

Cognitive Impact Evaluation of Multimodal Interfaces for Blind People: Towards a Systematic Review

Lana Mesquita^{*1}, Jaime Sánchez^{**} and Rossana M. C. Andrade^{*2}

^{*} Department of Computer Science,

Universidade Federal do Ceará (UFC), Fortaleza CE, Brazil

^{**} Department of Computer Science, Universidad de Chile, Santiago, Chile

lanamesquita@great.ufc.br, jsanchez@dcc.uchile.cl, rossana@ufc.br

Abstract. Visual disability has a major impact on people's quality of life. Although there are many technologies to assist people who are blind, most of them do not necessarily guarantee the effectiveness of the intended use. Then, we have conducted a systematic literature review concerning the cognitive impact evaluation of multimodal interfaces for blind people. We report in this paper the preliminary results of the systematic literature review with the purpose of understanding how the cognitive impact is currently evaluated when using multimodal interfaces for blind people. Among twenty-five papers retrieved from the systematic review, we found a high diversity of experiments. Some of them do not present the data clearly and do not apply a statistical method to guarantee the results. Besides this, other points related to the experiments are analyzed. We conclude that there is a need to better plan and present data from experiments on technologies for cognition of blind people. Moreover, as the next step in this research, we will investigate these preliminary results with a qualitative analysis.

Keywords: Impact Evaluation; Cognitive Evaluation; Multimodal Interfaces; Blind People.

1 Introduction

Visual disability has a major impact on the quality of life of people who has it, including their ability to study, work and to develop personal relationships [1]. In this aspect, technologies, such as serious games [2], have been designed to assist people who are blind to support daily life activities. These technologies work as aids to facilitate their independence, autonomy, and safety. Thus, such technologies improve the quality of life of people with visual disabilities and could stimulate and develop several skills, such as cognitive skills [3].

Even though there is technology specialized for blind people (i.e., visually impaired), they are still using applications that are similar to older applications for the sighted population. For example, Battleship was one of the earliest games to be produced as a computer game with its release in 1979 [4]. AudioBattleShip, a version for both blind

¹ Scholarship (MDCC/DC/UFC) sponsored by FUNCAP

² Researcher scholarship - DT Level 2, sponsored by CNPq

children and sighted playing together came many years later [5]. In general, blind people have particular human-computer interaction needs, and the user interface should be suitable for them.

There are many efforts towards to develop accessible multimodal interfaces for visually impaired, especially in multimodal games [6, 7]. Despite this effort and in contrast to the visual interface evolutions of games and applications for sighted people, interfaces for people who are blind explore other ways to interact with the user. In general, technologies for people with visual disabilities combine different sources of perceptual inputs and outputs. The modes (sources of perceptual inputs and outputs) combined, typically audio and haptics [5, 8], provide multimodal interfaces that enable multimodal channels for the combination of different user senses [2]. Although multimodal interfaces could help to improve the learning skills of people with visual disabilities, most of these technologies have been not completely validated; mostly, they remain in the prototype phase without being integrated into the people's everyday life [9].

In relation to the quality of applications, the No Child Left Behind (NCLB) Act defined that research in inclusive education must (a) utilize the scientific method, (b) be replicated in more than one set by more than one investigator, and, (c) result in findings that converge to a clear conclusion [10]. Thus, some studies in this area use Evidence-Based Practice [5, 11], which meets prescribed criteria related to the research design, quality, quantity, and effect size of supporting research [12, 13]. Thus, this method provides the measurement of the effectiveness of using technology.

Considering these multimodal interactions with interfaces for blind people, it is necessary to verify if the technologies thought for them are effective and how they impact users in cognitive dimensions [14]. Effective impact evaluation, which has been used as evidence-based in other domains for users with and without disabilities, should, therefore, be able to assess precisely the mechanisms by which people with visual disabilities are developing or enhancing cognitive skills [3]. According to Darin et al. (2015), there is a gap of studies proposing instruments and methods for evaluating the cognitive impact in the context of multimodal video games for cognitive enhancement of people who are blind. In general, the literature studies do not follow guidelines to support cognitive impact evaluation of multimodal interfaces for blind people.

In pursuing to shed some light concerning this, our work presents the state-of-the-art study on cognitive impact evaluation of multimodal interfaces for blind people with profound inability to distinguish light from dark, or the total inability to see [15]. This state-of-the-art study consists of a systematic review to analyze how studies evaluate the cognitive impact in this context. The systematic review is part of an ongoing work on the design of guidelines for evaluating the cognitive impact development and enhancement in multimodal interfaces for blind people. As related literatures, some works also propose guidelines for other concepts in the context of accessibility. As an example, the study [16] defines games accessibility guidelines helping developers to design their products with the interface and the parameters adapted to the needs of users with disabilities.

The remainder of this paper is organized as follows: Section 2 includes the Theoretical Background; Section 3 presents the methodology used in this work; Section 4

covers the results of the systematic review; Section 5 discusses the results; and, finally, Section 6 concludes the study.

2 Background

2.1 Cognitive Impact Evaluation

There are technological aspects to these systems which are investigated from the perspective of how they affect use [17]. Cognitive impact concerns the interaction between humans and systems. The field of human-computer interaction has pioneered in the formal study of the cognitive relationship between a person's activities, the artifact of the computer, and the task [18]. The technologies could enhance human cognitive capabilities.

Darin et al.'s (2015) literature mapping study shows the cognitive process analyze applications according to a four-dimensional classification (Interface, Interaction, Cognition, and Evaluation). The evaluation dimension includes two main aspects: usability and cognitive impact. This last one assures that an application can develop or enhance any cognitive skills for people with visual disabilities.

Still on this study, the cognition dimension comprises six skills: mental models, mental maps, spatial structures, Orientation and Mobility (O&M), problem-solving, and social collaboration. Such approach addresses the main cognitive skills developed and enhanced for impact evaluation purposes. These dimensions could provide directions to define tasks in an experiment to measure the cognitive impact as detection of some obstacles, a useful data for evaluating O&M [19].

The study also shows that most papers classified in main cognitive skills are about Mental Map and O&M. The O&M skill is a broad concept that is also related to way-finding and navigation. According to Pissaloux and Velázquez (2018, p. 1), "Human mobility is one of the most important cognitive tasks. Indeed, independent and secure mobility in real physical space has a direct impact on the quality of life, on well-being, and on integration in the numeric society."

