons as an alternative to the slide gesture. For the

**Table 4.** Task to measure the drawbacks of the added scroll against the incremented size of elements and space with surrounding elements.

| Task | Description | Website |
|---|---|---|
| CA and CA′ | Select a small link surrounded by others at the bottom of the page | Discapnet |
| ACT and ACT′ | Select three standard links, surrounded by others | Discapnet |

**Table 5.** Navigation search tasks.

| Task | Description | Website | Depth |
|---|---|---|---|
| SD1 | Search information about the special need in education of people with disabilities | Discapnet | 3 |
| SD2 | Search information about Type A flu | Discapnet | 3 |
| SD3 | Search the urban transportation guide | Discapnet | 3 |
| SD4 | Search information about the state of art of research in assistive technology | Discapnet | 3 |
| SB1 | Search for flyovers | Bidasoa Tursim | 2 |
| SB2 | Search for routes around Jaizkibel | Bidasoa Tursim | 2 |
| SB3 | Search information about the 'Faro de Higer 3ª' Camp sites | Bidasoa Tursim | 2 |
| SB4 | Search information about the 'J.Sebastian Elkano' youth hostel | Bidasoa Tursim | 2 |

**Table 6.** Target acquisition tasks description.

| Task | Description |
|---|---|
| TA1 | Select the highlighted link (9 times). The link to be selected has no other links in the surroundings. |
| TA2 | Select the highlighted link (9 times). The link to be selected is bordered by two links with a standard separation (Fig. 14) |
| TA3 | Select the highlighted link (9 times). The link to be selected is bordered by two links with a 20px separation. |

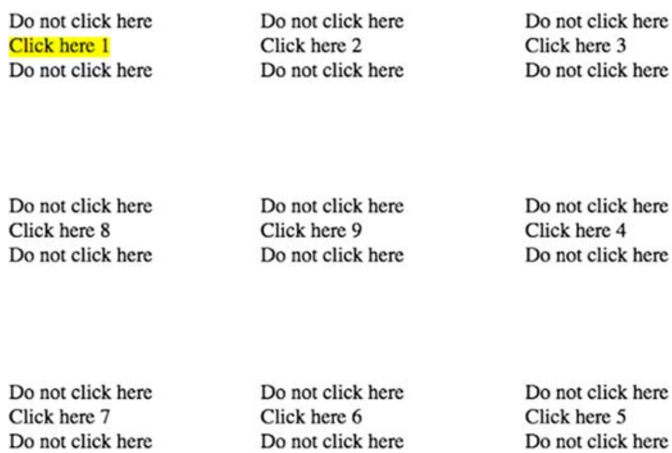


**Figure 14.** Target acquisition task with target links surrounded by other selectable elements with standard separation (TA2).

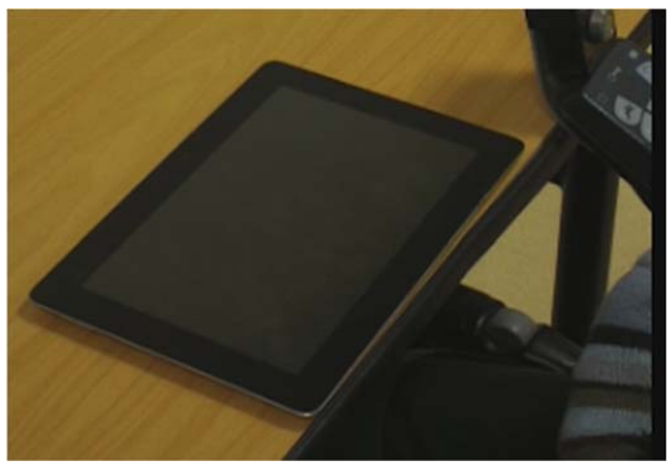

**Figure 15.** Tablet on the table with Velcro.

remaining two participants (User1 and User3) this setting was essential as their physical characteristics prevented them from making the 'slide' gesture. For this reason, the intervention of a researcher was necessary to perform the scroll in the sessions with the original version.

The experimental session was divided into three parts, Trial, Phase 1 and Phase 2. The Trial was used to enable participants to familiarize themselves with the task types and the system. Phase one was used to evaluate the design of the user interface after the application of the adaptation techniques. And finally Phase two aimed to evaluate the different interaction methods, 'standard', 'end tap', 'steady tap' and 'augmented tap'.

In Phase 1 of the session, each participant carried out nine tasks under each condition, adapted and original, in two websites Discapnet (4) and Bidasoa Turismo (5). Conditions were counterbalanced between users.

For each task type (ACT and ACT′, CA and CA′, etc.) one task was assigned randomly to a condition (ACT to the adapted) and the other task was assigned to the remaining condition (ACT′ to the original). The order of tasks was

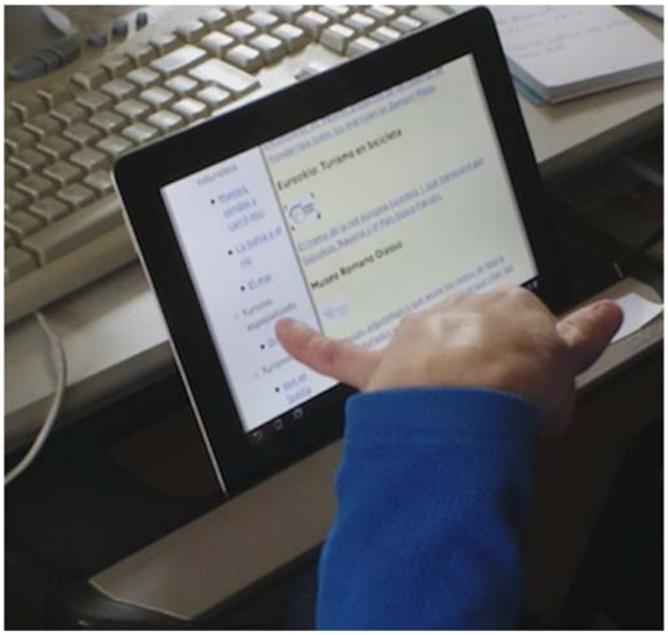

**Figure 16.** Tablet on a lectern.

**Figure 17.** Tablet on the table with a holder.

randomly assigned but the search tasks (SD1, SD2, …, SB4) were performed first. Once they completed each condition (original, adapted), participants rated it using the NASA TLX questionnaire.

Phase 2, consisted in three target acquisition tasks (TA1, TA2, TA3) that were carried out under four different interaction methods: 'standard', 'end tap', 'steady tap' and 'augmented tap'. The interaction methods were counterbalanced between users and the task order was randomly selected for each user.

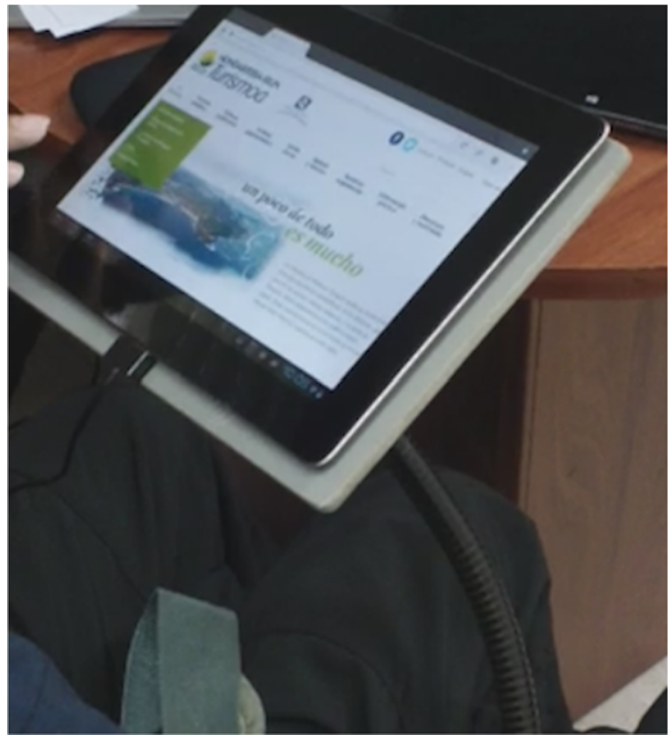

**Figure 18.** Tablet mounted on the wheelchair.

After finalizing the experimental session, a semi-structured interview was conducted in order to gather more information about the users' thoughts concerning the adapted or unassisted versions.

## 6. RESULTS

### 6.1. Phase 1—navigation tasks

#### *6.1.1. Adaptation system performance*

In order to calculate the efficiency of the adaptation system, the time elapsed from the page request to the page load was considered from all users for the adapted version and for the original version. The average value for the adapted version was 2315.61 and 1925.67 ms for the original, a difference of 389.94 ms.

#### *6.1.2. Enlarge elements and increase scroll*

To analyse how much 'enlarging the size of elements and therefore the site' benefits or harms the users, tasks ACT, CA, NAV1, NAV2 and NAV3, were analysed. The data obtained are presented in Table 7. Data were analysed using the Student's *t*-test for paired groups since the distribution of the data was normal except for CA and NAV1 tasks. In such cases, results were transformed logarithmically to have a normal distribution. In all cases the null or H0 hypothesis was 'the adapted and original versions produce similar results'.

In the ACT task (several selections of elements) no differences were found: $t\ (7) = 0.437$, $P = 0.675$. By contrast, in

**Table 7.** Time needed to complete task in both conditions (adapted and original).

| | ACT | CA | NAV1 | NAV2 | NAV3 |
|---|---|---|---|---|---|
| | | Adapted | | | |
| mob1 | 141.222 | 47.035 | 13.829 | 63.872 | 77.931 |
| mob2 | 134.318 | 48.752 | 57.621 | 89.740 | 59.231 |
| mob3 | 58.409 | 53.252 | 22.828 | 42.022 | 72.244 |
| mob4 | 70.936 | 28.499 | 34.404 | 23.711 | 34.457 |
| mob5 | 110.420 | 36.945 | 22.834 | 74.991 | 82.144 |
| mob6 | 57.693 | 20.041 | 11.857 | 22.522 | 24.780 |
| mob7 | 92.537 | 35.906 | 31.940 | 62.048 | 46.654 |
| mob8 | 120.543 | 27.661 | 12.079 | 43.322 | 62.944 |
| | | Original | | | |
| mob1 | 113.285 | 29.644 | 87.696 | 72.818 | 89.030 |
| mob2 | 179.573 | 74.760 | 24.907 | 70.850 | 29.875 |
| mob3 | 76.017 | 128.293 | 39.488 | 43.429 | 23.940 |
| mob4 | 60.255 | 38.332 | 33.866 | 39.731 | 23.435 |
| mob5 | 131.105 | 69.609 | 37.552 | 63.170 | 29.981 |
| mob6 | 68.955 | 18.448 | 14.961 | 15.871 | 13.583 |
| mob7 | 66.066 | 56.716 | 50.999 | 21.533 | 15.268 |
| mob8 | 43.735 | 80.440 | 25.950 | 37.816 | 44.810 |

the CA task (selecting a very small element with scroll), the differences between the adapted version and the original are very close to being significant: $t\ (7) = -2.2944$, $P = 0.055$. In this case, the adapted version is better than the original by an average of 24.77 s, with a confidence interval of between −0.11 and −49.42 s and an effect size of 0.81.

On the other hand, in the tasks involving a comparison of the navigation menu (‘toggle menu’) with the open menu, no significant differences could be found in the case with little scroll NAV1: $t\ (7) = -\ 1.641$, $P = 0.145$ and an effect size of 0.58. This supposes an average enhancement of 13.5 s for the adapted version over the original.

In the case with medium scroll requirements, NAV2, no differences were found: $t\ (7) = 1.151$, $P = 0.287$ and an effect size of 0.40. The original was 7.13 s faster than the adapted version.

Finally, for the case with more scroll demands, NAV3, the adapted web page, showed significant differences: $t\ (7) = 3.22$, $P < 0.05$ and an effect size of 1.14. The original was, on average, 23.81 s faster with a confidence interval ranging from 41.30 to 63.13 s.

#### 6.1.3. *Search tasks*

In the original condition participants were able to reach more targets, as can be seen in Table 8. Only one user (User4) finished more tasks in the adapted version. User2 did not find any and User6 found two in each version. The rest found more targets in the original, the most prominent being User7 who found three in the original and none in the adapted version.

**Table 8.** Number of tasks completed in search tasks in both conditions.

| | Adapted | Original |
|---|---|---|
| mob1 | 2 | 3 |
| mob2 | 0 | 0 |
| mob3 | 0 | 1 |
| mob4 | 2 | 1 |
| mob5 | 0 | 1 |
| mob6 | 2 | 2 |
| mob7 | 0 | 3 |
| mob8 | 1 | 3 |

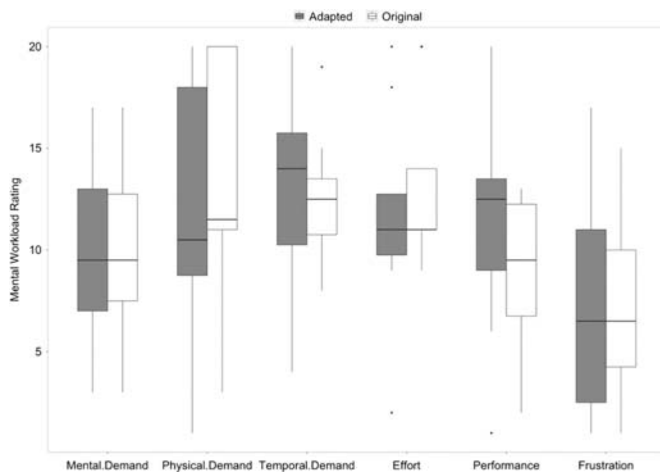


**Figure 19.** Boxplot for each NASA TLX questionnarie dimension.

#### 6.1.4. *NASA TLX questionnaire results*

Each of the dimensions of the NASA TLX questionnaire was analysed with the Wilcoxon test. The resulting dimension distribution can be seen in Fig. 19.

Unexpectedly, no significant differences were found for Mental Demand, even if in the adapted version both sites had the same structure ($W = 9.5$, $Z = -0\ 428$, $P = 0.688$). Neither were any differences found for the Effort dimension ($W = 0$, $Z = -1\ 720$, $P = 0.25$). By contrast, significant differences were found in Physical Demand ($W = 0$, $Z = -2.40$, $P < 0.05$, effect size = 0.849) in favour of the adapted version. Regarding Temporary Demand ($W = 6.5$, $Z = 0.835$, $P = 0.625$) and Performance ($W = 22$, $Z = 0.566$, $P = 0.656$), users tended to value the original page higher, although not significantly. Finally, Frustration Level ($W = 12.4$, $Z = 0.567$, $P = 0.625$) was very similar for both types of pages.

### 6.2. Phase 2—target acquisition tasks

#### 6.2.1. *How users select targets in a touch screen*

The target acquisition tasks (TA1, TA2, TA3) allowed the analysis of the different interaction methods: ‘standard’, 'augmented', 'end' and 'steady'. How participants select the targets was also analysed with the date obtained from the 'standard' method of interaction. We collected average distances from the centre of the target to the points where the finger touched the screen (TD) and left the screen (TU). The distance (*D*) travelled by the finger while selecting was also measured and its relation to the optimal distance (CI). In addition, the number of times (NF) that each user touched the screen with more than one finger was counted.

Table 9 presents the results gathered from the standard method of interaction. Most users did not move their fingers substantially during the selection except User8 who moved their finger 10.8 mm on average. Other users, such as User2, User5 and User7, also moved their fingers during target selection 4.90, 3.76 and 3.65 mm, respectively.

User8 had the largest CI, 2.07, followed by User7 (1.88), User2 (1.50) and User5 (1.27). This indicates that their fingers travelled longer than optimal distances indicating that they might have precision or control problems.

On the other hand, User1, User2 and User3 located their fingers quite far from the centre of the object 31.39, 29.21 and 31.39 mm, respectively, indicating difficulties in making the right selection.

With regard to the number of times they touched the screen with more than one finger, User2 did this eleven times. Evidently, touching the screen with more than one finger hindered the user from making an accurate target selection.

The three different scenarios discussed above were considered to assess the four different methods of interaction. Since data were not normal, it was analysed with the non-parametric Friedman test. In all scenarios the null hypothesis H0 was 'there are no differences between alternative methods and the standard one'.

### 6.2.2. TA1: *target selection with nothing around*

The Friedman's test found no significant differences between the methods of interaction, $\chi^2$ (3) = 1.819, $P$ = 0.610. Although the differences were not significant, the 'augmented' method achieved, on average, the lowest value (3516.78 ms), followed by 'end' (3783.77 ms), standard (4329.38 ms) and finally 'steady' (5055.23 ms). As can be seen in Fig. 20, the data are more compact for 'augmented'.

**Table 9.** Users' characteristics from target acquisition tasks with standard interaction.

| | TD | TU | *D* | CI | CTU |
|---|---|---|---|---|---|
| mob1 | 32.14 | 32.91 | 0.62 | 1.00 | 0 |
| mob2 | 29.21 | 28.56 | 4.90 | 1.50 | 11 |
| mob3 | 31.39 | 31.08 | 0.48 | 1.00 | 0 |
| mob4 | 3.51 | 3.51 | 0 | 1 | 0 |
| mob5 | 13.95 | 12.84 | 3.76 | 1.27 | 0 |
| mob6 | 5.28 | 5.28 | 0 | 1 | 0 |
| mob7 | 15.36 | 15.43 | 3.65 | 1.88 | 0 |
| mob8 | 18.54 | 19.02 | 10.80 | 2.07 | 0 |

For all the users, except for User6, at least one of the alternative methods of interaction produced a better average value. In some cases (User7, User4, User3, User1, User5), the difference was small: less than a second. Other users obtained higher differences on average: User2 obtained 3 s and User8 7 s.

### 6.2.3. TA2: *target selection with two bordering links with standard separation*

In this scenario significant differences were found: $\chi^2$ (3) = 41,732, $P < 0.01$. A Mann–Whitney test with a Bonferroni correction *post hoc* test, showed differences between standard and 'augmented' ($P < 0.01$), but not with others (standard-end and standard-steady, $P = 1$).

The 'augmented' method, which, having produced good results in the previous case, was the worst this time. Having to frequently disambiguate between probable targets significantly increases the time required to select an item. In general, 'end' (3930.08 ms) produced the best value on average, followed by 'steady' (4202.85 ms), 'standard' (4297.57 ms) and 'augmented' (10 597.92 ms) as shown in Fig. 21.

User1, User4, User5 and User6 obtained better values with the 'standard' interaction method. 'End' is similar to the 'standard' for the participants User3, User4 and User6 (<300 ms). The subject User2 obtained an average difference of 6 s with 'steady' and 3 s with the 'end'. User7 obtained a difference of about 2 s with 'end' and 'steady'. Finally, User8 obtained a difference of almost 5 s with 'steady' and nearly 3 s with 'end'.

### 6.2.4. TA3: *target selection with two bordering links with a 20 px separation*

In this case, the Friedman test found significant differences: $\chi^2$ (3) = 48,337, $P < 0.01$. The Mann–Whitney post hoc

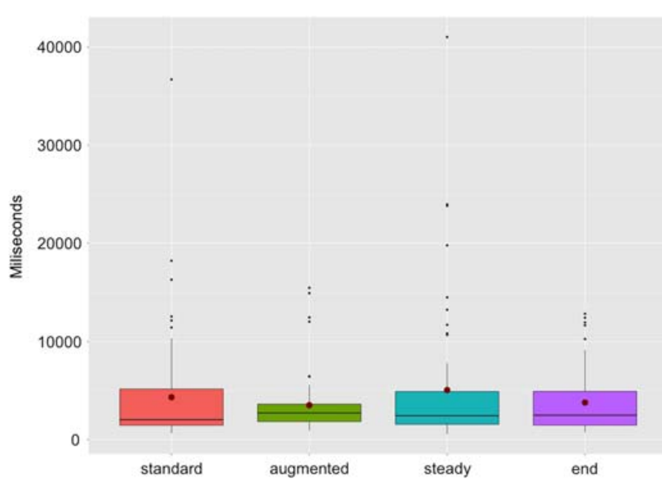


**Figure 20.** User selection time boxplots, by interaction method (links not surrounded by other selectable items).

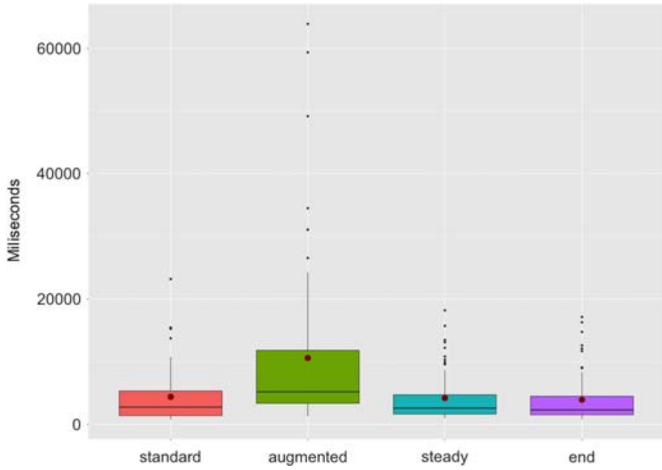


**Figure 21.** User selection time boxplots, by interaction method (links bordered by other selectable items).

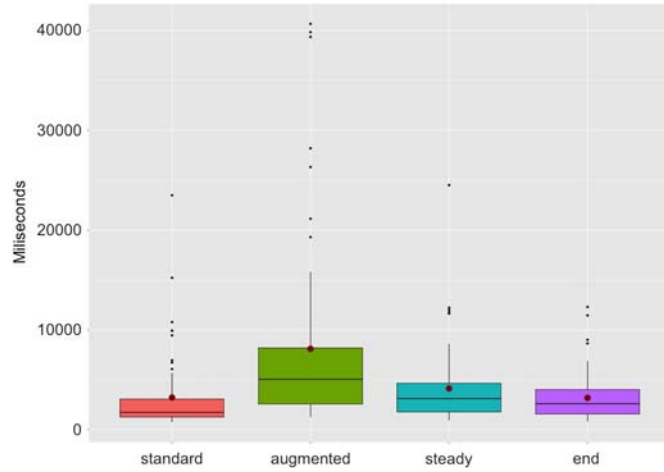


**Figure 22.** User selection time boxplots, by interaction method (links bordered with 20 px separation).

analysis with Bonferroni correction, showed differences between 'standard'–'augmented' ($P < 0.05$), and 'standard'–'steady' ($P < 0.05$). In both cases, the differences were in favour of the standard interaction method. Overall 'end' produced better average results (3192.83 ms) closely followed by 'standard' (3232.52 ms). Slightly further behind were 'steady' (4135.98 ms) and 'augmented' (8105.34 ms) as can be appreciated in Fig. 22.

On average, the 'standard' was better for User1, User3, User4, User7 and User8. However, 'end' produced similar values to the 'standard', except for User8. User6 obtained a minimal difference with 'end'. User5 reduced the time by 1 s with 'steady' while User2 reduced the time by almost half with 'end'.

#### 6.2.5. *Was any interaction method helpful for any participant?*

To find out whether, in any of the cases, the alternative methods were of any help, we analysed the users with the highest CI (User2, User7 and User8).

Firstly, we analysed User2, for whom significant differences were found in TA2: $\chi^2$ (3) = 14.6, $P < 0.01$, although the post hoc test could not clarify which pairs were implied ('standard'–'augment' 1, 'standard'–'steady' 0.50, 'standard'–'end' 1). Although for TA1 no significance was obtained it was quite close: $\chi^2$ (3) = 6.6, $P = 0.086$. The same result was obtained with the post hoc test.

For User7, no differences were found in TA1 and TA3 but there were differences in TA2: $\chi^2$ (3) = 13,133, $P < 0.01$ and in the post hoc values: 'standard'–'augmented' 1, 'standard'–'steady' 0.291, 'standard'–'end' 0.085. Although not significantly, 'end' is very near to differentiating (0.085). Therefore, it would appear that this user could benefit from 'end' when the selected target is surrounded by other targets.

Finally, User8 obtained significant results in all three cases: $\chi^2$ (3) = 13.93, $P < 0.01$, $\chi^2$ (3) = 17.4, $P < 0.01$ and $\chi^2$ (3) = 13.4, $P < 0.01$. The post hoc test could not find differences between the groups in TA1 ('standard'–'augmented' $P$ = 0.767, 'standard'–'steady' $P$ = 0.291 and 'standard'–'end' $P = 1$). In TA2 the results were better for 'steady' ('standard'–'augmented' $P$ = 1, 'standard'–'steady' $P$ = 0.0488, 'standard'–'end' $P$ = 0.3249). In TA3, the differences were found only for 'augment' worsening the time needed ('standard'–'augmented' $P = 0.012$, 'standard'–'steady' $P = 1$, 'standard'–end' $P = 0.182$).

### 6.3. Interview

Following the sessions, a final interview was carried out to obtain the thoughts and preferences of the users on both the adapted and the original websites. The following questions were asked in the interviews:

- Overall which is your preferred interface (Adapted, Original)?
- Which one is more comfortable for reading and selecting links (Adapted, Original)?
- Which do you prefer, having more space between links and an increased page size or having less space and a reduced page size?
- Which do you prefer, an open menu or a toggle menu?
- Do you think it is useful to maintain the structure of pages within websites?
- Can you perform the Zoom gesture?
- Can you perform the Slide gesture?

All but two participants (user6 and user8) preferred the adapted version to the original one, both from an overall point of view and specifically for reading or selecting links. Six participants preferred to have more space between links, one less space (user2) and the last one did not answer.

With regard to the open menu vs. toggle menu, five participants selected the toggle menu while the other three (user2, user5, user7) preferred the open menu. All but one thought that maintaining the same structure across websites could be useful. The remaining participant was indifferent once they had become accustomed to the different structure of the website.

With regard to the questions about the Zoom or the Slide gestures, only one participant (user6) was able to perform these without significant problems. The slide gesture was easier for most users although two participants (user1, user3) were not able to do it.

## 7. DISCUSSION

As far as design issues are concerned, it seems clear that having to perform a considerable amount of scrolling increased the time required to perform the task (NAV3). However, the larger size of the links and the space between elements added in the adapted page seems to have been helpful (NAV1). The absence of significant differences in NAV2 is quite promising, as performing the task in the adapted page entailed a certain amount of scrolling. On the other hand, when the links are very small and they are surrounded (CA) better results are obtained with the adapted page, despite the need to do more scrolling.

However, it should be pointed out that faster is not always better (a vision focused on productivity or business). Users do not always prefer to be faster, as indeed they reported for our study: six of eight preferred the adapted page. Comfort, reduction in the number of mistakes or easier item selection can be factors that lead to forming a preference for the adapted pages, even though task performance times may be longer.

The results in the search tasks might be explained because participants were probably already used to the structure of the sites. Another issue would be the use of the breadcrumbs as a method for identifying the page or section in the adapted page. Most of the users who took part in the study do not usually use—and are unfamiliar with—breadcrumbs. Therefore, a possible improvement would be the use of colours for the sections. Moreover, the information about the section in which the users find themselves should be more evident (and not based on breadcrumbs).

From the NASA TLX questionnaire it can be seen that adapted pages generated less physical demand, which is very important for people with motor disabilities. Results (although not significant) of the adapted pages in performance, frustration or time demand, seem to be explained by the probable relation with the number of completed tasks in each condition (adapted and original).

Regarding the target acquisition tasks, interestingly most users needed, on average, more time for the task with only one link (TA1) than for the other tasks, using the 'standard' interaction. The explanation could be that when the links are at the edges of the screen unintentional pressing or interaction with buttons on the navigation bar, back button, the watch, etc. can happen. On the other hand, in TA3 the best results were obtained in contrast to TA2 and TA1 by the different methods of interaction except 'steady'. This highlights the importance of active elements being maintained at a minimum distance from each other.

The 'augmented' technique applied to targets which were not surrounded appears to provide some help to many users, though not significantly. For other cases, having to disambiguate increases the target selection time significantly, making it unsuitable for surrounded interactive elements.

Some users can gain an advantage from alternative methods of interaction under certain conditions. Examples of this are the subjects User2 and User8, in the cases TA1 and TA2, with the steady method, or User7, with 'end', in TA2. In order to help other users, such as User1 and User3, it would be necessary to detect their pattern to select elements in order to preview the objective they want to click (Montague *et al.*, 2012; Mott *et al.*, 2016). Remember that, even if they do not drag their fingers for selection, they set them down slightly away from the target (>3 cm).

Finally, from the results of the users' feedback regarding the menus, it is important to provide a customization option for the type of menu ('open' or 'toggle'). Although the use of these types of menus for people with physical disabilities is not recommended, five users preferred it.

## 8. CONCLUSIONS

The evaluation carried out showed that the implemented transcoding system is able to adapt websites to touch screen mobile devices used by people with motor impairments. As a result, most users prefer transcoded pages, although several improvements in the user interface and in the interaction methods are required. Some of the required improvements are discussed below.

### 8.1. User interface

Regarding the user interface, more customization features are needed. Such as, for example, letting the users set their preferred font size or the minimum size of interaction elements.

The final interview showed that a number of users preferred a 'toggle' over the 'open' menu. This choice can be also provided by the system as a user preference. The provision of these options would allow the page size to be adjusted adequately—and therefore the need for scrolling—to the user requirements. This, in consequence, would help to reduce the time needed to accomplish their tasks.

Finally, the importance of providing buttons for scrolling was highlighted. While performing the 'slide' gesture is possible for most users, for some other users it is a very difficult—or impossible—gesture. The simple act of providing buttons for scrolling can make the difference between being able to surf the web or not.

### 8.2. Interaction methods

Due to the high variability of the characteristics and needs of users with motor disabilities, finding an optimal alternative method of interaction for everyone was not possible. However, some methods work well for specific people under particular circumstances. For instance, participants whose fingers move without control during the selection of targets can benefit from 'end tap' or 'steady tap' interaction methods.

The lack of more universal interaction methods could be resolved by a more thorough longitudinal study that would enable us to determine when—and for whom—one interaction method is better than another. This knowledge can be used to provide a path towards dynamically adaptive interaction. For instance, in an adaptive system, 'augmented tap' would be applied when a selectable element is alone and 'steady tap' or 'end tap' when the element is surrounded by other elements.

Nevertheless, it should not be overlooked that interaction methods must be changed with caution. Changes in the interaction techniques should not interfere with consolidated gestures, such as 'slide'. Therefore, in addition to testing interaction methods with target selection tasks, these should also be evaluated while surfing the web. This can help to determine how useful the alternative interaction method really is.

## ACKNOWLEDGEMENT

We would like to thank all the participants that took part in the study.

## FUNDING

This research work was developed within the project eGovernability, funded by the Spanish Government, Ministry of Economy, Industry and Competitiveness (MINECO), and the European Regional Development Fund (ERDF), under grant (TIM2014-52665-C2-1-R). J.E.P. holds a PhD Scholarship from the University of the Basque Country (UPV/EHU). Some of the authors are members of the EGOKITUZ/ADIAN research team, supported by the Basque Government, Department of Education, Universities and Research under grant (IT980-16).

## REFERENCES

Abram's, M., Phanouriou, C., Batongbacal, A.L., Williams, S.M. and Shuster, J.E. (1999) UIML: an appliance-independent XML user interface language. Comput. Netw., 31, 1695–1708.

Akpınar, E. and Yeşilada, Y. (2015) Old habits die hard!: eye tracking based experiential transcoding: a study with mobile users. In Proc. 12th Web for all Conf. (12 pages). ACM

Asakawa, C. (2005) What's the web like if you can't see it? In Proc. 2005 Int. Cross-Disciplinary Workshop on Web Accessibility (W4A) (pp. 1–8)

Asakawa, C. and Takagi, H. (2008) Transcoding. In Web Accessibility. pp. 231–260. Springer, London.

Belatar, M. and Poirier, F. (2008) Text entry for mobile devices and users with severe motor impairments: handiglyph, a primitive shapes based onscreen keyboard. In Proc. 10th Int. ACM SIGACCESS Conf. Computers and accessibility, pp. 209–216.

Bidasoa Turismo (2016) http://www.bidasoaturismo.com/

Discapnet (2016) http://www.discapnet.es/Castellano/Paginas/default.aspx

Findlater, L., Jansen, A., Shinohara, K., Dixon, M., Kamb, P., Rakita, J. and Wobbrock, J.O. (2010) Enhanced area cursors: reducing fine pointing demands for people with motor impairments. In Proc. 23nd annual ACM symposium on User interface software and technology, pp. 153–162.

Gajos, K. and Weld, D.S. (2004) SUPPLE: automatically generating user interfaces. In Proc. 9th Int. Conf. Intelligent user interfaces, pp. 93–100.

Ghiani, G., Manca, M., Paternò, F. and Porta, C. (2014) Beyond responsive design: context-dependent multimodal augmentation of web applications. In Int. Conf. Mobile Web and Information Systems, pp. 71–85.

Greasmonkey (2016) https://addons.mozilla.org/es/firefox/addon/greasemonkey/

Grossman, T. and Balakrishnan, R. (2005) The bubble cursor: enhancing target acquisition by dynamic resizing of the cursor's activation area. In Proc. SIGCHI Conf. Human factors in computing systems, pp. 281–290.

Gruber, T.R. (1993) A translation approach to portable ontology specifications. Knowl. Acquis., 5, 199–220.

Guerreiro, Y., Nicolau, H., Jorge, J. and Gonçalves., D. (2010) Towards accessible touch interfaces. In Proc. 12th Int. ACM SIGACCESS Conf. Computers and accessibility (ASSETS '10). pp. 19–26, ACM, New York, NY, USA.

Harper, S. and Bechhofer, S. (2007) SADIe: structural semantics for accessibility and device independence. ACM Trans. Comput. Hum. Interact., 14, 10 pages.

Hart, S.G. and Staveland, L.E. (1988) Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Adv. Psychol., 52, 139–183.

IMS ACCLIP (2016) http://wiki.cetis.ac.uk/ACCLIP_Briefing

Ivory, M.Y., Mankoff, J. and Le, A. (2003) Using automated tools to improve web site usage by users with diverse abilities. IT Soc., 1, 195–236.

Kane, S.K., Jayant, C., Wobbrock, J.O. and Ladner, R.E. (2009) Freedom to roam: a study of mobile device adoption and accessibility for people with visual and motor disabilities. In Proc. 11th

Int. ACM SIGACCESS Conf. Computers and accessibility, pp. 115–122.

Knutov, E., De Bra, P. and Pechenizkiy, M. (2009) AH 12 years later: a comprehensive survey of adaptive hypermedia methods and techniques. New Rev Hypermedia Multimed., 5, pp. 5–38.

Koutkias, V., Kaklanis, N., Votis, K., Tzovaras, D. and Maglaveras, N. (2016) An integrated semantic framework supporting universal accessibility to ICT. Univ. Access Inform. Soc., 15, 49–62. 0.

Li, L. and Gajos, K.Z. (2014) Adaptive click-and-cross: adapting to both abilities and task improves performance of users with impaired dexterity. In Proc. 19th Int. Conf. Intelligent User Interfaces, pp. 299–304.

Limbourg, Q., Vanderdonckt, J., Michotte, B., Bouillon, L. and López-Jaquero, V. (2004) USIXML: a language supporting multi-path development of user interfaces. EHCI/DS-VIS, 3425, 200–220.

Mirri, S., Salomoni, P., Prandi, C. and Muratori, L.A. (2012) GAPforAPE: an augmented browsing system to improve Web 2.0 accessibility. New Rev Hypermedia Multimed., 18, pp. 205–229.

Mobile Accessibility (2016) https://www.w3.org/WAI/mobile/

Mobile Web Application Best Practices (2016) http://www.w3.org/TR/mwabp/

Mobile Web Best Practices (2016) https://www.w3.org/TR/mobile-bp/

Montague, K., Hanson, V.L. and Cobley, A. (2012) Designing for individuals: usable touch-screen interaction through shared user models. In Proc. 14th Int. ACM SIGACCESS Conf. Computers and accessibility, pp. 151–158.

Mott, M.E., Vatavu, R.D., Kane, S.K. and Wobbrock, J.O. (2016) Smart Touch: Improving Touch Accuracy for People with Motor Impairments with Template Matching. In Proc. 2016 CHI Conf. Human Factors in Computing Systems, pp. 1934–1946.

Nicolau, H., Guerreiro, T., Jorge, J. and Gonçalves, D. (2014) Mobile touchscreen user interfaces: bridging the gap between motor-impaired and able-bodied users. Univ. Access Inform. Soc., 13, 303–313.

OWL (2016) https://www.w3.org/TR/owl-features/

Paterno, F., Santoro, C. and Spano, L.D. (2009) MARIA: a universal, declarative, multiple abstraction-level language for service-oriented applications in ubiquitous environments. ACM Trans. Comput. Hum. Interact., 16, 19 pp.

Pérez, J.E., Arrue, M., Valencia, X. and Moreno, L. (2014) Exploratory study of web navigation strategies for users with physical disabilities. In Proc. 11th Web for All Conf. pp. 4

Pérez, J.E., Valencia, X., Arrue, M. and Abascal, J. (2015) Depicting the keypad mouse interface: exploring pointing behaviors of three input devices. In Proc. 17th Int. ACM SIGACCESS Conf. Computers & Accessibility, pp. 413–414.

