Designing for Designers: Towards the Development of Accessible ICT Products and Services using the VERITAS Framework

Abstract

Among the key design practices which contribute to the development of inclusive ICT products and services is user testing with people with disabilities. Traditionally, this involves partial or minimal user testing through the use of standard heuristics, employing external assisting devices, and the direct feedback of impaired users. However, efficiency could be improved if designers could readily analyse the needs of their target audience as part of a highly iterative design process. The VERITAS framework simulates and systematically analyses how users with various impairments interact with the use of ICT products and services. Thus, facilitating an efficient approach to design and testing. This article reports qualitative insights into the use of the framework by 72 evaluators drawn from five application domains: infotainment-games; workplace design; smart living spaces; healthcare; and automotives. The findings show that the VERITAS framework is useful to designers, offering an intuitive approach to inclusive design. However, several key areas present challenges to designers; notably, their lack of technical knowledge made the interface difficult to comprehend and their lack of familiarity with virtual user modeling or simulation software made the workflow difficult to follow. Furthermore, designers had a number of expectations in terms of features and feedback which were not fulfilled. This article reviews these concerns and presents recommendations which will inform the design of future inclusive design tools.

Keywords: Inclusive Design, Universal Access, Accessibility Requirements, Virtual User Modelling, Tools, Software, Simulation, Impaired Users.

1. Introduction

Designing software and products with accessibility requirements in mind entails that designers should include these from the outset—or at least, in the early stages—of the design process. This is especially important given the significant proportion of people who have some form of cognitive, motor, or sensory impairment, whether these be on a temporary or permanent basis Barbotte et al. (2001); Union (2004); Yuan et al. (2011). Unfortunately, accessible design is complicated in practice by a host of factors, such as: the difficulty of gathering requirements and feedback from this segment of the population; the difficulty in simultaneously designing for more than one type of impairment; and, indeed, the difficulties that designers encounter when trying to understand the issues raised as a result of an inclusive design process (c.f. Keates et al. (2000); Choi et al. (2006); Law et al. (2006); Stephanidis and Akoumianakis (2001)).

The Virtual and Augmented Environments and Realistic User Interactions to Achieve Embedded Accessibility Designs Project[1] (VERITAS) provides designers with a framework of design and simulation tools which will help them overcome such challenges. Accordingly, the VERITAS framework provides designers with the capability of choosing from a wide range of disabilities defined within the VERITAS repository and generating a virtual user model (VUM) based on the particular disabilities a potential user might have Navarro et al. (2012). Using this VUM, the designer is then shown a simulated experience of how that particular user will perform a given task—of the designer's choice—with the latest iteration of the graphical user interface they are designing.

Previous research has shown that the design of the VERITAS framework is adequate in terms of acceptance and usability Spyridonis et al. (2014); Scott et al. (2015). However, with the completion of the second iteration of the framework, greater insight is being sought through a more detailed analysis driven by evaluators from a broad range of design backgrounds, including: infotainment-games; workplace design; smart living spaces; healthcare; and automotives. Hence, the following research question: *what are the key challenges that designers encounter while using the VERITAS framework to improve the accessibility of their ICT products and services?*

As part of a broader collaborative effort to develop and evaluate the

---

[1]Please refer to `http://veritas-project.eu` for more information.

VERITAS framework, this articles presents an analysis conducted by the authors on three specific tools: the Virtual User Model Generator (VerGen), the GUI Simulation Editor (VerSEd-GUI); and the GUI Simulation Viewer (VerSim-GUI). Section II provides additional background on the motivation for VERITAS and its related work. Section III then describes the framework and each tool in more detail. Section IV reviews the methodology and Section V describes the key findings. The article then closes with a discussion in Section VI, illustrating several implications for the design of tools that aim to address accessibility requirements and several recommendations on how to adapt tools that are intended for use by designers.

2. Background

Accessibility requirements are becoming increasingly broad and important Hull (2004). Typically, these are underpinned by legal drivers such as Section 508 of the 1998 Rehabilitation Act in the USA and 1998 Disabilities Act in the UK. For example, ensuring that the workplace is appropriately designed to accommodate the needs of all employees Eriksson et al. (1995). However, this is not always the case.

Many requirements focus on universal access, enabling those with permanent— and, often severe—impairments to engage with latest technological innovations. In healthcare many tools must be designed to address a spectrum of conditions during rehabilitation. For example, interactive therapies that involve restoring partial paralysis Leder et al. (2001) and balance Lange et al. (2011) must accommodate new patients (i.e., those with the most severe forms of a condition), as well as recovering patients. Thus, enabling all patients to benefit from these innovative therapies. Extending this notion, many serious games (see Janarthanan (2012) for a definition and review) also enhance learning (e.g. Papastergiou (2009)) or otherwise provide opportunities for enrichment (e.g., Scott and Ghinea (2013a)) and so it is considered inappropriate to unduly exclude individuals with impairments from receiving the benefits of such innovations Smith (2011).

Commercial considerations are also important because those with impairments form a considerable population of potential customers Yuan et al. (2011); Smith (2011). As an example, several games have demonstrated that those with impairments are interested in play (e.g. Westin (2004)) and there are many examples where traditional games have been adapted to provide direct access to them (e.g. Scott and Ghinea (2013b)) or to support new

technologies that provide indirect access Westin et al. (2011). It is also increasingly the case that those with impairments be involved in the design of products and services that are targeted at them. For example, using virtual environments such as HabiTest to enable those with impairments to evaluate new living spaces that aim to improve their quality of life Palmon et al. (2004).