Darin et al [19] define mobility in a four-dimensional problem: *walking*, *wayfinding* (or orientation), *space awareness* (or space knowledge) and *navigation*. According to this definition, walking is a low conscious cognitive task and involves displacement in the near space. It takes in account obstacle detection and localization. The wayfinding is a set of processes to know one's current position in space to reach one's target. The space awareness requires a high consciousness level. It includes forming mental maps, e.g., know the name of the street on a plan. The navigation, the highest level cognitive task, is a result of the implementation of all listed above functions while traveling.

2.2 Experiment in software engineering

The impact evaluation of a software is an experiment process and includes several steps: Scoping; Planning; Operation; Analysis and interpretation; Presentation and package [20]. This provides a high level of control, using a formal, rigorous and controlled investigation.

The main concepts involved in the experiment, shown in Table 1, are used to understand how the cognitive impact is evaluated and to conduct the designing of the guidelines. Figure 1 shows how these concepts are related to the experimental process.

Table 1. The main concepts of experimental design [20]

Concepts	Description
Measure	A mapping from the attribute of an entity to a measurement value.
Instrumentation	The instruments for an experiment are of three types, namely objects, guidelines and measurement instruments.
Dependent Variables	The dependent variables are those we want to see the effect; the independent variables are those controlled and manipulated.
Independent Variables	The independent variables are those controlled and manipulated.
Factors	The independent variables which the experiment changes to verify the effect. Treatment is one value of a factor.

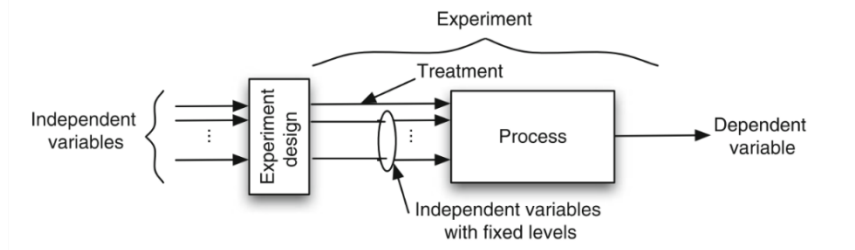


Fig. 1. Variable relationship in the experiment process (Source: [20]).

As an example of experiment process, we consider the measurements, instrumentation, and variables from the experiment conducted in [21]. In this study, the authors evaluate the navigational performance of a virtual environment called Audio-based Environment Simulator (AbES) that can be explored for the purposes of learning the layout of an unfamiliar, complex indoor environment. The dependent variable evaluated was the navigation performance (Orientation & Mobility, O&M).

Some information about the participants is controlled and works as independent variables such as etiology of blindness, age, gender, hand preference, and verbal memory (assessed by using the Wechsler Memory Scale). These variables are controlled and fixed to ensure the correct measurement. The factors in the experiment are the age of blindness onset and the interaction condition with AbES. The factors are randomly distributed into groups: early blind and late blind; and gamers, directed navigators, and control group. The measurements variables used are task success, navigation time, and shortest possible path score.

3 Methodology

The methodology consists in a Systematic Review [22], which aims to review the existing evidence concerning the impact evaluation of multimodal interfaces and also seeks to summarize the empirical evidence concerning the strengths and limitations of a specific evaluation method [22]. In contrast to an ad hoc literature review, the systematic review is a methodologically rigorous analysis and study of research results.

To achieve our goal, the main research question for this first part of the proposal was “How is the cognitive impact evaluated on multimodal interfaces for people who are blind?” For a better understanding, as a second goal question, we aim to learn the challenges regarding impact evaluation on this scenario.

The process of a systematic review includes three main phases: *planning the review*; *conducting the review* and *reporting the review* [22]. During all process of the systematic review, we used the tool StArt [23] and the software Microsoft Excel³ as a support to create the protocol, apply the filters, select the papers and show the results. We organize all references on software Mendeley⁴. As the papers retrieved from PubMed Central are in MEDLINE format, we developed the tool Medline2bibtex⁵. It works as a parser to permit the list to be read by both StArt and Mendeley.

The next subsections describe the planning (the study selection criteria, the research sources selected) and the conducting phase (the search Process, the data extraction form fields and the studies quality evaluation). The entire process was stored in an excel worksheet available online⁶.

3.1 Planning: definition of the protocol

In the planning phase, we define a review protocol that specifies the research question being addressed and the methods that will be used to perform the review [22]. Based on the goal of systematic review we define the search string as shown in the Figure 2.

```
( ((evaluat* OR assessment)
  AND (cognitive NEAR (impact or effectiveness)))
AND (design OR development)
AND (blind* OR "visually impaired" OR "visual disability")
AND (multimodal OR haptic OR audio OR auditory OR vibrotactile)
AND (interface OR "user interact*") )
```

Fig. 2. Search string

³ Microsoft Excel - <https://products.office.com/pt-br/excel>

⁴ Mendeley - <https://www.mendeley.com/>

⁵ Medline2Bibtex - <https://github.com/lanabia/Medline2Bibtex>

⁶ <https://www.dropbox.com/s/4wiiqnwyd3cquw/Systematic%20Review%20v5.xlsm?dl=0>

Study selection criteria. We define the search criteria as studies that present technology for people who are blind or visually impaired and has applied a cognitive impact evaluation. The inclusion and exclusion criteria are according to the goal of the systematic review. These requirements comply the general objective of this study, but also aim to have a broader view of the assessment in various technologies in the area. Table 2 presents the inclusion (*I*) and exclusion (*E*) selection criteria. To be accepted, a scientific paper must cover all inclusion criteria and none exclusion criterion.

Table 2. Inclusion and exclusion criteria

Code	Inclusion Criteria
<i>I.01</i>	The study must be published after 1998 and written in English.
<i>I.02</i>	The study has a technology for people who are blind or visually impaired according to the search criteria.
<i>I.03</i>	The study evaluates the technology by using some approach that involves the user.
<i>I.04</i>	The study evaluates the technology by using a method to evaluate the impact the cognitive impact
Code	Exclusion Criteria
<i>E.01</i>	Title and abstract out the search criteria (<i>I.02, I.03, I.04</i>).
<i>E.02</i>	Entire text out the search criteria (<i>I.02, I.03, I.04</i>).
<i>E.03</i>	The document is a book, a congress' abstract, an extended abstract, a poster, an oral communication, proceedings of a conference, a seminar, a research plenary, a dictionary or an encyclopedia.

The *I.01* criterion defines the scientific articles must be in English, because it is the mandatory language for the main events and scientific journals in the search area. And they must be published between 1st January 1998 and 2nd August 2017. The year 1998 was a milestone due to the paper Lumbreras et Sánchez (1998) which works with 3D acoustic interfaces for blind children and is the last study known [24].