Protégé (2016) http://protege.stanford.edu

Richards, J.T. and Hanson, V.L. (2004) Web accessibility: a broader view. In Proc. 13th Int. Conf. World Wide Web, pp. 72–79.

SWRL (2016) https://www.w3.org/Submission/SWRL/

Takagi, H. and Asakawa, C. (2000) Transcoding proxy for nonvisual web access. In Proc. 4th Int. ACM Conf. Assistive technologies, pp. 164–171.

Takagi, H., Asakawa, C., Fukuda, K. and Maeda, J. (2002) Site-wide annotation: reconstructing existing pages to be accessible. In Proc. 5th Int. ACM Conf. Assistive technologies, (pp. 81–88).

Trewin, S., Swart, C. and Pettick, D. (2013) Physical accessibility of touchscreen smartphones. In Proc. 15th Int. ACM SIGACCESS Conf. Computers and Accessibility (ASSETS '13). ACM, New York, NY, USA, Article 19, (8 pages).

Trewin, S., Keates, S. and Moffatt, K. (2006) Developing steady clicks: a method of cursor assistance for people with motor impairments. In Proc. 8th Int. ACM SIGACCESS Conf. Computers and accessibility, pp. 26–33.

Valencia, X., Arrue, M., Pérez, J.E. and Abascal, J. (2013) User individuality management in websites based on WAI-ARIA annotations and ontologies. In Proc. 10th Int. Cross-Disciplinary Conf. Web Accessibility, pp. 10.

Valencia, X., Pérez, J.E., Muñoz, U., Arrue, M. and Abascal, J. (2015) Assisted interaction data analysis of web-based user studies. Hum. Comput. Interact., LNCS 9296, 1–19.

WAI-ARIA (2016) https://www.w3.org/TR/wai-aria/

WCAG 2.0 (2016) https://www.w3.org/TR/WCAG20/

Wobbrock, J.O., Fogarty, J., Liu, S.Y.S., Kimuro, S. and Harada, S. (2009) The angle mouse: target-agnostic dynamic gain adjustment based on angular deviation. In Proc. SIGCHI Conf. Human Factors in Computing Systems, pp. 1401–1410.

Wobbrock, J.O., Myers, B.A. and Kembel, J.A. (2003) EdgeWrite: a stylus-based text entry method designed for high accuracy and stability of motion. In Proc. 16th annual ACM symposium on User interface software and technology, pp. 61–70.

XPATH (2016) https://www.w3.org/TR/xpath/

Yesilada, Y., Brajnik, G. and Harper, S. (2011) Barriers common to mobile and disabled web users. Interact. Comput., 23, 525–542.

## 3.4. Appendix 4

This appendix contains the article entitled “An exploratory study of web adaptation techniques for people with low vision” that was published in the international journal Universal Access in the Information Society (UAIS) in June 2020. This journal reached an impact factor of 1.815 according to Journal Citation Reports in 2019, ranking 10th out of 22 journals (Q2) in the “Computer Science, Cybernetics” category.

In this work we carried out two successive user tests with people with low vision, a first one in order to study the navigation strategies of this group of people and to select the appropriate adaptation techniques to design accessible web interfaces for their needs. In a second user test, the set of adaptation techniques implemented were evaluated by means of an exploratory study with the participation of twelve users with low vision. Other works related to the paper in this appendix were presented in various international conferences, each focused on studying the behaviour of users with visual impairments accessing diverse types of interfaces:

- LOURDES MORENO, XABIER VALENCIA, J. EDUARDO PÉREZ, AND MYRIAM ARRUE, “Exploring the Web navigation strategies of people with low vision,” In *Proceedings of the XIX International Conference on Human Computer Interaction* (Interacción 2018), Palma, Spain, Sep. 12–14, 2018, article no. 13. **[ ⋆ Jesús Lorés Award to the best research work ]**

- J. EDUARDO PÉREZ, MYRIAM ARRUE, MASATOMO KOBAYASHI, HIRONOBU TAKAGI, AND CHIEKO ASAKAWA, “Assessment of semantic taxonomies for blind indoor navigation based on a shopping center use case,” In *Proceedings of the 14th Web for All Conference* (W4A ‘17), Perth, Australia, Apr. 2–4, 2017, article no. 19.

- ARITZ SALA, MYRIAM ARRUE, J. EDUARDO PÉREZ, AND XABIER VALENCIA, "Accessibility-in-use of public e-services: an exploratory study including users with low vision," In *Proceedings of the XX International Conference on Human Computer Interaction* (Interacción '19), Donostia-San Sebastián, Spain, Jun. 25–28, 2019, article no. 15.

- ARITZ SALA, MYRIAM ARRUE, J. EDUARDO PÉREZ, AND SANDRA M. ESPÍN-TELLO, "Measuring complexity of e-government services for people with low vision," In *Proceedings of the 17th International Web for All Conference* (W4A '20), Taipei, Taiwan, Apr. 20–21, 2020, article no. 21. **[ ⋆ Best Communication Paper Award ]**

Universal Access in the Information Society
https://doi.org/10.1007/s10209-020-00727-6

LONG PAPER

# An exploratory study of web adaptation techniques for people with low vision

Lourdes Moreno[1] · Xabier Valencia[2] · J. Eduardo Pérez[2] · Myriam Arrue[2]



**Abstract**
People with low vision may experience accessibility barriers when they interact with the web. The navigation strategies of low-vision users are explored in this article in order to select the appropriate accessibility techniques needed to design web interfaces for their benefit. First, a literature study and an observational study involving low-vision users were carried out. From these studies, a set of adaptation techniques were obtained, which were then evaluated by means of an exploratory study with the participation of twelve users with low vision. The results show that the advantages of some adaptation techniques varied depending on the type of assistive technology used by participants to access the web. Some of the applied adaptation techniques seem turned out to be helpful only for users who utilized screen magnifier software, but not for those using the browser zoom feature. New research hypotheses for a future experimental study have been obtained based on the results of the study presented in this article.

**Keywords** Low vision · User interfaces · Adaptation techniques · Assistive technology · Web accessibility

## 1 Introduction

There are web accessibility barriers which deny the right of people with disabilities to access content on websites, even though equal access is mandatory in most countries [27]. Although web accessibility standards provide resources in order to achieve accessible web pages [45], most approaches to this target are based on complying with the accessibility standards without considering certain individual characteristics of people with disabilities.

✉ Lourdes Moreno
lmoreno@inf.uc3m.es

Xabier Valencia
xabier.valencia@ehu.eus

J. Eduardo Pérez
juaneduardo.perez@ehu.eus

Myriam Arrue
myriam.arrue@ehu.eus

1 Computer Science Department, Universidad Carlos III de Madrid, Madrid, Spain

2 EGOKITUZ: Laboratory of HCI for Special Needs, University of the Basque Country (UPV/EHU), Informatika Fakultatea, Donostia, Spain

The web is far less accessible for people with vision impairments than it is for sighted people. In particular, interacting with the web is often problematic for people with low vision. The number of people with a visual impairment is significant; the World Health Organization (WHO) estimates that there are 246 million people worldwide who have low vision [40].

There are different terms to refer to people with low vision, such as partially sighted or sight impaired. The WHO determines exactly what constitutes low vision and its subsequent categories, basing its classifications on levels of visual acuity and field of vision. "Low vision is a condition caused by eye disease, in which visual acuity is 20/70 or poorer in the better-seeing eye and cannot be corrected or improved with regular eyeglasses" [32].

People with low vision prefer to make use of their residual vision as much as possible [6]. However, they may encounter difficulties in accessing the information presented on web pages. Normally, the main difficulties are related to small font sizes, font colours that make reading even more difficult and background images on web pages that decrease legibility. Additionally, visual clutter, such as multiple columns, is also problematic [8, 13, 31].

People with visual impairments employ assistive technologies such as screen magnifiers or screen readers in order to access the web. People with low vision normally utilize screen magnifiers, applications that increase the size of visual elements by zooming in on the content. Screen magnifiers normally also include colour inversion and cursor enhancement features, among others. On the other hand, screen readers are employed mostly by blind individuals; however, they are also used by numerous people with low vision. These tools transform web content into auditory output. The group of the visually impaired user is very heterogeneous as regards their interaction with the web [39]. The behaviour of the user group which utilizes screen readers is different from those using magnification technologies. Additionally, the behaviour of those who use a combination of both is also different [2, 15].

A large proportion of research regarding web accessibility has been devoted to the topic of screen readers for blind users. Similarly, the majority of the literature has been oriented towards total blindness rather than low vision [18]. Because of this, research opportunities exist for improving low-vision accessibility tools [34].

In this work, a comprehensive literature review of adaptation techniques for people with low vision is presented as well as an analysis of the navigation strategies of people with low vision by means of an observational study. In addition, the resulting adaptation techniques were evaluated with users with low vision, and the results have provided findings for future research.

## 2 First phase. Literature review and Observational study

An exploratory research methodology is used for this study, which is divided into two stages. The first stage includes a literature review (see Sect. 2.1) in conjunction with an observational study of the behaviour of people with low vision in an effort to explore and understand their navigation strategies (see Sect. 2.2). A set of potential adaptation techniques to improve accessibility are obtained from this stage (see Sect. 2.3).

### 2.1 Literature review

In order to analyse the accessibility and navigation strategies of people with low vision when interacting with the web, a review of the related work was carried out.

#### 2.1.1 Accessibility standards

With regard to web accessibility standards, we can find the Web Content Accessibility Guidelines (WCAG) from the World Wide Web Consortium (W3C)'s Web Accessibility Initiative (WAI) [45]. The WCAG 2.0 [41] is internationally recognized as the benchmark for providing accessible web content, being referenced in the regulations of several countries that mandate the accessibility of public websites. Since 2012, the WCAG 2.0 is also considered as an international ISO standard.

The WCAG establishes a broad range of recommendations for making web content more accessible. However, according to the opinions of many accessibility experts, users and developers, the standard has deficiencies and fails to meet web accessibility requirements for people with low vision [42]. Given this fact, the Low Vision Accessibility Task Force (LVTF) was created to draft works concerning the accessibility of people with low vision. One of these documents is the "Accessibility Requirements for People with Low Vision" [43] in which sensitivity, field of vision and colour vision are considered. This document specifies the needs people with low vision have with regard to electronic content, tools and technologies being accessible. Due to this initiative, new guidelines for people with low vision are being included in the W3C Recommendation, WCAG 2.1 [44].

The new success criteria (SC) included in WCAG 2.1 are SC 1.4.12 (Text Spacing) that allows the adjustment of the line height, spacing between paragraphs, the letter spacing and the word spacing, SC 1.4.11 (Non-text Contrast) relating to the graphics contrast, it indicates that the visual presentation needs to have contrast between adjacent coloured user interface components and graphical objects where parts of graphics required to understand the content are conveyed to the user, SC 1.4.10 (Reflow) that tries to prevent horizontal scrolling on left to right or right to left, and SC 1.4.13 (Content on Hover or Focus) that tries to avoid the popups that the user cannot control things, hover and focus, when the default magnification is used.

#### 2.1.2 Adaptation techniques

There is a plethora of works found in the literature which are focused on adapting web interfaces for people with visual impairments. These works have been analysed in order to obtain the adaptation techniques included in Table 1.

These works include personalization techniques and approaches for personalized user interfaces which are used to meet visual needs. Much of the literature is focused either specifically on blind users [6, 22, 35] or people with visual disabilities in general, without distinguishing between the specific needs relating to blind or low-vision users [7, 8, 10, 25, 47]. Additionally, works can be found that deal with the elderly for whom visual impairment is latent [21, 24, 29].

The works in the literature focussing on users with low vision propose basic transformations (larger print, wide spacing and control over font family, specific text and colour) needed by these users [2, 5, 26, 31]. Given the heterogeneity

**Table 1** Information of participants in the exploratory study

| User ID | Gender | Age | Assistive technology |
|---|---|---|---|
| U1 | M | 44 | ZoomText 9.1 |
| U2 | F | 54 | Browser Zoom |
| U3 | F | 52 | ZoomText 10 + Text Reader + Browser Zoom |
| U4 | F | 65 | Browser Zoom |
| U5 | M | 45 | ZoomText 10 + Browser Zoom |
| U6 | F | 41 | Browser Zoom |

within the group of people with low vision [12], it is necessary to indicate that one single method of adapting the presentation of text may not be sufficient to meet the various needs all computer users with low vision have.

In the field of usability, adaptations for people with low vision are focused on improving legibility and readability. These cover high contrast, clean typeface and immediately communicate relevant information at the top of the page [30]. In [19], a colour model is proposed in order to improve the usability of web interfaces for people with low vision.

Concerning technology that supports the adaptation process, the work [8] offers a process that is performed through cascading style sheets. This adaptation process works best on sites that conform to W3C mark-up standards. Additionally, it provides total control of style choices. However, it only allows for the personalization of issues regarding content presentation. In [2], an exploratory study of the improvements in users' performance and overall experience with an adapted version of a website is presented. This study suggests that providing interface-level audio and visual help options that offer users additional information about links may potentially improve performance. Further, the ability to scale the text size and view the site in reverse contrast allows visually impaired users to interact more easily and confidently with the interface.

A conceptual approach can be found in [20], presenting a technique to make web pages accessible for people with low vision and proposing a model for this purpose (WILI—Web Interface for People with Low Vision Issues). This approach automatically replaces the existing display style of a web page with a new skin that adheres to the guidelines established by the Royal National Institute of Blind People. Other approaches present solutions to improve web accessibility for colour-blind individuals. The web page adaptations (including the alteration of text font, size, spacing and colour preservation) are carried out through an automatic process driven by metrics [9, 16, 17].

Along these lines, the authors have developed a system that supports the adaptation processes of user interfaces according to the needs of a user profile using different techniques [35] [37]. In this work, the aforementioned system is used to provide support for those individuals found in the low-vision profile.

#### 2.1.3 Using assistive technology

Some research works that question the support of the assistive technology were found in the literature [48]. Participants preference for accessing information visually (e.g. magnification tool) rather than aurally (e.g. screen readers) was shown in [34] through the results of a contextual research study. Some outcomes of these works indicate that accessibility tools did not provide participants with appropriate support because they had to constantly execute numerous actions in order to see the content properly.

Furthermore, there are works that specifically consider the interaction between users with low vision and screen magnification technology [1, 11, 13, 46]. All these sources indicate that users experience difficulties when interacting with screen magnification tools. The shortcomings of using screen magnifiers are that users only have a partial view of the page they are interacting with and, therefore, may experience loss of context due to the fact that interacting with this software requires frequently moving their field of vision.

Although a significant issue is that the magnification methods often involve the need for horizontal scrolling, the potential benefits of horizontally scrolling text as a reading aid for those with central vision loss, such as macular disease, are indicated in [14]. In [12], assistive technologies were compared with a responsive web design that reduces the need for horizontal scrolling. As a result, participants with low vision found this responsive design to be more usable than using screen magnifiers. A JavaScript system for magnifying web pages on average 60% larger without introducing the most common negative effects of magnification is presented in [3]. This technology improves experienced accessibility for a wide range of people.

Navigation strategies differ greatly in users with visual impairments depending on the type of assistive technology used [15]. Studies have indicated that several problems exist with the use of assistive technologies by people with low vision.

Other works have been published which present alternative assistive technologies besides the traditional ones used to improve accessibility for people with low vision. In [18], haptic-incorporated multimodal user interfaces in relation with low-vision individuals are studied. A user interface software called iBrowse is described in [45]. This tool, rather than utilizing the traditional magnification technique, allows low-vision users to adjust specific style parameters in order to maximize reading efficiency when they access websites. Some of these parameters include font size, font family, font colour, letter spacing, image magnification, hyperlink colours and a screen reader, among others. A magnification interface (called SteeringWheel) that leverages content semantics to preserve local context, in combination with a physical dial, supporting simple rotate and press gestures is presented in [4]. It aims to overcome some of the difficulties arising from the use of screen magnification tools by allowing users to quickly navigate different webpage sections, easily locate desired content, get a quick overview and seamlessly customize the interface.

### 2.2 Observational study

An observational study with six participants with low vision was carried out in order to analyse their navigation strategies [28]. The objective of this study was to detect new navigation strategies applied by people with low vision as well as to confirm those strategies useful for them from the ones gathered in the literature review referred as for people with visual impairments in general.

Experimental sessions were performed with participants from the Spanish Association of People with Low Vision in laboratories located at the University Carlos III of Madrid (UC3M) and in a news agency that employs people with visual disabilities located in Madrid.

#### 2.2.1 Participants

Six participants were involved in this study. All participants belonged to the group of people with low vision and performed the experiment using their usual assistive technology, known generically as screen magnification tools. Table 1 shows information about each participant, including individual gender, age and assistive technologies used during the experimental session.

#### 2.2.2 Stimuli

The stimuli proposed was the Discapnet website (Fig. 1), which focuses on providing information to people with disabilities. The website officially conforms to the AA level of the WCAG accessibility guidelines. This gave us the opportunity to investigate the navigation strategies participants applied and the accessibility barriers participants had to face even on an accessible website and propose adaptation techniques for improving their experience.

#### 2.2.3 Procedure

First, participants were briefed on the purpose of the experiment and signed a consent form before proceeding with the test. Participants were asked to browse the stimuli for 5 min in order to get familiar with the website. Participants were then asked to perform one search task on the website without using a search engine or any other similar facilitator. This


**Fig. 1** The homepage of Discapnet website

was done in order to ensure they were interacting with the web content. The target of the search task was three links away from the homepage and the time for performing the task was limited to 10 min. These browsing activities provided us with enough user interaction data in order to analyse the different types of navigation strategies applied by participants.

Additionally, all the participants' interactions were recorded with a camera located behind them during experimental sessions. These recordings were visually analysed in order to study navigation strategies.

#### 2.2.4 Results

The interaction data collected were analysed and Table 2 shows some of the results. Three of the six participants (U1, U5 and U6) found the target. The time required for completing the task varies among them as does the result of the Lostness formula proposed by Smith [33].

$$L = \sqrt{\left(\frac{N}{S-1}\right)^2 + \left(\frac{R}{N-1}\right)^2}.$$

The lostness formula was used to measure participants' disorientation during the search task, where $S$ is the total

**Table 2** Information about the search task part 1(TCT-task completion time, NPV-total number of pages visited during the task, NPR-number of pages revisited during the task, LNV-Lostness value by user

| User ID | TCT (s) | NPV | NPR | LNV |
|---|---|---|---|---|
| U1 | 309 | 7 | 1 | 0.36 |
| U2 | – | 13 | 4 | – |
| U3 | – | 10 | 4 | – |
| U4 | – | 37 | 15 | – |
| U5 | 101 | 4 | 0 | 0 |
| U6 | 191 | 10 | 3 | 0.57 |

**Table 3** Information about the search task part two (NZM-number of zoom magnification operations, NZD-number of zoom demagnification operations, VS-number of vertical page scrolling actions, HS-number of horizontal page scrolling actions)

| User ID | NZM | NZD | VS | HS |
|---|---|---|---|---|
| U1 | 2 | 1 | 34 | 0 |
| U2 | 2 | 2 | 71 | 4 |
| U3 | 1 | 2 | 56 | 0 |
| U4 | 0 | 0 | 37 | 0 |
| U5 | 0 | 0 | 7 | 4 |
| U6 | 2 | 4 | 43 | 4 |

number of pages, $N$ is the number of different web pages visited and R is the optimal number of page visits. 0 means not lost, from 0 to 0.4 not observably lost, and from 0.5 to 1 the user is lost.

The minimum time for reaching the objective was obtained by U5 who only needed 101 s. The path followed by U5 was also the optimal one (Lostness value is 0). U1 visited fewer pages (7 pages, Lostness value 0.36) than U6. User U4 seemed to be the most disoriented during the task. This user visited 37 webpages with 15 revisited webpages.

Roughly, the observations obtained are as follows. Regarding the webpage layout components, it was observed that the navigation bar component was the most explored element by participants since they used it as a reference point. They encountered some difficulties with decreased legibility such as small font sizes and contrast between font and background colours. Moreover, columns were problematic for users, because they tracked from the end of a line in one column to the beginning of the other. In relation to links, sometimes users did not identify what text was a link, and they had trouble accessing them.

Scrolling actions performed by participants were annotated (see Table 3). Vertical scrolling did not present any difficulties for them and was performed by all participants except U5 during the task (average number of vertical scrolling actions is 41.33). The common strategy for performing this action was to use the mouse wheel. Only U4 used scrollbar arrow buttons for scrolling (37 vertical scrolling actions). This strategy required moving the field of vision for locating these buttons. This action may cause loss of context. In fact, this participant seemed to be the most disoriented revisiting 15 web pages despite not reaching the target.

On the other hand, horizontal scrolling required a more complex sequence of actions from them: moving the field of vision to the bottom of the browser, clicking and dragging the scroll bar and going back to the content. Horizontal scrolling occurred depending on the magnification applied to the webpage. In the analysis of the recorded videos, we detected that all participants except U1 were required to perform at least one horizontal scrolling action if they wanted to access to the content of the entire webpage. However, the horizontal scrolling action was only performed by three participants (U2, U5 and U6). The number of horizontal scrolling actions was considerably lower than the vertical ones.

Two participants (U2 and U3) did access the web page that contained the target link during the task although they did not manage to locate it. Analysing the videos, we found that the link was visible on the screen for those users. However, participants were focused on the main content of the webpage at the centre of the screen, while the link was out of their sight on the left side of the screen.

We analysed the objective of these magnification/demagnification actions. This analysis revealed the following observations that show that a magnification-level increment was performed before the following actions:

- Exploring vertical/horizontal navigation bars on web pages (U1, U2, U6).
- Clicking on a link located on the vertical navigation bar (U1, U2, U3, U6).
- Using the vertical scroll bar (U1, U2, U6).

On the other hand, demagnifications were performed before the following actions:

- Exploring the content of the web page (U1, U2, U3, U6).
- Clicking on the browser's back button (U1, U6).
- Exploring the horizontal navigation bar (U3).
- Clicking on a link located on the horizontal navigation bar (U2).
- Using the horizontal scroll bar (U6).

The common strategy applied for zoom magnification/demagnification actions was using the "ctrl" key together with the mouse wheel.

### 2.2.5 Discussion

Results of the observational study reveal that there is no single solution that solves the problems faced by people with low vision and that there are a variety of strategies which depend on who or what the problem concerns (users, assistive technologies, designers, etc.).

In this study, it was observed that users who use the screen magnifier technology can access web content, albeit not satisfactorily. They experience hindrances such as the number of actions required to perform horizontal scrolling. Furthermore, users only have a partial view of the page and they experience loss of context due to the fact that interacting with screen magnification requires frequently moving their field of vision.

As a possible solution for this issue, we propose adaptations in the web interface so that the adapted web interface can improve the experience of users with low vision when utilizing assistive technology: adaptations such as techniques for minimizing the number of magnification and demagnification actions needed for accessing information on the Web or for avoiding the use of horizontal scrolling.

According to the results from this study, a proposal of a set of accessibility techniques to design web user interfaces for people with low vision are presented below.

- *Technique (T) (1)* Including the navigation bar component near the top of the page improves orientation, as it can be used as a reference point.

In relation to minimizing the magnification/demagnification operations:

- *T(2)* Presenting important information with a large font size so fewer zoom magnification operations are required to explore the content.
- *T(3)* Presenting a coherent structure to the entire website and organizing text in small blocks can help users to focus on the target content before performing magnification/demagnification actions.
- *T(4)* Presenting important information close to the centre of the screen could minimize movements in the field of vision that cause loss of context. It would improve the performance of users who were unable to reach the target as it was placed out of their sight.
- *T(5)* Adding a visible hot area around links: The hot area is visualized when the cursor is on the link. By applying this technique, magnification/demagnification actions for exploring links could be minimized.
- *T(6)* Using presentation properties that distinguish the links, for example using a specific colour and underlining for links. This enables links to be located on the screen effortlessly. In addition, navigation bar components and links can be easily identified without requiring a series of magnification actions.

With regard to avoiding the horizontal scroll:

- *T(7)* Linearize the page, providing a narrow-page presentation that minimizes content in the borders, increasing vertical scroll and avoiding horizontal scroll which requires a set of complex actions to be performed and provides a one-column page layout as well.

Also, from general observations obtained such as difficulties with contrast with font and background colours, the following technique is proposed:

- *T(8)* Applying enough information/background contrast.

From the results of the observational study, a set of techniques was defined which can improve accessibility and user experience for people with low vision when interacting with the web.

Some of the resulting techniques were found in the literature review as it is shown in Table 4. However, some of them were not specifically designed for people with low vision. This indicates that there are techniques for other user groups (older people [21, 29], individuals with visual impairments [26, 31, 44]) that could improve the accessibility for people with low vision. Table 4 shows a correspondence between the techniques resulting from the observational study and these references.

## 3 Second phase. Exploratory Study

The set of adaptations (see Table 4) obtained in the observational study have been developed and integrated in a previously developed automated web adaptation system [36, 38]. This system generates an adapted web interface from any previously annotated web interface. Study participants explored both the original web interface and the adapted version. The user interaction data were gathered in the experimental sessions and analysed for both versions (original and adapted versions).

Experimental sessions were performed with participants from the Spanish Association of People with Low Vision in laboratories located at the University Carlos III of Madrid (UC3M), the Computer Science School at the University of the Basque Country (EHU-UPV) and in a news agency that employs people with visual disabilities located in Madrid.

### 3.1 Participants

Twelve participants with low vision took part in this study. All reported having a high level of web navigation expertise. Table 5 presents the users' data.

**Table 4** Techniques obtained from the observational study

| Id Tech | Name | Description | Observational study | Literature review |
|---|---|---|---|---|
| 1 | Sorting Elements | 1.1. Sorting links: placing the strongest recommendations at the top<br>1.2. Sorting content elements: place important content areas near the top of the page<br>1.3. Using an effective format of white spaces and presenting text in small blocks | T(1)<br>T(3) | [21, 43, 44] |
| 2 | Text Adaptation | 2.1. Presenting important information in 12 to 14-point font size<br>2.2 Using Sans Serif fonts<br>2.3 Increasing line spacing<br>2.4 Using left justification<br>2.5 Using underlining for links | T(2)<br>T(6) | [2, 8, 21, 26, 29, 43, 44] |
| 3 | Layout Adaptation | 3.1. Presenting important information as close to the centre of the screen as possible<br>3.2. Using narrow-page presentation. Single-column pages<br>3.3. Using the page linearization | T(4)<br>T(7) | [2, 21, 31, 43, 44] |
| 4 | Altered Elements | 4.1. Adding hot area around a hyperlink | T(5) | [2, 21] |
| 5 | Contrast Colours | 5.1. Maximizing the information/background contrast in critical content areas<br>5.2. Presenting information in reds, oranges and yellows and use blues, greens and violets for background<br>5.3. Using a light-colour text on dark coloured background (negative polarity) | T(5)<br>T(8) | [2, 5, 21, 26, 43, 44] |

**Table 5** Information about the participants of the study: gender, age, their experience with the experiment website (Wex) in years (+7 means more than 7 years and 1–3 means between 1 and 3 years of experience), the AT and configuration used (Bz-browser zoom, Sm-screen magnification software, the screen resolution (Sr) and the operating system used (OS)

| ID | Gender | Age | Wex | Bz | Sm | Sr | OS |
|---|---|---|---|---|---|---|---|
| P1 | F | 39 | +7 | X | | 1600×900 | MS Windows |
| P2 | F | 50 | +7 | X | | 1600×900 | MS Windows |
| P3 | M | 37 | +7 | X | | 1600×900 | MS Windows |
| P4 | M | 37 | +7 | | X | 1440×900 | Mac OS |
| P5 | M | 46 | +7 | | X | 800×600 | MS Windows |
| P6 | F | 55 | +7 | X | | 1366×768 | MS Windows |
| P7 | F | 43 | +7 | X | | 1067×600 | MS Windows |
| P8 | M | 45 | +7 | | X | 1440×900 | Mac OS |
| P9 | M | 23 | +7 | | X | 2880×1800 | Mac OS |
| P10 | F | 19 | +7 | X | | 1440×900 | MS Windows |
| P11 | F | 24 | +7 | X | | 1440×900 | MS Windows |
| P12 | M | 72 | 1–3 | X | | 1440×900 | MS Windows |

## 3.2 Instrument

Two systems were used in this study. These systems were developed with the objective of automatically creating the adapted version of the web interfaces and conducting experimental sessions with participants and gathering user interaction data.

An automated web interface adaptation system was applied to obtain the adapted version of the web interfaces [36, 38]. This tool has been developed following a methodological approach based on the extension of the WAI-ARIA language. WAI-ARIA is a technical specification of the W3C that provides a framework to improve the accessibility and interoperability of web content and applications by defining a way of making web content and web applications more accessible to people with disabilities. WAI-ARIA specification provides an ontology of roles, states and properties that define the elements of a web interface. In this research work, an extension of the ontology was used by incorporating new roles and properties in order to be able to model user low-vision characteristics, adaptation techniques and the relationships between them for the low-vision profile.

The applied adaptation techniques were deduced from rules that associate the characteristics of the users with the adaptation techniques for specific interactive elements located on the website.

In addition, the *RemoTest* platform to design and perform user tests [1, 37] was used to conduct the experimental sessions and gather the interaction data. A secure server was in charge of running the remote modules of the *RemoTest* platform, both to gather interaction data from each participant and to provide them with the search tasks as well as the questionnaires that had to be completed in the session. The "participant module" of the *RemoTest* platform was locally installed in the computers used for the study. This module was in charge of communicating with the remote modules of *RemoTest* to identify each participant, present them the proposed tasks, gather and send to a remote server all user interaction data and present and store the data introduced in the online questionnaire after the sessions. The gathered interaction data include spatial data about selected links as well as additional interactions (keystroke, extra clicks, page scroll, cursor trajectory, etc.) together with browser and experiment-related events (page load, start and end of tasks, etc.). All the data are stored in a *MongoDB* database placed in the remote secure server.

The functionality of both systems was accessed by specifically developed add-ons which were installed in the *Mozilla Firefox* web browser in the computers used for this study.

### 3.3 Stimuli

The stimuli presented to participants were two different websites based on the Discapnet website: the adapted version and the original version (see Fig. 2). All the participants were familiar with this website as it provides information and news about legislation, rights and grants for people with disabilities. However, the website is huge and the content is frequently updated.

### 3.4 Procedure

First, participants were informed about the study and required to sign a consent form. Each participant was asked to complete search tasks in both scenarios: the original version (Scenario 1) and the adapted version (Scenario 2).

Twelve tasks were completed by each participant in this study, six in each scenario. Once the participants were informed and had given their consent, they were presented with one scenario in which they were asked to complete six search tasks. Subsequently, the other scenario was evaluated. The order of which scenario was evaluated first was counterbalanced. The search tasks were located at different depths (number of links from the homepage). In this way, the target of the tasks was one, two or three links away from the homepage. The study was designed to present two tasks at each depth in one scenario and another two in the other scenario. Both the tasks assigned to each scenario as well as the order in which the six tasks were presented to participants were randomized.

The time limit for each task was 4 min and the maximum length of the experimental session was around 1 h. A message was displayed when participants either found the target or the time allotted had expired. The next task was then presented and the homepage of the corresponding scenario was shown. Once the participants had finished all of the tasks in the scenario, an online questionnaire was presented to participants.


**Fig. 2** Adapted web interface (left) and original web interface (right) of the Discapnet homepage

The sessions were recorded via a camera located behind the user in order to obtain information about their interaction.

### 3.5 Measurements

The parameters measured were efficiency, effectiveness and satisfaction.

- Efficiency and effectiveness:

– To measure efficiency, the following parameters were calculated for tasks 1–12 during both scenarios: TCT (task completion time) and LNS (lostness values).
– To measure effectiveness, the following parameters were calculated for each task completion rate in each scenario: NTC (number of tasks completed), TPT (total pixels travelled with cursor by page), Scroll X (total scroll performed horizontally by page), Scroll Y (total scroll performed vertically by page), NZM (total number of zoom magnification operations), NZD (total number of zoom demagnification operations).

- Satisfaction: in order to measure this, a questionnaire based on the ASQ [23] was filled out by the participants. The parameters measured were ease, time and pleasantness.

### 3.6 Results

*Efficiency and Effectiveness* The results concerning efficiency and effectiveness obtained after analysing the interaction from the recorded video and the data gathered from each participant by the Remotest tool are shown in the following figures. Figure 3 shows the efficiency measures in both scenarios. A slight improvement can be appreciated in Scenario 2 in relation to Scenario 1, but with a high dispersion of the data.

Analysing the data, it is observed that most participants who obtained benefits in Scenario 2 used specific screen magnification software. Measurements grouped by the assistive technology used by participants (screen magnification software and browser zoom) were calculated in order to compare the results. Boxplots with the differently grouped measurements can be seen in Figs. 4, 5 and 6.

Participants using the browser zoom obtained better results with Scenario 1: more finished tasks, less time, on

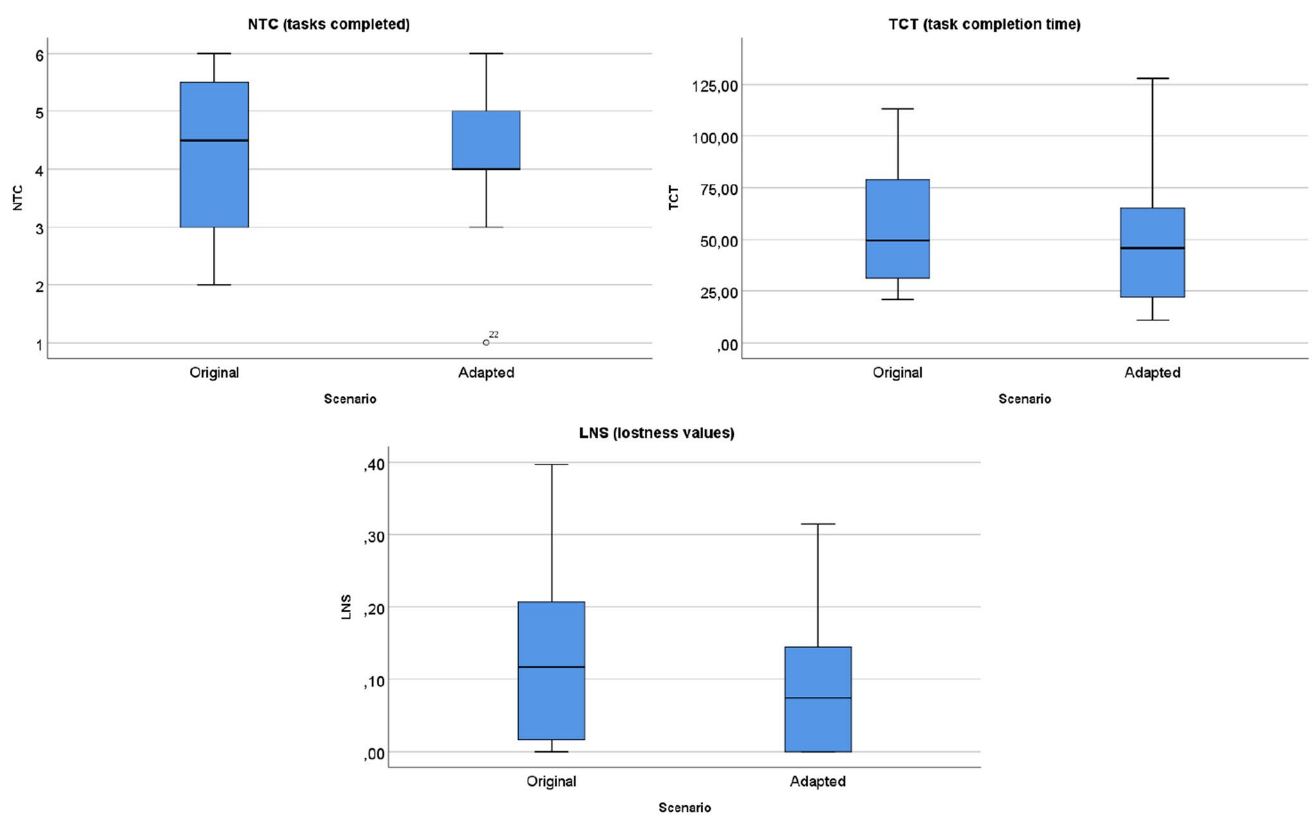


**Fig. 3** Boxplot graphs of NTC (number of tasks completed), TCT (task completion time) and LNS (lostness values)

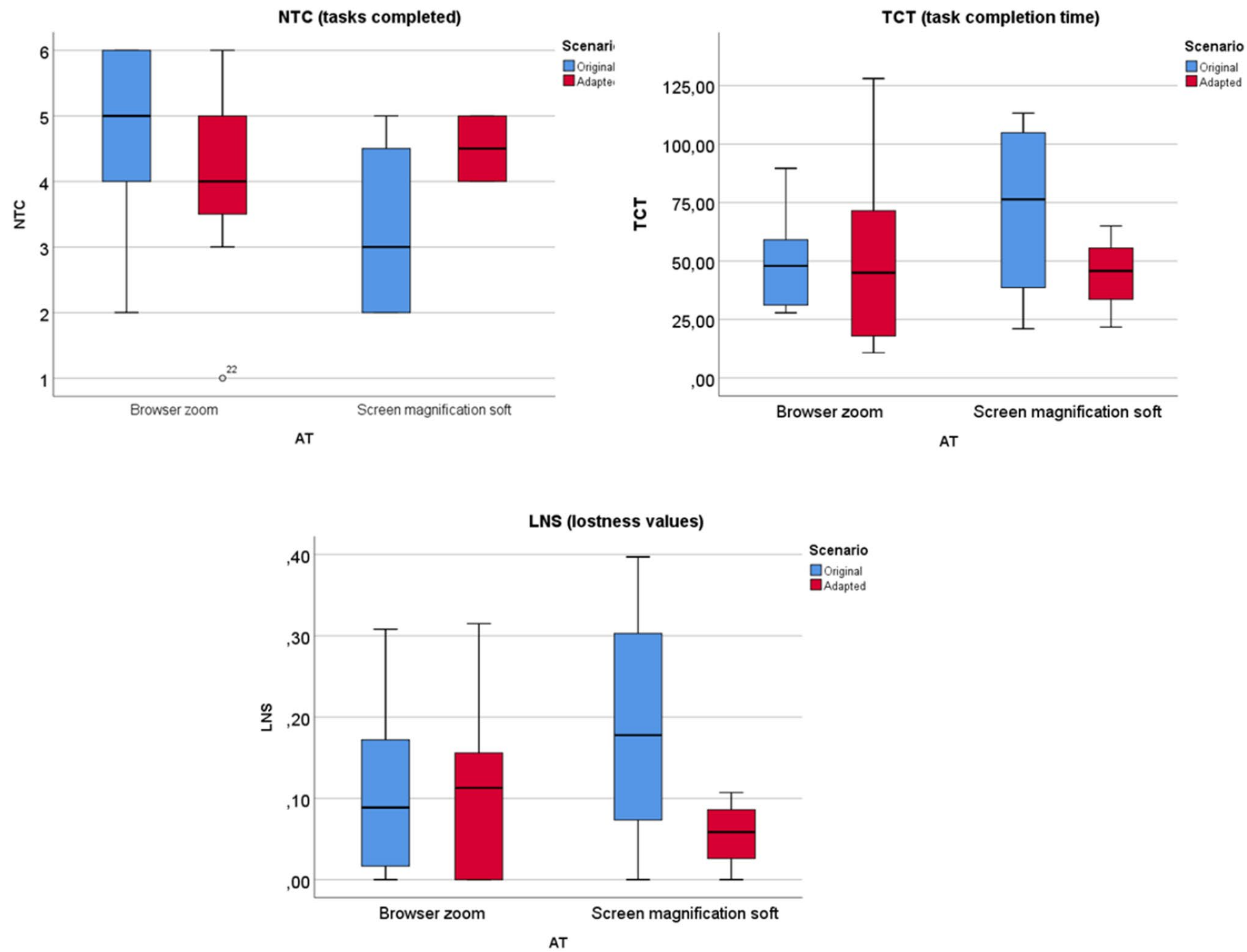


**Fig. 4** Boxplot graphs of NTC (number of tasks completed), TCT (task completion time) and LNS (lostness values) grouped by assistive technology used (Browser Zoom and Screen Magnifier)

average, to complete tasks and better lostness values (see Fig. 4).