Another important consideration is that many impairments are situational. That is, any individual could experience impaired ability to interact with a computer system as a result of their circumstances. As an example, driving an automotive vehicle. This activity creates constraints that affect motor and cognitive skills which will need to be considered when designing user interfaces that can be used safely in the car Schmidt et al. (2010).

In order to incorporate these different accessibility requirements into the design of ICT products and services, a range of inclusive design practices are often used by practitioners. These can include: *developing personas*, creating fictional characters to understand and empathise with a particular audience Cooper (1999); Picking et al. (2010); *standards review*, using a set of guidelines to ensure that constraints are accommodated within a design (e.g. Chisholm et al. (2001)); *automated checking*, using tools which evaluate designs against set guidelines automatically Abascal et al. (2004); Kasday (2000); and *user testing*, typically involving both, experts conducting a heuristic analysis to identify problems in a design Nielsen and Molich (1990), and, potential end-users providing general feedback on the use of a prototype Rubin and Chisnell (2008).

None of these approaches are mutually exclusive, but each has a number of weaknesses. Personas are not always believable, they are sometimes designed arbitrarily rather than using real-world data, they may not be communicated well, designers may not understand how to accommodate their requirements, and resources are required for their development Pruitt and Grudin (2003). Standards may be too restrictive, overly-complex, may not accommodate all impairments, and sometimes require interpretation by a designer which leads to errors in conformity Choi et al. (2006); Law et al. (2006); Milne et al. (2005). Automated checking tools may not be in sync with the latest guidelines and may be limited to assessing a particular type of product or component. User testing can be costly and time-consuming. Additionally, it is seldom used in the early stages of the design process so many good designs may be discarded before experts or potential users can comment on them.

These challenges form practical limitations that inhibit the design of ac-

cessible ICT products and services. Of particular note is that these approaches do not support systematic analysis and testing of accessibility requirements in the early stages of a project. Thus, a core contribution of the VERITAS framework is the level of systematic analysis and testing it offers. By using virtual users as testers, designers are given the opportunity to observe, quickly and first hand, the impact of their designs during the simulations. Thus, facilitating testing much earlier in the design process and with an efficiency that encourages fast iterative development. In some circumstances, this can even become an immersive experience that designers can use to guide their design thinking. For example, in the case of a virtual user with a motor impairment, the designer could attempt to complete a task themselves while the erratic cursor movement of the virtual user is simulated. Accordingly, providing insight and guidance to improve the accessibility of a design.
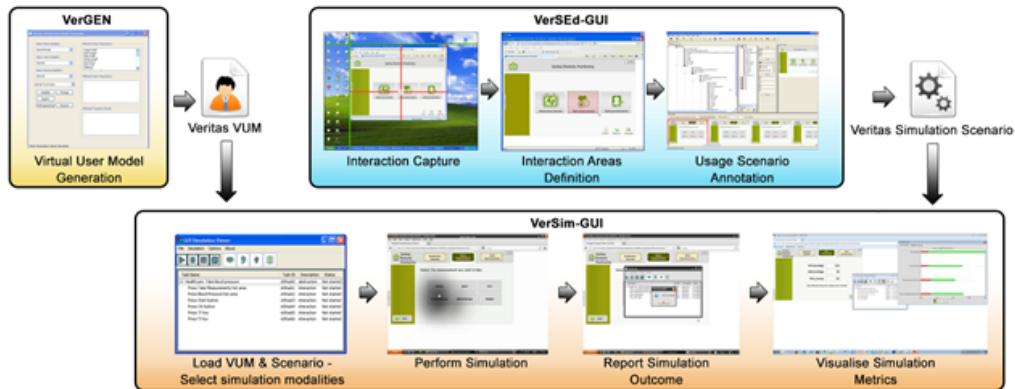

Figure 1: Workflow of the VERITAS Framework for Evaluating Accessibility Requirements in GUI-Based ICT Products and Services.

3. The Veritas Framework

The goal of the VERITAS framework is to provide support to designers as they evaluate the accessibility of their designs. To this end, the framework generates a report which highlights key problems and presents relevant use-statistics based on a simulation. To create the report, designers use three core tools within the VERITAS framework: *VerGen*, where designers specify

the nature of the impairments they wish to simulate in terms of a virtual user; *VerSEd-GUI*, where designers define and configure a series of actions to test on the user interface of their product or service; and *VerSim-GUI*, where the various impairments are simulated to reproduce the experience of an impaired user. It follows, then, that the VERITAS accessibility assessment is passed through three phases: (i) virtual user modelling; (ii) simulation scenario definition; and (iii) simulation of the virtual user actions.