The technologies defined in the *I.02* criterion include mobile application, computer software, IoT systems, virtual environments or a video game with multimodal interfaces. Also, we accept technologies that are not specifically for people who are blind or visually impaired with the goal of expanding the results, but the studies present the technology focused on users with visual disabilities. We exclude from all technologies that uses Sensory Substitution Devices [19], which substitutes a sense. The device SSD (Sensory Substitution Devices), out of our scope, includes sensory replacement, haptics as sensory augmentation, bionic eyes, retinal visual prosthesis, cortical implants and others. This definition is important to plan the methodology proposed and to delimit the focus.

We define the studies type in the *E.03*. This criterion excludes all studies type different from primary studies that present technology for people who are blind and its evaluation. We accept articles, conference papers, short papers, and book chapters. This

criterion includes documents that have the minimum information to understand the evaluation. We did not cover books because the information is dispersed inside them.

3.2 Conducting

In the conducting phase, firstly, we identify and select studies. To identify, we did a manual string research in five scientific bases: Scopus, Springer Link, PubMed, PubMed Central, and Web of Science. We chose the main research bases in the research area or the bases that index them [25]. Other bases were not included because they are indexed by the bases considered.

Search Process. The conducting phase starts with the initial search in the scientific bases proposed. The string was applied on the metadata of papers, which includes abstract, index terms, and bibliographic citation data (such as document title, publication title, etc.). It was retrieved 2136 papers. Figure 3 resumes the conducting phase process.

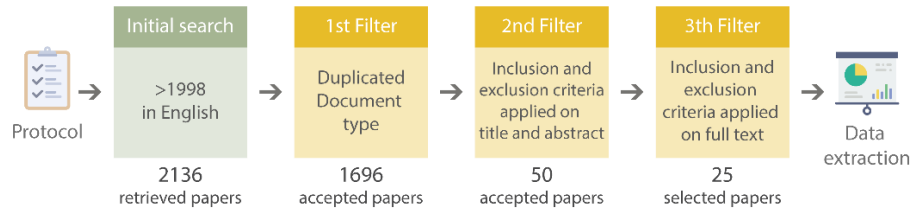


Fig. 3. Filters in the conducting phase

The first filter excluded papers duplicated and document types out the scope due their format (*E.03*). The second filter identifies which paper is in and out the scope by reading their titles and abstracts (*E.01*). A lot of papers were excluded in the first filter because the scientific base PubMed Central (PMC) brings a lot of medical papers focused on disease effectiveness and specific medical statements. Even though the area of this study is computer science, we decided to insert the PMC in the bases' list due to the nature of the subject.

Next, in the third filter, we evaluated each retrieved paper in its entirety (*E.02*). If necessary, besides the entire text, we search more about the technologies and processes described, as project and institutional websites, videos, newspaper articles and others.

Once we have chosen the select papers, we extract all data required (detailed in the protocol) to achieve the objective. The organization of the data generates data synthesis, which will be shown in the Section 5.1. It remains to be done the study quality assessment of each paper retrieved. Table 3 shows the quantity of papers selected in each filter per scientific bases. The main reason for withdrawing papers in the last filter was the evaluation performed is out the search and at most times related to the system performance, e.g., sensor performance evaluation.

Table 3. Quantity of papers accepted per filter

<i>Initial search</i>	<i>Filter 1</i>	<i>Filter 2</i>	<i>Filter 3</i>	<i>Acceptance rate</i>
-----------------------	-----------------	-----------------	-----------------	------------------------

<i>Scopus</i>	217 (10,2%)	135	9	5 (20%)	2,3%
<i>Springer</i>	424 (19,85%)	235	16	9 (36%)	2,1%
<i>WoS</i>	1 (0,05%)	1	1	1 (4%)	100,0%
<i>VIB</i>	7 (0,33%)	7	2	1 (4%)	14,3%
<i>PMC</i>	1486 (69,57%)	1317	21	8 (32%)	0,5%
<i>Manual</i>	1 (0,05%)	1	1	1 (4%)	100,0%
Total	2136	1696	50	25	-

Data Extraction Form Fields. The data extraction was designed to answer the main and second questions and to understand the context in which each paper is inserted. We divide the data collected into three categories: (i) General, (ii) Research and (iii) Empirical. The general category comprises bibliographic information.

The research category comprises the attributes of the research and the technology present in the paper. In this category we appraise the scientific paper into two classifications. The first one fits the paper according to the research type [26], which can be validation research, evaluation research, solution research, philosophical research, opinion paper or experience papers. The second classification fits the technology presented in the key features of multimodal interfaces for the cognition of people who are blind [3]. This classification is divided into 4-dimension: Interface, Interaction, Cognition, and Evaluation; and it is applied to video games and virtual environments (Figure 4).

For our purpose, we classify only in the interaction, interface and cognition dimensions; and we cover, in the classification, more than video games and virtual environments, since we also found these features present in the technologies selected. These features provide necessary insights for the practical understanding of the issues involved in their design and evaluation [3]. They are useful in our research for giving a comprehensive overview of technologies and evaluations regarding the multimodal interfaces. The research category still shows more information about the research, as other strategies used to evaluate.

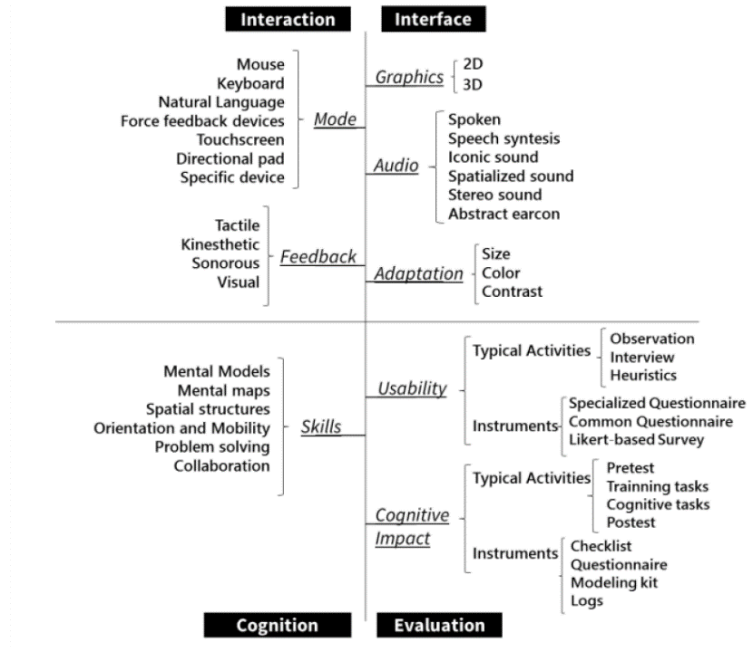


Fig. 4. Key features in multimodal interfaces (Image from [3])

The *empirical category* provides information specifically about how the empirical method evaluates the impact of the cognitive impact. The empirical method classification [27], which is retrieved in this category, classifies the empirical method in three types: Experiment, Survey and Case Study.