The average cursor trajectory distance by page was lower in Scenario 2 but it required them to perform more vertical scrolls due to the narrower page. This narrow-style page was mainly intended to avoid horizontal scroll. However, browser zoom users did not experience significant horizontal scrolling, since they tended to magnify the web page until just before the point in which the horizontal scroll appears. Therefore, added vertical scrolling in Scenario 2 hindered navigation without providing the benefit of avoiding horizontal scrolling for this group of users (see Fig. 5).

As for the zoom changes, people using browser zoom did not tend to perform a noticeable number of zoom changes. They usually enlarged the web page once in order to make it as large as possible without breaking the web page layout. Regarding Scenario 2, fewer zoom changes were performed since the size of the web elements were already sufficiently sized (see Fig. 6).

As far as using screen magnification software as assistive technology was concerned, Scenario 2 was helpful for participants who used it. All of them found almost four targets with Scenario 2, while, in Scenario 1, some participants found only two targets. Moreover, they needed less time to complete the tasks with Scenario 2 and also had better lostness values. Conversely to participants using the browser zoom, they suffered the horizontal scroll in Scenario 1, while it was not noticeably present in Scenario 2. This, in conjunction with the narrower page, allowed them to reduce the cursor trajectory distance needed to explore the page. Even if the vertical scrolling values were higher in Scenario 2, the benefits of avoiding horizontal scrolling and reducing the cursor trajectory counteracted the difficulties introduced by vertical scrolling. As to the magnification

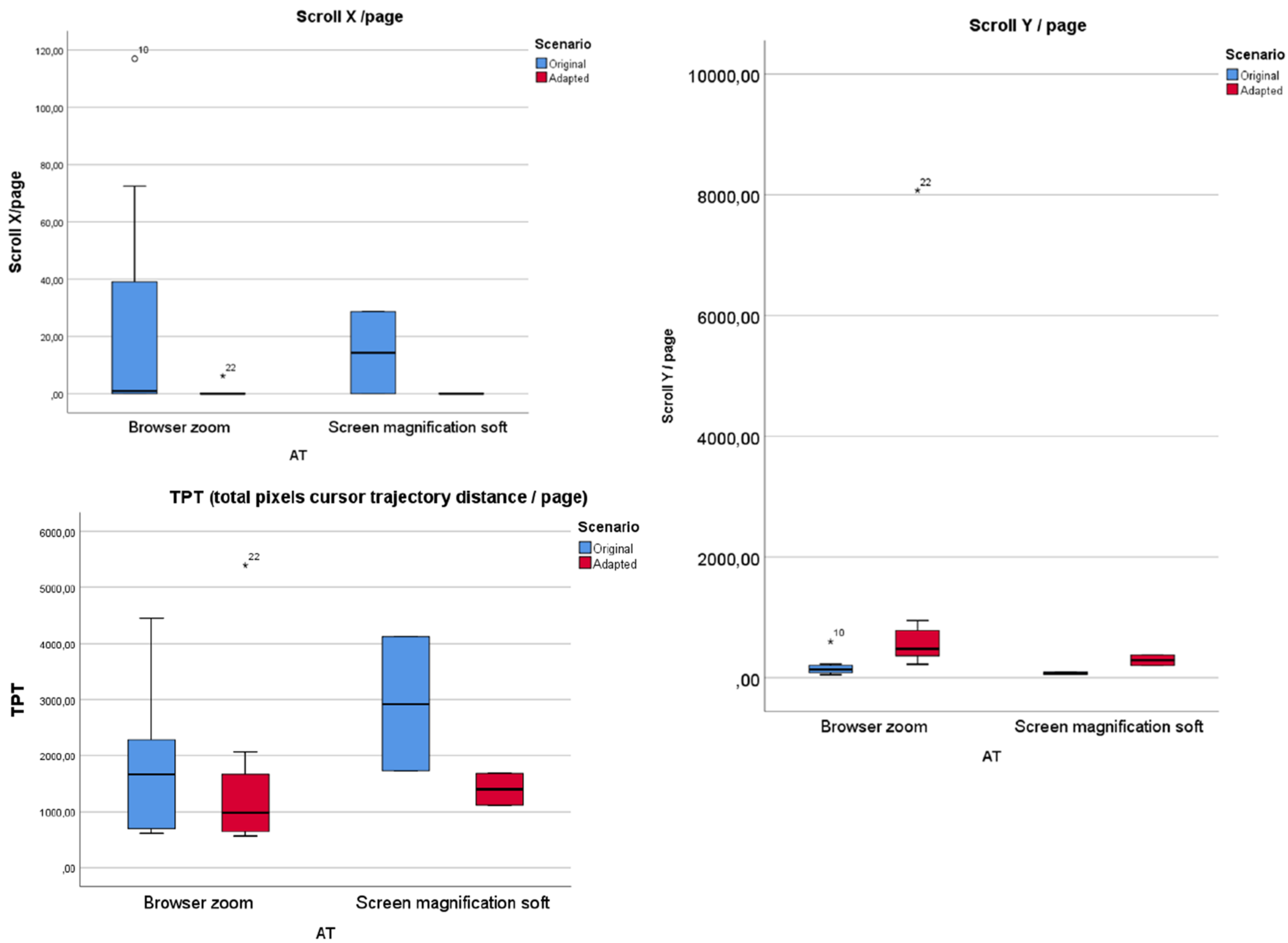


**Fig. 5** Boxplot graphs of TPT (total pixels travelled with cursor by page), Scroll X (total scroll performed horizontally by page), Scroll Y (total scroll performed vertically by page) grouped by assistive technology used (Browser Zoom and Screen Magnifier)

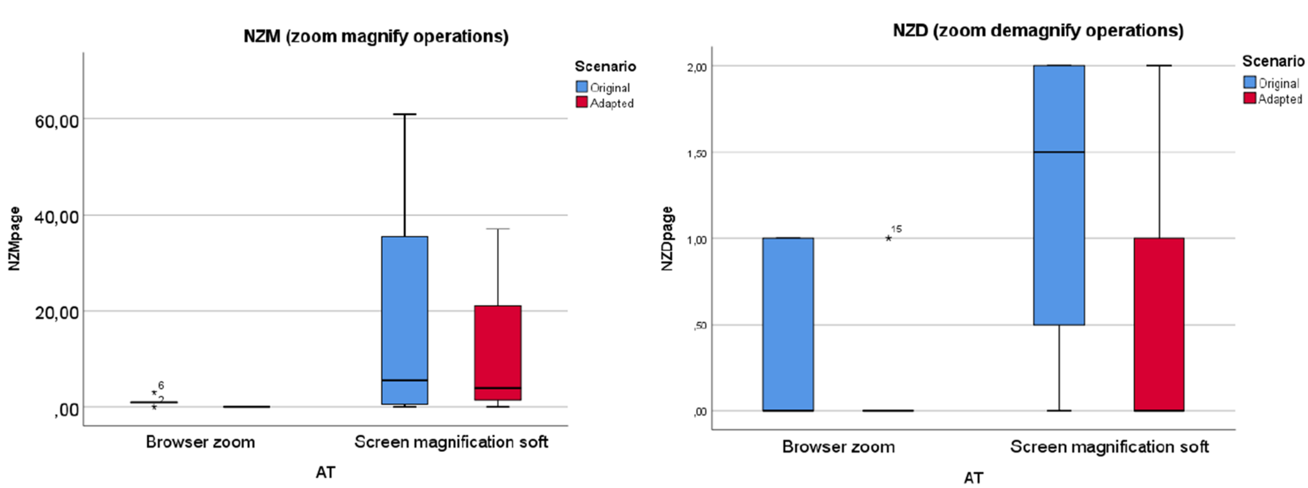


**Fig. 6** Boxplot graphs of NZM (total number of zoom magnification operations), NZD (total number of zoom demagnification operations) grouped by assistive technology used (Browser Zoom and Screen Magnifier)

and demagnification operation changes, fewer zoom changes were made in Scenario 2.

*Satisfaction* Central tendency measurements for the three features (ease, time, pleasantness) asked about during the post-scenario questionnaires were quite low in both cases (see Table 6).

Participants using screen magnifiers perceived Scenario 2 to be easier. There is a noteworthy difference between the average values and the standard deviation is smaller. On the other hand, participants using the browser zoom scored higher values for Scenario 1.

Regarding the perceived time needed to complete tasks, participants using the screen magnifier scored Scenario 2 with higher values as opposed to participants who used the browser zoom. Lastly, as regards the pleasantness of each scenario, screen magnifier users gave notably higher scores to Scenario 2. Conversely, participants using the browser zoom preferred Scenario 1, but, in this case, the difference between each scenario was small.

### 3.7 Discussion

Analysis of the different results revealed that some of the adaptation techniques considered in this work might be beneficial for the participants using screen magnifier software. Contrarily, these same techniques were not found to be advantageous for others using the browser zoom. While participants using specific magnification software improved their experience, users utilizing the browser zoom found navigating the original version easier. They seemed to take advantage of the visual information (structure, colours and icons) that allowed them to browse faster.

For instance, the navigation behaviour analysis for the participants using specific magnification software revealed that the total distance the cursor travelled over the web pages during the session significantly decreased for some participants. In the videos, we observed that they had to constantly move the cursor in the original version in order to place the software window over the part of the web page they wanted to explore. These cursor movements significantly decreased in the adapted version due to the modified narrow-page presentation. This leads us to believe that these participants may have been more comfortable during their interaction with the adapted version and the values obtained by satisfaction questionnaires (ease and pleasantness) reflect this.

Thus, the techniques “3.1. Presenting important information as close to the centre of the screen as possible”, “3.2. Using narrow-page presentation. Single-column pages” and “3.3. Using page linearization” offered benefits to the participants using specific screen magnification software. However, benefits were not observed for participants using the browser zoom as they were required to apply more vertical scroll due to the narrower page.

Participants who used screen magnifier software also benefited from the adaptation technique “1.1. Sorting links: placing the strongest recommendations at the top”. Users with low vision applied magnification increment/decrement operations to detect and explore vertical and horizontal navigation bars. They located these navigation bars and used them as reference points for their navigation, returning to them after a loss of context. The adapted version required fewer zoom-level change actions. We believe that placing all of the navigational elements in a single vertical navigation bar located at the top of the page influenced these results. In the adapted version, users utilizing screen magnifier software could come back to it by moving the software window directly up to the top of the page. In addition, increasing the size of the web elements allowed some participants to dispense with the screen magnifier software during the session with the adapted version.

In relation to the techniques that minimize the zoom changes (“1.1. Sorting links: placing the strongest recommendations at the top”; “1.3. Using an effective format of white spaces and presenting text in small blocks”; “2.1. Presenting important information in 12–14-point font size”; “2.2 Using Sans Serif fonts”; “2.3 Increasing line spacing”; “4.1. Adding hot area around a hyperlink” and “2.5 Using underlining for links”), in the adapted version both groups of participants performed fewer zoom changes since the size of the web elements was already sufficient. However, it must be noted that people who used the browser zoom did not tend

**Table 6** Measures of central tendency and dispersion for user satisfaction features of both scenarios

| | Ease | | Time | | Pleasantness | |
|---|---|---|---|---|---|---|
| | Scenario 1 | Scenario 2 | Scenario 1 | Scenario 2 | Scenario 1 | Scenario 2 |
| Browser zoom | | | | | | |
| □ | 5.15 | 4.29 | 4.86 | 4.43 | 4.71 | 4.43 |
| $\sigma$ | 1.12 | 1.91 | 1.55 | 1.92 | 1.75 | 1.99 |
| Screen magnifier | | | | | | |
| □ | 2.8 | 3.8 | 2.2 | 2.8 | 3.4 | 4.8 |
| $\sigma$ | 1.60 | 1.17 | 0.75 | 1.60 | 1.85 | 2.23 |

to perform many zoom changes with either the original or the adapted version.

The remainder of the applied techniques, such as those related to the contrast colour and using left justification, seemed to satisfy both groups. So, it seems that these techniques are not dependent on the assistive technology used.

From these findings, we can define the following hypotheses to validate in a future experimental study:

**Hypothesis 1** The adaptation techniques related to minimizing the magnification/demagnification and the horizontal scroll operations are beneficial to people with low vision who use specific screen magnification software, but not to those using the browser zoom, or, in other words, there is a dependency on which type of assistive technology is used and the effectiveness of this type of techniques.

**Hypothesis 2** The techniques related to contrast with font and background colours benefit all low-vision users, or, in other words, there is not a dependency on the type of assistive technology used and this type of techniques is effective for the majority of people with low vision.

## 4 Study limitations

Studies with people with disabilities such as the one presented in this article enclose many difficulties in recruiting participants for experimental studies. Twelve users have participated in this study. This small sample has not allowed obtaining representative results or evidence with statistical significance; however, suspicions and enrichment of qualitative information have been obtained as a result of the observational study. These suspicions have guided us to define two research hypotheses that need to be validated through a future experimental study.

## 5 Conclusions

People with low vision experience accessibility barriers when they interact with the web, despite using their assistive technology. In order to address this issue, an observational study was conducted for discovering the navigation strategies of people with low vision. With the knowledge obtained, adaptation techniques that optimize web interfaces according to these navigation strategies were defined.

The results obtained in the study have provided us with enriched information for selecting techniques beneficial for this group of users. The application of these accessibility techniques may improve the experience of low-vision users by minimizing the number of magnification/demagnification actions needed. This could also make it easier to locate navigation components and the main content of web pages as well as decreasing the number of complex actions such as horizontal scrolling.

The techniques have been included in an automated adaptation system. The system, based on a web interface, generates an adapted web interface. This adapted web interface was evaluated by an exploratory study carried out with twelve people with low vision to assess the efficiency, effectiveness and satisfaction of the approach. The outcome indicates that some adaptation techniques proposed and tested are beneficial for users who utilize screen magnifier software, but not for users who use the browser zoom feature.

Research hypotheses of a future experimental study were obtained based from results of the exploratory study shown in this article. Future work will be focused on fine-tuning the user profile devoted to people with low vision depending on assistive technology used during the web navigation, and studying possible benefits of other personalization techniques.

**Acknowledgments** The authors would like to thank all study participants for their collaboration on the experiment. J. Eduardo Pérez held a Ph.D. scholarship from the University of the Basque Country UPV/EHU (grant PIF13/284) when the study was conducted. Some of the authors are members of the EGOKITUZ/ADIAN research team, supported by the Basque Government, Department of Education, Universities and Research under grant (IT980-16). This work was partially funded by the Ministry of Economy and Competitiveness of the Spanish Government by the European Regional Development Fund – ERDF (eGovernAbility, grant TIN2014-52665-C2, MINECO/FEDER), PhysComp project (TIN2017-85409-P) and by the DeepEMR project (TIN2017-87548-C2-1-R).

## References

1. Arrue, M., Valencia X., Pérez, J.E., Moreno L., Abascal J.: Inclusive web empirical studies in remote and in situ settings: a user evaluation of the RemoTest platform, International Journal of Human-Computer Interaction (in press) (2018)
2. Bergel, M., Chadwick-Dias, A., LeDoux, L., & Tullis, T. (2005). Web accessibility for the low vision user. In Usability Professionals Association Annual Conference, Montreal, Canada
3. Bigham, J.P.: Making the web easier to see with opportunistic accessibility improvement. In: Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST ‘14). ACM, New York, NY, USA, pp 117–122, 2014
4. Billah, S.M., Ashok, V., Porter, D., & Ramakrishnan, I.V.: SteeringWheel: a locality-preserving magnification interface for low vision web browsing. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI ‘18). ACM, New York, NY, USA, Paper 20, 13 pages, 2018
5. Bonavero, Y., Huchard, M., Meynard, M.: Reconciling user and designer preferences in adapting Web pages for people with low vision. In: Proceedings of the 12th Web for All Conference – W4A ‘15 (p. 10). Florence, Italy, 2015
6. Borodin, Y., Bigham, J.P., Dausch, G., Ramakrishnan, I.V. More than meets the eye: a survey of screen-reader browsing strategies. In Proceedings of the 2010 International Cross Disciplinary

Conference on Web Accessibility—W4A ‘10 (p. 13). Raleigh, North Carolina, USA, 2010
7. Correani, F., Leporini, B., Paternò, F.: Supporting Web Usability for Vision Impaired Users. In: User-Centered Interaction Paradigms for Universal Access in the Information Society, pp 242-253 (2004))
8. Dick, W.E.: Using cascading style sheets to accommodate Websites for individuals with low vision. ACM SIGACCESS Access. Comput. **84**, 13–19 (2006)
9. Foti, A., & Santucci, G. (2009). Increasing Web accessibility through an assisted color specification interface for colorblind people. Interaction Design and Architecture(s) Journal – IxD&A, 5, 41-48
10. Goble, C., Harper, S., & Stevens, R.: The travails of visually impaired Web travellers. In Proceedings of the eleventh ACM on Hypertext and hypermedia—HYPERTEXT ‘00 (pp. 1-10). San Antonio, Texas, USA, 2000
11. Gowases, T., Bednarik, R., & Tukiainen, M.: Text highlighting improves user experience for reading with magnified displays. In CHI’11 Extended Abstracts on Human Factors in Computing Systems—CHI EA ‘11 (pp. 1891-1896). Vancouver, BC, Canada, 2011
12. Hallett, E.C., Arnsdorff, B., Sweet, J., Roberts, Z., Dick, W., Jewett, T., & Vu, K.P.L.: The Usability of Magnification Methods: A Comparative Study between Screen Magnifiers and Responsive Web Design. In International Conference on Human Interface and the Management of Information – HIMI ‘15 (pp. 181–189). Los Angeles, California, USA, 2015
13. Harland, S., Legge, G.E., Luebker, A.: Psychophysics of reading XVII Low-vision performance with four types of electronically magnified text. Optomet. Vis. Sci. **75**(3), 183–190 (1998)
14. Harvey, H., Walker, R.: Reading with peripheral vision: a comparison of reading dynamic scrolling and static text with a simulated central scotoma. Vis. Res. **98**, 54–60 (2014)
15. Jacko, J.A., Barreto, A.B., Marmet, G.J., Chu, J.Y., Bautsch, H.S., Scott, I.U., & Rosa Jr, R.H.: Low vision: the role of visual acuity in the efficiency of cursor movement. In: Proceedings of the fourth International ACM Conference on Assistive technologies—ASSETS ‘00 (pp. 1–8). Arlington, Virginia, USA, 2000
16. Jefferson, L., Harvey, R.: Accommodating color blind computer users. In: Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility—ASSETS ‘06 (pp. 40–47). Portland, Oregon, USA
17. Jenny, B., Kelso, N.V.: Color design for the color vision impaired. Cartogr. Perspect. **58**, 61–67 (2007)
18. Kim, H.N., Smith-Jackson, T.L., Nam, C.S.: Elicitation of haptic user interface needs of people with low vision. Int. J. Hum. Comput. Interact. **29**(7), 488–500 (2013)
19. Kulpa, C.C., Teixeira, F.G., Da Silva, R.P.: A color model in the usability of computer interface applied to users with low vision. In: International Conference of Design, User Experience, and Usability—DUXU ‘13 (pp. 330-339). Las Vegas, NV, USA, 2013
20. Kuppusamy, K.S., Francis L.M., & Aghila G. (2012). WILI—‘Web Interface for people with low vision issues. Int. J. Comput. Sci. Appl. (IJCSA) **2**(2)
21. Kurniawan, S.H., King, A., Evans, D.G., Blenkhorn, P.L.: Personalising Web page presentation for older people. Interact. Comput. **18**(3), 457–477 (2006)
22. Lazar, J., Allen, A., Kleinman, J., Malarkey, C.: What frustrates screen reader users on the Web: a study of 100 blind users. Int. J. Hum. Comput. Interact. **22**(3), 247–269 (2007)
23. Lewis, J.R.: Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ. ACM Sigchi Bull. **23**(1), 78–81 (1991)
24. Leonard, V.K., Jacko, J.A., Pizzimenti, J.J.: An exploratory investigation of handheld computer interaction for older adults with visual impairments. In: Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility—ASSETS ‘05 (pp. 12-19). Baltimore, MD, USA, 2005
25. Lunn, D., Harper, S., Bechhofer, S.: Identifying behavioral strategies of visually impaired users to improve access to Web content. ACM Trans. Access. Comput. (TACCESS) **3**(4), 13 (2011)
26. Mirri, S., Prandi, C., Salomoni, P.: Experiential adaptation to provide user-centered Web content personalization. In Proceedings of the 6th International Conference on Advances in Human oriented and Personalized Mechanisms, Technologies, and Services – CENTRIC ‘13 (pp. 31-36). Venice, Italy, 2013
27. Moreno, L., Martínez, P.: The harmonization of accessibility standards for public policies. IEEE Comput. **52**(7), 57–66 (2019)
28. Moreno, L., Valencia, X, Pérez, J. E and Arrue, M.: Exploring the Web navigation strategies of people with low vision. In Proceedings of the XIX International Conference on Human Computer Interaction (Interacción 2018). ACM, New York, NY, USA, 2018
29. National Institute on Aging: Making your Website senior friendly: A checklist. http://www.nlm.nih.gov/pubs/checklist.pdf, (2001). Accessed Dec 2019
30. Nielsen, J.: Legibility, Readability, and Comprehension: Making Users Read Your Words. https://www.nngroup.com/articles/legibility-readability-comprehension/, (2015). Accessed Dec 2019
31. Richards, J.T., Hanson, V.L.: Web accessibility: a broader view. In Proceedings of the 13th international conference on World Wide Web – WWW ‘04 (pp. 72-79). New York, USA, 2004
32. Scheiman, M., Scheiman, M., Whittaker, S.: Low Vision Rehabilitation: A Practical Guide for Occupational Therapists. Slack Incorporated, Thorofare (2007)
33. Smith, P.A.: Towards a practical measure of hypertext usability. Interact. Comput. **8**(4), 365–381 (1996)
34. Szpiro, S.F.A., Hashash, S., Zhao, Y., Azenkot, S.: How People with Low Vision Access Computing Devices: Understanding Challenges and Opportunities. In Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility – ASSETS ‘16 (pp. 171-180). Reno, Nevada, USA, 2016
35. Takagi, H., Saito, S., Fukuda, K., Asakawa, C.: Analysis of navigability of Web applications for improving blind usability. ACM Trans. Comput. Hum. Interact. (TOCHI) **14**(3), 13 (2007)
36. Valencia, X., Arrue, M., Pérez, J.E., Abascal, J.: User individuality management in websites based on WAI-ARIA annotations and ontologies. In Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility (W4A ‘13). ACM, New York, NY, USA,, Article 29, 10 pages, 2013
37. Valencia, X., Pérez, J.E., Muñoz, U., Arrue, M., Abascal, J.: Assisted interaction data analysis of web-based user studies. INTERACT **1**, 1–19 (2015)
38. Valencia, X., Pérez, J.E., Arrue, M., Abascal, J., Duarte, C., Moreno, L.: Adapting the Web for people with upper body motor impairments using touch screen tablets. Interact. Comput. **29**(6), 794–812 (2017)
39. Vigo, M., Harper, S.: Coping tactics employed by visually disabled users on the Web. Int. J. Hum Comput Stud. **71**(11), 1013–1025 (2013)
40. World Health Organization. Visual impairment and blindness. Fact Sheet N°282. http://www.who.int/mediacentre/factsheets/fs282/en/. (2014). Accessed Dec 2019
41. World Wide Web Consortium.: Web content accessibility guidelines (WCAG) 2.0. http://www.w3.org/TR/WCAG20/ (2008). Accessed Dec 2019
42. World Wide Web Consortium.: Low Vision Accessibility Task Force Work Statement. https://www.w3.org/WAI/GL/low-vision-a11y-tf/work-statement (2015). Accessed Dec 2019
43. World Wide Web Consortium.: Accessibility Requirements for People with Low Vision. http://w3c.github.io/low-vision-a11y-tf/requirements.html (2019). Accessed Dec 2019

44. World Wide Web Consortium.: Web Content Accessibility Guidelines (WCAG) 2.1 W3C Recommendation 05 June 2018, from https://www.w3.org/TR/WCAG21/ (2018). Accessed Dec 2019
45. World Wide Web Consortium.: Web Accessibility Initiative (WAI). https://www.w3.org/WAI/ (2018). Accessed Dec 2019
46. Xiao, G., Xu, G., & Lu, J.: iBrowse: Software for low vision to access Internet. In Biomedical Engineering and Informatics (BMEI). In: 2010 3rd International Conference on (Vol. 5, pp. 2062–2066). Yantai, China
47. Yesilada, Y., Stevens, R., Harper, S., Goble, C.: Evaluating DANTE: semantic transcoding for visually disabled users. ACM Trans. Comput. Hum. Interact. (TOCHI) **14**(3), 14 (2007)
48. Zhao, Z., Rau, P.-L.P., Zhang, T., Salvendy, G.: Visual search-based design and evaluation of screen magnifiers for older and visually impaired users. Int. J. Hum. Comput. Stud. **67**(8), 663–675 (2009)

## 3.5. Appendix 5

This appendix includes the article entitled *"Evaluation of two virtual cursors for assisting web access to people with motor impairments"* that was published in the International Journal of Human-Computer Studies (IJHCS) in December 2019. That year the journal reached an impact factor of 3.163 according to Journal Citation Reports, ranking 7th out of 22 journals (Q2) in the "Computer Science, Cybernetics" category.

As people with motor impairments face difficulties with pointing and clicking targets in graphical user interfaces, we developed two cursor aids based on previous research works and on the literature for assisting link selection on the Web. We conducted a web-based user study with 15 participants with the objective of evaluating the suitability (in terms of performance and satisfaction) of both virtual cursors to assist web browsing for people with motor impairments who were users of alternative pointing devices. Other works related to behavioural studies of people with motor impairments in point and click tasks were presented in various international conferences:

- J. EDUARDO PÉREZ, XABIER VALENCIA, MYRIAM ARRUE, AND JULIO ABASCAL, "Depicting the keypad mouse interface: exploring pointing behaviors of three input devices," In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility* (ASSETS '15), Lisbon, Portugal, Oct. 26–28, 2015, pp. 413–414.

- J. EDUARDO PÉREZ, XABIER VALENCIA, MYRIAM ARRUE, AND JULIO ABASCAL, "A usability evaluation of two virtual aids to enhance cursor accessibility for people with motor impairments," In *Proceedings of the 13th Web for All Conference* (W4A

‘16), Montreal, Canada, Apr. 11–13, 2016, article no. 20. **[ ⋆ Best Communication Paper Award ]**

- J. EDUARDO PÉREZ AND MYRIAM ARRUE, “Virtual cursors to enhance web accessibility for people with limited dexterity: usability test results and future directions,” in *The Newsletter of ACM SIGACCESS Accessibility and Computing*, no. 115, pp. 3–11, Jun. 2016.

International Journal of Human-Computer Studies 132 (2019) 81–98

Contents lists available at ScienceDirect

International Journal of Human-Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

# Evaluation of two virtual cursors for assisting web access to people with motor impairments

J. Eduardo Pérez*, Xabier Valencia, Myriam Arrue, Julio Abascal

*University of the Basque Country (UPV/EHU), 20018 San Sebastián, Spain*

ARTICLE INFO

*Keywords:*
Web accessibility
Link selection
People with motor impairments
Alternative pointing devices
Virtual cursors
Empirical measurements

ABSTRACT

People with motor impairments (MI) may face accessibility barriers when using computers due to their health conditions and therefore need to use alternative devices to a standard mouse for pointing and clicking in graphical user interfaces (GUI). In this study with users of different pointing devices, we evaluate 2 virtual cursors (the novel cross cursor and the standard area cursor) implemented for assisting link selection on the Web by reducing respectively cursor displacement and the precision required. Both cursor adaptations were developed for this work based on previous research, and have been compared with the original unassisted cursor in a web-based study with fifteen regular computer users applying their usual pointing device. Nine participants with MIs participated, including 4 using keyboards as an alternative pointing device, 4 joystick users and 1 trackball user. Six participants without MIs also participated in the study applying a standard mouse to complete the same experimental tasks. User interactions with the pointing device, as well as subjective assessments about the usability of the cursor variants tested were gathered from study participants. An in-depth analysis of point and click trajectories showed that virtual cursors improved the effectiveness and efficiency of most participants with MIs in link selection. Subjective assessments about cursor variants tested showed that a majority of participants with MIs generally preferred one of either the two virtual cursors to the original one for web navigation.

## 1. Introduction

People with motor impairments (MI) may have problems using standard input devices to access computers (e.g., mouse) due to lack of dexterity in their upper limbs (Trewin and Pain, 1999), and therefore need to use assistive technologies (AT). MIs can hinder user interactions with computers in different ways (Sears et al., 2008), including, among others: poor coordination, slow movements, low strength, tremor, spam, rapid fatigue, or difficulty controlling direction or distance. Some diseases (WHO, 2001) resulting in MIs hindering the use of computers are: cerebral palsy, spinal cord injury, multiple sclerosis, muscular dystrophy, Parkinson's disease, arthritis or missing limbs and digits. In order to facilitate access to computers to this heterogeneous group of people, many different ATs have been developed (Cook and Polgar, 2014), such as, for example, mouse alternatives enabling direct interaction in graphical user interfaces (GUI). In a similar vein, some ATs that allow people with MIs to interact with an on-screen cursor are: specific alternative pointing devices such as a joystick or trackball, or software applications such as mouse keys to use the numeric keypad on a keyboard as a pointing device.

Despite these ATs people with MIs still find challenges when interacting with standard GUIs, such as, for example, selecting links on the Web (Trewin, 2008), so further research is needed to continue enhancing computer access for all of these users. Many studies have investigated difficulties faced by people with MIs when pointing and clicking targets on GUIs (Almanji et al., 2014; Hwang et al., 2004; Keates et al., 2002; Payne et al., 2017; Pérez et al., 2015; Valencia et al., 2017). Also other works have proposed new selection methods with on-screen cursor to assist point and click interactions on GUIs for people with MIs (Grossman and Balakrishnan, 2005; Harada et al., 2006; Hwang et al., 2003; Mott and Wobbrock, 2014; Payne et al., 2016; Salivia and Hourcade, 2013; Trewin et al., 2006; Wobbrock and Gajos, 2008). Nevertheless, these works mainly carry out tests with users of similar pointing devices, on repetitive tapping tasks within closed experimental environments, and focused on studying user performance. To improve both performance and experience of people with MIs accessing GUIs with any AT, new cursor enhancements have to be studied, with different users of alternative pointing devices, and on tasks recreating activities from their everyday computer use.

Based on previous work (Pérez et al., 2014) about web navigation strategies of users with MIs, we developed two virtual cursors as browser add-ons for assisting point and click interactions on the Web to

* Corresponding author.
*E-mail address:* juaneduardo.perez@ehu.eus (J.E. Pérez).

https://doi.org/10.1016/j.ijhcs.2019.08.001
Received 5 September 2018; Received in revised form 1 August 2019; Accepted 1 August 2019
Available online 02 August 2019
1071-5819/ 

two groups of users of alternative pointing devices. The novel cross cursor aimed at keyboard users to reduce cursor displacement when pointing, and the standard area cursor aimed at joystick and trackball users to reduce the accuracy needed to select links. The objective of this work was to study empirically the suitability of both virtual cursors for assisting web browsing with users of different pointing devices, by means of performance and satisfaction measures. For this purpose a web-based test was conducted with people with MIs users of the aforementioned alternative devices, as well as with mouse users without MIs in order to compare the performance and acceptance of cursors tested by the different groups of users. Several measurements from literature were calculated from point and click trajectories recorded during user tests (MacKenzie et al., 2001; Keates et al., 2002; Hwang et al., 2004) to compare performance achieved by participants with each cursor variant. We studied the usability of each cursor variant tested from subjective assessments of study participants, gathered by means of questionnaires based on reference works (Brooke, 1996; Hart and Staveland, 1988). Study findings were promising in terms of performance and satisfaction achieved by participants with MIs, and showed that users of alternative pointing devices can clearly benefit from point and click facilitators for accessing the Web. The study results also suggested that the pointing device used by the user is a good indicator in order to provide a better cursor assistance for people with MIs.

The rest of the paper is composed of the following six sections. Related work about enhanced cursors for assisting point and click interactions for people with MIs are described in Section 2. The two virtual cursors implemented to assist link selection on the Web for people with MIs using alternative pointing devices are presented in Section 3. Experimental evaluation methodology is explained in Section 4. Results obtained from this research are detailed in Section 5. The discussion about results of the study is included in Section 6. Conclusions about this research and future work are presented in Section 7.

## 2. Related work

Much research has been done into facilitating pointing and clicking in GUIs even though little of it has been specifically focused on assisting people with MIs on the Web. The following works proposed some renowned cursor enhancements to assist pointing and clicking interactions, although these were not always initially aimed at people with MIs. Even if these works were not focused on assisting web browsing, some can be directly translated to this scenario and served to define the basis of our research.

The steady clicks assistance (Trewin et al., 2006) suppresses accidental clicks and slipping when clicking by freezing the cursor during mouse clicks, preventing overlapping button presses and cancelling clicks made while the mouse is moving at a high speed. The evaluation showed that this option improves time performance and enables users with MIs to select targets using fewer attempts; moreover, participants expressed their preference for this assistance (9 out of 11) over the unassisted condition. This alternative aims to assist the clicking task for people with low dexterity in their upper limbs, but its usefulness for keyboard-only users is not so obvious, since these are more affected by distance to target than accuracy.

The angle mouse (Wobbrock et al., 2009) is a pointing facilitation method that attempts to improve target acquisition by adjusting the mouse control-display gain based on the deviation angles of the cursor path sampled during movement. Thus, unlike most cursor enhancements, this technique (like the previous one) is based solely on the user's behaviour and requires no knowledge of targets on the GUI. Study results proved that this alternative improved pointing performance for users with MIs while remaining unobtrusive for people without impairments. However, all participants from the study (both with MIs and without MIs) were using the same standard mouse to complete experimental tasks and no alternative input device was tested.