The overall workflow for the VERITAS framework and the links between the tools within the toolset is illustrated on the previous page in Figure 1. The workflow consists of a sequence of tasks which must be conducted across the three tools in order to generate the relevant files for the simulation and then to run the simulation itself. These tasks are described below in Table 1:

Table 1: Tasks Involved in the VERITAS Workflow

ID	Task	Tool
1.1	Initialize the VUM	VerGen
1.2	Select an Appropriate Population Distribution	
1.3	Adjust the Disability Parameters	
1.4	Generate the VUM	
1.5	Export the VUM	
2.1	Select the GUI Design to be Captured	VerSEd-GUI
2.2	Capture the Sample Product using the GUI Design	
2.3	Set the Hot Areas	
2.4	Set the Before and After Images	
2.5	Set Flags and Export the Simulation Scenario	
3.1	Import the Virtual User Model	VerSim-GUI
3.2	Open Simulation Scenario	
3.3	Perform Simulation	

The following walk-through describes each tool in more detail, explaining how each tool fits within the workflow for testing the accessibility of a sample product. Additionally, each key task in the process will be highlighted.

3.1. VerGen

User modelling is based on measurement parameters described in the medical literature and the VERITAS Multisensorial Framework, which was

deliberately created for this purpose Navarro et al. (2012). As such, models are derived from a database of profiles of parameters (many based on the WHO impairement definitions Barbotte et al. (2001)). This results in a very large number of parameters. For example, motor parameters include: weight shift; step length; step width; stride length; gait cycle; cadence; typical velocity; knee flexion; hip flexion; hip extension; and many more. To simplify the setup of user models, profiles are available based on generic specifications for known impairment groups (e.g., people with a cataract, people with parkinsons, people with presbyacusis, etc.). Due to the availability of relevant standards and quantitative data needed to drive the simulations, a medical view of disability is implicitly assumed. However, psychological and behavioural aspects of users have also been included. These are parameterized using the Adaptive Control of Thought-Rational Model Navarro et al. (2012). This permits, for example, parameters such as visual-attention-latency to vary across conditions which may induce emotion, stress or fatigue. As a result of the broad range of parameters and these modifiers, an extremely rich range of disabilities and contexts can be simulated.

The generation of the VUM is handled by the VerGen, with which the designer selects what impairments the virtual test user will have. VerGen can be used to define the severity of the impairment or even combine two or three impairments into a single, more sophisticated model. The tool exports a VUM, which contains the specification of an indicative virtual user selected from a population percentile, with one or many impairments of that severity.

The first stage is to initialise the user model by selecting a particular profile of impairments (Task 1.1) and the population distribution (Task 1.2), for example the Parkinsons motor impairment shown in Figure 2.

The model can be manipulated to create increasingly complex and sophisticated models. So, the next stage involves the modification of individual parameters (Task 1.3) as shown in Figure 3.

Each parameter is associated with a population density function, showing the prevalence of the condition within the target population. During the simulation, different values from the model can then be selected based on a probabilistically representative set of virtual users rather than a single unrepresentative extreme case. The severity of a particular parameter that a designer is interested in can be modified to increase its prevalence in the simulation or to, perhaps, simulate multiple impairments of the same type simultaneously (i.e., combining two forms of visual impairment such as cataracts and astigmatism).
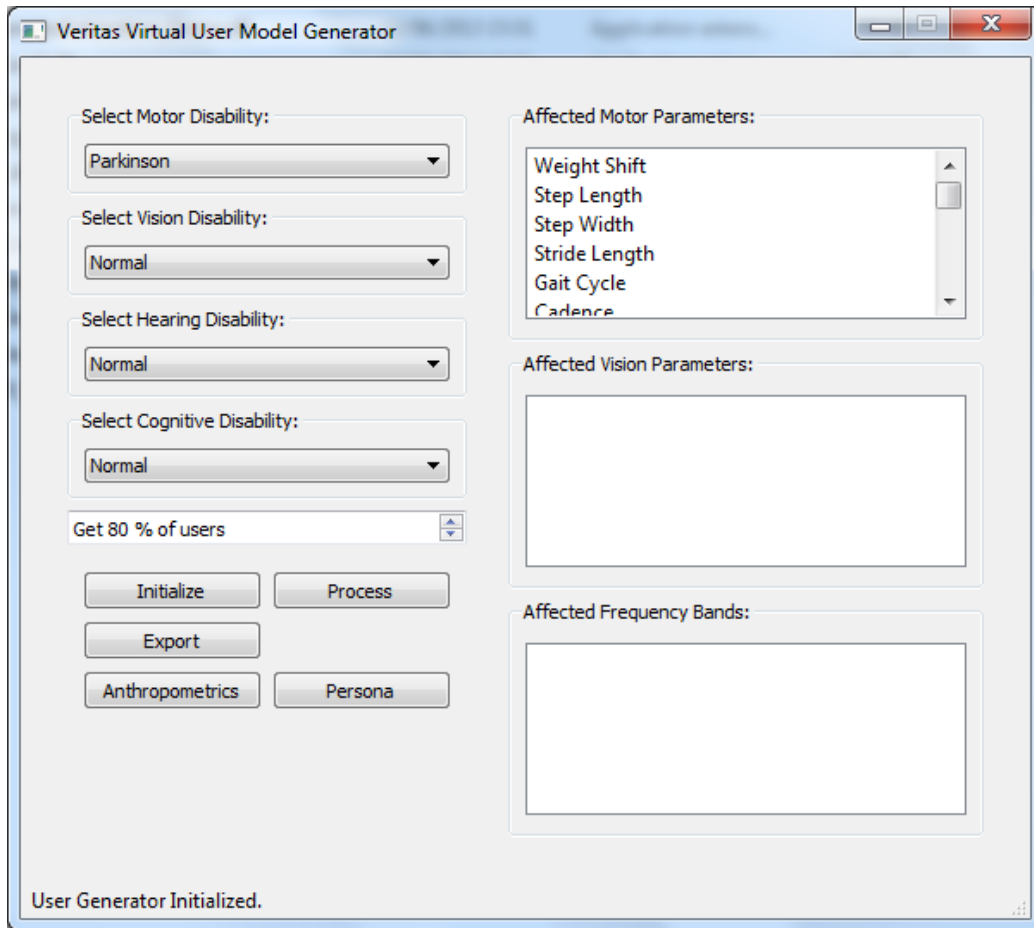
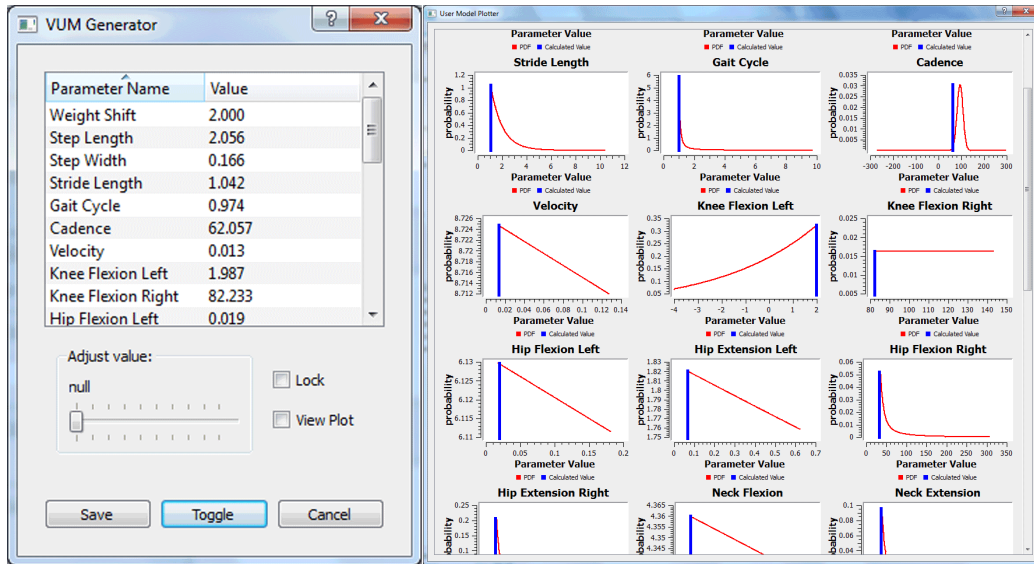Figure 2: Virtual User Model Generator.