To conclude the achievements, all data pass by a manual analysis for acquiring qualitative and quantitative results. The section 4 of this paper brings the compilation of data extraction.

Studies quality evaluation. Study quality assessment. The quality evaluation was based on [28]. Although the areas are different, we apply adaptation that resulted in the quality checklist of Table 4. The checklist assesses the studies and measures the weight of each study found in the final results. Each question subtracts or adds points that give a general score to the empirical method. This score was defined according to the importance of the requested data and they are all linked to the data extraction form fields.

Table 4. Quality assessment form

ID	Question
<i>Q1</i>	Was it possible to extract all data regarding the data the Key features in multimodal interfaces (Classification category from extraction form? (-0.1 pts per missing input; min value: -4.1 pts)
<i>Q2</i>	Is there a complete description of how the evaluation has been applied? (1.0 pt per complete input in Empirical category; max value: 8,0 pts)

Q3	Are the groups of participants in the experiment randomly assigned? (0.5 pts)
Q4	Is the description of the impact evaluation understandable? (0.5 pts)
Q5	How many steps does the experiment take (e.g., Pretest, Training Tasks, Main Tasks and Posttest)? (0.08 pts per step)
Q6	Does the article present different evaluation types of the proposal? (0.25 pts x number of evaluations types)
Q7	How many experiments does the paper present? (0.5 pts per experiment, if more than two experiments)
Q8	Is the goal of the evaluation cleared defined? (0.5 pts)
Q9	Is the hypothesis (null and alternative) explicitly described in the study? (1 pt)

4 First Results of the systematic review

The systematic review brought 25 papers that have 28 experiments because three papers have two experiments per paper ([29][19][30]). Among these 25 papers, some of them do a cognitive impact evaluation in the same technology. Then the papers present 23 technologies for people who are blind or visually impaired as shown in Table 5. The next subsections explain all results obtained in each category of data.

Table 5. Technologies encountered

Technology	Technology description
AbES [21, 31, 32]	An environment to be explored for the purposes of learning the layout of an unfamiliar and complex indoor environment.
aMFS [29]	Audible Mobility Feedback System.
ATM machine [33]	ATM usage fees are the fees that many banks and interbank networks charge for the use of their automated teller machines (ATMs).
Audio Haptic Maze (AHM) [34]	AHM allows a school-age blind learner to be able to navigate through a series of mazes from a first-person perspective.
AudioBattle-Ship [5]	A sound-based interactive and collaborative environment for blind children. This system is a similar version of the traditional battleship game for sighted people but including both a graphical interface for sighted users and an audio-based interface for blind people.
AudioGene [35]	A game that uses mobile and audio-based technology to assist the interaction between blind and sighted children and to learn genetic concepts.
AudioLink [36]	An interactive audio-based virtual environment for children with visual disabilities to support their learning of science.
AudioMath [37]	An interactive virtual environment based on audio to develop and use short-term memory, and to assist mathematics learning of children.
AudioNature [38]	An audio-based virtual simulator for science learning implemented in a mobile device (PocketPC) platform.
Audiopolis [8, 39, 40]	A video game designed for developing orientation and mobility.
AudioStory-Teller [11]	A tool for PocketPC to support the development of reading and writing skills in learners with visual disabilities (LWVD) through storytelling, providing diverse evaluation tools to measure those skills.

Audi-oTransantiago [41]	A handheld application that allows users to plan trips and provide contextual information during the journey through the use of synthesized voices
BlindAid [42]	The system allows the user to explore a virtual environment.
Building Navigator [43]	A digital-map software with synthetic speech installed on a cell phone or PDA.
EyeCane [44]	A hand-held device which instantaneously transforms distance information via sound and vibration.
Méto	A software solution for blind users that represents a subway system.
Méto Mobile (mBN)	A software solution for blind users that represents a subway system in a mobile version.
Mobile devices	Mobile devices in general.
R-MAP	Reconfigured Mobile Android Phone is a fully integrated, stand-alone system, that has an easy-to-use interface to reconfigure an Android mobile phone.
TactiPad	An academic prototype of a tangible refreshable multimodal interface that has an interface similar to a Braille matrix.
Tower Defense	A video game that allows users who are blind to gradually build up a mental model based on references between different points on a Cartesian plane, in a way that is both didactic and entertaining.
UnrealEd	A 3D virtual environment.
vMFS	A vibrotactile mobility feedback system.

General Data. Among all papers, 14 are conference papers ([45][8][34][39][11][46][5][47][40][38][36][33][30][48]), 1 paper ([19]) is a book chapter and the others (10 papers) are from scientific journals. Figure 5 shows the papers selected regarding the timeline that started in 2004 and has peaked in 2011 and 2014. The affiliation most present is University of Chile with 15 papers.

Research Data. The papers selected present two types of research [26]: one paper is validation research ([19]), and 24 papers are evaluation research. The only validation research paper is defined in this way due to the technology presented is novel and have not yet been implemented in practice.

The interfaces, interactions and cognition skills of technologies presented in the papers are classified according to their key features of multimodal interfaces [3]. These characteristics are important to understand which type of interfaces are assessed and how the impact is evaluated on them. Figure 6 shows that the most notable mode of interaction is the keyboard and the least used is the mouse. The keyboard does not generate more complexity and expenses as the Novint Falcon device, used in [34]. The Novint Falcon is one of the Force Feedback Device that promotes a Tactile and Kinesthetic Feedback. Mouse and Natural language are less used. The mouse is replaced by buttons or other specific devices for better interaction as shown in [33]. The most common Feedback is the Sonorous and the main Audio Interface used is the Iconic Sound, which are sounds associated with each available object and action in the environment [8].