Wobbrock and Gajos (2008) claimed that the difficulties faced by people with MIs could be alleviated in a different target acquisition paradigm called goal crossing where users do not aim at a restricted area, but instead pass over a target line to select it. Empirical results indicated a preference for goal crossing among people with MIs, although error rates were higher with this alternative. Authors also introduced some design principles for this new target acquisition paradigm, but these are not usable on standard web interfaces.

Hwang et al. (2003) studied the performance of users with and without MIs in a point and click task with force feedback applied to targets modelled as virtual gravity wells. Their results showed the greatest improvements for the users with the most severe impairments, even when multiple on-screen targets were haptically enabled. This technique looks promising for complex GUIs with numerous targets such as the Web, although the study did not include any subjective perception from participants about tested enhancements, or test different alternative pointing devices.

Worden et al. (1997) studied the effectiveness of two interaction techniques: the area cursor and the sticky icons, for improving the performance of older adults (with declined motor abilities) in basic selection tasks. The area cursor, successfully tested before with people without MIs (Kabbash and Buxton, 1995), is a cursor with a larger activation area than normal. The latter technique makes an icon "sticky" by automatically reducing the cursor's gain ratio (number of pixels moved in response to a single increment of movement by the physical device) when it is over a target icon. Both techniques improved pointing time, especially the area cursor when the target icon was not in close proximity to another icon and for smaller target sizes. Results also showed that neither technique impeded performance in problematic cases (e.g., differentiation between closely spaced targets). Other works have studied enhancements for the area cursor, by dynamically resizing the cursor's activation area (Grossman and Balakrishnan, 2005), or with different combinations of visual and motor magnification or goal crossing (Findlater et al., 2010; Mott and Wobbrock, 2014; Payne et al., 2016). Results from these works generally revealed improvements in performance, although they were mainly based on users without MIs, the pointing device used was generally the mouse, or participants sometimes did not prefer the proposed methods.

Felzer et al. (2016) compared two different methods for mouse emulation with a numeric keypad called DualPad. The first method was called CKM and allows moving the mouse pointer in cardinal directions and clicking similarly as with the mouse keys application. The second method was the DualMouse, and does not rely on mouse movement at all, but directly clicks at a destination location following a step-by-step locating process. Evaluation based on a case study with a single user with MIs revealed a higher throughput with the CKM method than with the DualMouse. Surprisingly, no cursor enhancement has been investigated to assist pointing and clicking interactions of keyboard-only users applying the mouse keys application included in every major operative system. There are applications such as VimVixen[1] (a Mozilla Firefox add-on) that enables web browsing by using only the keyboard. These applications label every link of a web page with shortcuts, so that a user can select any link keystroking the corresponding sequence of letters from his/her keyboard and without having to move the cursor pointer at all. The cross cursor that we developed for this study aims to reduce shortcuts to only one letter key each by combining cursor movements to label just those links at reach of this virtual cursor (Fig. 1 Right).

Pierson and Magee (2017) present a browser plug-in that implements a predictive link following algorithm for assisting link selection to people with MIs. Their algorithm analyses mouse movement and erroneous clicks before instructing the browser to follow a link.

[1] https://addons.mozilla.org/en-US/firefox/addon/vim-vixen

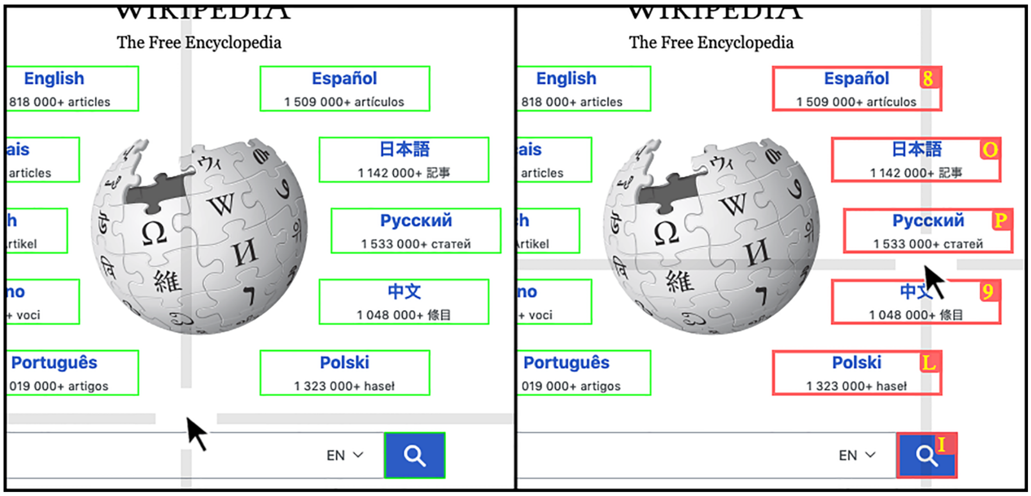


**Fig. 1.** (Left) The cross cursor without any link at reach, and (Right) with 6 links at reach via keyboard shortcuts after diagonally moving the cursor up to the right.

Experimental evaluations of their system with different pointing devices reported an improved performance with the proposed approach.

The Fitts' law paradigm has been widely applied to the comparison and optimization of pointing devices and interaction techniques as in Wobbrock and Gajos (2007). However, there are conflicting reports that differ on whether Fitts' law can be applied to pointing movements of people with motor impairments or not. While some works contribute with evidence in favour of the suitability of Fitts' law (Rao et al., 2000), others affirm just the opposite (Gump et al., 2002). Considering that pointing trajectories of keyboard user do not follow a ballistic movement supposed by Fitts' law, our evaluation was based on cursor measures as detailed in Section 4.

## 3. Virtual cursors

In previous works (Pérez et al., 2015, 2014; Valencia et al., 2015) concerning pointing and clicking behaviours of people with MIs, we observed that main difficulties faced by participants varied depending on the pointing device alternative used. Thus, users of keyboard as pointing device were more affected by the total distance to the link and by the pointing trajectory until reaching the target, whereas users of specific alternative pointing devices such as the joystick or trackball tended to have more problems near targets to stop the cursor over them. In order to study empirically if these web browsing issues faced by people with MIs can be alleviated by means of pointing and clicking assistances, we developed 2 virtual cursors as browser add-ons and tested them with real users on real web environments. Both virtual cursors, the novel cross cursor (Fig. 1) and the already existent area cursor (Fig. 2) were design to assist web browsing to people with MIs, by implementing different techniques to modify standard pointing and clicking. In the following 2 subsections we present both virtual cursors, how the new pointing and clicking assistances work, as well as some technical details.

### 3.1. The cross cursor

This virtual cursor aims to assist target acquisition on the Web by reducing cursor displacement required for pointing. This is achieved by combining cursor movement and providing single-letter shortcuts to every link at reach of the cross cursor (Fig. 1 Left and Right). Links at reach of this virtual cursor are those traversed by the cross cursor lines. The cross cursor continuously displays along its movement a horizontal and a vertical line (that respectively extend over the entire width and height of the web page) crossing perpendicularly below its current position (Fig. 1). Shortcuts are automatically assigned and displayed next to every link reached by the cross cursor each time the virtual cursor stops motion, and disappear whenever the cursor starts moving again. Single-letter shortcuts (together with number keys) are automatically assigned in order of proximity to the cursor pointer, starting from the right of the keyboard, with the closest keys to the numeric keypad first and the furthest to the left at the end. If all letter and number keys (36 in our case) have already been assigned in this way, the additional links within reach of the cross cursor will not have any shortcut assigned, having to approach them with the cursor pointer to display a shortcut. We used fixed values (10 px width and 90% translucent grey colour) for the visual appearance of the cross cursor lines.

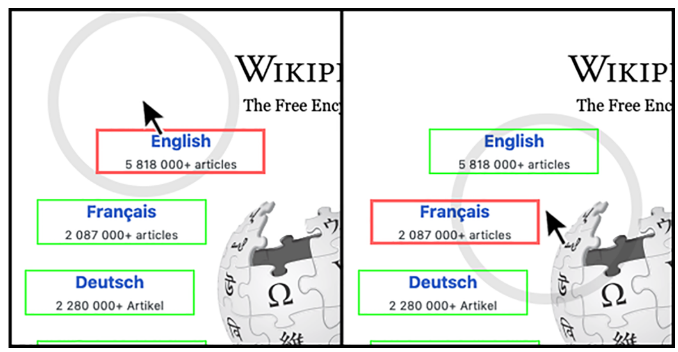


**Fig. 2.** (Left) The area cursor allows clicking a nearby link without needing to hover the cursor over it. Link at reach is highlighted in a different colour. (Right) If the activation area of this cursor reaches more than one link, the nearest one to the cursor pointer is highlighted and can be selected.

The mouse keys feature, included on every major operative system, allows keyboard users as alternative pointing device to use the numeric keypad as a mouse alternative (Fig. 3 left) by pressing the central '5' key for cursor clicking, and the surrounding number keys for moving it in vertical, horizontal and diagonal directions (Fig. 3 right).

### 3.2. The area cursor

This virtual cursor corresponds to the standard area cursor and aims to assist target acquisition on the Web by reducing the accuracy required to click a link. The area cursor continuously displays as it moves a circle of fixed size that is always centred with its current position (Fig. 2) and which corresponds to its activation area. In this way, this virtual cursor enables the closest link within its activation area to be clicked (highlighted targets in red in Fig. 2) without needing to hover over it. We used fixed values for the visual appearance of the area

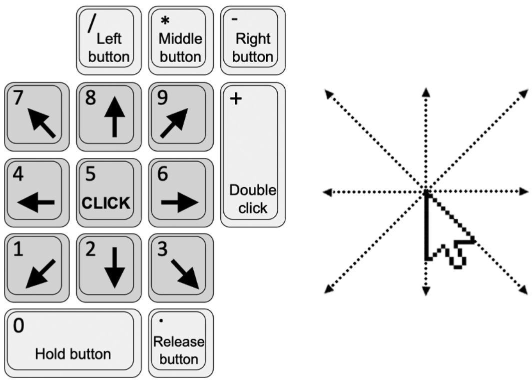


**Fig. 3.** (Left) Numeric keypad located on the right side of a keyboard highlighting the set of keys used by mouse keys for cursor pointing and clicking. (Right) Eight possible paths that cursor can travel from its current position when using mouse keys.

cursor (10 px width and 90% translucent grey colour) as well as for the activation area diameter (130 px).

Both virtual cursors were implemented using Scalable Vector Graphics (SVG) to add visual elements (lines, circles, rectangles and letters), along with JavaScript to handle users' interactions with the cursor and the web content. The add-on that implements each virtual cursor is in charge of parsing every visited web page to find all the links included. The information about the location and size of each visible link within a page is processed by the add-on, which also handles users' interactions (mouse moves, clicks, and keystrokes) to modify standard pointing and clicking and assist target selection with the virtual cursors as presented here.

## 4. Experiment method

In order to compare both virtual cursors (area and cross) with the original cursor, we carried out a web-based experiment with people with MIs and people without disabilities, applying their usual pointing device. For the purpose of analysing cursor movements on point and click interactions, two kinds of task were defined: searching tasks and target acquisition tasks. Cursor trajectories, as well as other related events were recorded during experimental sessions within an interaction log for later analysis. After participants completed experimental tasks, their subjective assessments were also gathered to measure the usability of each cursor variant tested. This study was approved by the Ethics Committee for Research Involving Human Beings from the University of the Basque Country.

### *4.1. Participants*

A total of 15 participants took part in this study, 9 of which were people with MIs involving reduced mobility in their upper limbs. The other 6 participants were people without impairments, recruited as control subjects. All participants were regular computer users, accustomed to accessing the Web frequently, and were specifically chosen to participate in the study based on their usual input device alternative for mouse pointing in graphical user interfaces. In this way, the following 3 groups were defined:

- **KU group** with 4 keyboard users (3 females, mean = 49 years, SD = 7.8)
- **JU group** with 5 participants, 4 joystick users and 1 oversized trackball user (2 females, mean = 50.2 years, SD = 15.1)
- **MU group** with 6 mouse users (2 females, mean = 33.2 years, SD = 5.2)

All participants from the KU and JU groups were people with limited dexterity in their upper limbs that prevented them from using a standard mouse. All of them were experienced users with the pointing device alternative used during the experiment, and were mainly recruited from the Elkartu association of people with physical disabilities from our local area. Two participants from the KU group (KU1 and KU2) were able to push keyboard keys directly with their hand, while the other two (KU3 and KU4) needed a head wand. By contrast, participants from the MU group were people without disabilities and with over 7 years of experience using the mouse as their usual pointing device for computer access. Table 1 shows detailed information about the 15 participants of the study, grouped by the alternative pointing device used, including: each person's gender, age, regularity of use and years of experience with the corresponding pointing device, health condition (if applicable), and location where the experimental session was carried out (at their home, in a laboratory of the University of the Basque Country-UPV/EHU or on the premises of the Elkartu association).

### *4.2. Apparatus*

All participants used the same equipment to complete the experiment, except for the pointing device. A Dell Precision M6700 laptop running a 64 bits version of the Windows 7 OS was used alongside an additional 24 in. widescreen LCD monitor (Dell model U2412MB) to

**Table 1**
Demographic data about study participants grouped by pointing device used.

| Id | Gender | Age | Pointing device | Use | Exp. | Health condition | Location |
|---|---|---|---|---|---|---|---|
| **Keyboard users group (KU)** | | | | | | | |
| KU1 | F | 58 | Keyboard | Daily | +7 | Glutaric aciduria t1 | Home |
| KU2 | F | 53 | Keyboard | Daily | +7 | Glutaric aciduria t1 | Home |
| KU3 | M | 42 | Keyboard + head wand | Daily | +7 | Cerebral palsy | Home |
| KU4 | F | 43 | Keyboard + head wand | Daily | 1–3 | Cerebral palsy | Home |
| **Joystick & trackball users group (JU)** | | | | | | | |
| JU1 | M | 45 | Oversized trackball | Daily | +7 | Cerebral palsy | Lab |
| JU2 | M | 42 | Joystick | Daily | +7 | Cerebral palsy | Home |
| JU3 | M | 46 | Joystick | Daily | +7 | Cerebral palsy | Lab |
| JU4 | F | 41 | Joystick | Daily | +7 | Cerebral palsy | Home |
| JU5 | F | 77 | Joystick | Weekly | 4–6 | Spinal cord injury | Elkartu |
| **Mouse users group (MU)** | | | | | | | |
| MU1 | F | 30 | Mouse | Daily | +7 | – | Lab |
| MU2 | F | 33 | Mouse | Daily | +7 | – | Lab |
| MU3 | M | 30 | Mouse | Daily | +7 | – | Lab |
| MU4 | M | 28 | Mouse | Daily | +7 | – | Lab |
| MU5 | M | 36 | Mouse | Daily | +7 | – | Lab |
| MU6 | M | 42 | Mouse | Daily | +7 | – | Lab |

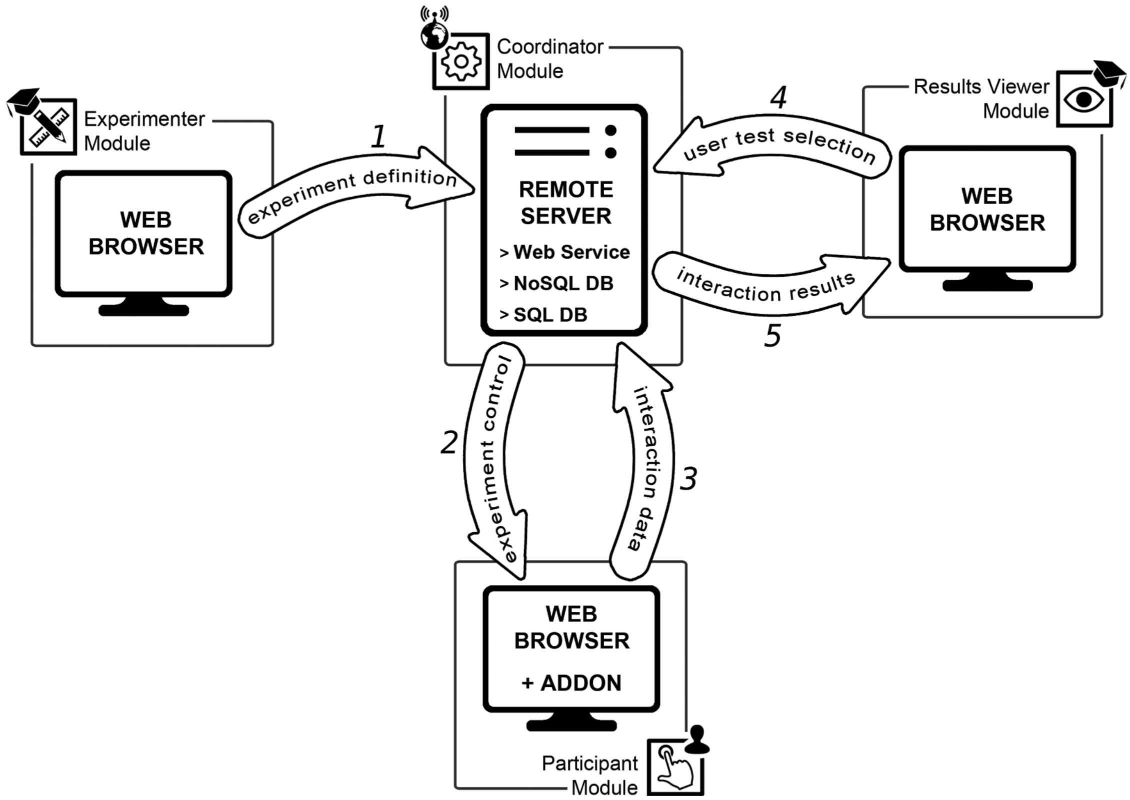


**Fig. 4.** The RemoTest platform general architecture and interactions between modules.

present the stimuli to participants. Participants from the KU and JU groups used their own personal input device alternative for mouse pointing while participants from the MU group used the same optical USB mouse (Dell model M-UVDEL1). Before starting the study, participants were encouraged to adjust the pointer motion options on the Windows control panel to fit their own preferences, and thus ensure their best performance during the experimental tasks.

The RemoTest platform (Arrue et al., 2018; Valencia et al., 2015) was used to specify and conduct the experimental sessions of this study. The architecture of the platform is based on a hybrid architecture model that includes some functionality in a client-side module and the other ones in some server-side modules. The platform is split into four modules: Experimenter Module (EXm), Participant Module (PAm), Coordinator Module (COm) and Results Viewer Module (RVm). Fig. 4 shows the general architecture and interactions between these modules.

Each module has specific functions and uses different technologies. The EXm module is responsible of assisting experimenters in the experiment definition process. The experiment definition is stored in a XML file based on the vocabulary of the specifically developed Experiment Specification Language. This file is the input for the COm module (Step 1 in Fig. 4). The COm module transforms the experiment specifications into personalized experimental sessions specified in XML format. These personalized sessions are transferred to the corresponding PAm modules (Step 2 in Fig. 4). The PAm module guides participants during the experimental sessions and gets the required user interaction data. This data is sent to the COm where it is stored for future analysis (Step 3 in Fig. 4). The RVm module organizes and presents the abundant interaction data gathered in an experiment, provided by the COm module (Step 4 and 5 in Fig. 4).

The experiment described in this paper was defined using functionalities of the EXm module of the RemoTest platform and an add-on was created for each participant with all the necessary data for presenting the experiment tasks to participants. This add-on is the PAm module of the platform, which was installed in the Mozilla Firefox web browser, used in the experimental sessions in conjunction with another add-on that implemented both of the virtual cursors. The user interaction data gathered during the experimental sessions was transmitted to the COm module and stored in a MongoDB database. This interaction data consisted of on-screen cursor trajectories with a sampling frequency of 100 Hz (X and Y cursor coordinates recorded each 10 ms approximately), as well as selected link (top, bottom, left and right coordinates) in order to later compare participants' performance with each cursor variant. Other user interactions with input devices (keystroke, click, page scroll, etc.) were also gathered, as well as browser and experiment events (page load, start and end of tasks, etc.) in order to identify valid cursor trajectories and delimit data for analysis. For this purpose, a separate Java application was implemented to parse interaction data recorded from participants and calculate a variety of measures for each trial, which were later analysed with the RStudio statistical tool.

### 4.3. Tasks and materials

Participants were asked to perform a set of tasks during the experimental session within two different websites: Discapnet (2015) and Gipuzkoa (2015). Discapnet (Fig. 5 left) is a website which provides information aimed at people with disabilities, organizations or relatives of people with disabilities. They provide news, information about the rights of people with disabilities, etc. Gipuzkoa (Fig. 5 right) is an institutional website of the Gipuzkoa provincial council with news relating to the council, institutional information about local governments, etc. A third website about touristic information of the Bidasoa local area (Bidasoa Turismo, 2015) was used for training purposes, so participants could learn how to use the virtual cursors being tested, as well as to inform them about the experimental tasks they were going to carry out next. All 3 websites claimed, within their accessibility sections, to conform to a certain level of the WCAG 1.0 guidelines (Discapnet to Level AA, Guipuzkoa and Bidasoa to the Level A).


**Fig. 5.** Home page of both websites that study participants had to navigate to complete experiment tasks: Discapnet (left) and Gipuzkoa (right).

Two types of tasks were defined to be performed with each cursor variant tested (original, area and cross):

- *Search tasks*: in which study participants had to navigate through both proposed websites searching for different content, starting each time from the home page of the corresponding site. All searches had a similar level of difficulty, with the objective content located at 3 levels from the home page. A total of 12 different search tasks were defined, 6 within Discapnet (2015) and 6 within Gipuzkoa (2015), each to be completed within a 3 min time limit. These 12 search tasks were distributed between the 3 cursors tested (2 searches from each website by cursor variant) counterbalancing the order between participants. The goal was twofold: firstly that participants practiced and became familiar with how to select links with cursor variants before the following tasks, and secondly to gather their subjective assessments after a natural usage of cursors. In addition, 3 more search tasks were defined within the Bidasoa Turismo (2015) website for training purposes, each of which had to be completed by participants with a cursor variant before carrying out the actual experimental tasks. During search tasks the current objective was displayed at the bottom of the browser continuously, so that participants did not forget what they were looking for (Fig. 6 bottom).
- *Target acquisition tasks*: in which study participants had to sequentially select highlighted links on the browser screen. In order to cover different approaching angles in each trial, chosen targets were evenly distributed between quadrants of imaginary Cartesian axes centred on the screen. In this way, a total of 48 targets were defined (half from each website) that had to be selected by participants with each tested cursor variant. Before each trial, participants had to position the cursor over a home button located in the centre of the screen (Fig. 7). Upon selecting it, the home button disappeared and a new trial started, in which participants had to select the highlighted link on the screen as fast as possible. After completing each trial, the home button re-appeared and next target was highlighted. The goal was to record cursor trajectories on intentional movements of target acquisition (avoiding unintended moves that might occur during search tasks) in order to compare performance achieved with each cursor variant. In addition, 5 more targets were defined within the Bidasoa Turismo (2015) website for training purposes, that participants had to select with every tested cursor variant before carrying out actual experimental tasks.

### 4.5. Procedure

First, participants were briefed on the purpose of the study and then signed a consent form before starting with the experimental session. Information on demographics and about expertise with the corresponding pointing device was collected through a brief pre-session interview. Prior to testing, experimental tasks and virtual cursors were introduced to participants through demonstration and practice during a training session of between 10–20 min (5–10 min in the case of participants without MIs). Participants had to complete 3 consecutive training blocks corresponding to the 3 cursor variants being tested, each one including a search task followed by 5 target acquisition tasks. The Bidasoa Turismo (2015) website was exclusively used for this purpose.

After concluding the training session, participants had to complete actual experimental tasks grouped in 3 consecutive blocks corresponding to each cursor variant tested (original, area and cross). Cursor blocks were presented in counterbalanced order to participants, and each included 4 search tasks followed by 48 target acquisition tasks,


**Fig. 6.** Browser status bar at the botom displaying enlarged text of the content to look for within Gipuzkoa website during a search task.

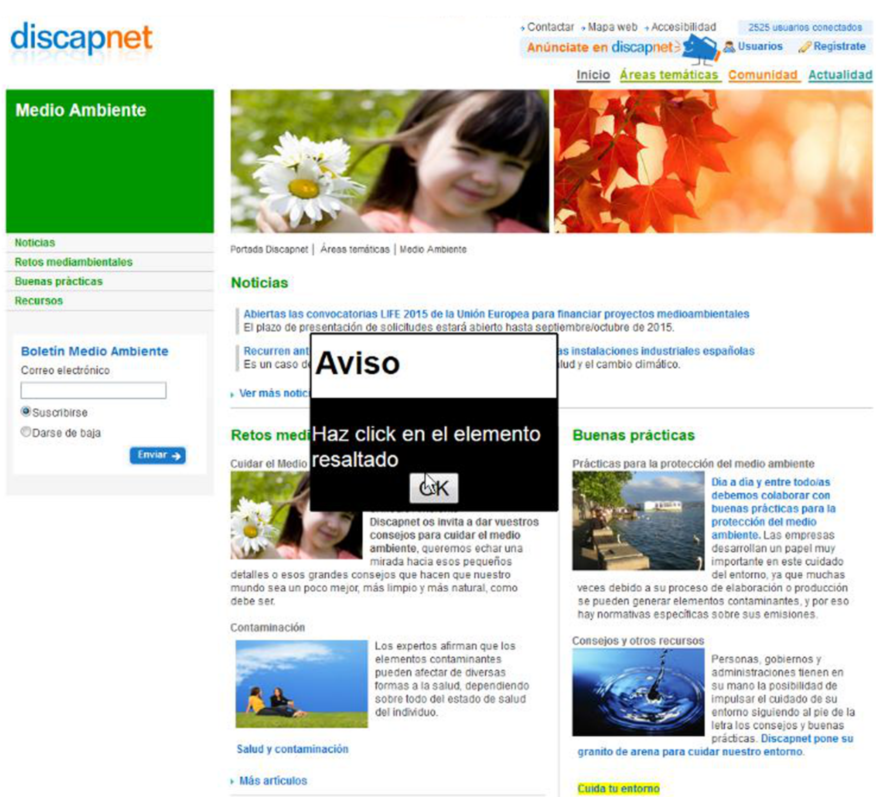


**Fig. 7.** Screen capture of a target acquisition trial, with cursor over the home button and the 14th target from Discapnet website highlighted on the bottom right side.

distributed equally between the Discapnet (2015) and Gipuzkoa (2015) websites. After completing each block of tasks for a particular cursor variant, a semi-structured interview was conducted in order to gather subjective assessments from participants about that cursor variant.

After completing all 3 blocks of experimental tasks with each cursor variant, participants were asked to rank them from most to least favourite for web browsing purposes. Experimental sessions lasted between 1 and 2 h (30–50 min in the case of participants without MIs). After concluding experimental sessions, participants were rewarded with a voucher worth 25€ for their collaboration in the study.

### 4.6. Measuring cursors usability

Two different methods were used to measure the usability of the 3 cursor variants tested in this study. On one hand, performance achieved by participants with each cursor variant on target acquisition tasks was studied by means of several cursor path evaluation measures. In this way, the following cursor measurements proposed in the literature (MacKenzie et al., 2001; Keates et al., 2002; Hwang et al., 2004) were used to calculate the efficiency and efficacy achieved on each trial of target acquisition tasks:

- *Movement time (MT)*: the time interval from clicking the home button until target link is selected. The MT corresponds to total time needed to complete a trial, and was calculated based on timestamps of events recorded during experimental sessions.
- *Pointing time (PT)*: the time interval from when the on-screen cursor starts moving until it finally stops before the target link is selected. The PT corresponds to time needed to move the on-screen cursor to complete a trial. The PT is a portion of the total MT, and was also calculated based on timestamps of events recorded during experimental sessions.
- *Clicking time (CT)*: the time interval from when the on-screen cursor finally stops moving until the target link is selected. The CT corresponds to the time needed to perform the click to complete a trial. The CT is a portion of the total MT, and was also calculated based on timestamps of events recorded during experimental sessions.
- *Distance Travelled (DT)*: the total distance traversed (in pixels) by the on-screen cursor along the pointing trajectory. The DT was computed for each trial as the sum of distances from each point to the next point. The distance between two consecutive points ($X_1$,$Y_1$) and ($X_2$,$Y_2$) is given by:

$$d = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2}$$

- *Curvature index (CI)*: the ratio of *DT* and the straight-line distance between the starting and ending points of cursor trajectory. A value of one indicates the cursor has followed a straight line, while growing values shows increasing deviations. The CI was calculated based on computed DT for a trial and the corresponding distance between first and last cursor location.
- *Number of pauses (NP)*: the times the cursor stops along the pointing trajectory. The NP represents the number of corrections made by the user in order to select a target. Low values indicate fewer corrections and therefore fewer problems on the pointing trajectory, while a high number means the user has more difficulties to select that target. The NP was calculated based on the time interval between consecutive cursor motion events, considering intervals equal to or greater than 100 ms as a pause.

On the other hand, satisfaction of study participants with each cursor variant was collected by means of questionnaires based on the System Usability Scale (Brooke, 1996) and the NASA TLX (Hart and Staveland, 1988). In this way, participants provided their subjective assessments for each cursor variant based on the following 8 categories that were rated on a 7-point Likert scale (from strongly positive to strongly negative):

- *Learnable*: How easy was it to learn to use this cursor variant?
- *Memorable*: How easy was it to remember how to use this cursor variant?
- *Accurate*: How accurate was this cursor variant in selecting links?
- *Easy to use*: How easy was it to use this cursor variant to select links?
- *Effortless*: How would you describe the level of effort you need to make with this cursor variant to select links?
- *Natural*: How natural was it to use this cursor variant to select links?
- *Fun*: How much fun was it to use this cursor variant for browsing the Web?
- *Not frustrating*: How would you describe your level of frustration when using this cursor variant to select links?

To conclude, after completing experimental tasks with all 3 cursor variants, participants were also asked to rank them from most to least preferred choice for browsing the Web. Promising preliminary results were obtained from these subjective assessments of the participants (Pérez et al., 2016).

## 5. Results

In the following subsections we analyse the data collected from participants during experimental sessions. In the first subsection we study how participants leveraged both virtual cursors, explain the filtering process to remove invalid trials, and discuss some implications found. In the following subsection we analyse target acquisition tasks based on several cursor measurements in order to compare the effectiveness and efficiency of the different user groups with each cursor variant tested. Finally, we present a qualitative analysis about participants' satisfaction with each cursor variant based on their responses gathered in the interviews.

### 5.1. Use of cursor variants and data cleaning

Each of the 15 participants in the study completed 48 target acquisition trials with each of the 3 cursor variants tested (cross, area and original cursor), resulting in a total of 2160 trials. Cursor trajectories gathered in the experimental sessions were analysed in order to filter invalid trials, and thus obtain meaningful measurements of targets acquisition tasks. Below, we analyse the use of each virtual cursor along target acquisition task to understand the acceptance of both assistances by participants.

As we have explained before, the cross cursor enables to shorten link selection time by reducing pointing trajectory and presenting single-letter shortcuts for links at reach of the virtual cursor. Considering the keyboard approach of the cross cursor and the difficulties that some participants had to access this device, they could also use this virtual cursor with standard pointing and clicking (i.e., without leveraging shortcuts) if it was easier for them this way. Table 2 shows for each user group the usage of shortcuts with the cross cursor for assisting link selection, in comparison with standard pointing and clicking. While all participants from the KU group leveraged the cross cursor help to

**Table 2**
Usage of cross cursor shortcuts on target acquisition task, ordered by user group. The distribution within each group is displayed below the total and percentage values.

| | Cross cursor | | |
|---|---|---|---|
| *Group* | *Participants* | *Total trials* | *Assistance usage* |
| KU | 4 | 192 | 191 (99.5%) |
| JU | 5 | 240 | 50 (20.8%) |
| MU | 6 | 288 | 53 (18.4%) |

**Table 3**
Target selections completed from distance for each virtual cursor and user group. The distribution within each group is displayed below the total and percentage values.

| Group | Participants | Total trials | Outside selections |
|---|---|---|---|
| **Cross cursor** | | | |
| KU | 4 | 192 | 184 (95.8%) |
| **Area cursor** | | | |
| KU | 4 | 192 | 147 (76.6%) |
| JU | 5 | 240 | 124 (51.7%) |
| MU | 6 | 288 | 101 (35.1%) |

**Table 4**
Number of invalid trials filtered from target acquisition task, ordered by user group, cursor variant, and type of issue. The distribution within each group is displayed below the total and percentage of filtered trials.

| Cursor | Misses | Erroneous | Outliers |
|---|---|---|---|
| **Keyboard users group (KU)** | | | |
| Original | 7 (3.6%) | 3 (1.6%) | 11 (5.7%) |
| Area | 5 (2.6%) | 4 (2.1%) | 6 (3.1%) |
| Cross | 11 (5.7%) | 8 (4.2%) | 7 (3.6%) |
| **Joystick & trackball users group (JU)** | | | |
| Original | 18 (7.5%) | 1 (0.4%) | 10 (4.2%) |
| Area | 10 (4.2%) | 2 (0.8%) | 11 (4.6%) |
| **Mouse users group (MU)** | | | |
| Original | 22 (7.6%) | 0 (0%) | 10 (3.5%) |
| Area | 4 (1.4%) | 1 (0.3%) | 12 (4.2%) |

complete target acquisition tasks (99.5%), participants from JU and MU groups generally preferred to move the virtual cursor pointer over links to perform standard pointing and clicking (respectively 79.2% and 81.6%) as with the original unassisted cursor. Only one participant from the JU group (JU1) and another from the MU group (MU2) used repeatedly the cross cursor shortcuts to complete target acquisition tasks. For this reason, subsequent analyses presented about the cross cursor refer only to the keyboard users group (KU).

Both virtual cursors studied (the cross and area cursor) enable link section without needing to hover the cursor pointer over targets to perform a click. In this respect, Table 3 shows the number of target acquisition trials completed by selecting the link from outside the target area (i.e., taking advantage of the cursor assistance), ordered by virtual cursor variant and user group. The KU group of keyboard users as pointing device achieved the highest rates of target selection from distance, both with the cross cursor (95.8%) and the area cursor (76.6%). Joystick and trackball users (JU group) completed on average slightly more than half of the trials with the area cursor (51.7%) by clicking outside of the target link. On the contrary, the MU group participants without MIs were the ones that, on average, most frequently clicked over the target links with the area cursor without leveraging the virtual cursor assistance (64.9%). Results on Tables 2 and 3 show a good acceptance and use of the novel cross cursor by the KU group participants, whereas participants of the other 2 groups (JU and MU groups) generally avoided using this assistance. On the other hand, on average, the 3 groups of participants took advantage of the area cursor to assist link selection, although with different results (Table 3). All participants of the KU group leveraged regularly the area cursor assistance, with users KU2 and KU1 achieving the highest (83.3%) and the lowest (66.7%) rates of this group, respectively. The participants of the JU group also leveraged the area cursor assistance, although unevenly, with user JU1 achieving the highest rate (81.3%) and JU4 the lowest (27.1%). In contrast, several MU group participants barely used the area cursor assistance and clicked over the target area in most of the trials, as for instance user MU6 (95.8%) and MU2 (91.7%). These results, as expected, show a greater preference for the virtual cursors by participants with MIs (KU and JU groups) than by participants without MIs (MU group).

Invalid trials were removed for subsequent analyses of cursor trajectories and corresponded to misses on target acquisition tasks, erroneous trials including unexpected events on pointing and clicking interactions, as well as outlier trials. For the cross cursor, a missed trial occurred when one or more additional letter keystrokes were registered before target selection. For the area and original cursors, a missed trial occurred when one or more additional clicks were registered before target selection. Erroneous trials were defined as those including user interactions not related with the target acquisition task, and corresponded to any event other than the cursor move followed by a target selection (a standard click for the area and original cursors, or a letter keystroke for the cross cursor shortcut). Events that allowed identifying erroneous trials include the use of the Control, Shift, Escape, and Arrow keys (by the KU group), the use of the Enter key and the reload page button (by the JU group), or scrolling the page (by the MU group). We also removed outlier trials corresponding to cases where the movement time (MT) divided by the index of difficulty (ID) was two standard deviations or more away from participant's mean with the corresponding cursor variant. To calculate the ID of each trial we used the following equation for bivariate pointing (1) which considers both target width W and height H, as well as distance D from starting point to target (Accot and Zhai, 2003). A similar calculation was made to normalize other studied cursor measurements related to target acquisition as detailed in the following subsection. Table 4 shows the number of invalid trials filtered this way from target acquisition tasks, for each cursor variant, user group, and issue type.

$$ID = \log_2\left(\sqrt{(D/W)^2 + \eta\,(D/H)^2} + 1\right) \quad (1)$$

For keyboard users as pointing device (KU group), the data cleaning process resulted in removing 10.9% of the trials completed with the original cursor, 7.8% with the area cursor, and 13.5% with the cross cursor. Taking into account only missed and erroneous trials, the filtered data was 5.2%, 4.7%, and 9.9%, respectively. These results show a slight improvement with the area cursor compared to the original cursor, although this was not repeated by all KU group participants (KU2 and KU3 in Table 4). With the cross cursor, in contrast, the number of invalid trials increased compared to the other two cursors tested.