Figure 3: Setup and Modification of Individual Parameters.

The VUM can then be generated by selecting the appropriate button once the parameters have been configured (Task 1.4) and exported while taking into account any further, more general modifications that a designer may want to make with respect to the persona of the virtual user (Task 1.5). For example, the designer may want to further restrict the model to being either male or female.

3.2. VerSEd-GUI

Once the user model has been set up, the product must be initialised for simulation and testing. This takes the form of a simulation scenario. This simulation scenario is defined using VerSEd-GUI, which produces a file listing the expected actions and contexts that the user is expected to conduct during the simulation.

First, a GUI design is selected so the simulation recorder knows what to expect from the sample product (Task 2.1). Then, the sample product is launched and a window is defined in order to capture interactions with it (e.g., mouse clicks in particular areas of the interface) (Task 2.2). Each interaction, alongside success and failure criteria, must be defined within VerSEd-GUI. This involves defining: the hot areas where interactions are

expected to occur (Task 2.3); the expected order in which events (and subsequent screen transitions) are expected to occur (Task 2.4); and the flags associated with the event (Task 2.5) (e.g. optional, critical task, etc.). This is done through annotating events (each with an associated screenshot) that was collected during the previous capture task using the editor window. This is shown below in Figure 4. A simulation scenario file is finally exported.
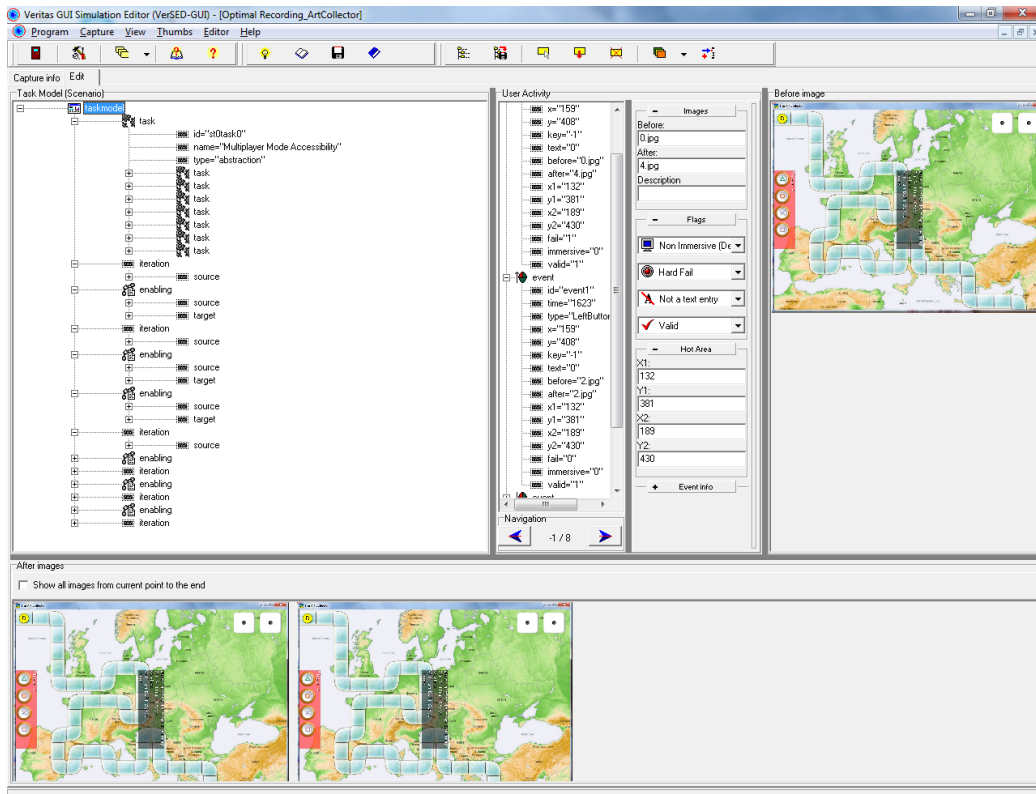