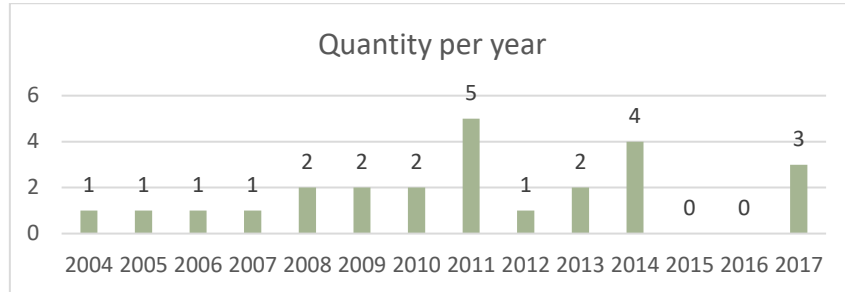


Fig. 5. Timeline papers

Many papers apply another approach to evaluate other criteria not covered in the cognitive evaluation. Concerning other strategies used to evaluate the interface, 14 papers ([42][29][34][35] [37][39][11][5][47][41][38][36][30][48]) applied a usability evaluation together with cognitive evaluation. Other strategies used were: evaluation of a tactile perception ([19][46]), system performance ([38]) that tested the hardware used, evaluation based on HCI heuristics ([33]), recognition of pattern ([19]), obstacle awareness goal ([19]), homing and obstacle avoidance ([19]) and iconic evaluation ([39]). Eight papers do not apply other strategy to evaluate the system or user interaction.

Usability evaluation is the most used assessment besides cognition impact evaluation. The usability evaluation is mainly used to obtain information about the user's acceptance of the software and the match between his or her mental model and the representation included in the software [36].

Empirical Data. Concerning the 28 empirical strategies that evaluate the cognitive impact, most papers applied experiments and only one applied a case study ([38]) due to the small number of participants. As our focus is the experiments, we treat in both types only the experiment criteria. Due to this fact, we call all empirical methods as experiments. The data that came from the Empirical category gave a substantial part of the comprehension of this research. Figure 7 presents all data extracted from each experiment. It is important to note that not all the papers presented the data sought. These faults were included in the quality criteria and then considered in the discussions and conclusions. The empirical data is counted according to the number of experiments (28 experiments), and not papers.

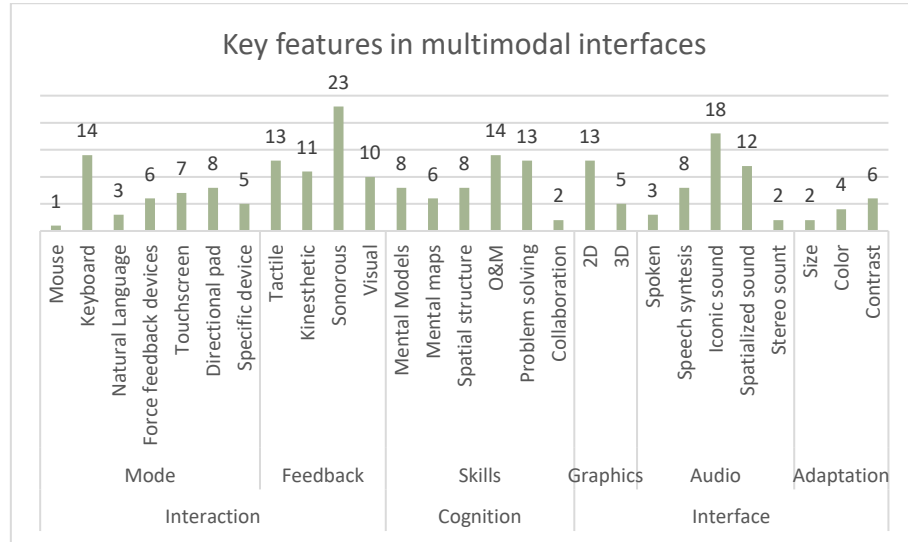


Fig. 6. Results of key features in multimodal interfaces

Instruments. The instruments are means for data collection in the cognitive impact evaluation with the objective of identifying some user ability controlled on the Experiment (as an independent variable), e.g., the mathematics knowledge test in [37] or is used to guide the evaluation process, e.g., observation guideline to assess O&M skills in [39]. Among them, there are 15 checklists ([45][8][34][37][39][11][46] [5][32][40][41]) which include guidelines and specific tests, 11 questionnaires ([32][42][29][43][35][11][46][47][38][36]), 7 interviews ([42][29][8][39] [41][33]), 6 modeling kits ([42][37][5][32][19][30]) and 9 logs ([42][29][5][31][32][36] [33][30][48][21]) which include, in addition to the system log, the video and audio logs. Many studies produce their instruments (7 experiments), others work with instruments found in scientific literature.

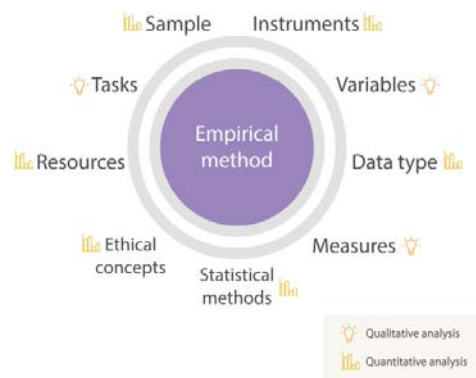


Fig. 7. Data extracted in Empirical category

Statistical methods. Figure 8 shows the statistical methods used in the experiment data analysis. Neither we take account of simple statistical methods as averages and gain percent, that is the only method used by 11 experiments ([42][35][37][31][32][38][19][33][30]). Three experiments ([45][39][36]) do not specify the statistical method used, neither in the references. T-test, which uses statistical concepts to reject or not a null hypothesis is the most used, followed by ANOVA and Pearson's Correlation Coefficients. These two methods are used to analyze the variation between groups. ANOVA procedure was applied in one, two and three-way.

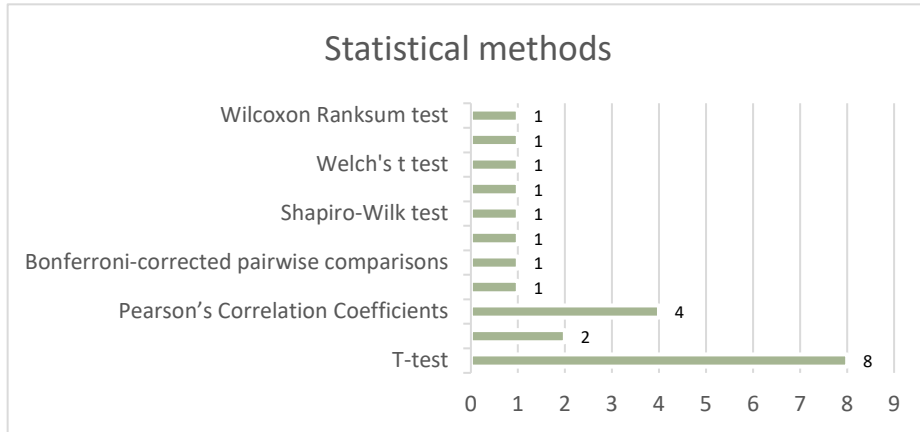


Fig. 8. Statistical Methods

Resources. Concerning the resources, 7 experiments ([37][11][41][38][36][30]) give information about the time spent in the evaluation and 2 experiments in the same paper ([29]) specify the human resource. The mean of the time among these is 3.4 months and the longest time spent is 6 months. None of them talk about financial resources.