For joystick and trackball users (JU group), the data cleaning process resulted in removing 12.1% of the trials completed with the

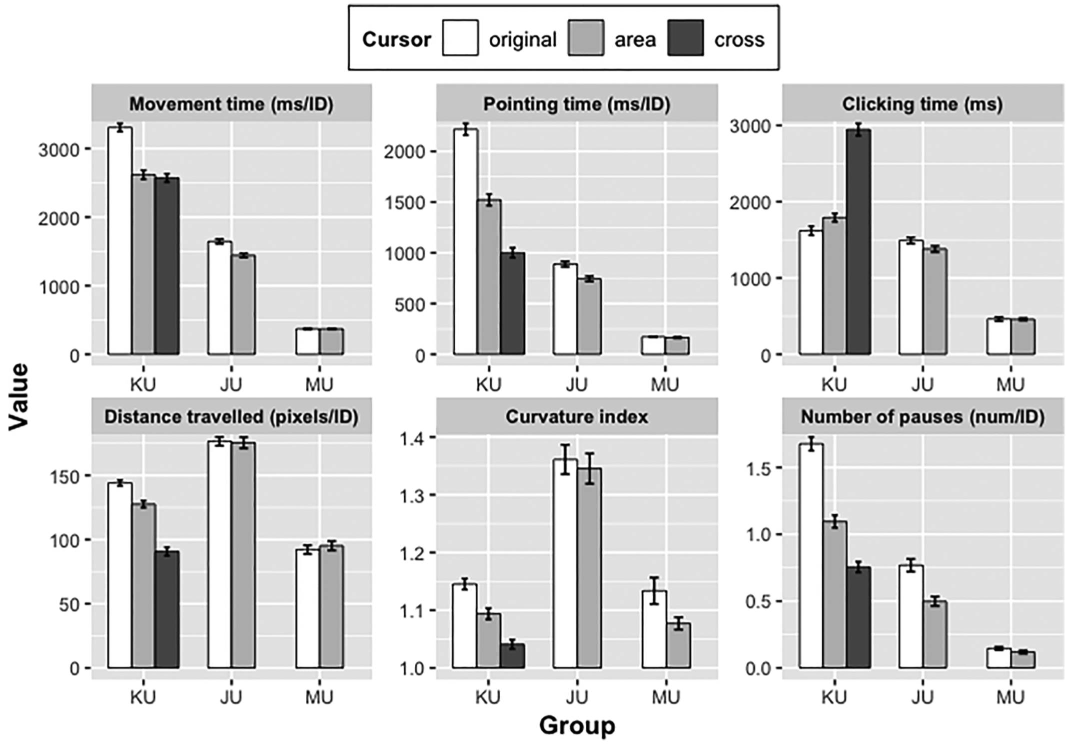


**Fig. 8.** Mean values of each cursor measure studied by users group and cursor variant tested. Error bars represent ± 1 standard error (SE).

original cursor and 9.6% with the area cursor. Taking into account only missed and erroneous trials, the filtered data was 7.9% and 5%, respectively. These results show a decrease of invalid trials with the area cursor compared to the original cursor.

For participants without MIs (MU group), the data cleaning process resulted in removing 11.1% of the trials completed with the original cursor, and 5.9% with the area cursor. Taking into account only missed and erroneous trials, the filtered data was 7.6% and 1.7%, respectively. These results show an improvement with the area cursor compared to the original cursor, which was repeated by all MU group participants.

In general terms, the 3 groups of participants reduced the miss rate on target acquisition trials with the area cursor in comparison with the original cursor. With the cross cursor, in contrast, KU group participants increased the number of missed and erroneous trials.

### 5.2. *Cursor accuracy measures*

We used the following 6 features to measure effectiveness and efficiency of participants on pointing & clicking tasks with each cursor tested: movement time (MT), pointing time (PT), clicking time (CT), distance travelled (DT), curvature index (CI) and number of pauses (NP). These accuracy measurements have only been applied to interaction data gathered from target acquisition tasks, in order to study intended cursor movements of link selection and avoid unintended moves that appear during free navigation in searching tasks.

Firstly, measures about MT, PT, DT and NP were normalized as trials of the target acquisition task had different link sizes and distances to starting point. To do this, we divided results calculated for each trial by its index of difficulty (ID) (1) as mentioned before. Results concerning CI and CT were not normalized, since the former is a ratio that considers distances (travelled and straight-line between starting and ending points) of each trial, while the latter depends on pointing device and cursor variant used rather than on presentation of the GUI.

To compare the 3 cursors tested (original, area and cross) we studied performance of each participant separately, as well as average values achieved by each group of users: the keyboard users as pointing device (KU group), the joystick and trackball users (JU group), and the mouse users without MIs (MU group). In the following 6 subsections we address each cursor measure studied, comparing the average results achieved by each user group (Fig. 8), as well as analysing general trends on per-participant boxplots calculated for each user group (Fig. 9 for KU group, Fig. 10 for JU group, and Fig. 11 for MU group). Considering the small size of our sample (something generally inherent in studies involving people with disabilities), statistical analyses were avoided.

#### 5.2.1. *Movement time (MT)*

On average, the KU group improved the MT with both virtual cursors (Fig. 8 top left), achieving a slightly better result with the cross cursor than with the area cursor. Mean values with the cross and area cursor were 2571 ms/ID and 2618 (SD = 765 and 870) respectively, whereas 3306 (SD = 773) with the original cursor. The JU group, on average, achieved the best MT result with the area cursor (mean = 1442 ms/ID, SD = 458) followed by the original cursor (mean = 1644, SD = 484). The MU group, on average, got the same MT results with the area cursor (mean = 370 ms/ID, SD = 131) and the original cursor (mean = 370, SD = 167).

All participants from KU group achieved best median values of MT with the cross cursor (Fig. 9 top left), followed by the area cursor and the original cursor, second and third respectively for all the group participants. All JU group participants (Fig. 10 top left), and two of the six participants from MU group (MU2 and MU4 in Fig. 11 top left) got lowest median values of MT with the area cursor than with the original cursor.

#### 5.2.2. *Pointing time (PT)*

On average, the KU group got the best PT result (Fig. 8 top centre) with the cross cursor (1000 ms/ID, SD = 752), followed by the area cursor (1522, SD = 753) and the original cursor (2217, SD = 626). The JU group, on average, achieved the best PT result with the area cursor (mean = 744 ms/ID, SD = 381) followed by the original cursor (mean = 889, SD = 381). The MU group, on average, got similar PT results with the area cursor (mean = 165 ms/ID, SD = 97) and the

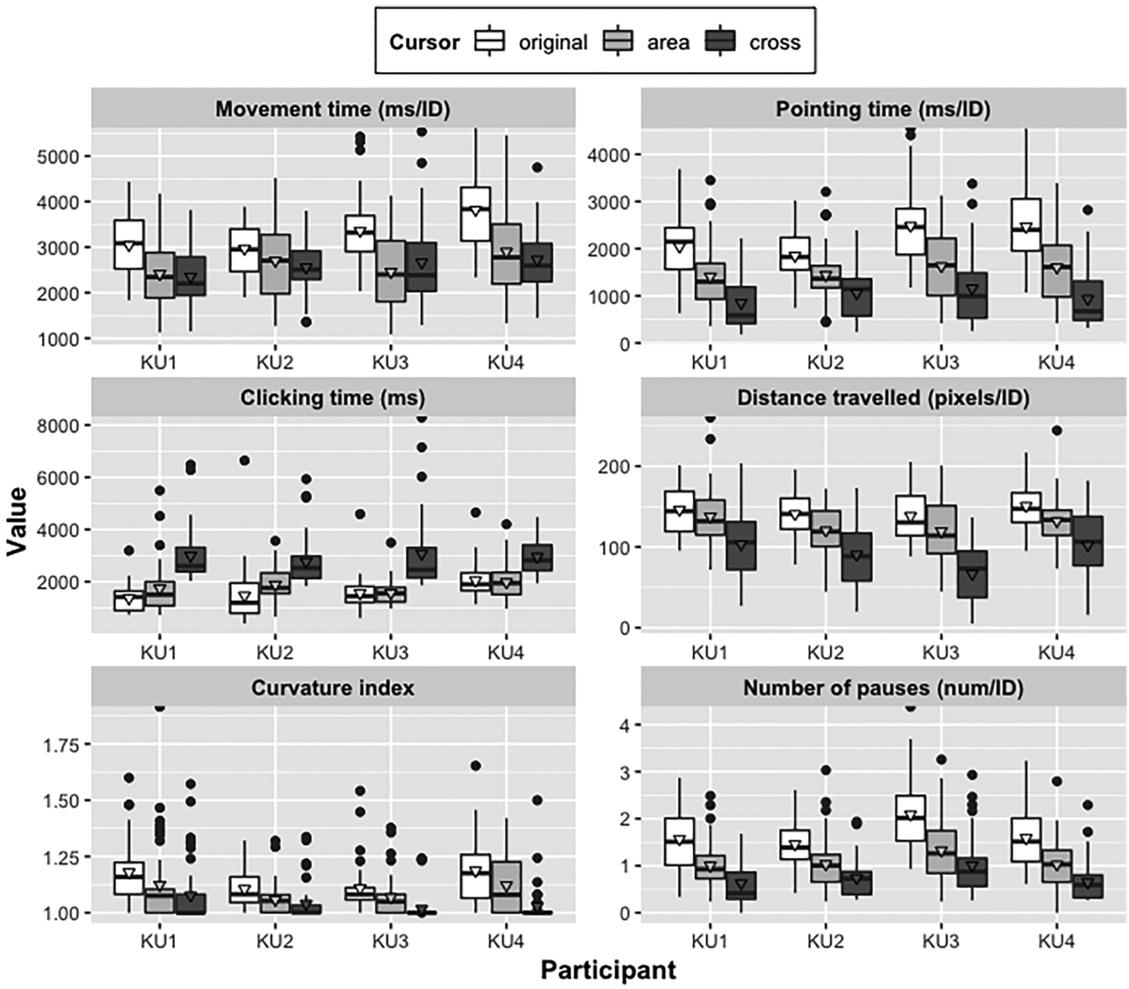


**Fig. 9.** Boxplots collection from KU group showing distribution of measures studied for each participant and cursor tested. Band and triangle inside each box represent the corresponding median and mean values respectively.

original cursor (mean = 172, SD = 78).

All participants of the KU group achieved best median PT values with the cross cursor (Fig. 9 top right), followed by the area cursor and the original cursor, second and third respectively for all the group participants. Four of the five JU group participants (JU1, JU2, JU4 and JU5 in Fig. 10 top right), and three of the six MU group participants (MU1, MU2 and MU4 in Fig. 11 top right), got lower median values of PT with the area cursor than with the original cursor.

### 5.2.3. *Clicking time (CT)*

On average, the KU group obtained the worst CT result (Fig. 8 top right) with the cross cursor (2945 ms, SD = 1032), followed by the area cursor (1791, SD = 717), and the original cursor (1618, SD = 773). For the JU group, a slightly better CT value was achieved with the area cursor (mean = 1380 ms, SD = 590) than with the original cursor (mean = 1491, SD = 593). The MU group got, on average, similar CT results with the area cursor (mean = 459 ms, SD = 269) and the original cursor (mean = 463, SD = 419).

According to median values of CT, all KU group participants obtained the worst results with the cross cursor (Fig. 9 middle left), whereas best results every time corresponded to the original cursor, followed by the area cursor. Four of the five JU group participants (JU1, JU2, JU4 and JU5 in Fig. 10 middle left), and three of the six MU group participants (MU2, MU4 and MU6 in Fig. 11 middle left) got best median values of CT with the area cursor than with the original cursor.

### 5.2.4. *Distance travelled (DT)*

On average, the KU group obtained the best DT result (Fig. 8 bottom left) with the cross cursor (mean = 91 px/ID, SD = 42), followed by the area cursor (mean = 128, SD = 35) and the original cursor (mean = 144, SD = 30). The JU group got almost identical DT results with the area cursor (mean = 175 px, SD = 64) and the original cursor (mean = 176 px, SD = 49). Also the MU group got almost identical DT results with the area cursor (mean = 95 px/ID, SD = 59) and the original cursor (mean = 92, SD = 54).

All participants from KU group obtained best median values of DT with the cross cursor (Fig. 9 middle right), followed by the area cursor and the original cursor, second and third respectively for all the group participants. Three of the five JU group participants (JU1, JU2 and JU5) got lower median values of DT with the area cursor than with the original, whereas JU4 got identical median values with both variants (Fig. 10 middle right). Three of the six MU group participants (MU1, MU2 and MU4) got lower median values of DT with the area cursor than with the original cursor (Fig. 11 middle right).

### 5.2.5. *Curvature index (CI)*

On average, the KU group obtained the best CI result (Fig. 8 bottom centre) with the cross cursor (mean = 1.04, SD = 0.1), followed by the area cursor (mean = 1.09, SD = 0.13) and the original cursor (mean = 1.15, SD = 0.13). The JU group got similar CI results with the area cursor (mean = 1.35, SD = 0.39) and the original cursor (mean = 1.36, SD = 0.37). The MU group, on average, achieved the

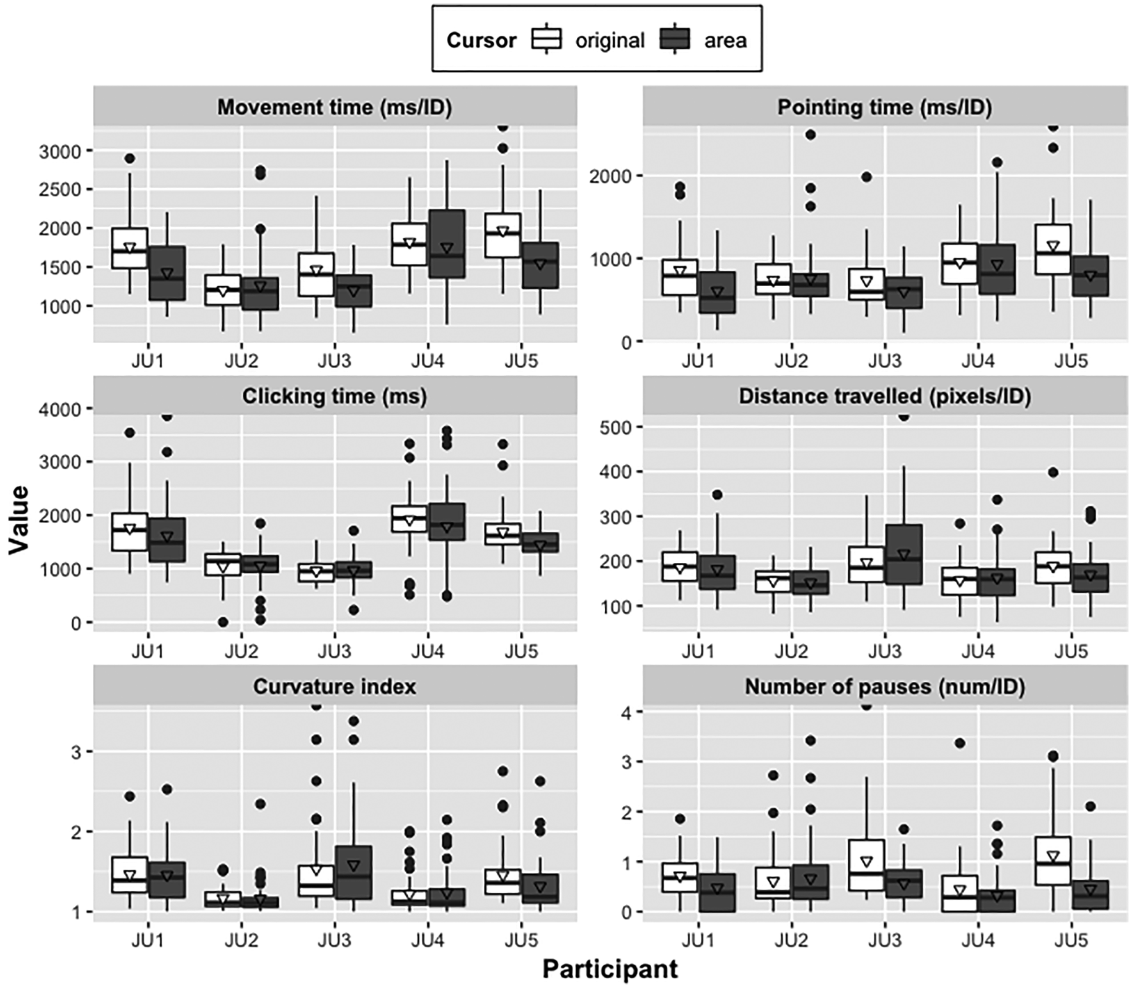


**Fig. 10.** Boxplots collection from JU group showing distribution of measures studied for each participant and cursor tested. Band and triangle inside each box represent the corresponding median and mean values respectively.

best CI result with the area cursor (mean = 1.08, SD = 0.18), followed by the original cursor (mean = 1.13, SD = 0.37).

All participants of the KU group achieved best median values of CI with the cross cursor (Fig. 9 bottom left), followed by the area cursor and the original cursor, second and third respectively for all the group participants. Two of the five JU group participants (JU4 and JU5) got better median values of CI with the area cursor than with the original, whereas JU2 got identical median values with both variants (Fig. 10 bottom left). Three of the six MU group participants (MU1, MU3 and MU4) got better median values of CI with the area cursor than with the original cursor (Fig. 11 bottom left), whereas the other 3 participants obtained same median values with both cursor variants.

#### *5.2.6. Number of pauses (NP)*

On average, the KU group achieved the best NP result (Fig. 8 bottom right) with the cross cursor (0.75 sum/ID, SD = 0.51), followed by the area cursor (1.1, SD = 0.62) and the original cursor (1.68, SD = 0.67). The JU group achieved the best NP result with the area cursor (mean = 0.5 sum/ID, SD = 0.51), followed by the original cursor (mean = 0.77, SD = 0.69). The MU group got a slightlty better NP result with the area cursor (mean = 0.12 sum/ID, SD = 0.21) than with the original cursor (mean = 0.15, SD = 0.2).

All KU group participants achieved best median values of NP with the cross cursor (Fig. 9 bottom right), followed by the area cursor and the original cursor, second and third respectively for all the group participants. Four of the five JU group participants (JU1, JU3, JU4 and JU5) got better median values of NP with the area cursor than with the original (Fig. 10 bottom right). One of the six MU group participants (MU2) got better median value of NP with the area cursor than with the original (Fig. 11 bottom right), whereas the other 5 participants obtained same median values with both cursor variants.

### *5.3. Participants satisfaction*

Fig. 12 includes bar graphs for each of the 8 categories of the satisfaction questionnaire, showing distribution of responses by each group of participants (KU, JU and MU) about cursor variants tested (original, area and cross).

The cross cursor obtained the worst results in the learnable category from JU and MU groups whereas it was highly rated by participants of the KU group. The area cursor obtained the best values from the JU group.

The original cursor obtained the best responses for the memorable category, which is not surprising as all the participants were already accustomed to it. However, the values obtained by both enhanced cursors are worthy of further attention, especially the ratings given to the cross cursor by the KU group participants and the ones given to the area cursor by the JU group participants.

Both groups of participants with disabilities (KU and JU) gave higher ratings to the area cursor in the accurate category. Nevertheless, half of the participants in the KU group preferred the cross cursor when asked to rank the cursor variants. Three participants without disabilities (MU group) also indicated their preference for the area cursor in the accurate category.

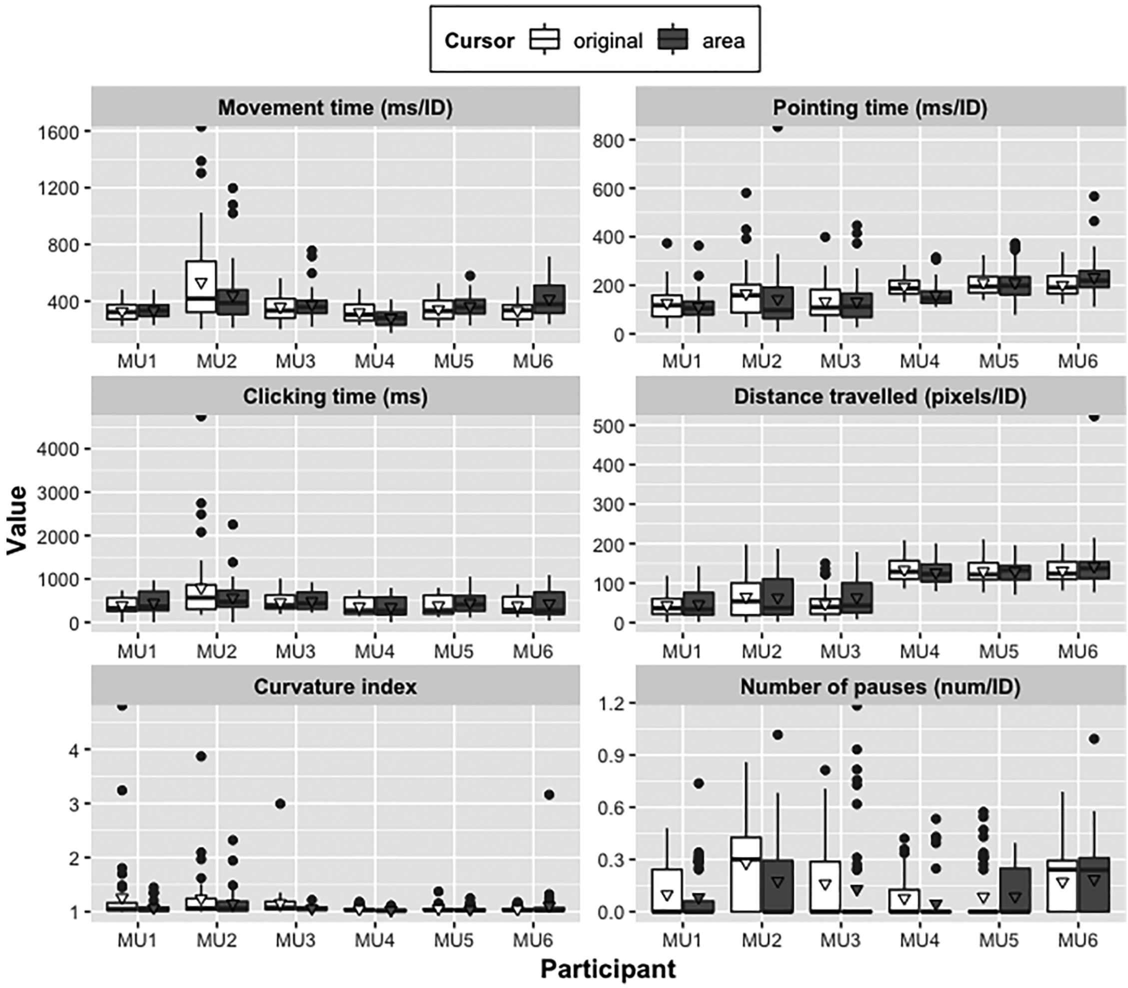


**Fig. 11.** Boxplots collection from MU group showing distribution of measures studied for each participant and cursor tested. Band and triangle inside each box represent the corresponding median and mean values respectively.

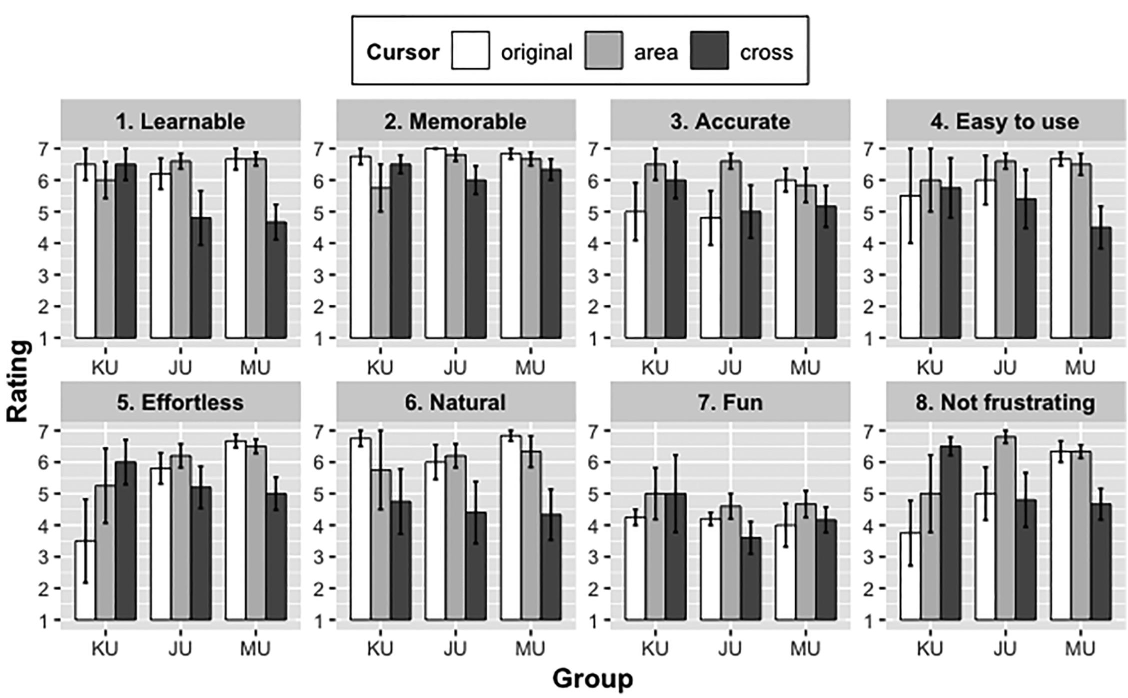


**Fig. 12.** Average ratings by users group for each cursor tested (Likert scale from 1 – strongly negative to 7 – strongly positive). Error bars represent ± 1 standard error (SE).

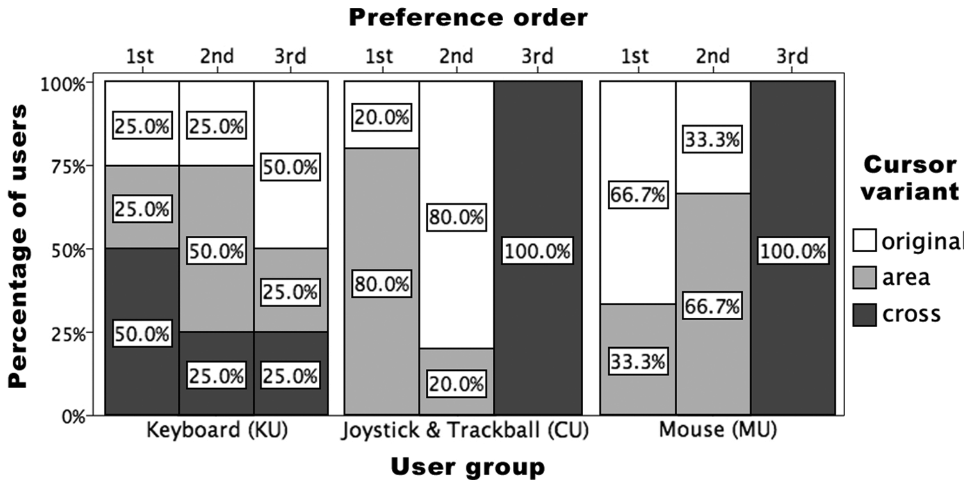


**Fig. 13.** Ranking by user group for preferred cursor variant for web browsing.

Regarding the easy to use category, the area cursor is the best rated option by participants with disabilities. The cross cursor is the last choice in the ranking for participants from JU and MU groups, whereas the first one for half of participants in the KU group.

Values obtained for the effortless category show a clear preference for the area cursor from the JU group, for the cross cursor from the KU group and for the original cursor from the MU group. However, the area variant also obtained high values from participants without disabilities.

The original cursor is highly rated in the natural category by the different user groups. Half of the participants in the KU group ranked the cross cursor second, whereas it was the last ranked option for participants in the other groups.

Participants felt some insecurity when rating the fun category and values given to the cursor variants did not differ significantly. However, the area cursor obtained the best results from the three groups, as it was something new they were trying and they found it to be a friendly cursor option. Regarding the rankings obtained in this category, participants in the KU group ranked the cross cursor as the first option (75%) followed by the area cursor (50%). 50% of the JU group selected the area cursor as the first option whereas the cross cursor was more often the lowest ranked option (80%). Participants in the MU group mainly selected the area cursor in first place and the cross cursor as last option.

There are clear differences between values given by participants with disabilities in the not frustrating category. The cross cursor was highly rated by users in the KU group (6.5), the area cursor was the one obtaining best values from users in the JU group (6.8), and the area and original cursors obtained the same mean value (6.3) for users without disabilities. Regarding the ranking of cursor variants in this category, the original cursor was the last option for 75% of the KU group and the cross cursor was the lowest ranked option for the entire JU group.

Fig. 13 shows the overall values of cursor preferences by user group. As can be seen, 50% of participants in the KU group preferred the cross cursor, 80% of participants in the JU group preferred the area cursor and 66.7% of participants in the MU group preferred the original cursor.

## 6. Discussion

### *6.1. Keyboard users group (KU)*

Results presented in Figs. 8 and 9 show that the KU group participants improved almost every performance measure studied (except the clicking time) using either of the 2 virtual cursors tested (cross or area) instead of the original cursor. In addition, all participants of the KU group obtained, on average, better results with the cross cursor than with the area cursor for every measure (except CT), highlighting the pointing time, distance travelled, curvature index, and number of pauses. These results proved that the cross cursor was the most beneficial variant for the KU group participants on target acquisition tasks, reducing both distance and difficulties along cursor trajectory to a greater extent than the area cursor.

As expected, the 4 participants of the KU group worsen the CT using the cross cursor instead of the original or area cursor (Fig. 9 middle left). Unlike these 2 cursors (original and area) with which keyboard users always use the same key to select a link, the cross cursor requires to keystroke the appropriate letter to leverage the virtual cursor assistance. This way of interacting with the cross cursor entailed larger CT compared to the other 2 variants tested, but as results show, all KU group participants still achieved better total MT results (which also includes the CT) with the cross cursor (Fig. 9 top left).

Considering that the cross cursor implements a novel target selection mechanism for the participants, and that experimental sessions lasted no more than 2 h, we think that keyboard users could improve their performance with this virtual cursor after a longer learning period. For instance, Fig. 14 shows how KU group participants performed differently on target acquisition tasks, when pointing with the original cursor (Fig. 14 top row) compared with both virtual cursors (area – middle row and cross – bottom row). It is striking that even though participants KU1, KU2 and KU4 leveraged the cross cursor selecting links from distance (Fig. 14, 1st, 2nd and 4th column, bottom row), each of them carried on pointing towards targets similarly as with the original and area cursors. By contrast, participant KU3 was able to reduce travelled distances even further with the cross cursor, by changing his pointing behaviour with respect to the original cursor and aiming across the targets' width with only horizontal cursor movements (Fig. 14, 3rd column, bottom row). We can affirm that participant KU3 was able to point with the cross cursor optimally and travel shorter paths (Fig. 9 middle right), and that other participants from this group might approach this method if they were provided with a longer learning period.

The study of invalid trials filtered from target acquisition task (Table 4) provided additional insights about the virtual cursors tested. For the KU group, the number of missed and erroneous trials increased with the cross cursor (9.9%) compared to the original (5.2%) and area cursor (4.7%). This was expected, as target selection based on shortcuts that implements the cross cursor is more cognitively and physically demanding than the other 2 variants tested, and thus error-prone. The highest rate for the KU group corresponded to misses with the cross cursor (5.7%), in comparison with the original (3.6%) and area cursor (2.6%). Missed selections in the Web can bring the user unintentionally to unwanted web pages, making recovery back to the original page particularly tedious and time consuming for people with MIs. Despite

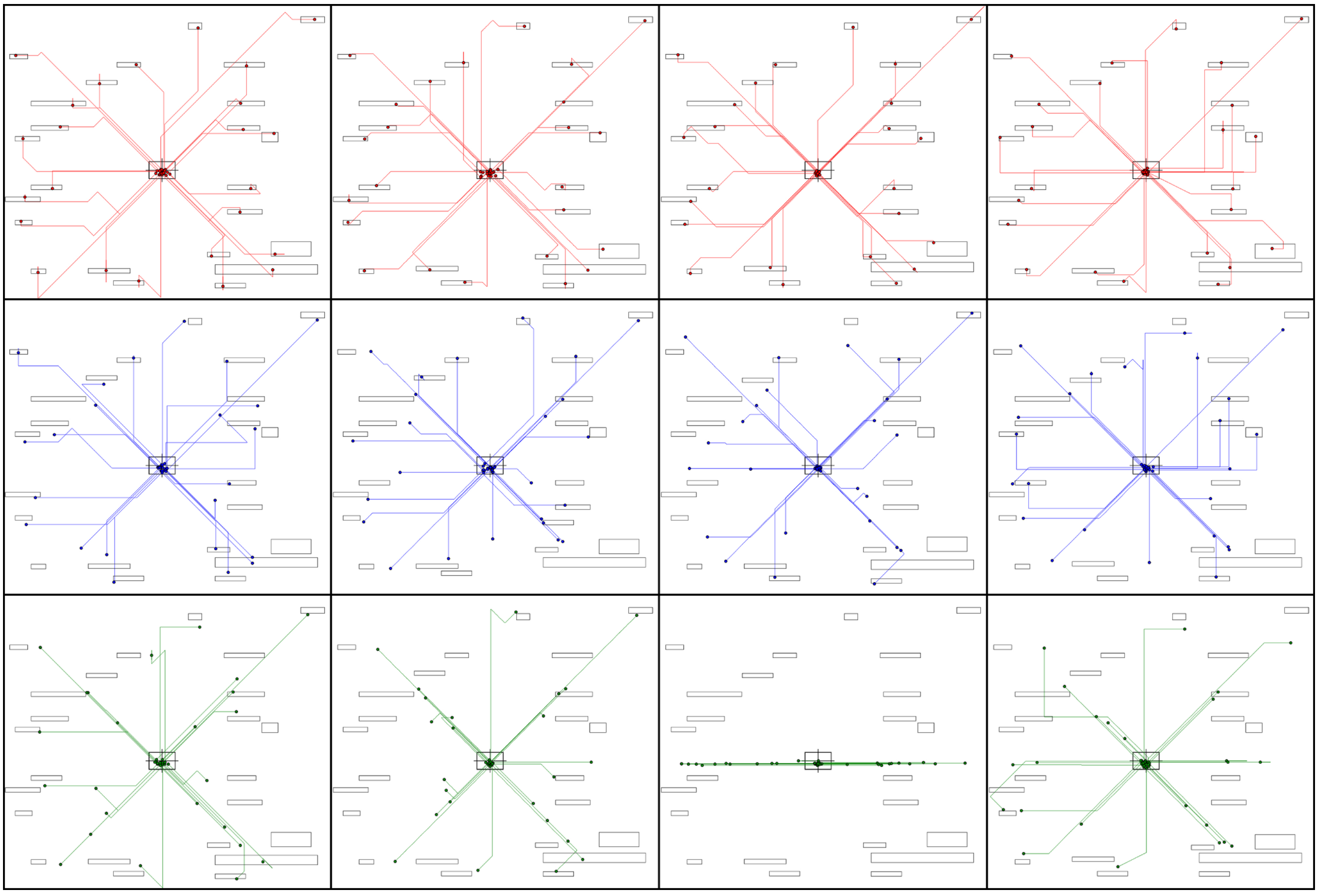

**Fig. 14.** Cursor trajectories of each KU group participant (each column corresponds to a user: KU1 to KU4 from left to right) with cursors studied (top row – original, middle – area and bottom – cross) in target acquisition trials of the first website (Discapnet).

these results about missed trials, the KU group participants positively assessed the cross cursor and, in general terms, felt comfortable with this new virtual cursor. Although the miss rate obtained by the KU group with the cross cursor was not high, we think that keyboard users could also improve at this point with a longer period of use of the virtual cursor. As pointed out before, the lowest miss rate of the KU group corresponded to the area cursor. Even though KU group participants did not achieved the best results with the area cursor, further research should study an improved version of the cross cursor with a larger than normal activation area.

According to subjective assessments provided by KU group participants, the cross cursor was, on average, the preferred option for browsing the Web from among the 3 variants tested (Fig. 13). Despite that, one keyboard participant (KU2) ranked the cross cursor as the least useful variant. She mentioned not remembering she could click distant links by automatically provided keystrokes. Nevertheless, she affirmed that she found the cross cursor very useful and was of the opinion that it could be beneficial to her in the long term once she got accustomed to it. The cross cursor also obtained the best average subjective assessments from keyboard participants (Fig. 12) for significant categories such as effortless and not frustrating, as well as generally good average scores for the remaining categories.

Participant KU2 declared, *"with the original cursor it is difficult to aim at small targets, whereas with the area variant this problem was reduced. Although it is easier to aim targets with the area variant, you have to pay attention on which link is highlighted to leverage this virtual cursor"*. Participant KU3 said he *"would like to use the cross variant as you do not have to move the cursor so much, you only have to press one letter and that's all"*. Participant KU4 declared *"I do not like the area variant as I found issues when links are close together"*. On the contrary, KU4 said she *"would use the cross variant, although I found tiring to select each time a shortcut letter and I would need time to get used to"*.