Figure 4: Setup and Modification of the Simulations.

3.3. VerSim-GUI

The simulations themselves, and subsequent analysis of the data, are conducted within VerSim-GUI. This section of the framework is designed to facilitate the evaluation of the accessibility of a GUI. In order to start a simulation, the VUM (Task 3.1) and the simulation scenario (Task 3.2) are

loaded into the simulation. Once this is done, the simulation will begin as shown below in Figure 5.
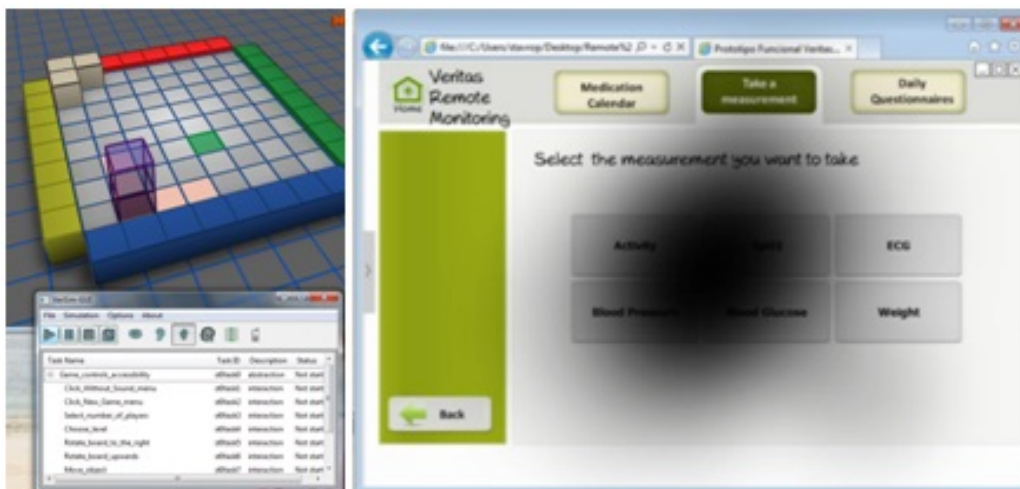

Figure 5: Examples of a Simulated Game and a Simulation of a Visual Impairment within a Sample Healthcare Product.