Ethical concepts. 5 experiments ([29][43][21][44]) mentioned signing consents. One of them also applies stop rules to enforcing ethical concept [21] and present the ethics council that it approved.

Sample. From the number of users to the onset age of blindness, there are many sample combinations in the selected experiments. The sample choice could be based on level of experience required to do the task, age or level of blindness. Some characteristic controlled in the sample are related to the disabilities, as the onset of blindness, the etiology of a visual impairment or the presence of another disability. The quantity of users varies as shown in Figure 9. Most of the experiments (75%) are applied to 3 to 12 users. About the range of age, the most of experiments (9 experiments) are applied in teenagers (10 to 15 years old). The range and ages are shown in Figure 10.

The gender distribution in samples is equilibrated in most cases, but we not count the 11 experiments which do not describe this information. The mean of gender proportion is 50% for women and men, with 2% of variance.

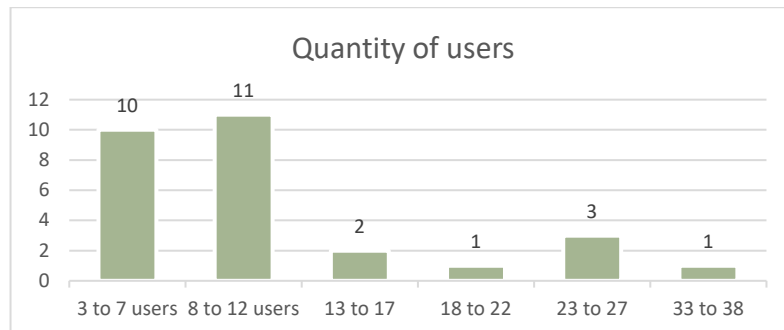


Fig. 9. Quantity of users in the samples

The distribution between the blindness level is varied. There are experiments where the sample is all formed by people who are blind [49] and there are samples formed only by people who are blindfolded, as the experiments in [19]. Figure 10 shows the user age range. Figure 11 shows the blindness level distribution between the samples.

Tasks. The tasks explored in the experiments are related to the technology assessed. The tasks to assess the development of O&M skills are based on virtual and real environments [30][34]. Some of them use modeling kits to represent the virtual environment in the real world and to analyze the space awareness of each participant and the cognition improvement [30]. Another example of cognitive task is to read of a text with the guidance of virtual sounds. Some works have used levels of complexity in the tasks to quantify the cognition impact, for example, the work [45] estimates the task performance in 5 levels.

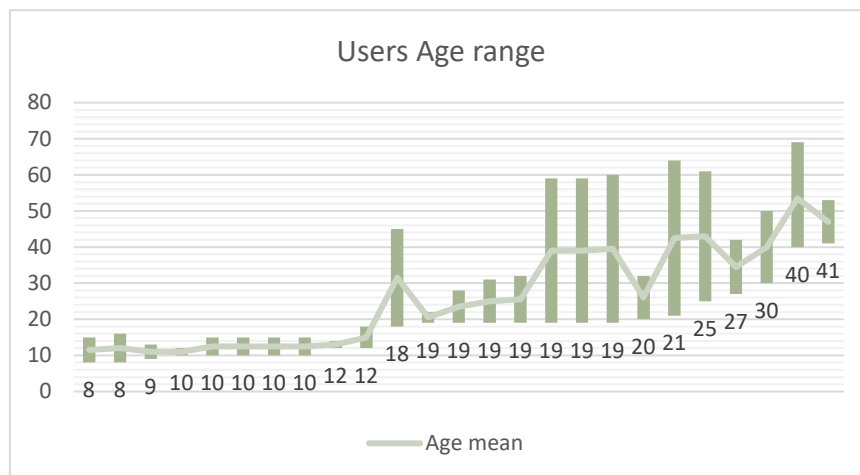


Fig. 10. Users age range per experiment

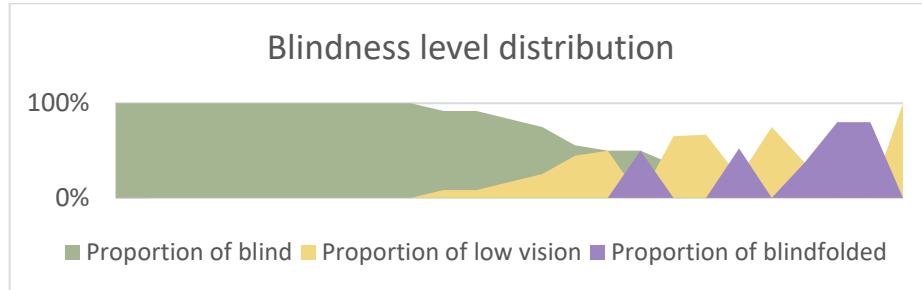


Fig. 11. Blindness level distribution

Variables. The most current independent variables, which are controlled in the experiment, are related to the sample choice, the characteristics defined as etiology of blindness, age, blindness level and gender [41]. The dependent variables, in which we want to see the effect, are related to the measures and the impact. The factor, which is a type of variable controlled is modified in the experiment to see the cognitive effect. For example, the experiment proposed in [8] uses as factor different outputs of the multi-modal interface (audio group, haptic group, and haptic-audio groups).

Measures. The measure to evaluate the cognitive impact focuses on the performance to compare before and after the use of the technology assessed, e.g., [37], or between two groups with and without the technology, e.g. [21]. The measures are strongly related to the instruments used. For example, the checklists of [34] assess the task using scores for sensory perception, tempo-spatial development, and O&M skills. In some works, each measure has a property scale or options, as the Likert scale used in [41]. Eleven experiments ([42][8][34][37][39][11][40][38][30][48]) apply instruments as pretest and posttest to measure the impact.

5 Discussion

As shown in the previous section there is consistent research that evaluates the cognition impact since 2004, showing the importance of this type of study on technologies for blind people. Among all papers retrieved from initial search in the systematic review, we could see in general, that the papers do not evaluate the cognition impact. The preference of evaluation lies to the system, be it hardware or software. The same result is seen in papers that are retrieved in the third filter, one of the most used strategies to evaluate the technologies is the system evaluation besides cognitive evaluation. Despite this finding, the leading role of these technologies is often to help a user's cognitive activity. Thus, assessing if the cognitive purpose was achieved is an important part of constructing software for blind people.