### 6.2. Joystick and trackball users group (JU)

The JU group improved, on average, all cursor measures with the area cursor compared to the original (Fig. 8), although larger benefits corresponded to movement time, pointing time, clicking time and number of pauses. According to the distribution of results per-participant (Fig. 10), all 5 JU group participants improved the MT with the area cursor, whereas only 4 participants improved the PT, CT and NP, 3 improved the DT, and 2 improved the CI. In addition, these results show that only one participant of the JU group (JU5) improved, on average, all cursor measures with the area cursor compared to the original cursor.

According to subjective assessments provided by JU group participants (Fig. 12), the area cursor received, on average, better scores than the other 2 variants tested for 7 out of 8 categories (the exception being memorable). Best results corresponded to the categories accurate, effortless and not frustrating, obtaining, on average, bigger differences with respect to the original cursor. As for the preferred variant for web browsing (Fig. 13), 4 out of 5 participants from the JU group rated the area cursor as their favourite cursor. Concerning the cross cursor, it was unanimously selected as the least preferred variant for web browsing by all 5 participants from the JU group (Fig. 13), as well as receiving, on average, the worst rating for all satisfaction categories (Fig. 12). In addition, only one participant from the JU group (JU1) used the cross cursor shortcuts to complete target acquisition task (Table 2).

Results on performance and satisfaction of participants from the JU group revealed that the area cursor, initially proposed to assist joystick and

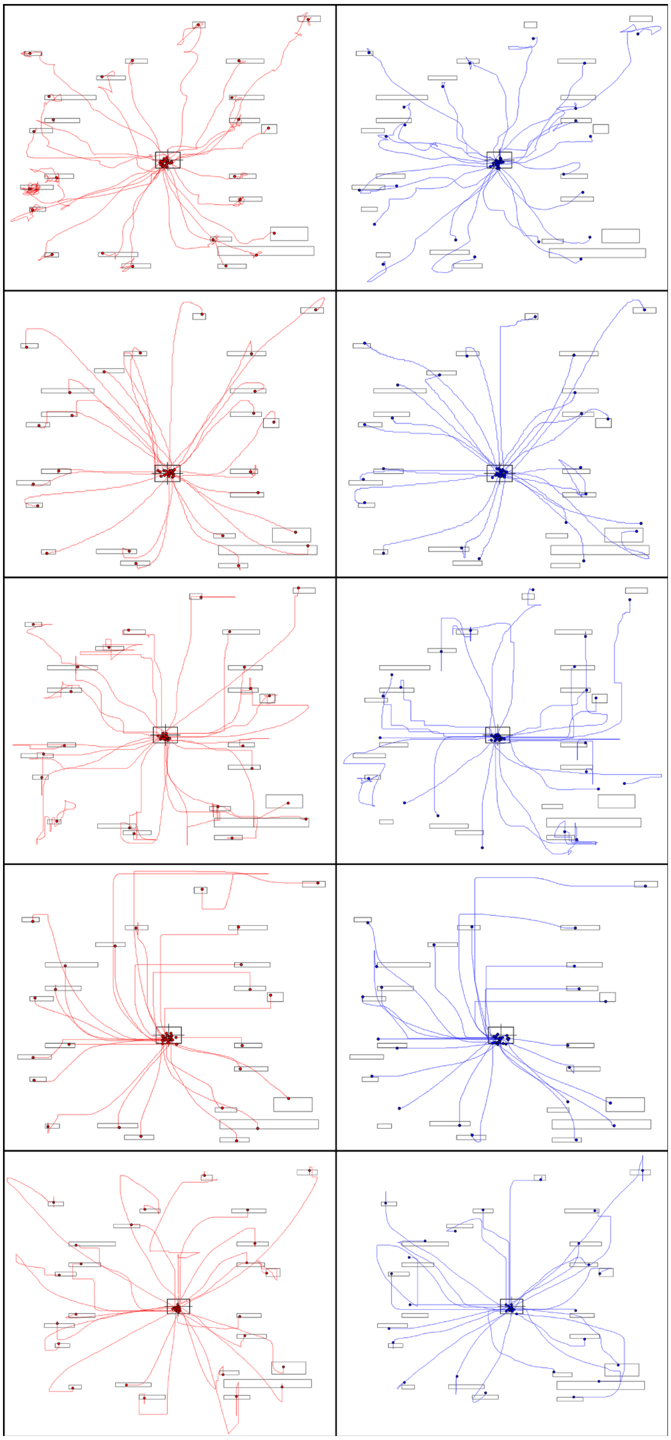

**Fig. 15.** Cursor trajectories of each participant of the JU group (each row corresponds to a user: JU1 to JU5 from top to bottom) with cursors studied (left – original and right – area) in target acquisition trials of the first website (Discapnet).

trackball users, was the most beneficial and preferred variant on average for this group of participants. Despite positive results with the area cursor, participants from the JU group were able to take advantage of this cursor variant, on average, on around half of the target acquisition trials (51.7%) by clicking outside of target link (Table 3). Furthermore, the JU group participants leverage the area cursor assistance unevenly, with user JU1 achieving the highest rate (81.3%) and JU4 the lowest (27.1%).

The right column of Fig. 15 shows how the 5 participants from the JU group performed differently with the area cursor. While participants JU1 and JU3 (Fig. 15, 1st and 3rd row, right) selected the majority of links by clicking outside the targets, JU2 and JU4 (Fig. 15, 2nd and 4th row, right) behaved the opposite. Participants JU2 and JU4 followed a similar point and click behaviour with both the original and the area cursors (Fig. 15, 2nd and 4th row), and on many occasions kept on aiming for the target link with the virtual cursor despite this being unnecessary. By contrast, JU1, JU3 and JU5 were able to leverage the area cursor (Fig. 15, 1st, 3rd and 5th row, right) and mitigate difficulties around the target (Fig. 15, 1st, 3rd and 5th row, left). Considering that the area cursor implements a novel target selection mechanism for participants, and that sessions of this study lasted no more than 2 h, we think that joystick and trackball users could improve their performance with this virtual cursor with the benefit of a longer learning period.

Moreover, several participants stated during interviews that pointing with the area cursor was more difficult when various links were close to the target link, which reduced the cursor activation area and forced them to click within the target area. These distractors around the target reduce the area cursor assistance the closer they are to the target link, resulting in the unassisted original cursor behaviour if the target link is surrounded by distractors at a minimum distance. As information about distractors was not gathered in the interaction log this time, we were not able to take the distractor factor into account in this study when analysing participant performance. In order to consider distractors it would be necessary to have information about their sizes and locations, in a similar way as we do here for the target link.

For participant JU1 *"it was more fun to use the area variant than the other two cursor tested, since this required less precision to select links"* and *"the area variant was also less tiring to use than the original cursor"*. Participant JU1 said about the area cursor that: *"it was easy to use, but it is necessary to be alert on the highlighted link in order to leverage this assistance"*. Participant JU5 declared *"the area cursor was easy to use, and useful for people like us with motor impairments"*. Participant JU4, although she had no significant improvement with the area cursor, declared about the original cursor: *"I usually have difficulties aiming to small links"*.

Regarding the invalid trials filtered from target acquisition task (Table 4), the JU group reduced, on average, the number of missed trials with the area cursor (4.2%) compared to the original cursor (7.5%), which is very desirable to assist link selection to this group of people with MIs. However, to enable more users with MIs to benefit from this assistance, further research should study how to highlight target at reach more conveniently to make it more perceivable for users, and to reduce the cognitive tiredness after prolonged use.

### 6.3. Mouse users group (MU)

As expected, participants without MIs (MU group) were those that perceived less improvement or deterioration in performance when using the area cursor instead of the original cursor (Fig. 8 and Fig. 11). The average results of the MU group (Fig. 8) show that the performance with the area cursor improved compared to the original cursor only for the curvature index, whereas for the rest of cursor measurements, results obtained with both cursor variants (original and area) were very similar. According to the distribution of results per-participant (Fig. 11), 3 of the 6 MU group participants improved with the area cursor in comparison with the original cursor on the PT, CT, DT and CI, whereas 2 participants improved the MT, and only 1 improved the NP.

On average, the area and the original cursor received similar subjective assessments from MU group participants (Fig. 12), although with a slight preference for the latter. The original cursor was selected by 4 of the 6 MU participants as the preferred variant for web browsing (Fig. 13), followed by the area cursor. On the other hand, the cross cursor was unanimously selected as the least preferred variant by all 6 participants from the MU group (Fig. 13), and, on average, received worse ratings for all satisfaction categories except for the fun category

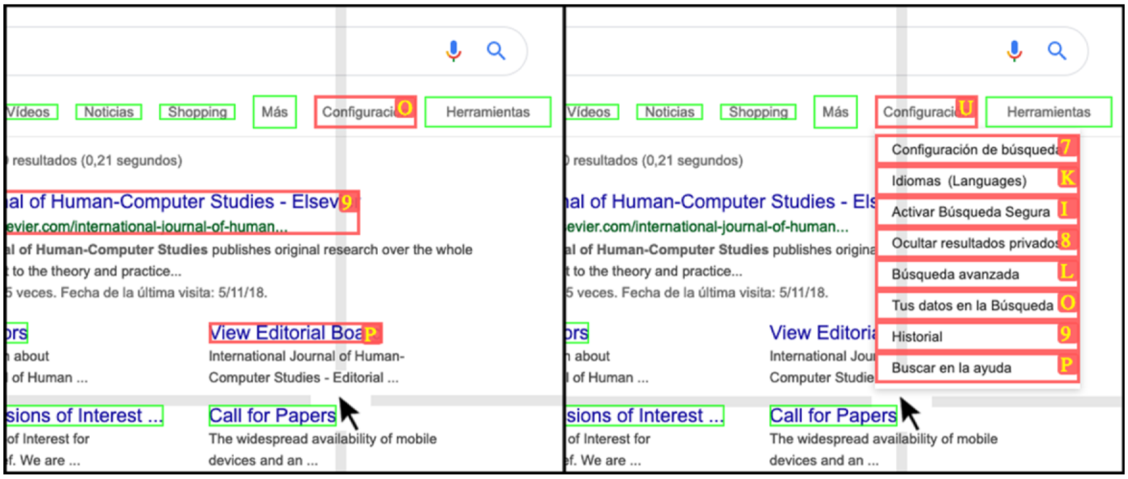


**Fig. 16.** Accessing the content of a drop-down menu with the cross cursor, without additional mouse movements (left – menu hidden before selection, right – menu expanded after selection).

(Fig. 12). Its keyboard related interaction conflicted with fluent mouse usage, and it was mainly avoided for this reason.

According to participants' comments during following interviews, they proposed some improvements for virtual cursors. Participant MU2 declared about the area cursor that she *"did not like how links were highlighted, and this made me tired during web navigation"*. Participant MU5 said about the area cursor that he *"would reduce its activation area to select this variant as favourite"*, and that *"this assistance was useful for small and isolated links, but not for close links"*. About the cross cursor, MU5 stated, *"it is distracting both the letters displayed next to links and how links are highlighted"*.

Regarding the invalid trails filtered from target acquisition task (Table 4), the MU group of participants without MIs surprisingly reduced the number of missed trials with the area cursor (1.4%) compared to the original cursor (7.6%), getting even greater improvement than the other 2 groups of participants with MIs. In some cases (MU3, MU4 and MU5), the accuracy improvement was achieved by leveraging the area cursor assistance. In other cases (MU6), in contrast, the accuracy improvement was not achieved by leveraging the area cursor assistance, but by reducing the speed of cursor movement.

### 6.4. Other practical applications of the virtual cursors

In addition to assist standard link selection, we implemented both virtual cursors (area and cross) to handle further web content. In this work we only focused on evaluating cursor variants on standard target acquisition tasks, but the following practical applications on the Web can be of interest for people with MIs.

The use of vertical drop down menus within navigation bars has been generalized on the Web, as these allow organizing and accessing the content of a website. However, interacting with these moving menus can be difficult for people with MIs. In this regard, the cross cursor was implemented to allow accessing navigation bars (Fig. 16), which involves handling with hidden and overlapping content, as well as the generation of appropriate shortcuts.

Selecting web forms elements such as radio buttons, check boxes or text boxes can also be challenging for people with MIs due to their small size. Although not all these web elements were implemented for this work, both virtual cursors (area and cross) could handle all them to assist access to web forms to people with MIs.

## 7. Conclusions and future work

The proposed virtual cursors (the cross and area cursors) proved to be beneficial for participants with MIs. Results obtained in the experimental sessions showed improvements in their performance when selecting links on web interfaces. Participants using a keyboard as an alternative pointing device (KU group), benefited from both virtual cursors to select links, although as expected, the best results were achieved with the cross cursor. On the other hand, users of specific alternative pointing devices such as the joystick or trackball (JU group), only benefited from the area cursor while the cross cursor was unanimously rated as the least preferred, behind the original cursor. These results are promising despite the limited number of participants in the study and support the idea that people with MIs need personalized adaptations in order to assist point and click interactions in GUIs, such as, for example, to access the Web.

However, considering the novelty of these virtual cursors for participants of the study and the duration of experimental sessions, we believe that with a longer learning period both groups of users with MIs would improve their performance. Therefore, we plan to conduct a longitudinal study, in which keyboard users on one hand and joystick and trackball users on the other, regularly perform point and click tasks with the cross and area cursor respectively, as well as with the original cursor. Besides analysing the learning effect on usage of each virtual cursor variant, we also seek to understand how the presence of distractor links around the target may influence performance (especially with the area cursor). It will also be valuable to see if participant satisfaction with virtual cursors declined or not over time. In addition, we will endeavour to recruit more participants in order to get more reliable insights.

## Declaration of Competing Interest

None.

## References

Accot, J., Zhai, S., 2003. Refining Fitts' law models for bivariate pointing. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 193–200. https://doi.org/10.1145/642611.642646.

Almanji, A., Davies, T.C., Stott, N.S., 2014. Using cursor measures to investigate the effects of impairment severity on cursor control for youths with cerebral palsy. Int. J. Hum. Comput. Stud. 72 (3), 349–357. https://doi.org/10.1016/j.ijhcs.2013.12.003.

Arrue, M., Valencia, X., Pérez, J.E., Moreno, L., Abascal, J., 2018. Inclusive web empirical studies in remote and in-situ settings: a user evaluation of the remotest platform. Int. J. Hum. 1–16. https://doi.org/10.1080/10447318.2018.1473941.

Bidasoa. https://www.bidasoaturismo.com/ (Accessed July 2015).

Brooke, J., 1996. SUS: a 'Quick and Dirty' usability scale. In: Jordan, P.W., Thomas, B., McClelland, I.L., Weerdmeester, B. (Eds.), Usability Evaluation in Industry. CRC Press, pp. 189–194.

Cook, A.M., Polgar, J.M., 2014. Assistive Technologies: Principles and Practice, fourth ed. Elsevier Health Sciences.

Discapnet. https://www.discapnet.es/ (accessedJuly 2015).
Felzer, T., MacKenzie, I.S., Magee, J., 2016. Comparison of two methods to control the mouse using a keypad. In: International Conference on Computers Helping People with Special Needs. Cham. Springer, pp. 511–518. https://doi.org/10.1007/978-3-319-41267-2_72.
Findlater, L., Jansen, A., Shinohara, K., Dixon, M., Kamb, P., Rakita, J., Wobbrock, J.O., 2010. Enhanced area cursors: reducing fine pointing demands for people with motor impairments. In: Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology. ACM, pp. 153–162. https://doi.org/10.1145/1866029.1866055.
Gipuzkoa. https://www.gipuzkoa.eus (Accessed July 2015).
Grossman, T., Balakrishnan, R., 2005. The bubble cursor: enhancing target acquisition by dynamic resizing of the cursor's activation area. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 281–290. https://doi.org/10.1145/1054972.1055012.
Gump, A., LeGare, M., Hunt, D.L., 2002. Application of Fitts' law to individuals with Cerebral Palsy. Percept. Mot. Skills 94 (3 Pt 1), 883–895 https://doi.org/10.2466%2Fpms.2002.94.3.883.
Harada, S., Landay, J.A., Malkin, J., Li, X., Bilmes, J.A., 2006. The vocal joystick: evaluation of voice-based cursor control techniques. In: Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility. ACM, pp. 197–204. https://doi.org/10.1145/1168987.1169021.
Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Advances in Psychology 52. North-Holland, pp. 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9.
Hwang, F., Keates, S., Langdon, P., Clarkson, J., 2004. Mouse movements of motion-impaired users: a submovement analysis. ACM SIGACCESS Accessibility and Computing. ACM, pp. 102–109. No. 77-78. https://doi.org/10.1145/1029014.1028649.
Hwang, F., Keates, S., Langdon, P., Clarkson, P.J., 2003. Multiple haptic targets for motion-impaired computer users. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 41–48. https://doi.org/10.1145/642611.642620.
Kabbash, P., Buxton, W.A., 1995. The "prince" technique: fitts' law and selection using area cursors. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM Press/Addison-Wesley Publishing Co., pp. 273–279. https://doi.org/10.1145/223904.223939.
Keates, S., Hwang, F., Langdon, P., Clarkson, P.J., Robinson, P., 2002. The use of cursor measures for motion-impaired computer users. Univers. Access Inf. Soc. 2 (1), 18–29. https://doi.org/10.1007/s10209-002-0033-7.
MacKenzie, I.S., Kauppinen, T., Silfverberg, M., 2001. Accuracy measures for evaluating computer pointing devices. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 9–16. https://doi.org/10.1145/365024.365028.
Mott, M.E., Wobbrock, J.O., 2014. Beating the bubble: using kinematic triggering in the bubble lens for acquiring small, dense targets. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 733–742. https://doi.org/10.1145/2556288.2557410.
Payne, A.R., Plimmer, B., McDaid, A., Luxton-Reilly, A., Davies, T.C., 2016. Expansion cursor: a zoom lens that can be voluntarily activated by the user at every individual click. In: Proceedings of the 28th Australian Conference on Computer-Human Interaction. ACM, pp. 81–90. https://doi.org/10.1145/3010915.3010942.
Payne, A.R., Plimmer, B., McDaid, A., Luxton-Reilly, A., Davies, T.C., 2017. How can adding a movement improve target acquisition efficacy? In: IFIP Conference on Human-Computer Interaction. Cham. Springer, pp. 496–514. https://doi.org/10.1007/978-3-319-67687-6_34.
Pérez, J.E., Valencia, X., Arrue, M., Abascal, J., 2016. A usability evaluation of two virtual aids to enhance cursor accessibility for people with motor impairments. In: Proceedings of the 13th Web for All Conference. ACM, pp. 20. https://doi.org/10.1145/2899475.2899489.
Pérez, J.E., Valencia, X., Arrue, M., Abascal, J., 2015. Depicting the keypad mouse interface: exploring pointing behaviors of three input devices. In: Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility. ACM, pp. 413–414. https://doi.org/10.1145/2700648.2811396.
Pérez, J.E., Arrue, M., Valencia, X., Moreno, L., 2014. Exploratory study of web navigation strategies for users with physical disabilities. In: Proceedings of the 11th Web for All Conference. ACM, pp. 20. https://doi.org/10.1145/2596695.2596715.
Pierson Stachecki, L., Magee, J., 2017. Predictive link following plug-in for web browsers. In: Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility. ACM, pp. 333–334. https://doi.org/10.1145/3132525.3134776.
Rao, R.S., Seliktar, R., Rahman, T., 2000. Evaluation of an isometric and a position joystick in a target acquisition task for individuals with Cerebral Palsy. IEEE Trans. Rehab. Eng. 8 (1), 118–125. https://doi.org/10.1109/86.830956.
Salivia, G., Hourcade, J.P., 2013. PointAssist: assisting individuals with motor impairments. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 1213–1222. https://doi.org/10.1145/2470654.2466157.
Sears, A., Young, M., Feng, J., 2008. Physical disabilities and computing technologies: an analysis of impairments. In: Jacko, J.A., Sears, A. (Eds.), The Human-Computer Interaction Handbook. Lawrence Erlbaum Associates, New York, pp. 829–852.
Trewin, S., 2008. Physical impairment. In: Harper, S., Yesilada, Y. (Eds.), Web Accessibility: A Foundation for Research. Springer Science & Business Media, London, pp. 37–46. https://doi.org/10.1007/978-1-84800-050-6_4.
Trewin, S., Keates, S., Moffatt, K., 2006. Developing steady clicks: a method of cursor assistance for people with motor impairments. In: Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility. ACM, pp. 26–33. https://doi.org/10.1145/1168987.1168993.
Trewin, S., Pain, H., 1999. Keyboard and mouse errors due to motor disabilities. Int. J. Hum. Comput. Stud. 50 (2), 109–144. https://doi.org/10.1006/ijhc.1998.0238.
Valencia, X., Pérez, J.E., Arrue, M., Abascal, J., Duarte, C., Moreno, L., 2017. Adapting the web for people with upper body motor impairments using touch screen tablets. Interact. Comput. 29 (6), 794–812. https://doi.org/10.1093/iwc/iwx013.
Valencia, X., Pérez, J.E., Muñoz, U., Arrue, M., Abascal, J., 2015. Assisted interaction data analysis of web-based user studies. Human-Computer Interaction. Springer, Cham, pp. 1–19. https://doi.org/10.1007/978-3-319-22701-6_1.
Wobbrock, J.O., Fogarty, J., Liu, S.Y.S., Kimuro, S., Harada, S., 2009. The angle mouse: target-agnostic dynamic gain adjustment based on angular deviation. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 1401–1410. https://doi.org/10.1145/1518701.1518912.
Wobbrock, J.O., Gajos, K.Z., 2008. Goal crossing with mice and trackballs for people with motor impairments: performance, submovements, and design directions. ACM Trans. Access. Comput. (TACCESS) 1 (1), 4. https://doi.org/10.1145/1361203.1361207.
Wobbrock, J.O., Gajos, K.Z., 2007. A comparison of area pointing and goal crossing for people with and without motor impairments. In: Proceedings of the Conference on Computers and Accessibility (ASSETS '07). New York. ACM Press, pp. 3–10. https://doi.org/10.1145/1296843.1296847.
Worden, A., Walker, N., Bharat, K., Hudson, S., 1997. Making computers easier for older adults to use: area cursors and sticky icons. In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 266–271. https://doi.org/10.1145/258549.258724.
World Health Organization (WHO), 2001. International Classification of Functioning, Disability and Health: ICF. WHO, Geneva.

**J. Eduardo Pérez** received the B.Sc. and the M.Sc. degrees in Computer Science from the University of the Basque Country (UPV/EHU), in 2007, and 2011, respectively. He is currently a PhD candidate and member of the Egokituz Laboratory of HCI for Special Needs at the UPV/EHU. His research interests focuses on studying web navigation behaviours within diverse groups of users with special needs and testing different GUI enhancements in order to improve their web browsing experience.

**Xabier Valencia** received the M.Sc. and the Ph.D. degrees in Computer Science from the University of the Basque Country (UPV/EHU), in 2008, and 2017 respectively. He is currently a researcher at the Egokituz Laboratory of HCI for Special Needs (UPV/EHU). His main research interest is focused on providing higher accessibility to the Web to people with diverse capabilities, principally applying transcoding techniques.

**Myriam Arrue** received the M.Sc. and the Ph.D. degrees in Computer Science from the University of the Basque Country (UPV/EHU), in 1999, and 2008 respectively. She is currently an Assistant Professor at the UPV/EHU and a member of Egokituz Laboratory of HCI for Special Needs. Her research focuses on web accessibility evaluation, web navigation behaviour analysis, and the application of transcoding techniques for improving UX in web interactions.

**Julio Abascal** BSD in Physics (U. de Navarra, 1978) and Ph.D. in Informatics (University of the Basque Country-Euskal Herriko Unibersitatea, 1987). He is currently a Professor of the Computer Architecture and Technology Department of the UPV/EHU, where he has worked since 1981. In 1985 he co-founded Egokituz Laboratory of Human-Computer Interaction for Special Needs, which has participated in several R&D project at national and international level.

## 3.6. Appendix 6

This appendix corresponds to the article entitled *"Longitudinal Study of Two Virtual Cursors for People with Motor Impairments: A Performance and Satisfaction Analysis on Web Navigation"* that was published in the international journal IEEE Access in June 2020. This journal reached an impact factor of 3.745 in 2019 according to Journal Citation Reports, ranking 35$^{th}$ out of 156 journals (Q1) in the "Computer Science, Information Systems" category.

In order to explore the long-term benefits of both cursor aids presented in the previous appendix, we carried out a longitudinal study on the Web with users with motor impairments. We conducted a six-week remote study in which eight users of alternative pointing devices participated from home, obtaining in this way, interaction data from everyday web navigation in an unsupervised fashion. The objective was to study their performance with both virtual cursors in comparison with the original cursor and the influence of distractors (i.e., other nearby links) on target acquisition on the Web, the learning effect on performance with each cursor variant, adoption of virtual assistance in free navigation, and participant satisfaction with the cursor aid being tested.

Received May 11, 2020, accepted June 8, 2020, date of publication June 11, 2020, date of current version June 25, 2020.
*Digital Object Identifier 10.1109/ACCESS.2020.3001766*

# Longitudinal Study of Two Virtual Cursors for People With Motor Impairments: A Performance and Satisfaction Analysis on Web Navigation

**J. EDUARDO PÉREZ, MYRIAM ARRUE, XABIER VALENCIA, AND JULIO ABASCAL, (Member, IEEE)**
EGOKITUZ Laboratory of HCI for Special Needs, Informatika Fakultatea, University of the Basque Country (UPV/EHU), 20018 Donostia, Spain
Corresponding author: J. Eduardo Pérez (juaneduardo.perez@ehu.eus)

This work was supported in part by the Department of Education, Universities and Research of the Basque Government (Adian Research Group, IT980-16), and in part by the Ministry of Economy and Competitiveness of the Spanish Government co-founded by the ERDF (PhysComp project, TIN2017-85409-P). J. Eduardo Pérez held a Ph.D. scholarship from the UPV/EHU (PIF13/284) when the study was conducted.

**ABSTRACT** The lack of dexterity in the upper limbs of people with motor impairments may prevent the use of standard pointing devices, such as mice, to access graphical user interfaces. In these cases, pointing and clicking are usually performed by means of alternative devices such as joysticks, trackballs or standard keyboards. However, target acquisition can still be challenging for this group of people due to their physical condition. Based on previous works, we developed two virtual cursors: the novel cross cursor and the standard area cursor. They are devoted to assist two different groups of users with link selection within web pages: keyboard-only users, and joystick and trackball users, respectively. Both virtual cursors have been evaluated and compared with the original unassisted cursor in a longitudinal study. Eight people with motor impairments participated in an unsupervised experiment from their own personal computers at home. For a period of six weeks, each participant used both a virtual cursor and the original unassisted cursor to freely navigate the Web, and to perform predefined target acquisition tasks. Interaction data was automatically logged throughout the study along with subjective assessments concerning the usability of the virtual cursor being tested. Results show significant improvements for both virtual cursors in six of the seven cursor parameters studied, albeit with performance variations between some participants. The virtual cursors were extensively used for free web navigation and in their subjective assessments both were positively endorsed by participants who also put forward improvement suggestions for future developments.

**INDEX TERMS** Alternative pointing devices, human performance, longitudinal study, people with motor impairments, user satisfaction, virtual cursors, web accessibility.

## I. INTRODUCTION

Computer access is often depicted as an act of pointing to and selecting graphical elements on the screen [1]. People with motor impairments (MIs) in their upper limbs may suffer from conditions such as poor coordination, slow movements, low strength, tremors, spasms, rapid fatigue, or difficulty controlling direction or distance, that hinder these actions in different ways [2]. These conditions may prevent the use of standard pointing devices [3] for activities such as navigating the Web.

In the last two decades various alternatives to the standard mouse have been developed to enable people with MIs to access graphical user interfaces (GUIs). These assistive technologies (ATs) seek to meet the special needs of this heterogeneous group of users who cannot grip or control standard mice [4], [5]. The most commonly used ATs to overcome this issue include specific alternative pointing devices (e.g., trackball or joystick), software applications such as mouse keys that enable the use of keyboards as an alternative to a mouse, or devices to support alternative manipulation (e.g., head wands or mouth sticks). Although these mouse alternatives allow people with MIs to interact with the on-screen cursor, performing point and click tasks is still difficult in many cases. Accordingly, Keates *et al.* [6] detected that users with MIs had to make an additional effort to plan and control physical movements resulting in them being 50% slower than their counterparts without motor disabilities.

The associate editor coordinating the review of this manuscript and approving it for publication was Young Jin Chun.

  


(a) (b)

**FIGURE 1. a) The cross cursor with no links within reach from its current location, therefore no shortcuts are available to assist target selection. b) After diagonally moving the pointer up to the right, 6 links are within reach of the cross cursor via single-key shortcuts.**

Other authors also found pointing and clicking actions to be less precise and more time consuming for people with MIs [7]–[17].

Pointing actions require the user to move the cursor to a particular target on the screen. These actions consist of an initial phase of ballistic movements followed by a slower homing phase until the cursor is positioned on the desired target [18]. Positioning errors [19] may occur, including additional sub-movements, movement direction changes, and indirect motion towards a target. In addition, targets can also be missed when the mouse pointer enters and leaves a target multiple times. Clicking actions require the user to press and release a button while holding the cursor over the target [9] thereby completing the selection process [18]. As with pointing actions, errors may also occur in clicking tasks. These errors include moving the cursor during an ''attempted click'' and ''extended clicks'' [19].

Several research works have been carried out to mitigate the effects of pointing and clicking errors. Ivory *et al.* [20] opined that most developments were focused on vision related issues leaving the needs of users with MIs insufficiently supported.

Various authors approached the development of alternative hands-free interaction mechanisms in order to reduce the effort required from people with MIs when carrying out pointing and clicking tasks in GUIs. Some examples are vision-based user interfaces that automatically recognized facial gestures [21], head mice that allow head movements to control the cursor [22], voice-based mouse pointer controls [23]–[25] or head-operated devices [26]. There are also works oriented towards implementing new alternative input devices such as the 2-D haptic device [27], an assistive robotic aid to minimize the absence of motor control in the upper limbs. However, Almanji *et al.* [7] argued that the use of AT for computer access has encountered barriers that have led to the use of standard mice or touch screens for practical reasons, mainly a lack of training and the elevated costs of some solutions. Virtual enhancements that modify the standard behaviour of the cursor to assist pointing and clicking tasks [1], [28]–[35] would appear to be more affordable solutions. Our work is underpinned by the conviction that further research is needed to improve assistive virtual tools due to the heterogeneous characteristics and needs of people with MIs.

People with MIs may face similar difficulties in pointing and clicking tasks depending on the mouse alternative being used. For instance, previous behavioural studies in GUIs by users with MIs [14]–[16] showed that keyboard-only users are more affected by total distance to the target, whereas joystick and trackball users tend to have accuracy issues when bringing the cursor to a halt over the target. Based on those findings, we developed two specific virtual cursors for assisting link selection on the Web: the novel cross cursor (Fig. 1) for keyboard-only users, and a standard area cursor (Fig. 2) for joystick and trackball users. Initially, both virtual cursors were tested with real users in a single-session supervised web-based experiment. It showed that both groups benefited from their respective assistance method for link selection, both in terms of user satisfaction [36], [37] and performance [38]. This study also suggested there was an improvement in performance with frequent usage of the virtual cursors. These results confirmed the need to continue the search for solutions to improve assistance on point and click tasks for this group of users. The importance of the Web for the personal autonomy of people with MIs underscores the need to continue researching web browsing assistance.

The main contribution of this paper is to explore the long-term benefits of both the aforementioned virtual cursors (the cross and area cursor) in a longitudinal study on the Web with real users with MIs. We conducted a six-week remote study in which eight users of alternative pointing devices (including keyboard, joystick and trackball) participated from home, obtaining in this way more naturalistic interaction

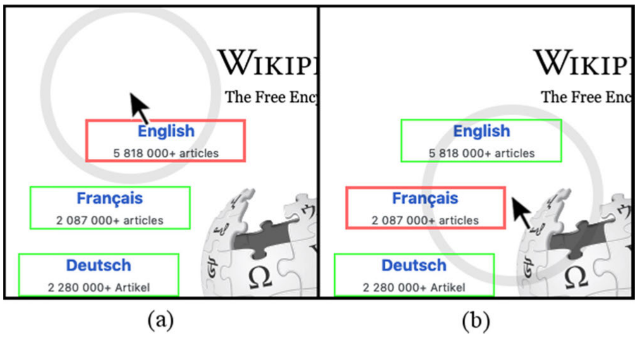


**FIGURE 2.** **a) The area cursor being tested enables the highlighted link to be clicked without needing to hover the cursor pointer over it. b) If the activation area includes more than one link, the nearest one to the pointer can be selected from a distance.**

data from everyday computer use. Our goal was to study their performance with both the virtual cursors in comparison with the original unassisted cursor (from now on referred to as "original cursor") and the influence of distractors (i.e., other nearby links) on target acquisition on the Web. We also reported on the learning effect on performance with each cursor, user behaviour in free navigation, and participant satisfaction with the virtual cursor being tested.

The rest of the paper is organized in this way: Section II presents first several longitudinal studies that explore user behaviour and learning effects in interaction with computers, and then describes some well-known cursor enhancements for point and click assistance. Section III explains the implementation and functioning of both the virtual cursors being evaluated. Section IV details the methodology followed in the study. Section V describes the results regarding the performance and satisfaction of participants with the different variants tested. Section VI includes a discussion of the results. Section VII presents the conclusions.

## II. RELATED WORK

### A. LONGITUDINAL USER STUDIES

Longitudinal studies provide a feasible way to record users' interactions over extended periods of time, allowing users' behaviours and their evolution over time to be studied. Longitudinal studies also allow naturalistic data to be obtained when carried out remotely and unobtrusively without the participants moving from their familiar environment [39]–[41]. For this reason, we carried out a longitudinal study with real users in order to explore the learning effect on their performance and satisfaction with the two virtual cursors.

Longitudinal studies have not been extensively used because they are time consuming and more difficult to perform. Nevertheless, we adopted a longitudinal study methodology for analysing the behaviour of users when interacting with the Web, as laboratory experiments do not always provide reliable results about the use and adoption of assistive tools. The number of participants and the duration of longitudinal studies vary notably depending on their purpose and characteristics. Longitudinal studies found in the literature focusing on people with special needs usually include far fewer participants than studies with people without impairments. This is due to the difficulty of recruiting samples of users with the required characteristics and to the complexity of the experiments.

Some longitudinal studies have analysed web accessibility problems from the perspective of blind users, and explored their performance in different situations. Bigham *et al.* [42] conducted a remote study over a period of one week with 10 blind and 10 sighted participants to evaluate differences in the browsing behaviour of these two groups. To this end a tracking proxy was used to remotely record both the visited pages and the actions taken by users on the web pages that they visited. For our experiment we adopted a similar data gathering scheme, but we used a client tool and a remote server to manage logged data instead of a proxy. Nicolau *et al.* [43] carried out a longitudinal study with five blind novice smartphone users to develop a richer characterisation of everyday typing performance on touch-screens. For eight weeks, in-situ device usage data was collected and weekly laboratory text-entry evaluations were conducted. Obtained performance measures include touch behaviours (e.g., touch contact points, exploration movements, and lift positions), character-level errors, and learning experience. In our experiment we also analysed performance and behaviour data focused on users with MIs.

Longitudinal studies have also been carried out with users with MIs to evaluate alternative input mechanisms. Mahmud *et al.* [24] conducted a comparative longitudinal study of two voice-based cursor control systems to get better understanding of novice users' experience over time. Ten participants were recruited for a longitudinal experiment over five consecutive days. In each experimental session participants had to complete 96 target acquisition trials with each cursor control system, as well as providing subjective ratings of each cursor modality in terms of their ease of learning, ease of use, level of fatigue, level of frustration, satisfaction and confidence. In order to characterize the pointer movement under each modality six different measures were analysed including target re-entry, task axis crossing, and movement error. Results showed that quantitative measurements as well as subjective ratings improved with time. This study was extended to people with MIs by Harada *et al.* [25]. They evaluated the learning curve for one of the systems involved in the previous study, the vocal joystick, in another longitudinal study with five participants with MIs and four participants without MIs. For our experiment we designated predefined tasks (adding free navigation) and also gathered subjective assessments. Sporka *et al.* [44] investigated the usability of a novel text entry application for users who cannot access a manual keyboard by gauging first impressions and how users adapted over time to the new system. To this end, a longitudinal study was conducted with five users with MIs. Participants were asked to use the tool for a minimum of 30 minutes each day, over the course of seven days. Their performance was measured on example phrases, and their subjective assessments were collected in several interviews (pre-test, first-impression, post-test). The study reported that all participants improved their text entry rates with the tested system during the course of the experiment. In this experiment the peak performance did not outperform other solutions. In our study, the participants achieved better results with the proposed virtual cursors than with the original one.

### B. ASSISTED POINT & CLICK

Much research has been done to facilitate target acquisition in GUIs, even though little of it has been specifically focused on assisting people with MIs to navigate the Web. The following works proposed some well-known cursor enhancements to assist point and click interactions, although these were not always initially aimed at people with MIs. Even if these works were not focused on assisting web browsing, some can be directly translated to this scenario and have served to define the basis of our research.