In this phase, motor, vision, hearing and/or cognitive simulation will reproduce interactions while taking the defined impairments into account. This provides designers with both an experiential analysis (by simulating the various effects of impairments, such as sound distortion or jittery mouse movement) allowing the designer to spot any flaws in the design. The designer may then complete the tasks defined for the scenario (Task 3.3) and, once the simulation is over, the performance is recorded as an XML file. The designer then has the opportunity to review metrics associated with this XML file, which will inform them about the success or failure of the design, while also helping them to evaluate particular design decisions. Many metrics can be drawn from the data, with examples including the number of actions required to complete the task, the time required to complete the task, as well as average distance between an action and its associated hot area. Visualisations, including charts and graphs, are also available. Figure 6 illustrates one such example where mouse presses and drag-actions of a virtual user are compared against the expected paths and hot areas defined in the simulation scenario:
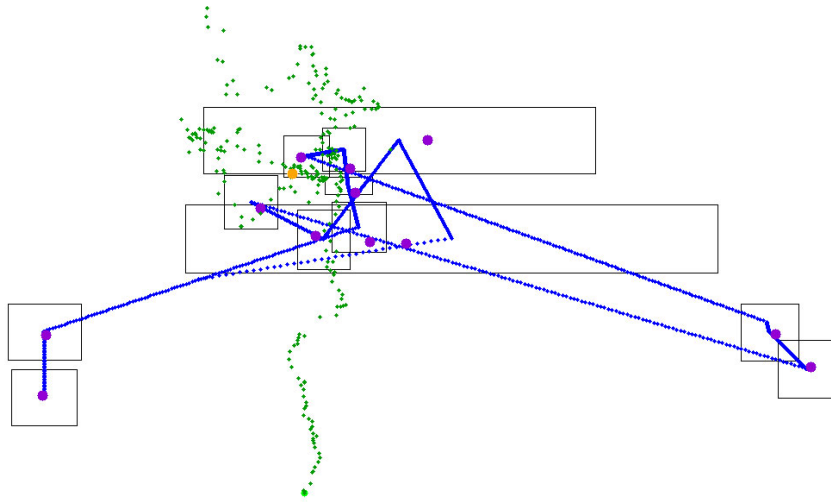
Figure 6: Examples of a Visualisation Created from the Output XML File from a Simulation of the Primix Game Showing Expected and Simulated Action Coordinates.
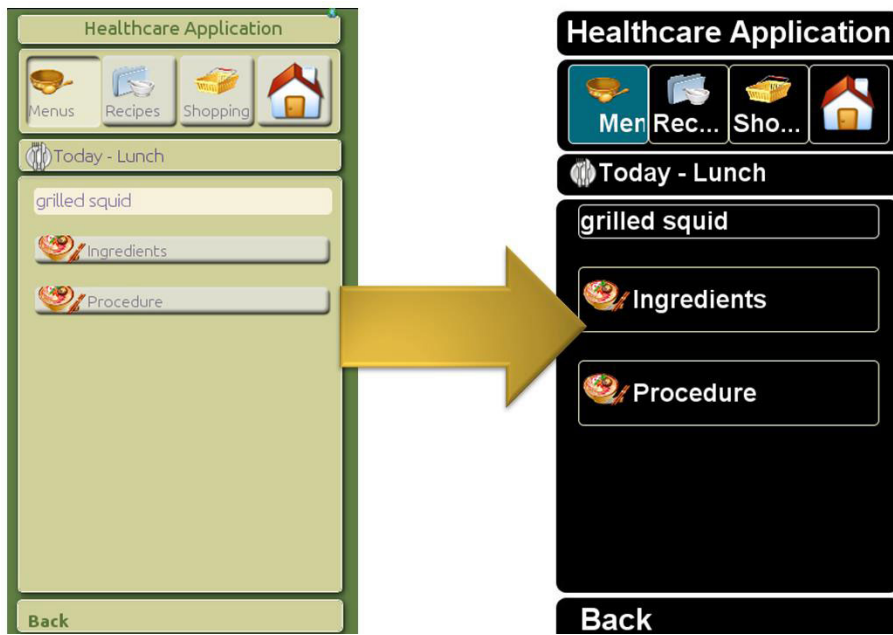

Figure 7: A Healthcare Mobile App Developed at CERTH/HIT Showing How the Initial Version (Left) Was Improved Using VERITAS (Right)

In practical terms, these data and visualisations offers designers the opportunity to assess and improve their designs. Figure 7 shows how the application of the results from a report prompted designers to change their design which, during pilot testing, demonstrated increased usability for people with disabilities in three test scenarios using real users[2].

4. Methodology

Each tool in the VERITAS framework was assessed using the expert team method. This method uses a group of individuals who are members of the software development communities with a high level of task related expertise. Each team produced empirical evidence on the usability of the tools in the VERITAS framework after a hands-on experience. In order to review the acceptance and usability of the tool in a way that would provide greater depth compared to the previous results Spyridonis et al. (2014), a qualitative approach based on thematic analysis, which provides rich insights into users concerns, was adopted Braun and Clarke (2006). As such, participants evaluated each tool on a product drawn from their respective application area and provided comments by means of an open-ended questionnaire. The thematic analysis was then conducted on the comments provided by the evaluators.

All sites received ethical approval from their own regional ethics committee in addition to the EU VERITAS ethics committee.

4.1. Participation & Recruitment

The participants were representative users for the VERITAS tools. All were professional designers working in their relevant fields and all were involved in work where the value of applying accessibility tools could be identified. The requirements for a designer to take part in the evaluation included a professional orientation with ICT software and previous experience in GUI development.

The participants formed a convenience sample, whom were recruited through: recruiting employees who were working in unrelated projects, and were therefore unaware of the specifics of VERITAS project; contracting those enrolled on databases or were involved in previous unrelated research efforts; through adverts and websites; through recommendations by colleagues

---

[2]Figure 7 shows an example drawn from page 93 of the Pilot Results Report (D 3.8.2) authored by S. Edwards. This report is not currently available to the public.

from other establishments; and, more generally, word-of-mouth recruitment. In most cases, participants were reimbursed financially for their time. The characteristics of the participants are presented below in Table 2:

Table 2: Participant Demographics

Application Area	Count	Mean Age	Gender Ratio
Automotive	15	31.6	13:2
Healthcare	16	34.1	7:1
Infotainment-Games	10	34.8	4:1
Smart Living Spaces	20	29.5	13:7
Workspace	11	34.3	7:4
TOTAL	**72**	**32.4**	**12:67**

Ages ranged between 23 and 58 years old. It is evident that male participants were over-represented in the whole sample. However, such a distribution is not unusual within the professional population. Although difficult to confirm Hafkin and Huyer (2007), several surveys carried out during the last decade show that women account for a small portion of ICT professionals in many parts of the world Huyer et al. (2005); UNESCO (2007).

4.2. Data Collection

The evaluation of the VERITAS framework was organized across six sites within the European Union (including Germany, Greece, Italy, United Kingdom, France, and Spain), with each site taking responsibility for two application areas; such that each application area was evaluated in more than one site. Each evaluation included two cycles of training and tool assessment; of which, only the findings from the final iteration are reported here. Each site used an identical set of tools and used the same protocol.