The research data give us an understanding of how the interaction works on multi-modal interfaces of the technologies for blind people. The interaction, cognition skills, and interface characteristics encountered model the planning of the experiment. The combination of a keyboard as the mode of interaction, the sonorous feedback and iconic sound to the interface stands out for the interactions encountered.

In the empirical data, we found a huge diversity of experiments. Although we understand these differences between the experiments, we come across differences in the process of planning and presenting the experiment and data acquired. Instruments, tasks, variables, and measures are strongly related to the technology assessed and the cognitive purpose. Almost no work presented the variables explicitly according to the classification of dependent variables, independent variables, and factors. To understand the experiments, we define these three kinds of variables in each experiment. However, this adds a cost to fully understand the experiment.

From the number of users to the onset age of blindness, there are many sample combinations in the selected experiments. One point that stands out in the selection of the sample is the number of blind users who perform the experiments and how these data are computed. There is, for example, one work that uses the only blindfolded sample, which refocus the conclusion of the experiment.

Statistical power is an inherent part of empirical studies that ensure the results found and to the study conclusions [50]. However, 11 experiments analyzed only the percentage of the gain or the average. Wohlin et al. [20] mention that one of the advantages of an experiment is the ability to perform statistical analysis using hypothesis testing methods and opportunities for replication

Almost no experiment dealt with the resources and ethical concepts. Although there is information about the resources in some papers, in general, this information is not clear in the text and is often incomplete. The missing information makes difficult to repeat the experiment. The ethical concepts also are not well covered in the papers; even it is an essential step to produce an experiment with people who have disabilities [20].

6 Conclusion

The goal of a state-of-the-art is to review the existing evidence concerning the impact evaluation of multimodal interfaces and also seeks to summarize the empirical evidence concerning the strengths and limitations of a specific evaluation method [22]. With this work, we expected to have created a bibliographic review on the cognitive impact evaluation based on the steps of the systematic review approach. Our scope was bounded by multimodal interfaces for people who are blind (in this paper a synonym of visually impaired).

The technologies for people who are blind have many needs due to the target audience and special characteristics with multimodal interfaces, moreover, lots of applications for blind people aims to improve cognitive skills, such as enhancement in O&M, wayfinding, and navigation skills, and thus supporting the user in daily lives. As so, it is important to point out that the use of evidence-based is essential to measure the real impact of the technologies.

After compiling the data from the systematic review and analyzing theoretical foundations, we conclude that there is a need to better plan and present data from experiments on technologies for blind people. With this, we guarantee the quality of the experiment itself and the interaction of the technology with respect to the cognitive

objective. Faced with this nego, we propose as future work to better explore the preliminary results found to improve the data analysis using Grounded Theory and create a set of guidelines that appropriately guide experiments to evaluate tools for blind people. These guidelines will provide a way to evaluate the cognitive impact to development and enhancement in blind people or visually impaired, considering the main aspects of multimodal interfaces.

Acknowledgments. This research is funded by Chilean FONDECYT #1150898; Basal Funds for Centers of Excellence, Project FB0003 - CONICYT. We would like to thank FUNCAP for sponsoring a master scholarship (MDCC/DC/UFC) for the author Lana Mesquita.

References

1. Bamac, B., Aydin, M., Ozbek, A.: Physical fitness levels of blind and visually impaired goalball team players. *Isokinet. Exerc. Sci.* 12, 247–252 (2004).
2. Sánchez, J., Darin, T., Andrade, R., Viana, W., Gensel, J.: Multimodal Interfaces for Improving the Intellect of the Blind. *XX Congr. Informática Educ. - TISE*. Vol. 1, 404–413 (2015).
3. Darin, T., Andrade, R., Sánchez, J.: Dimensions to Analyze the Design of Multimodal Videogames for the Cognition of People Who Are Blind Dimensions to Analyze the Design of Multimodal Videogames for the Cognition of People Who Are Blind. (2015).
4. Hinebaugh, J.P.: A board game education. Rowman & Littlefield Education (2009).
5. Sánchez, J., Baloian, N., Hassler, T.: Blind to Sighted Children Interaction Through Collaborative Environments. Presented at the (2004).
6. Buaud, A., Svensson, H., Archambault, D., Burger, D.: Multimedia Games for Visually Impaired Children. Presented at the (2002).
7. Grammenos, D., Savidis, A., Georgalis, Y., Stephanidis, C.: Access Invaders: Developing a Universally Accessible Action Game. Presented at the (2006).
8. Sánchez, J., De Borba Campos, M., Espinoza, M., Merabet, L.B.: Audio haptic videogaming for developing wayfinding skills in learners who are blind. Presented at the (2014).
9. Gorla, M., Cappaglia, G., Tonellia, A., Baud-Bovyb, G., Finocchietti, S.: Devices for visually impaired people: High technological devices with low user acceptance and no adaptability for children. *Neurosci. Biobehav. Rev.* 69, 79–88 (2016).
10. Alicyn Ferrell, K. a.: Evidence-Based Practices for Students With Visual Disabilities. *Commun. Disord. Q.* 28, 42–48 (2006).
11. Sánchez, J., Galáz, I.: AudioStoryTeller: Enforcing Blind Children Reading Skills. In: *Universal Access in Human-Computer Interaction. Applications and Services*. pp. 786–795. Springer Berlin Heidelberg, Berlin, Heidelberg (2007).
12. Weber, A.: Credentialing in assistive technology. *Technol. Disabil.* 9, 59–63 (1998).
13. Dalton, E.M.: Assistive Technology Standards and Evidence-Based Practice: Early Practice and Current Needs. Presented at the June 10 (2015).
14. Dumas, B., Lalanne, D., Oviatt, S.: Multimodal Interfaces : A Survey of Principles ,