The steady clicks assistance [34] suppresses accidental clicks and slipping when clicking by freezing the cursor during mouse clicks. In this way, it prevents overlapping button presses and cancels clicks made while the mouse is moving at a high speed. Evaluations showed that this option improves time performance and that users with MIs required fewer attempts to select targets; moreover, participants expressed their preference for this assistance (9 out of 11) over the unassisted condition. This assistance aims to reduce clicking errors of people with MIs; however, it may not be useful for keyboard-only users as it was designed for a pointing device such as the mouse.

The angle mouse [45] is a pointing facilitator that attempts to improve target acquisition by adjusting the mouse control-display gain based on the deviation angles of the cursor path during movement. Thus, unlike most cursor enhancements, this technique (like the steady clicks) is based solely on the user's behaviour and requires no information from the targets on the GUI. Published results proved that this alternative improves the pointing performance of users with MIs while remaining unobtrusive for people without impairments. However, no alternative input device was tested as all the study participants (both with and without MIs) used a standard mouse to complete the experimental tasks.

Wobbrock and Gajos [35] suggested that difficulties faced by people with MIs could be alleviated by a different target acquisition paradigm called goal crossing. In this proposal users do not aim at a restricted area, but instead pass over a target line to perform a selection. Empirical results indicated a preference for goal crossing among people with MIs, although error rates were higher with this alternative. The authors also presented some design principles for this new target acquisition paradigm, but these are not usable on standard web interfaces.

Hwang *et al.* [30] studied the performance of users with and without MIs in a point and click task with targets modelled as virtual gravity wells. Their results showed the greatest improvements for the users with the most severe impairments, even when multiple on-screen targets were haptically enabled. This technique looks promising for complex GUIs with numerous targets, as are the majority of web interfaces. More testing would be required in order to adopt it because it was only tested with a standard mouse, and no subjective perception from participants about the tested enhancements was published.

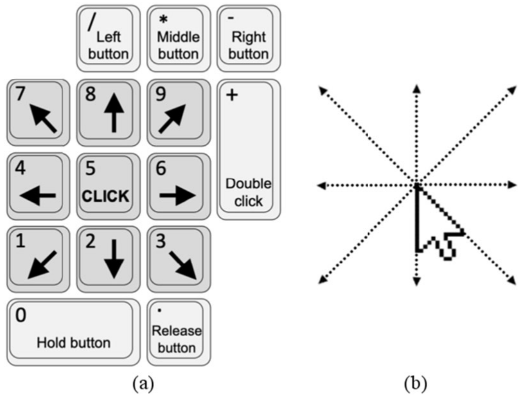


**FIGURE 3.** a) The mouse keys application allows the user to point and click with the on-screen cursor by using the highlighted keys of the numerical keypad. b) Eight possible paths that the cursor can travel from its current position with mouse keys.

Worden *et al.* [46] studied the effectiveness of two interaction techniques: the area cursor (our circular version for the Web was inspired by this squared version) and the sticky icons, for improving the performance of older adults (with declined motor abilities) in basic selection tasks. The area cursor, successfully tested previously with people without MIs [47], uses a cursor with a larger than usual activation area. The latter technique makes an icon ''sticky'' by automatically reducing the cursor's gain ratio (number of pixels moved in response to a single increment of movement by the physical device) when it is over a target icon. Both techniques improve pointing time. The area cursor is especially useful when the icons are not too close together and also when the target size is small. The results showed that neither technique hindered performance in difficult situations (e.g., closely spaced targets). Other authors have studied variations of the area cursor, by dynamically resizing the cursor's activation area [28], or with different combinations of visual magnification or goal crossing [31], [32], [48]. The results from these works generally revealed improvements in performance, although the trials were mainly based on users without MIs, the only pointing device used was the mouse and sometimes participants did not prefer the proposed method.

Felzer *et al.* [49] compared two different methods for mouse emulation with the DualPad numeric keypad. The first method, called CKM, allows the mouse pointer to be moved in cardinal directions and to be clicked in a similar way to the mouse keys application (Fig. 3). The second method, called the DualMouse, does not rely on mouse movement at all, but directly clicks at a destination location following a step-by-step locating process. Evaluation based on a case study with a single user with MIs revealed that the CKM method produced higher throughput than the DualMouse. Surprisingly, no cursor enhancement has been studied to assist pointing and clicking interactions of keyboard-only users applying mouse keys (Fig. 3) included in every major operative system. A Mozilla Firefox add-on, Vim Vixen[1], enables web browsing by using only the keyboard. This application labels every link on the visited web page with different shortcuts. Links are selected by typing the corresponding sequence of letters without having to use the cursor. For our experiment, we designed an enhanced version of this method. The cross cursor reduces the length of shortcuts to only one key by combining cursor movements to label just those links within reach of the virtual cursor (Fig. 1b).

## III. VIRTUAL CURSORS FOR THE WEB

As we observed in previous works [14]–[16] about pointing and clicking behaviours of people with MIs, difficulties faced by participants varied depending on the alternative pointing device used. Thus, keyboard-only users were especially affected by the total distance from the starting point to the target (i.e., the pointing trajectory), whereas users of specific alternative pointing devices, such as joysticks or trackballs, had problems to halt the cursor over the target due to a lack of accuracy. In order to study empirically if these issues can be alleviated in web browsing, we developed two virtual cursors, implemented as browser add-ons, in order to test them with real users: the novel cross cursor (Fig. 1) and a standard area cursor (Fig. 2).

### A. THE CROSS CURSOR

This virtual cursor aims to assist link selection on Web interfaces by reducing cursor displacement required for pointing. This is achieved by combining cursor movement and providing single-key letter shortcuts to every link within reach of the cross cursor (Fig. 1a and Fig. 1b). Links within reach of this virtual cursor are those traversed by the cross cursor lines. During its movement the cross cursor continuously displays a horizontal and a vertical line (that respectively extend over the entire width and height of the web page) crossing perpendicularly below its current position (Fig. 1). Shortcuts are automatically assigned and displayed next to every link within reach of the cross cursor when the virtual cursor comes to a halt and disappear whenever the cursor starts moving again. Single-letter shortcuts (together with number keys) are automatically assigned in order of proximity to the cursor pointer, starting from the right of the keyboard, with the closest keys to the numeric keypad first and the furthest to the left at the end. If all letter and number keys (36 in our case) have already been assigned in this way, the additional links within reach of the cross cursor will not have any shortcut assigned, requiring the cursor pointer to be brought nearer to display a shortcut. We used fixed values (10 px width and 90% translucent grey colour) for the visual appearance of the cross cursor lines.

The mouse keys feature, included on every major operative system, allows keyboard-only users to use the numeric keypad as a mouse alternative (Fig. 3a) by pressing the central '5' key for cursor clicking, and the surrounding number keys for moving it in vertical, horizontal and diagonal directions (Fig. 3b).

### B. THE AREA CURSOR

This virtual cursor corresponds to the standard area cursor and aims to assist target acquisition on the Web by reducing the accuracy required to click a link. As it moves the area cursor continuously displays a circle of fixed size that is always centred on the current position of the pointer (Fig. 2) and which corresponds to its activation area. In this way, the closest link within its activation area can be clicked (see highlighted targets in red in Fig. 2) without needing to hover over it. We used fixed values for the visual appearance of the area cursor (10 px width and 90% translucent grey colour) as well as for the activation area diameter (130 px).

### C. IMPLEMENTATION

Both virtual cursors were implemented using Scalable Vector Graphics (SVG) to display the corresponding visual elements (lines, circles, rectangles and letters) on the browser window, along with JavaScript to handle users' interactions with the cursor and the web content. To this end, the add-on that implements each virtual cursor is in charge of parsing every visited web page to find every link. The information about the location and size of each visible link within a page is processed by the add-on, which also handles users' interactions (mouse moves, clicks, and keystrokes) to modify standard pointing and clicking and assist target selection. To explore the long-term benefits of both virtual cursors in this longitudinal study we implemented some upgrades with respect to the previous versions tested on a single-session supervised study [36]–[38]. These included processing hidden and overlapping links in different layers (e.g., dropdown menus of navigation bars), as well as a command to allow users to deactivate and reactivate the virtual cursor at any time during web browsing (Fig. 4).

## IV. METHOD

We conducted a longitudinal study on the Web with people with MIs in order to explore the satisfaction and performance achieved with the two virtual cursors in comparison with the original cursor. During a period of 6 weeks, we collected usage data of the three cursors tested (original, area and cross). During this period participants used their own personal computers at home to perform unsupervised tasks. This study was approved by the "Ethics Committee for Research Involving Human Beings" (CEISH) of the University of the Basque Country (UPV/EHU) that reviewed the purpose and methodology of the experiment and authorized us to collect and analyse the resulting data.

[1] https://addons.mozilla.org/en-US/firefox/addon/vim-vixen (accessed on April 10, 2020)

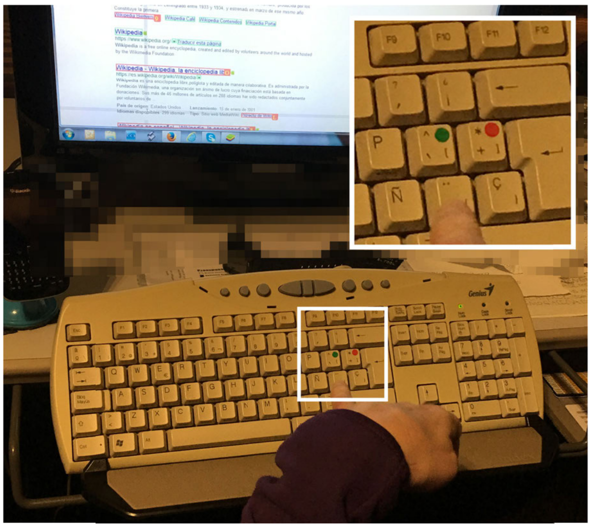

**FIGURE 4.** A keyboard user practicing with the cross cursor during the first visit to her house. Enlarged detail of the keyboard on the upper right corner shows the two keys tagged that were used on the experimental web browser to activate/refresh and deactivate the virtual cursor during free navigation.


**FIGURE 5.** Four participants performing a supervised task during the first home visit of the longitudinal study: a) a keyboard user interacting through a head wand with her laptop, b) a keyboard user interacting directly with her hand, c) a joystick user, and d) an oversized trackball user.

### A. PARTICIPANTS

Eight participants took part in this longitudinal study, all of them people with MIs involving limited dexterity in their upper limbs that prevent them from using a standard mouse (Fig. 5). According to Lazard *et al.* [50] this number of participants is generally acceptable for research focusing on users with a specific disability. In our study, all the participants were regular computer users and were selected based on the alternative pointing device they used for target selection in GUIs. In this way, the following 2 groups were defined, each as the target group of a virtual cursor:

- *Keyboard users group (KU group)*: 4 keyboard users (mean age = 52.5 years, SD = 8.3) to test the cross cursor.
- *Joystick and trackball users group (JU group)*: with 4 participants, 3 joystick users and 1 trackball user (mean = 46 years, SD = 2.9) to test the area cursor.

Two participants from the KU group (K1 and K2) needed a head wand to interact with the keyboard, whereas the other two were able to push the keys directly with their fingers. Table 1 shows detailed information about the participants.

### B. APPARATUS

All participants used their own personal computer at home. All of them used a desktop computer except participant K1 who owned a laptop computer. The operative system running on each computer was a version of Windows7 (K3, K4, J1, J2 and J4) or Windows10 (K1, K2, J3). Table 1 includes display sizes and resolution used by each participant.

The RemoTest platform to design and perform remote user tests [15], [51] was used to conduct this unsupervised longitudinal study. A secure server was in charge of running the remote modules of the RemoTest platform, both to gather interaction data from each participant and to provide them with the target acquisition tasks that had to be repeated regularly. Additionally, we installed the Mozilla Firefox web browser (version 44.0.1) on each participant's computer, personalized with two add-ons: one corresponding to the virtual cursor to be tested, and the other one implementing the Participant Module of the RemoTest platform. This module was in charge of communicating the experimental browser with the remote modules of RemoTest to identify each participant, present them with the proposed tasks, and log and send the data of the user interactions to a remote server. In this regard, a remote MongoDB database was used to store the set of events that occurred in the participants' browser, together with the corresponding timestamps. The set of events included the cursor location (X and Y coordinates) along its movement with a sampling frequency of 100Hz, in order to study different features of point and click trajectories. To identify invalid cursor trajectories and to delimit data for analysis, other events were also recorded, such as keystrokes, cursor clicks and page scrolls. Additionally, browser and experiment related events, such as page loads, or the start and end of tasks, were also gathered. To study other performance related features, each time a click event occurred we recorded the spatial information of the selected link (location, width and height) and of the other visible targets on the screen.

A separate Java application was implemented to parse interaction data gathered throughout the study and calculate a variety of measures from cursor trajectory that were later analysed with the RStudio statistical tool. This Java

**TABLE 1.** Information about study participants.

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Keyboard users group (KU)** | | | | | | | | | | | | | | | | | |
| | | | | *Self-reported impairments* | | | | | | | | | | | | *Display* | |
| *ID* | *G* | *A* | *Health condition* | *Fa* | *Co* | *St* | *Mo* | *Gr* | *Ho* | *Tr* | *Sp* | *Se* | *Dir* | *Dis* | *Pointing device* | *Inches* | *Resolution* |
| K1 | F | 46 | Cerebral palsy | ■ | | | | ■ | ■ | | ■ | | ■ | ■ | Keyboard + head wand | 17 | 1366×768 |
| K2 | M | 45 | Cerebral palsy | | | ■ | | ■ | ■ | ■ | | | ■ | | Keyboard + head wand | 26 | 1920×1080 |
| K3 | F | 57 | Glutaric aciduria t1 | | | | ■ | ■ | ■ | | ■ | | | | Keyboard | 22 | 1360×768 |
| K4 | F | 62 | Glutaric aciduria t1 | ■ | ■ | ■ | | ■ | ■ | | ■ | ■ | ■ | | Keyboard | 28 | 1920×1080 |
| **Joystick and trackball users group (JU)** | | | | | | | | | | | | | | | | | |
| | | | | *Self-reported impairments* | | | | | | | | | | | | *Display* | |
| *ID* | *G* | *A* | *Health condition* | *Fa* | *Co* | *St* | *Mo* | *Gr* | *Ho* | *Tr* | *Sp* | *Se* | *Dir* | *Dis* | *Pointing device* | *Inches* | *Resolution* |
| J1 | F | 43 | Cerebral palsy | ■ | | ■ | ■ | ■ | | | | | | | Joystick | 27 | 1536×864 |
| J2 | M | 49 | Cerebral palsy | ■ | ■ | | | ■ | | ■ | ■ | | | | Joystick | 27 | 1536×864 |
| J3 | M | 44 | Cerebral palsy | | | | | | ■ | ■ | ■ | | | | Joystick | 27 | 1920×1080 |
| J4 | M | 48 | Cerebral palsy | ■ | ■ | | | ■ | ■ | ■ | | | | | Trackball | 23 | 1920×1080 |

Table includes details about each participant's gender (G), age (A), health condition, pointing device used, as well as the size of the monitor and resolution setup. Categories for self-reporting impairments in upper limbs [48] were: rapid fatigue (Fa), poor coordination (Co), low strength (St), low movements (Mo), difficulty gripping (Gr), difficulty holding (Ho), tremor (Tr), spasm (Sp), lack of sensation (Se), difficulty controlling direction (Dir), difficulty controlling distance (Dis).

application initially allowed us to organize a huge amount of data gathered from user tests; for instance, delimiting point and click trajectories from start to end point based on sequences of cursor movement events ended with a click event, or detecting erroneous point and click cursor trajectories if these included more than one click event.

### C. TASKS AND MATERIALS

During the course of this longitudinal study the participants had to perform two different tasks on the browser installed on their computers: free navigation and target acquisition.

#### 1) FREE NAVIGATION TASK

The participants were asked to use regularly (according to their habits) the web browser on their computers for personal and autonomous web navigation. The objective was for them to practice and get used to the virtual cursor, in order to allow us to explore usage and acceptance of both variants. Usage data generated during free navigation was remotely collected similarly to the target acquisition task, and included mainly cursor trajectories, data related to cursor clicks, and on-screen distribution of links on visited pages. The participants were able to deactivate the virtual cursor assistance and to use the original cursor at any time of web navigation. Activation/deactivation actions were registered. On the other hand, sensitive data about participants' privacy during their web navigation was not recorded by our data gathering tools (e.g., any typed text on the keyboard or information identifying the visited web pages).

#### 2) TARGET ACQUISITION TASK

The participants were also asked to regularly repeat (but never more than once a day) the same target acquisition task until completing 15 sessions. This exercise consisted of a multidirectional point and click task with 12 targets arranged in a circular layout (Fig. 6) that was repeated for 3 target configurations (Fig. 7), and for both cursor variants tested by each participant (original cursor and corresponding virtual cursor). The participants had to complete a total of 72 trials per session. The order of both tested factors (cursor variant and target configuration) was counterbalanced between sessions, and the target sequence was randomized each time. Based on guidelines to assist problems with fine movements [52]–[54], target sizes were defined smaller than recommended (95 pixels width by 15 height) to test the benefit of the virtual cursors on more difficult cases. A constant radius of 250 pixels was used for the size of the circular layout, avoiding any horizontal or vertical scroll during this task. Before each trial, participants had to position the cursor over a home button located in the centre of the circular layout (Fig. 6). Upon selecting it, the home button disappeared and a new trial started in which participants had to select as fast as possible the displayed target labelled as "click here". After completing each trial, the home button re-appeared along with the next target. The goal of this task was to record cursor trajectories on intentional movements of target acquisition, avoiding unintended movements that might occur during free navigation.

### D. PROCEDURE

During a first home visit each participant was briefed on the purpose and details of the experiment and then signed a consent form before beginning the longitudinal study. Information on demographics and about the computing equipment of each participant was collected through a preliminary interview. Next, we installed the browser and both add-ons on the personal computer of each participant so they could perform the experimental tasks autonomously throughout the longitudinal study. Participants were then taught to use the virtual cursor (Fig. 4), and to perform the proposed tasks without supervision from an experimenter. All participants completed a first supervised session of the target acquisition

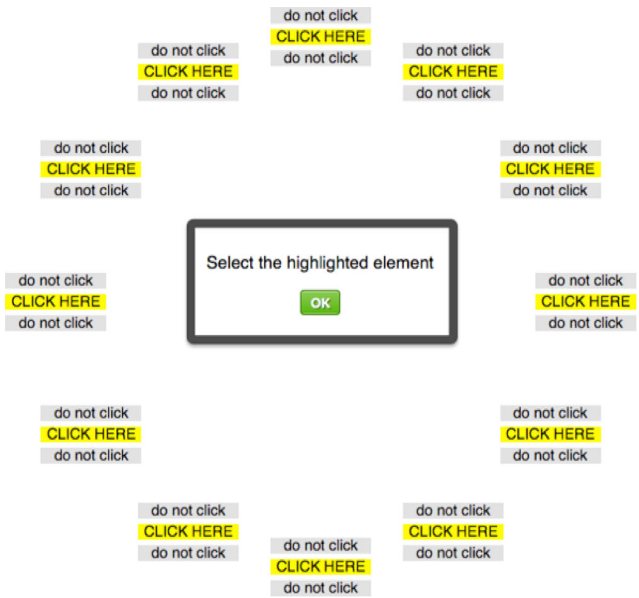


**FIGURE 6. Circular layout of the target acquisition task showing all 12 targets at once for a particular distractor configuration, and with the home button in the center as the starting point for each trial. During the test only the current target was displayed to participants, the other 11 remained hidden.**

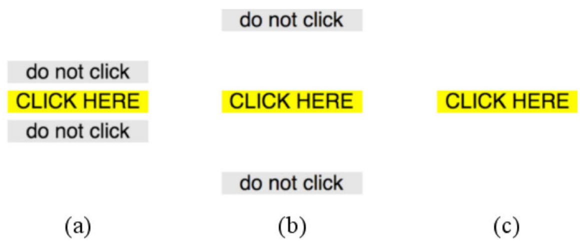


**FIGURE 7. The three target configurations tested on the target acquisition task correspond to: a) configuration 1 with distractors 5 pixels away from target, b) configuration 2 with distractors 40 pixels away, and c) configuration 3 without any distractor near the target link.**

task while we recorded their interactions with the pointing device using a video camera located behind them. To conclude this first home visit, the participants were invited to respond to a usability questionnaire to rate the virtual cursor just tested. Thereafter, participants were asked to use the installed browser autonomously for free navigation and to regularly perform the unsupervised target acquisition task.

Once participants had completed the target acquisition task fourteen times and enough data had been collected from the free navigation task, we visited them again to conclude the longitudinal study. Before that, participants had to complete the last supervised session of the target acquisition task (also recorded on video), followed by the same usability questionnaire that they had answered during the first visit, in order to compare both responses. After concluding the study, each participant was rewarded with a voucher worth 200 €in appreciation for their collaboration in the study.

### *E. MEASURES*

Two different methods were used to compare the two virtual cursors with the original cursor. Firstly, participants filled in the same questionnaire both at the beginning and end of the longitudinal study to measure their satisfaction with the virtual cursor tested. We used the System Usability Scale (SUS) questionnaire [55] for this purpose, which includes 10 items to be rated on a 5-point Likert scale (from strongly agree to strongly disagree). Secondly, seven cursor path evaluation measurements described in the literature [10], [12], [56] were used to study the performance of participants with the cursor variants tested on target acquisition tasks. Although Fitts's law [57], as described by MacKenzie [58], has been widely used to study target acquisition in GUIs, we did not apply this paradigm in our evaluation as there is evidence against the suitability of Fitts's law to model pointing and clicking movements of people with MIs [59]. Furthermore, cursor trajectories of keyboard-only users do not follow the ballistic movement supposed by Fitts's law.

#### 1) MOVEMENT TIME (MT)

The time interval from clicking the home button and until the target link is selected (Fig. 6). The MT corresponds to the total time needed to complete a trial, and was calculated based on the timestamps of the recorded events.

#### 2) POINTING TIME (PT)

The time interval from when the cursor starts moving until it finally stops before the target link is selected (i.e., a click event occurred). The PT corresponds to the time needed to move the on-screen cursor to complete a trial. The PT is a portion of the total MT, and was also calculated based on the timestamps of the recorded events.

#### 3) CLICKING TIME (CT)

The time interval from when the cursor finally stops moving until the target link is selected. The CT corresponds to the time needed to perform the click to complete a trial. The CT is a portion of the total MT, and was also calculated based on the timestamps of the recorded events.

#### 4) DISTANCE TRAVELLED (DT)

The total distance traversed (in pixels) by the on-screen cursor along the pointing trajectory. The DT was computed for each trial as the sum of Euclidian distances from each point to the next point. The distance between two consecutive points $(X_1, Y_1)$ and $(X_2, Y_2)$ is given by:

$$d = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2} \quad (1)$$

#### 5) CURVATURE INDEX (CI)

The curvature index is the ratio of the DT to the straight-line distance between the starting and ending points of cursor trajectory. A value of one indicates the cursor has followed a straight line to the target, while larger values show increasing deviations. The CI was calculated based on the computed distance travelled for each trial and the corresponding distance between the home position and the target location.

#### 6) NUMBER OF PAUSES (NP)
The times the cursor stops along the pointing trajectory to a target. The number of pauses indicates the number of trajectory corrections made by the user in order to select a target. Low values indicate fewer corrections and therefore fewer problems on the pointing trajectory, while high values means that the user has had more difficulties to reach a target. The NP was calculated based on the time interval between consecutive cursor motion events. According to interaction data gathered, time intervals equal to, or greater than, 100 milliseconds were considered as pauses.

#### 7) TARGET RE-ENTRY (TR)
A TR occurs when the pointer enters the target region, then leaves, and then it enters again. A result of zero indicates perfect accuracy on target acquisition, while growing values mean increasing accuracy issues. For both virtual cursors, the TR was calculated considering as the target region the extended area from where selection from distance was possible.

## V. RESULTS
The following subsections present results from qualitative and quantitative analyses about acceptance and performance achieved with the cursors tested.

### A. FREE NAVIGATION TASK
Various measures were calculated from the interaction data collected in order to understand the browsing activity of each participant during the free navigation task. Table 2 shows results from each participant after filtering inactivity periods during web browsing of over 15 minutes without any user interaction. Some measures, such as the total number of hours (TH), number of sessions (NS), or total number of pages visited (TPV), reveal disparate results among participants. Despite this, all the participants were able to navigate on their own and repeatedly with the experimental browser throughout the study, as well as to use the corresponding virtual cursor extensively as shown by the PVtr values (Table 2).

**TABLE 2.** Browsing habits of each participant on free navigation task.

| **KU group (Cross & original cursors)** | | | | | | | |
|---|---|---|---|---|---|---|---|
| | *Participants* | | | | | | |
| *Measure.* | *K1* | *K2* | *K3* | *K4* | *Distrib.* | *Mean* | *SD* |
| TH | 7 | 17.4 | 29.5 | 5.3 | | 14.8 | 9.7 |
| NS | 55 | 90 | 70 | 30 | | 61.3 | 21.9 |
| ASDm | 7.6 | 11.6 | 25.3 | 10.7 | | 13.8 | 6.8 |
| APDs | 95.7 | 44 | 125.2 | 103.1 | | 92 | 29.8 |
| TPV | 262 | 1423 | 848 | 186 | | 679.8 | 499.8 |
| APS | 4.8 | 15.8 | 12.1 | 6.2 | | 9.7 | 4.5 |
| ADPS | 4.2 | 12.3 | 9.1 | 5.2 | | 7.7 | 3.2 |
| PVtr | 93.7 | 99.4 | 99.4 | 96.2 | | 97.2 | 2.4 |
| **JU group (Area & original cursors)** | | | | | | | |
| | *Participants* | | | | | | |
| *Measure.* | *J1* | *J2* | *J3* | *J4* | *Distrib.* | *Mean* | *SD* |
| TH | 9.6 | 2.2 | 62.6 | 3.4 | | 19.5 | 25.1 |
| NS | 29 | 20 | 69 | 23 | | 35.3 | 19.8 |
| ASDm | 19.8 | 6.5 | 54.4 | 9 | | 22.4 | 19.1 |
| APDs | 132.4 | 60.5 | 168.4 | 65.9 | | 106.8 | 45.5 |
| TPV | 260 | 130 | 1338 | 188 | | 479 | 498.1 |
| APS | 9 | 6.5 | 19.4 | 8.2 | | 10.8 | 5.1 |
| ADPS | 7 | 5.3 | 11.7 | 6.1 | | 7.5 | 2.5 |
| PVtr | 80.3 | 99.2 | 99.3 | 97.3 | | 94 | 8 |

Measurements calculated for each participant include: total hours browsing the Web (TH), number of sessions accessing the browser (NS), average session duration in minutes (ASDm), average page duration in seconds (APDs), total number of pages visited (TPV), average number of pages visited per session (APS), average number of different pages visited per session (ADPS), and time percentage using the virtual cursor variant tested over the original cursor (PVtr).

### B. TARGET ACQUISITION TASK
This section presents the results of the quantitative analysis of cursor trajectories to compare each of the virtual cursors with the original cursor. First, we describe how interaction data was filtered for subsequent statistical analysis.

#### 1) DATA CLEANING AND USE OF VIRTUAL CURSORS
Each of the 8 participants repeated the same target acquisition task 15 times, which included 72 trials (half with each cursor variant tested), resulting in a total of 8640 trials (1080 trials by each participant). Interaction data gathered from participants was filtered before statistical analysis by removing invalid trials. Excluded trials were those corresponding to erroneous point and click interactions (i.e., any sequence of events different from a set of cursor movements followed by a target selection), as well as outlier trials with a movement time two standard deviations or more away from the participant's mean. In this way, 532 trials out of 4320 (12.3%) were filtered for the KU group, and 326 trials (7.5%) for the JU group, which in total corresponds to 858 (9.9%) of all target acquisition trials. Table 3 shows the distribution of invalid trials by cursor variant and filtering category for each group of participants.

Both the tested virtual cursors enable link selection without needing to hover the pointer over targets. Table 4 shows the percentage of valid trials completed by clicking from outside the target for each cursor variant and target configuration, as well as for each participant. Only half of JU group participants (J3 and J4) leveraged the area cursor on more than 50 percent of the trials (depending on target configuration)

**TABLE 3.** Distribution of filtered trials from target acquisition task.

| Category | Cursor | Filtered | Distrib. | Mean | SD |
|---|---|---|---|---|---|
| **Keyboard users group (KU)** | | | | | |
| Erroneous | Original | 236 (10.9%) | | 59 | 34.3 |
| | Cross | 213 (9.9%) | | 53.3 | 22.8 |
| Outliers | Original | 45 (2.1%) | | 11.3 | 2.7 |
| | Cross | 38 (1.8%) | | 9.5 | 5.7 |
| Total filtered | Original | 281 (13%) | | 70.3 | 34.9 |
| | Cross | 251 (11.6%) | | 62.8 | 27.7 |
| **Joystick & trackball users group (JU)** | | | | | |
| *Category* | *Cursor* | *Filtered* | *Distrib.* | *Mean* | *SD* |
| Erroneous | Original | 119 (5.5%) | | 29.8 | 38.7 |
| | Area | 73 (3.4%) | | 18.3 | 20.5 |
| Outliers | Original | 72 (3.3%) | | 18 | 5.2 |
| | Area | 62 (2.9%) | | 15.5 | 1.5 |
| Total filtered | Original | 191 (8.8%) | | 47.8 | 33.5 |
| | Area | 135 (6.3%) | | 33.8 | 19.9 |

**TABLE 4.** Percentages of links selected from outside the target.

| Conf. | K1 | K2 | K3 | K4 | Distrib. | Mean | SD |
|---|---|---|---|---|---|---|---|
| **Cross cursor – KU group** | | | | | | | |
| 1 | 143 (100%) | 168 (100%) | 152 (100%) | 167 (100%) | | 100 | 0 |
| 2 | 143 (100%) | 162 (100%) | 171 (100%) | 169 (100%) | | 100 | 0 |
| 3 | 144 (100%) | 158 (100%) | 167 (100%) | 165 (100%) | | 100 | 0 |
| Ov | 430 (100%) | 488 (100%) | 490 (100%) | 501 (100%) | | 100 | 0 |
| **Area cursor – JU group** | | | | | | | |
| *Conf.* | *J1* | *J2* | *J3* | *J4* | *Distrib.* | *Mean* | *SD* |
| 1 | 53 (31%) | 32 (21%) | 41 (25%) | 61 (36%) | | 28.3 | 5.7 |
| 2 | 76 (43%) | 68 (43%) | 76 (45%) | 121 (70%) | | 50.3 | 11.4 |
| 3 | 74 (42%) | 64 (40%) | 100 (58%) | 151 (85%) | | 56.3 | 18 |
| Ov | 203 (39%) | 164 (35%) | 217 (43%) | 333 (64%) | | 45.3 | 11.2 |

Table shows the distribution of results between the participants and the three target configurations tested (Fig. 7), as well as for the overall results of all configurations of the target acquisition task (Ov).

by clicking outside the target. Nonetheless, all JU group participants increased this ratio with the less restrictive target configurations (2 and 3). On the other hand, all participants from the KU group leveraged the cross cursor on every valid trial by tapping the corresponding shortcut key for target selection from distance.

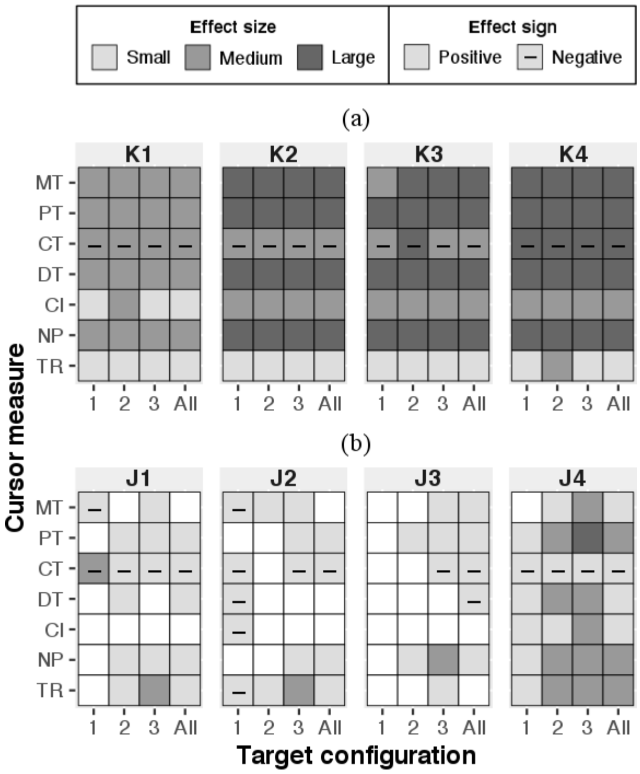


**FIGURE 8.** Summary of results from significance tests, showing on a tiled heatmap per participant the existence of significant differences between the original cursor and a) the cross cursor for KU group participants, or b) the area cursor for JU group participants. For each cursor measure and target configuration studied, a cell displays the effect size (small if $r \geq 0.1$, medium if $r \geq 0.3$, or large if $r \geq 0.5$) and the effect sign, if significant differences existed between cursors. A blank cell means no significance was found.

## 2) CURSOR MEASURES

To compare the participants' performance with the different cursors (original cursor and cross or circular cursor respectively by KU group or JU group participants) we performed a quantitative analyse of cursor trajectories on the target acquisition task. A Wilcoxon signed-rank test was used for each cursor measurement, each target configuration (as well as for all configurations together) and each participant, to find out if significant differences existed between the tested cursors. Fig. 8 summarizes the results of the 224 statistical tests carried out ($7 \times 4 \times 8$), showing by means of tiled heatmaps the existence of significant differences between cursors and the effect sizes in such cases. Positive effect sizes in Fig. 8 correspond to the virtual cursor (cross or circular cursor for the KU group or the JU group respectively) improving the performance of the original cursor, whereas negative values represent the opposite.

The 4 participants from the KU group (K1 to K4) found significant differences between the cross and the original cursor for all cursor measurements and target configurations (Fig. 8a). The cross cursor improved significantly the performance of each KU group participant for all cursor measurements, except for the clicking time (CT). Measurements for the movement time (MT), pointing time (PT), distance travelled (DT) and number of pauses (NP) reflected the highest effect sizes for every KU group participant, followed by the curvature index (CI) and target re-entry (TR). Concerning MT, PT, DT and NP, results from K2, K3 and K4 (Fig. 8a) showed consistently large effect sizes (0.50 to 0.57), except for MT (configuration 1) from K3 with a medium effect size ($W = 8594$, $Z = 8.7938$, $p < 0.01$, $r = 0.46$). Results from K1 for MT, PT, DT and NP showed consistently medium effect sizes (0.39 to 0.48). Regarding CI, results from K2, K3 and K4 corresponded to medium effect sizes (0.39 to 0.47), whereas K1 obtained small effect sizes (0.28 to 0.29) except for configuration 2 with a medium effect size ($W = 1967$, $Z = 5.8455$, $p < 0.01$, $r = 0.31$). Regarding TR, all participants from the KU group consistently obtained small effect sizes (0.11 to 0.29), except for K4 who obtained a medium effect size on configuration 2 ($W = 1061$, $Z = 6.483$, $p < 0.01$, $r = 0.34$). On the other hand, the CT worsened significantly with the cross cursor in comparison with the original cursor for all participants from the KU group. Results from K4 for CT showed consistently large effect sizes (−0.54 to −0.56), whereas participants K1, K2 and K3 obtained medium effect sizes (−0.35 to −0.49) except for K3 who obtained a large effect size on configuration 2 ($W = 531$, $Z = -9.5826$, $p < 0.01$, $r = -0.51$). In short, all participants from the KU group improved significantly on point and click tasks with the cross cursor on six of the seven measures studied, achieving similar results for each target configuration tested (Fig. 8a).

On the other hand, we found significant differences between the area cursor and the original cursor for all the measures studied, although results varied among JU group participants (J1 to J4) and the target configurations being tested (Fig. 8b). J4 obtained the most noteworthy benefits with the area cursor, achieving significant differences for all combinations studied (except for configuration 1 of MT), as well as the highest effect sizes. J1 obtained the second best results of the JU group with the area cursor compared to the original cursor, followed by J2 and J3 who obtained slightly less significant differences between cursor variants tested, as well as shorter effect sizes. Despite these differences, all JU group participants achieved the best results with the area cursor for measurements for MT, PT, NP and TR, on target configuration 3. Additionally, the highest effect sizes in favour of the area cursor corresponded to target configuration 3 and cursor measurement TR for J1 ($W = 1655$, $Z = 5.9894$, $p < 0.01$, $r = 0.32$) and J2 ($W = 1870$, $Z = 5.7277$, $p < 0.01$, $r = 0.3$), NP for J3 ($W = 531$, $Z = 9.5826$, $p < 0.01$, $r = 0.33$), and PT for J4 ($W = 13676$, $Z = 9.7518$, $p < 0.01$, $r = 0.51$). However, for all JU group participants the CT worsened significantly with the area cursor in comparison with the original cursor 9 out of 12 times. Results for CT showed small effect sizes for all participants (−0.1 to −0.27), except for J1 who obtained a medium effect size on configuration 1 ($W = 2507$, $Z = -6.7005$, $p < 0.01$, $r = -0.35$). In summary, participants achieved the best results with the area cursor on less restrictive target configurations. Participant J4, using a trackball, obtained the greatest benefits from the JU group with the virtual cursor.