In the training phase of the evaluation, participants familiarised themselves with the tools. They were first provided access to a Virtual Learning Environment[3] prior to their arrival. The training curriculum consisted of structured modules for the different tools accompanied by videos and reading materials. Care was taken so that all participants had enough time to

---

[3]Please refer to `http://training-veritas.atosresearch.eu` to examine these materials.

study the material and the tools themselves before their participation. In addition, training workshops were organised at each site to inform participants about the project's objectives, the tools to be tested, and the scenarios to be completed. Only generic scenarios were distributed to users at the workshop.

During the tool assessment phase of the evaluation, designers were provided access to each tool in the VERITAS framework. They were requested to perform tests on a sample product which had been drawn from a real-world problem solving situation. Different products were made available to different designers, depending on the application areas that designers were drawn from. Immediately after the designer had finished using a tool, they were given a questionnaire to complete before moving on to the next tool. A copy of this questionnaire is provided in Appendix A.

4.3. Data Analysis

Data collection process was guided using a set of structured questionnaires containing a number of open-ended questions, each tailored to the specific tool under evaluation. After data collection, the data was structured into a grid, such that rows represented cases and columns represented questions, for the coding process.

The nature of this data, the presence of a priori usability concerns, and relevant heuristics drawn from the literature promoted the adoption of a hybrid approach to coding: inductive coding, based on manual review of the data to create in vivo codes on-the-fly as new terms and issues appeared in the data, and deductive coding, based on standardising references to the user interface elements and workflow processes (i.e., nouns and verbs) in addition to anticipating the possible range of responses to the questionnaire items (i.e., valences, adverbs, adjectives) Fereday and Muir-Cochrane (2008); Crabtree and Miller (1992).

In order to analyse and interpret the qualitative data, the six-phase approach to thematic analysis was adopted, as advocated in Braun and Clarke (2006). This method attempts to synthesise a dataset in order to identify key commonalities. The process consists of six discrete stages: data familiarization; generation of the initial codes; search for themes; review of the themes; defining themes; and producing a thick description.

As the dataset was reasonably small, the data was analysed manually using nVIVO 10 and Microsoft Excel. However, based on the suggestion by Santos et al. (1999), VOSViewer Van Eck and Waltman (2011) was used to create visualisations. The technique applied examined keywords in terms of

frequency of occurrence, representing this as a heat-map with red showing commonly used terms, and 'relatedness', represented as the normalized distance between pairs of terms. This supported the process of establishing themes, focusing on summaries and theme recognition during stages two and three of the analysis, and indicated the prominence of those themes within the dataset.

Unlike in traditional thematic analysis, the sixth stage of 'thick description' was forgone in favour of a breakdown of the key themes identified. This was done to examine prevalence of each concern within the cohort of evaluators. In this procedure, textual analysis (as a form of content analysis) was used to count the positive and negative codes which arose in the coded qualitative data.

5. Findings

In general, 63% of the sentiments expressed in the questionnaire responses were positive in nature, with the remaining 37% being negative or having no clear valence (i.e., constructive feedback and suggestions). In addition, four key themes emerged across the tools that were evaluated as part of this study: comprehensibility for designers; simulation workflow; requirements and expected features; and system feedback. Figures 8—10 on the following pages illustrate how these themes emerged through the use of thematic maps. These maps show how the feedback provided on each tool were coded and how these codes were combined to form more general themes. A summary of the themes and their associated codes are also shown in Table 3. These four themes highlight usability concerns that tool developers need to consider when creating tools for designers. These are discussed in more detail in the sections below.