- Models and Frameworks. *Lect. Notes Comput. Sci.* 5440, 1–25 (2009).
15. The prevention of blindness. Report of a WHO Study Group. *World Health Organ. Tech. Rep. Ser.* 518, 1–18 (1973).
16. Ossmann, R., Miesenberger, K.: Guidelines for the Development of Accessible Computer Games. Presented at the (2006).
17. Ritter, F.E., Baxter, G.D., Churchill, E.F.: *Foundations for Designing User-Centered Systems.* (2014).
18. Norman, D.A., Donald, A., Carroll, M.: Cognitive artifacts. *Des. Interact. Psychol. Human-Computer Interface.* Cambridge, 17–38 (1991).
19. Pissaloux, E., Velázquez, R.: Model of Cognitive Mobility for Visually Impaired and its Experimental Validation. In: *Mobility of Visually Impaired People.* pp. 311–352. Springer International Publishing, Cham (2017).
20. Wohlin, C., Runeson, P., Höst, M., Ohlsson, M.C., Regnell, B., Wesslén, A., Anders, W.: *Experimentation in Software Engineering.* Kluwer Academic Publishers, Norwell, MA, USA (2000).
21. Connors, E., Chrastil, E., Sánchez, J., Merabet, L.: Virtual environments for the transfer of navigation skills in the blind: a comparison of directed instruction vs. video game based learning approaches. *Front. Hum. Neurosci.* 8, (2014).
22. Kitchenham, B., Charters, S.: Guidelines for performing Systematic Literature reviews in Software Engineering Version 2.3. *Engineering.* 45, 1051 (2007).
23. StArt, http://lapes.dc.ufscar.br/tools/start_tool, (2016).
24. Lumbreras, M., Sánchez, J.: 3D aural interactive hyperstories for blind children. *Virtual Real.* 119–128 (1998).
25. Harzing, A.-W., Alakangas, S.: Google Scholar, Scopus and the Web of Science: a longitudinal and cross-disciplinary comparison. *Scientometrics.* 106, 787–804 (2016).
26. Petersen, K., Feldt, R., Mujtaba, S., Mattsson, M.: Systematic Mapping Studies in Software Engineering. 12th Int. Conf. Eval. Assess. Softw. 1–10 (2008).
27. Ampatzoglou, A., Stamelos, I.: Software engineering research for computer games: A systematic review. *Inf. Softw. Technol.* 52, 888–901 (2010).
28. Santos, I. de S., Andrade, R.M. de C., Rocha, L.S., Matalonga, S., de Oliveira, K.M., Travassos, G.H.: Test case design for context-aware applications: Are we there yet? *Inf. Softw. Technol.* 88, 1–16 (2017).
29. Adebisi, A., Sorrentino, P., Bohlool, S., Zhang, C., Arditti, M., Goodrich, G., Weiland, J.D.: Assessment of feedback modalities for wearable visual aids in blind mobility. 1–17 (2017).
30. Sánchez, J., Maureira, E.: Subway Mobility Assistance Tools for Blind Users. In: *Universal Access in Ambient Intelligence Environments.* pp. 386–404. Springer Berlin Heidelberg, Berlin, Heidelberg (2007).
31. Connors, E.C., Yazzolino, L.A., Sánchez, J., Merabet, L.B.: Development of an Audio-based Virtual Gaming Environment to Assist with Navigation Skills in the Blind. 1–7 (2013).
32. Sánchez, J., Sáenz, M.: Enhancing Navigation Skills through Audio Gaming. (2014).
33. Shafiq, M., Choi, J.-G., Iqbal, M., Faheem, M., Ahmad, M., Ashraf, I., Irshad, A.: Skill Specific Spoken Dialogues Based Personalized ATM Design to Maximize Effective Interaction for Visually Impaired Persona. Presented at the (2014).

34. Sánchez, J., Espinoza, M.M., Sanchez, J., Espinoza, M.M.: Audio Haptic Videogaming for Navigation Skills in Learners Who are Blind. In: ASSETS 11: PROCEEDINGS OF THE 13TH INTERNATIONAL ACM SIGACCESS CONFERENCE ON COMPUTERS AND ACCESSIBILITY. pp. 227–228. ASSOC COMPUTING MACHINERY, 1515 BROADWAY, NEW YORK, NY 10036-9998 USA (2011).
35. Sánchez, J., Aguayo, F.: AudioGene: Mobile Learning Genetics through Audio by Blind Learners. In: Learning to Live in the Knowledge Society. pp. 79–86. Springer US, Boston, MA (2008).
36. Sánchez, J., Elías, M.: Science Learning in Blind Children through Audio-Based Games. , London (2008).
37. Sánchez, J., Flores, H., Sánchez, J.: AudioMath: Blind children learning mathematics through audio. 4, 311–316 (2005).
38. Sánchez, J., Flores, H., Sáenz, M.: Mobile science learning for the blind. Presented at the (2008).
39. Sánchez, J., Mascaró, J.: Audiopolis, Navigation through a Virtual City Using Audio and Haptic Interfaces for People Who Are Blind. Presented at the (2011).
40. Sánchez, J., Espinoza, M., De, B., Merabet, L.: Enhancing Orientation and Mobility Skills in Learners who are Blind through Video gaming. Conf. Creat. Cogn. Conf. 2013, 353–356 (2013).
41. Sánchez, J., Oyarzún, C.: Mobile audio assistance in bus transportation for the blind, (2008).
42. Lahav, O., Schloerb, D., Kumar, S., Srinivasan, M.: A Virtual Environment for People Who Are Blind – A Usability Study. J. Assist. Technol. 6, (2012).
43. Kalia, A., Legge, G., Roy, R., Ogale, A.: Assessment of Indoor Route-finding Technology for People with Visual Impairment. J. Vis. Impair. Blind. 104, 135–147 (2010).
44. Buchs, G., Simon, N., Maidenbaum, S., Amedi, A.: Waist-up protection for blind individuals using the EyeCane as a primary and secondary mobility aid. 35, 225–235 (2017).
45. Villane, J., Sánchez, J.: 3D Virtual Environments for the Rehabilitation of the Blind. Presented at the (2009).
46. Guerreiro, T., Oliveira, J., Benedito, J., Nicolau, H., Jorge, J., Gonçalves, D.: Blind People and Mobile Keypads: Accounting for Individual Differences. Presented at the (2011).
47. Hossain, G., Shaik, A.S., Yeasin, M.: Cognitive load and usability analysis of R-MAP for the people who are blind or visual impaired. Presented at the (2011).
48. Espinoza, M., Sánchez, J., de Borba Campos, M.: Videogaming Interaction for Mental Model Construction in Learners Who Are Blind. Springer Verlag (2014).
49. Lahav, O., Mioduser, D., Lahav, O.: Blind persons' acquisition of spatial cognitive mapping and orientation skills supported by virtual environment. Int. J. Disabil. Hum. Dev. 4, 231–238 (2005).
50. Dybå, T., Kampenes, V.B., Sjøberg, D.I.K.: A systematic review of statistical power in software engineering experiments. Inf. Softw. Technol. 48, 745–755 (2006).