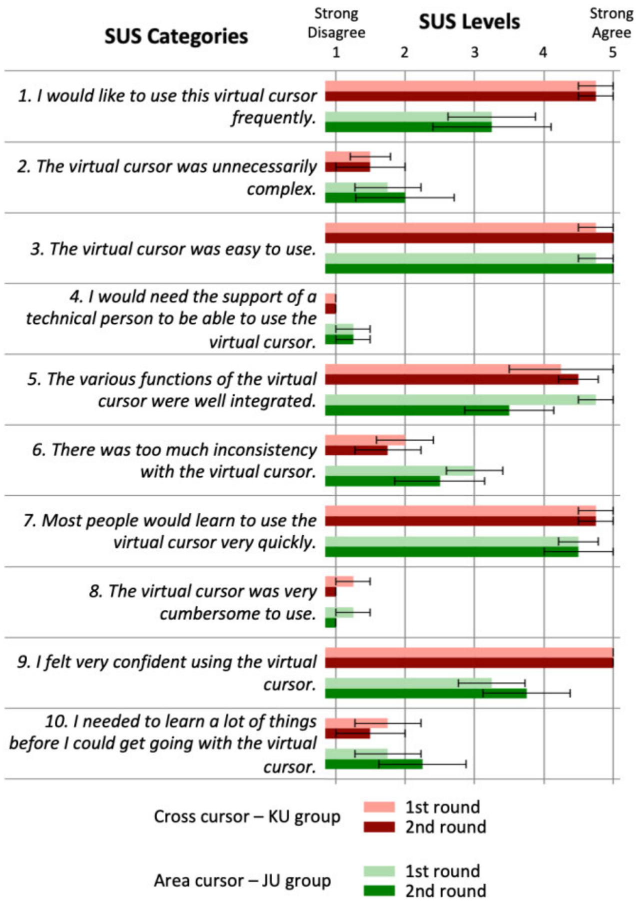

**FIGURE 9.** Average ratings by user group and round of questionnaire for the virtual cursor tested. Error bars represent ±1 standard error (SE).

### C. USABILITY QUESTIONNAIRE

Participants from both groups completed the SUS usability questionnaire twice, at different stages of the study, in order to gather their subjective assessments about the virtual cursor tested. Each participant responded to the 10 categories of the questionnaire, first at the beginning of the study after the training and first session of the target acquisition task, and second at the end after the last session of the target acquisition task. On average, both virtual cursors were positively rated for all categories of the SUS questionnaire, except the area cursor, which obtained a medium score of 3 points on the sixth item at the beginning of the study. In general, the cross cursor obtained better average scores than the area cursor, as well as less variability between subjects' responses. Fig. 9 includes bar graphs for each category of the SUS questionnaire, showing the distribution of responses by group of participants and questionnaire round.

The cross cursor got the worst results for categories 5, 6 and 10 of the questionnaire. Two participants from the KU group (K1 and K2) stated that they were confused at first about how to leverage the virtual cursor for link selection, and that they needed to practice to get used to it. Participant K3 reported that she visited a web site where she could not use the cross cursor for link selection, as the page contained a text bar that was continuously listening for the keyboard. After prolonged use of this virtual cursor, scores for categories 5, 6 and 10 tended to improve on the second round of the questionnaire. For the rest of categories, the cross cursor obtained very positive assessments from the beginning of the study.

The area cursor got the worst results for categories 1, 5, 6, and 9 of the questionnaire. J3, who made extensive use of the virtual cursor, stated that when there were a large number of links on the screen, visual cues about these could be tiring for his eyesight, and suggested highlighting only the target within reach. Two participants (J1 and J2) stated that sometimes the area cursor seemed to slow the web browser down, although this may have been due to using a less powerful computer. There were also opposing assessments, while J4 found the area cursor very useful for him and thought he would like to use it frequently, participant J3 said he did not need it for assisting his web browsing. After extensive use of the area cursor, some assessments improved (such as items 6 and 9) while others worsened (such as items 5 and 10).

## VI. DISCUSSION

Both virtual cursors tested in this longitudinal study proved to be beneficial for the participants according to the results from the quantitative and qualitative analyses carried out. The cross cursor and the area cursor improved performance and satisfaction of the participants of the keyboard users group (KU), and the joystick and trackball users group (JU) respectively, albeit to a different extent. The discussion of these results aspires to provide clues for further research on link selection assistance.

### A. THE CROSS CURSOR

The quantitative results indicated that the cross cursor was extensively used by the KU group during the free navigation task (active on average 97.2% of the time, SD = 2.4), resulting in an average of 61.3 sessions accessing the Web per participant (SD = 21.9). The amount of invalid trials on the target acquisition task was slightly reduced by KU group participants when using the cross cursor (Table 3), showing that this virtual cursor was not more error prone on link selection than the original cursor. Results from statistical tests showed that all participants from the KU group significantly improved link selection with the cross cursor in comparison with the original cursor according to 6 out of the 7 performance parameters studied (Fig. 8a). The cross cursor outperformed the original cursor in the following parameters: movement time (MT), pointing time (PT), distance travelled (DT), curvature index (CI), number of pauses (NP) and target re-entry (TR), which confirmed cross cursor benefits on pointing trajectories and on accuracy to reach targets. However, the clicking time (CT) significantly worsened with the cross cursor in comparison with the original cursor for any member of the KU group. This was due to the fact that, in order to leverage the cross cursor, the user had to identify on the screen the letter assigned to the reachable target and then type it on the keyboard. By contrast, for clicking links with the original cursor, keyboard-only users always stroke the same key (Fig. 3a). Despite this worse performance on CT with the cross cursor, all KU group participants significantly improved total MT (which includes CT) using the cross cursor in comparison with the original cursor. Considering that keyboard-only users can type some keys more easily than others depending on their physical condition, it should be studied how to map the shortcuts to links in order to optimize CT. For instance, it would be beneficial to identify the most efficient shortcuts for each user considering their keyboard use. Therefore, the easiest shortcuts may be assigned to the most relevant links within visited web pages.

The members of the KU group achieved similar statistical results regardless of the target configuration being tested (effect sizes in Fig. 8a). According to these results, the performance with the cross cursor of the keyboard users was not affected by how far apart links are presented from each other within a web page.

Qualitatively, similar and highly positive opinions about the usability of the cross cursor were given by KU group participants (Fig. 9). The average SUS score from the KU group for the cross cursor was 90 points out of 100 (SD = 10.6) at the beginning of the study, and 93.1 points (SD = 9.4) at the end, showing a small improvement after extensive use. Despite all KU group participants having been trained to use the cross cursor, several declared that it was confusing at first and that they needed to get used to this new assistance to be able to take better advantage of it. Progression of the KU group throughput (TP) during the course of the 15 target acquisition task sessions showed an improvement on performance with the cross cursor when compared with the original cursor (Fig. 10a). The corresponding linear functions of each cursor variant reveal that the KU group improved their TP almost 10 times more with the cross cursor than with the original cursor. More training with advanced guidance on key aspects of this new virtual cursor functioning may reduce the learning period.

### B. THE AREA CURSOR

Quantitative results showed that the area cursor was extensively used by the JU group during the free navigation task (active on average 94% of the time, SD = 8). Each participant accessed the Web for an average of 35.3 sessions (SD = 19.8) throughout the study. The amount of invalid trials on the target acquisition task was slightly reduced by JU group participants when using the area cursor (Table 3), showing that this virtual cursor was not more error prone on link selection than the original cursor. The results from the statistical tests showed that the JU group participants significantly improved

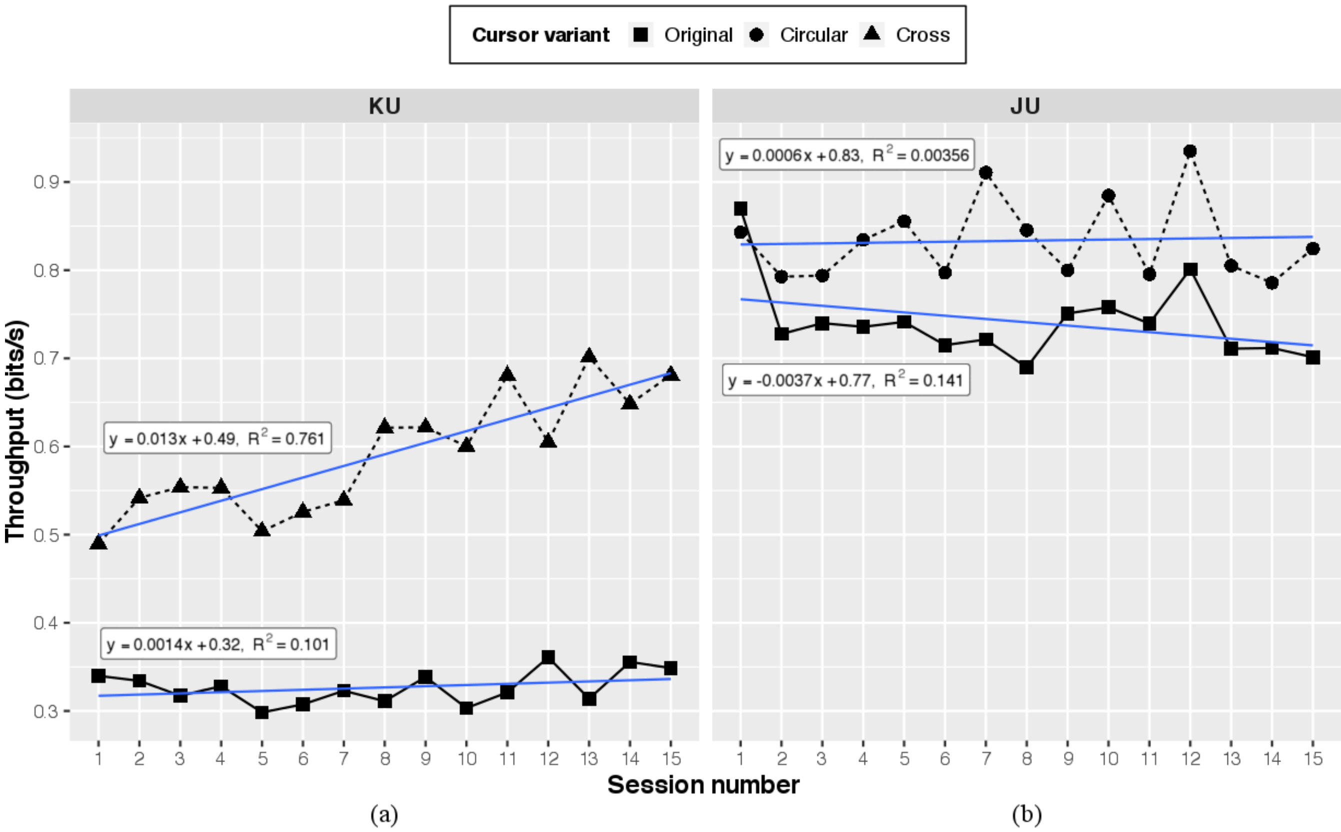


**FIGURE 10.** Relationship between the average throughput (TP) of each user group with each cursor variant, and the target acquisition task session (from 1 to 15): a) KU group with the original and cross cursor, and b) JU group with the original and area cursor. The TP was calculated by dividing the index of difficulty (bits) of trials as defined by [60], between the mean MT. Graphs also include linear regression lines and corresponding prediction equations.

link selection with the area cursor in comparison with the original cursor according to 6 of the 7 performance measurements studied, although there were variations between participants and target configurations tested (Fig. 8b). Statistical results for the MT, PT, NP and TR registered significant differences in favour of the area cursor for all JU group participants, confirming its benefits on pointing trajectories and on accuracy for reaching targets. By contrast, two JU group participants (J1 and J4) significantly improved with the area cursor in comparison with the original cursor on the DT, whereas only one (J4) on the CI. On the other hand, their CT significantly worsened with the area cursor, partly due to the implementation of this virtual cursor. Each time the pointer stopped moving, the area cursor needed a few milliseconds to calculate and highlight the closest target. Although this issue was not mentioned by participants, an improved implementation of the area cursor should reduce this response time. In addition, results from statistical tests (Fig. 8b) showed that for most cursor measurements, performance of JU group participants with the area cursor improved on less restrictive target configurations (Fig. 7b and Fig. 7c). On the other hand, the benefits of the area cursor were attenuated, or even disappeared for some participants, on the most restrictive target configuration (Fig. 7a).

According to these results and to the opinions of some participants, who argued that the area cursor added too much information to Web GUIs, this virtual cursor may be improved by adapting its assistance dynamically. Assistance may be automatically deactivated when moving over links that are very close to each other. Other cursor enhancements relying on magnification approaches should be tested in those cases [48]. However, the assistance may be activated again when moving over links that are further apart. Introducing an adaptive bubble cursor approach [61] that dynamically increases and decreases the selection area according to cursor proximity to the surrounding links may also be useful. In addition to these enhancements, a different approach can be adopted: that of adapting the virtual cursor to the type of web page being visited, its structure, or the areas detected within its layout (navigation bar, banner, content, etc.). In order to avoid the loss of performance from the standard area cursor, this could be combined with a transcoding system for adapting web content on the fly. For instance, a transcoding system that automatically adapts web pages, such as the one presented by Valencia *et al.* [62], could be used to increase the distance between close links.

Qualitatively, the usability of the area cursor was evaluated positively by participants from the JU group (Fig. 9), although on average it received lower scores than the cross cursor. Average SUS score from the JU group for the area cursor was 78.8 points out of 100 (SD = 4.3) at the beginning of the study, and 77.5 points (SD = 8.4) at the end, showing similar scores but a higher variability after prolonged used. Several participants from the JU group highlighted difficulties to leverage the area cursor when the links were too close together, as indicated by the answers to items 2, 5, 6 and 9 (Fig. 9). Although the functioning of the area cursor was assessed by participants as being easy to understand from the beginning, no improvement was appreciated on the TP progression over the course of the 15 sessions of the target acquisition task (Fig. 10b). Despite the fact that evolution of the TP shows a strong variation across sessions for both cursor variants being tested, participants of the JU group achieved, on average, better performance with the area cursor than with the original cursor. Performance improvements should be studied by testing the different enhancements proposed in the previous paragraph for selection of closely spaced links with the area cursor.

## VII. CONCLUSIONS

The longitudinal study we have presented here aims to explore the long-term benefits of two virtual cursors for assisting link selection on the Web to two different groups of people with MIs. Two groups of experienced users with different alternative pointing devices participated from their home on this remote 6-week unsupervised study to evaluate each virtual cursor: the novel cross cursor by keyboard-only users, and a standard area cursor by joystick and trackball users.

Interaction data and subjective assessments were collected over the six weeks. Generally, participants assessed both virtual cursors positively. Although they were able to deactivate the assistance during web browsing, participants extensively used both virtual cursors (more than 80% of the total time).

Quantitative results showed that the cross and area cursors improved the performance of both groups of participants compared to the original cursor. Except for the clicking time, the other six performance measurements studied were significantly improved with both virtual cursors. The group of joystick and trackball users improved their performance with the area cursor in comparison with the original cursor on the less restrictive target configurations (i.e., when the target link was further away from other links or distractors). On the other hand, the performance of keyboard-only participants did not worsen with the cross cursor compared to the original cursor on closely spaced links and was similar for all three target configurations tested.

Several improvements were proposed in order to reduce the clicking time both with the cross cursor and the area cursor. Other enhancements were also proposed to improve the performance of the standard area cursor on web environments with closely spaced links.

## ACKNOWLEDGMENTS

The authors thank all the individuals who participated in this study, as well as the participants' relatives who assisted the experimenter during supervised sessions.

## REFERENCES

[1] R. Balakrishnan, "'Beating–Fitts' law: Virtual enhancements for pointing facilitation," *Int. J. Hum.-Comput. Stud.*, vol. 61, no. 6, pp. 857–874, Dec. 2004, doi: 10.1016/j.ijhcs.2004.09.002.

[2] A. Sears, M. Young, and J. Feng, "Physical disabilities and computing technologies: An analysis of impairments," in *The Human-Computer Interaction Handbook*, J. A. Jacko A. Sears, Eds. New York, NY, USA: Lawrence Erlbaum, 2008, pp. 829–852.

[3] S. Trewin and H. Pain, "Keyboard and mouse errors due to motor disabilities," *Int. J. Hum.-Comput. Stud.*, vol. 50, no. 2, pp. 109–144, Feb. 1999, doi: 10.1006/ijhc.1998.0238.

[4] A. M. Cook and J. M. Polgar, "Control interfaces for assistive technologies," in *Assistive Technologies: Principles and Practice*, 4th ed. Amsterdam, The Netherlands: Elsevier, 2014, pp. 139–170.

[5] WEBAI. *Assistive Technologies*. Accessed: Aug. 23, 2019. [Online]. Available: https:webaim.org/articles/motor/assistive

[6] S. Keates, P. Langdon, P. J. Clarkson, and P. Robinson, "User models and user physical capability," *User Model. User-Adapt. Interact.*, vol. 12, nos. 2–3, pp. 139–169, Jun. 2002, doi: 10.1023/A:1015047002796.

[7] A. Almanji, T. Claire Davies, and N. Susan Stott, "Using cursor measures to investigate the effects of impairment severity on cursor control for youths with cerebral palsy," *Int. J. Hum.-Comput. Stud.*, vol. 72, no. 3, pp. 349–357, Mar. 2014, doi: 10.1016/j.ijhcs.2013.12.003.

[8] T. Surdilovic and Y.-Q. Zhang, "Convenient intelligent cursor control web systems for Internet users with severe motor-impairments," *Int. J. Med. Informat.*, vol. 75, no. 1, pp. 86–100, Jan. 2006, doi: 10.1016/j.ijmedinf.2005.07.042.

[9] F. Hwang, "Partitioning cursor movements in 'point and click' tasks," in *Proc. Ext. Abst. Hum. Fact. Comput. Sys.*, Fort Lauderdale, FL, USA, Apr. 2003, pp. 682–683, doi: 10.1145/765891.765927.

[10] F. Hwang, S. Keates, P. Langdon, and J. Clarkson, "Mouse movements of motion-impaired users: A submovement analysis," in *Proc. ACM SIGACCESS Conf. Comput. Accessibility*, 2004, pp. 102–109, doi: 10.1145/1028630.1028649.

[11] S. Trewin, "Physical impairments," in *Web Accessibility: A Foundation for Research*, S. Harper Y. Yesilada, Eds. London, U.K.: Springer, 2008, pp. 37–46, doi: 10.1007/978-1-84800-050-6_4.

[12] S. Keates, F. Hwang, P. Langdon, P. J. Clarkson, and P. Robinson, "Cursor measures for motion-impaired computer users," in *Proc. 5th Int. ACM Conf. Assistive Technol.*, 2002, pp. 18–29, doi: 10.1007/s10209-002-0033-7.

[13] A. R. Payne, B. Plimmer, A. McDaid, A. Luxton-Reilly, and T. C. Davies, "How can adding a movement improve target acquisition efficacy?" in *16th IFIP TC 13 Int. Conf. Hum.-Comput. Interact. (INTERACT)*, Mumbai, India, Sep. 2017, pp. 496–514, doi: 10.1007/978-3-319-67687-6_34.

[14] J. E. Pérez, M. Arrue, X. Valencia, and L. Moreno, "Exploratory study of web navigation strategies for users with physical disabilities," in *Proc. 11th Int. Cross-Discip. Conf. Web Access. (W4A)*, Seoul South Korea, 2014, Art. no. 20, doi: 10.1145/2596695.2596715.

[15] X. Valencia, J. E. Pérez, U. Muñoz, M. Arrue, and J. Abascal, "Assisted interaction data analysis of web-based user studies," in *15th IFIP TC 13 Int. Conf. Hum.-Comput. Interact. (INTERACT)*, Bamberg, Germany, Sep. 2015, pp. 1–19, doi: 10.1007/978-3-319-22701-6_1.

[16] J. E. Pérez, X. Valencia, M. Arrue, and J. Abascal, "Depicting the keypad mouse interface: Exploring pointing behaviors of three input devices," in *Proc. 17th Int. ACM SIGACCESS Conf. Comput. Accessibility (ASSETS)*, Lisbon, Portugal, Oct. 2015, pp. 413–414, doi: 10.1145/2700648.2811396.

[17] X. Valencia, J. E. Pérez, M. Arrue, J. Abascal, C. Duarte, and L. Moreno, "Adapting the Web for people with upper body motor impairments using touch screen tablets," *Interacting Comput.*, vol. 29, no. 6, pp. 794–812, Nov. 2017, doi: 10.1093/iwc/iwx013.

[18] M. A. Velasco, A. Clemotte, R. Raya, R. Ceres, and E. Rocon, "Human-computer interaction for users with cerebral palsy based on head orientation. Can cursor's movement be modeled by Fitts's law?" *Int. J. Hum.-Comput. Stud.*, vol. 106, pp. 1–9, Oct. 2017, doi: 10.1016/j.ijhcs.2017.05.002.

[19] J. Mankoff, H. Fait, and R. Juang, "Evaluating accessibility by simulating the experiences of users with vision or motor impairments," *IBM Syst. J.*, vol. 44, no. 3, pp. 505–517, 2005, doi: 10.1147/sj.443.0505.

[20] M. Y. Ivory, J. Mankoff, and A. Le, "Using automated tools to improve web site uage by users with diverse abilities," *Inf. Tech. Soc.*, vol. 1, no. 3, pp. 195–236, Jan. 2003.

[21] J. Varona, C. Manresa-Yee, and F. J. Perales, "Hands-free vision-based interface for computer accessibility," *J. Netw. Comput. Appl.*, vol. 31, no. 4, pp. 357–374, Nov. 2008, doi: 10.1016/j.jnca.2008.03.003.

[22] M. A. Velasco, A. Clemotte, R. Raya, R. Ceres, and E. Rocon, "A novel head cursor facilitation technique for cerebral palsy: Functional and clinical implications," *Interacting Comput.*, vol. 29, no. 5, pp. 755–766, Jul. 2017, doi: 10.1093/iwc/iwx009.

[23] J. A. Bilmes, X. Li, J. Malkin, K. Kilanski, R. Wright, K. Kirchhoff, A. Subramanya, S. Harada, J. A. Landay, P. Dowden, and H. Chizeck, "The vocal joystick: A voice-based human-computer interface for individuals with motor impairments," in *Proc. Conf. Hum. Lang. Tech. Emp. Meth. Nat. Lang. Proc. (HLT)*, Vancouver, BC, Canada, Oct. 2005, pp. 995–1002, doi: 10.3115/1220575.1220700.

[24] M. Mahmud, A. J. Sporka, S. H. Kurniawan, and P. Slavík, "A comparative longitudinal study of non-verbal mouse pointer," in *Proc. 11th IFIP TC 13 Int. Conf. Hum.-Comput. Interact. (INTERACT)*, Rio de Janeiro, Brazil, Sep. 2007, pp. 489–502, doi: 10.1007/978-3-540-74800-7_44.

[25] S. Harada, J. O. Wobbrock, J. Malkin, J. A. Bilmes, and J. A. Landay, "Longitudinal study of people learning to use continuous voice-based cursor control," in *Proc. SIGCHI Conf. Hum. Fact. Comput. Sys. (CHI)*, Boston, MA, USA, Apr. 2009, pp. 347–356, doi: 10.1145/1518701.1518757.

[26] G. J. Capilouto, D. J. Higginbotham, B. McClenaghan, H. G. Williams, and J. Dickerson, "Performance investigation of a head-operated device and expanded membrane cursor keys in a target acquisition task," *Technol. Disability*, vol. 17, no. 3, pp. 173–183, Aug. 2005, doi: 10.3233/TAD-2005-17305.

[27] N. G. Tsagarakis and D. G. Caldwell, "Improving mouse-based computer interaction in users with weak upper limb motion control using a haptic assistive system," *IEEE Trans. Human-Mach. Syst.*, vol. 43, no. 2, pp. 177–187, Mar. 2013, doi: 10.1109/TSMCC.2012.2204872.

[28] T. Grossman and R. Balakrishnan, "The bubble cursor: Enhancing target acquisition by dynamic resizing of the cursor's activation area," in *Proc. SIGCHI Conf. Hum. Fact. Comput. Sys. (CHI)*, Portland, OR, USA, Apr. 2005, pp. 281–290, doi: 10.1145/1054972.1055012.

[29] S. Harada, J. A. Landay, J. Malkin, X. Li, and J. A. Bilmes, "The vocal joystick: Evaluation of voice-based cursor control techniques," in *Proc. 8th Int. ACM SIGACCESS Conf. Comput. Access. (ASSETS)*, Portland, OR, USA, Oct. 2006, pp. 197–204, doi: 10.1145/1168987.1169021.

[30] F. Hwang, S. Keates, P. Langdon, and P. J. Clarkson, "Multiple haptic targets for motion-impaired computer users," in *Proc. SIGCHI Conf. Hum. Fact. Comput. Sys. (CHI)*, Fort Lauderdale, FL, USA, Apr. 2003, pp. 41–48, doi: 10.1145/642611.642620.

[31] M. E. Mott and J. O. Wobbrock, "Beating the bubble: Using kinematic triggering in the bubble lens for acquiring small, dense targets," in *Proc. SIGCHI Conf. Hum. Fact. Comput. Sys. (CHI)*, Toronto, Canada, Apr. d2014, pp. 733–742, doi: 10.1145/2556288.2557410.

[32] A. R. Payne, B. Plimmer, A. McDaid, A. Luxton-Reilly, and T. C. Davies, "Expansion cursor: A zoom lens that can be voluntarily activated by the user at every individual click," in *Proc. 28th Austral. Conf. Comput.-Hum. Interact. (OzCHI)*, Launceston, Australia, Nov./Dec. 2016, pp. 81–90, doi: 10.1145/3010915.3010942.

[33] G. Salivia and J. P. Hourcade, "PointAssist: Assisting individuals with motor impairments," in *Proc. SIGCHI Conf. Hum. Fact. Comput. Sys. (CHI)*, Paris, France, Apr./May 2013, pp. 1213–1222, doi: 10.1145/2470654.2466157.

[34] S. Trewin, S. Keates, and K. Moffatt, "Developing steady clicks: A method of cursor assistance for people with motor impairments," in *Proc. 8th Int. ACM SIGACCESS Conf. Comput. Access. (ASSETS)*, Portland, OR, USA, Oct. 2006, pp. 26–33, doi: 10.1145/1168987.1168993.

[35] J. O. Wobbrock and K. Z. Gajos, "Goal crossing with mice and trackballs for people with motor impairments: Performance, submovements, and design directions," *ACM Trans. Access. Comput. (TACCESS)*, vol. 1, no. 1, May 2008, Art. no. 4, doi: 10.1145/1361203.1361207.

[36] J. E. Pérez, X. Valencia, M. Arrue, and J. Abascal, "A usability evaluation of two virtual aids to enhance cursor accessibility for people with motor impairments," in *Proc. 13th Int. Cross-Discip. Conf. Web Access. (W4A)*, Montreal, QC, Canada, Apr. 2016, Art. no. 20, doi: 10.1145/2899475.2899489.

[37] J. E. Pérez and M. Arrue, "Virtual cursors to enhance web accessibility for people with limited dexterity: Usability test results and future directions," *Newslett. ACM SIGACCESS Access. Comput.*, vol. 115, pp. 3–11, Jun. 2016, doi: 10.1145/2961108.2961109.

[38] J. E. Pérez, X. Valencia, M. Arrue, and J. Abascal, "Evaluation of two virtual cursors for assisting Web access to people with motor impairments," *Int. J. Hum.-Comput. Stud.*, vol. 132, pp. 81–98, Dec. 2019, doi: 10.1016/j.ijhcs.2019.08.001.

[39] A. Apaolaza, S. Harper, and C. Jay, "Understanding users in the wild," in *Proc. 10th Int. Cross-Discip. Conf. Web Access. (W4A)*, Rio de Janeiro, Brazil, May 2013, Art. no. 13, doi: 10.1145/2461121.2461133.

[40] A. Apaolaza, S. Harper, and C. Jay, "Longitudinal analysis of low-level web interaction through micro behaviours," in *Proc. 26th ACM Conf. Hyper. Soc. Med. (HT)*, Guzelyurt, Northern Cyprus, Sep. 2015, pp. 337–340, doi: 10.1145/2700171.2804453.

[41] C. Tossell, P. Kortum, A. Rahmati, C. Shepard, and L. Zhong, "Characterizing web use on smartphones," in *Proc. SIGCHI Conf. Hum. Fact. Comput. Sys. (CHI)*, Austin, TX, USA, May 2012, pp. 2769–2778, doi: 10.1145/2207676.2208676.

[42] J. P. Bigham, A. C. Cavender, J. T. Brudvik, J. O. Wobbrock, and R. E. Ladner, "WebinSitu: A comparative analysis of blind and sighted browsing behaviour," in *Proc. 9th Int. ACM SIGACCESS Conf. Comput. Access. (ASSETS)*, Tempe, AZ, USA, Oct. 2007, pp. 51–58, doi: 10.1145/1296843.1296854.

[43] H. Nicolau, K. Montague, T. Guerreiro, A. Rodrigues, and V. L. Hanson, "Typing performance of blind users: An analysis of touch behaviors, learning effect, and *in-situ* usage," in *Proc. 17th Int. ACM SIGACCESS Conf. Comput. Access. (ASSETS)*, Lisbon, Portugal, Oct. 2015, pp. 273–280, doi: 10.1145/2700648.2809861.

[44] A. J. Sporka, T. Felzer, S. H. Kurniawan, O. Polácek, P. Haiduk, and I. S. MacKenzie, "CHANTI: Predictive text entry using non-verbal vocal input," in *Proc. SIGCHI Conf. Hum. Fact. Comput. Sys. (CHI)*, Vancouver, BC, Canada, Canada, May 2011, pp. 2463–2472, doi: 10.1145/1978942.1979302.

[45] J. O. Wobbrock, J. Fogarty, S. Y. Liu, S. Kimuro, and S. Harada, "The angle mouse: Target-agnostic dynamic gain adjustment based on angular deviation," in *Proc. SIGCHI Conf. Hum. Fact. Comput. Sys. (CHI)*, Boston, MA, USA, Apr. 2009, pp. 1401–1410, doi: 10.1145/1518701.1518912.

[46] A. Worden, N. Walker, K. Bharat, and S. Hudson, "Making computers easier for older adults to use: Area cursors and sticky icons," in *Proc. ACM SIGCHI Conf. Hum. Fact. Comput. Sys. (CHI)*, Atlanta, GA, USA, Mar. 1997, pp. 266–271, doi: 10.1145/258549.258724.

[47] P. Kabbash and W. A. S. Buxton, "The 'prince' technique: Fitts' law and selection using area cursors," in *Proc. SIGCHI Conf. Hum. Fact. Comput. Sys. (CHI)*, Denver, CO, USA, May 1995, pp. 273–279, doi: 10.1145/223904.223939.

[48] L. Findlater, A. Jansen, K. Shinohara, M. Dixon, P. Kamb, J. Rakita, and J. O. Wobbrock, "Enhanced area cursors: Reducing fine pointing demands for people with motor impairments," in *Proc. 23th Symp. User Interface Soft. Tech. (UIST)*, New York, NY, USA, Oct. 2010, pp. 153–162, doi: 10.1145/1866029.1866055.

[49] T. Felzer, I. S. MacKenzie, and J. Magee, "Comparison of two methods to control the mouse using a keypad," in *Proc. Int. Conf. Comput. Help. Peop. Spec. Needs (ICCHP)*, Linz, Austria, Jul. 2016, pp. 511–518, doi: 10.1007/978-3-319-41267-2_72.

[50] J. Lazar, J. H. Feng, and H. Hochheiser, "Working with research participants with disabilities," in *Research Methods in Human-Computer Interaction*. San Mateo, CA, USA: Morgan Kaufmann, 2017, pp. 493–522.

[51] M. Arrue, X. Valencia, J. E. Pérez, L. Moreno, and J. Abascal, "Inclusive Web empirical studies in remote and *in-situ* settings: A user evaluation of the remoTest platform," *Int. J. Hum.-Comput. Interact.*, vol. 35, no. 7, pp. 568–583, May 2018, doi: 10.1080/10447318.2018.1473941.

[52] World Wide Web Consortium. *Web Content Accessibility Guidelines 2.02008*. Accessed: Sep. 3, 2019. [Online]. Available: www.w3.org/TR/WCAG20

[53] S. H. Kurniawan, A. King, D. G. Evans, and P. L. Blenkhorn, "Personalising web page presentation for older people," *Interact. Comput.*, vol. 18, no. 3, pp. 457–477, May 2006, doi: 10.1016/j.intcom.2005.11.006.

[54] G. Brajnik. (2004). *Barrier Walkthrough Method*. Accessed: Sep. 3, 2019. [Online]. Available: www.users.dimi.uniud.it/ giorgio. brajnik/projects/bw/bw.html

[55] J. Brooke, "SUS: A 'quick and dirty' usability scale," in *Usability Evaluation in Industry*, P. W. Jordan, B. Thomas, I. L. McClelland, and B. Weerdmeester, Eds. London, U.K.:, CRC Press, 1996, ch. 21, pp. 189–194, doi: 10.1201/9781498710411.

[56] I. S. MacKenzie, T. Kauppinen, and M. Silfverberg, "Accuracy measures for evaluating computer pointing devices," in *Proc. SIGCHI Conf. Hum. Fact. Comput. Sys. (CHI)*, Seattle, WA, USA, Mar. 2001, pp. 9–16, doi: 10.1145/365024.365028.

[57] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *J. Express Psyc.*, vol. 47, no. 6, pp. 381–391, Jun. 1954, doi: 10.1037/h0055392.

[58] I. S. MacKenzie, "Fitts' law as a research and design tool in human-computer interaction," *Hum.-Comput Interact.*, vol. 7, no. 1, pp. 91–139, 1992, doi: 10.1207/s15327051hci0701_3.

[59] A. Gump, M. Legare, and D. L. Hunt, "Application of fitts' law to individuals with cerebral palsy," *Percept. Motor Skills*, vol. 94, no. 3, pp. 883–895, Jun. 2002, doi: 10.2466/pms.2002.94.3.883.

[60] J. Accot and S. Zhai, "Refining Fitts' law models for bivariate pointing," in *Proc. SIGCHI Conf. Hum. Fact. Comput. Sys. (CHI)*, Fort Lauderdale, FL, USA, Apr. 2003, pp. 193–200, doi: 10.1145/642611.642646.

[61] A. Martin-Hammond, F. Hamidi, T. Bhalerao, C. Ortega, A. Ali, C. Hornback, C. Means, and A. Hurst, "Designing an adaptive web navigation interface for users with variable pointing performance," in *Proc. 15th Int. Cross-Discip. Conf. Web Access. (W4A)*, Lyon, France, Apr. 2018, Art. no. 31, doi: 10.1145/3192714.3192818.

[62] X. Valencia, M. Arrue, J. E. Pérez, and J. Abascal, "User individuality management in websites based on WAI-ARIA annotations and ontologies," in *Proc. 10th Int. Cross-Discip. Conf. Web Access. (W4A)*, Rio de Janeiro, Brazil, May 2013, Art. no. 29, doi: 10.1145/2461121.2461128.

**J. EDUARDO PÉREZ** received the B.Sc. and M.Sc. degrees in computer science from the University of the Basque Country (UPV/EHU), in 2007 and 2011, respectively, where he is currently pursuing the Ph.D. degree.

He is also a member of the Egokituz Laboratory of HCI for Special Needs, UPV/EHU. His research interests focus on studying web navigation behaviors within diverse groups of users with special needs and testing different GUI enhancements in order to improve their web browsing ability.

**MYRIAM ARRUE** received the M.Sc. and Ph.D. degrees in computer science from the University of the Basque Country (UPV/EHU), in 1999 and 2008, respectively.

She is currently an Assistant Professor with the UPV/EHU and a member of the Egokituz Laboratory of HCI for Special Needs. Her research focuses on web accessibility evaluation, web navigation behavior analysis, and the application of transcoding techniques for improving UX in web interactions.

**XABIER VALENCIA** received the M.Sc. and Ph.D. degrees in computer science from the University of the Basque Country (UPV/EHU), in 2008 and 2017, respectively.

He is currently a Researcher with the Egokituz Laboratory of HCI for Special Needs (UPV/EHU). His main research interest is focused on providing greater accessibility to the Web for people with diverse capabilities, principally by applying transcoding techniques.

**JULIO ABASCAL** (Member, IEEE) received the BSD degree in physics from the Universidad de Navarra, in 1978, and the Ph.D. degree in informatics from the University of the Basque Country-Euskal Herriko Unibersitatea, in 1987.

He is currently a Professor of the Computer Architecture and Technology Department, UPV/EHU, where he has been working, since 1981. In 1985, he co-founded Egokituz Laboratory of Human-Computer Interaction for Special Needs, which has participated in several Research and Development project at national and international levels.