5.1. Comprehensibility for Designers

The comprehensibility theme that emerged represents the sentiment that designers felt they did not understand many elements of a particular tool. Across the three tools, this theme emerged most prominently in the form of three sub-themes: (i) interpreting the different parameters and numbers presented in the model generators and simulations; (ii) understanding how a disability is represented in the simulations, and specifically how severe the representation is; and (iii) understanding the different terms used throughout the user-interface. Although designers from some application areas positively
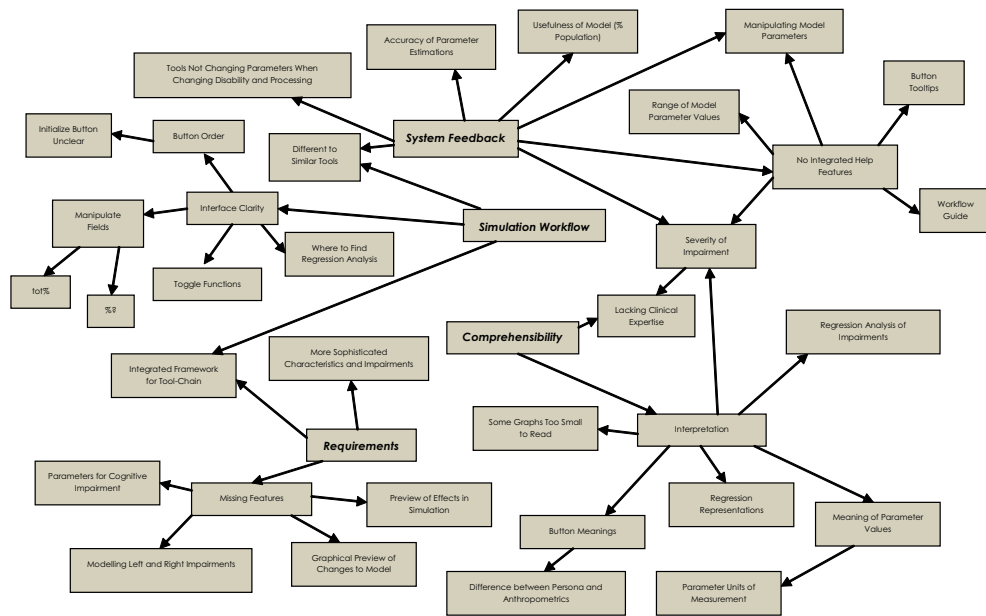
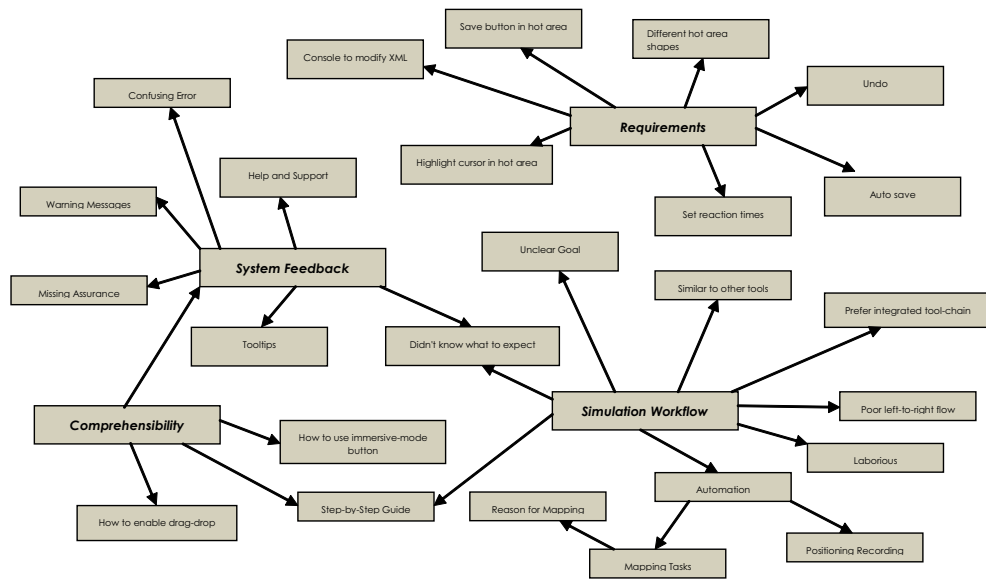Figure 8: Thematic Map for VerGen

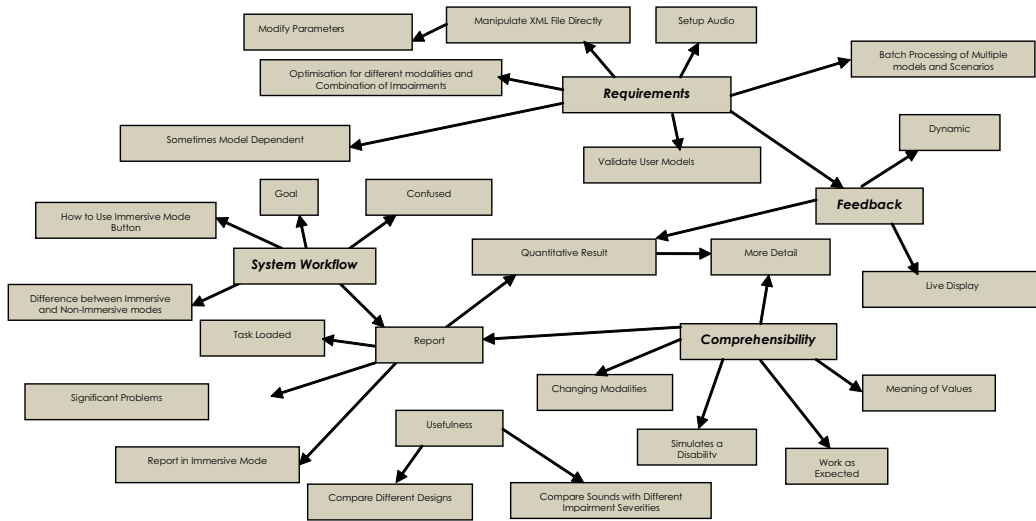Figure 9: Thematic Map for VerSEd-GUI

Figure 10: Thematic Map for VerSim-GUI

endorsed these themes, the sentiments were generally quite negative. Typically, these were accompanied by statements about designers' 'familiarity with simulations' or their level of 'expertise'. For example: "I did not understand the meaning of the parameters"; "As a designer (and not an expert of hearing impairments) I expected to set the severity of the disability, and not some numeric parameters. Besides, I do not know how the parameters affect the simulation, and which severity they correspond to'.'; "The report is not clear (i.e., what does *task succeeded* mean? I had *task succeeded* even though I moved the application in a wrong place and the simulator could not actually complete the task)".

5.2. Simulation Workflow

The workflow theme that emerged represents the sentiment that designers felt towards the sequences required to interact with the tools. Across the three tools, this theme was most prominent with respect to: (i) the clarity of different functions and how to use them; (ii) the amount of work that was needed to use the tools; (iii) the need for automation; and (iv) a need for improved feedback and support, most notably in the form of step-by-step guides or videos. Generally, designers desired much more support using the tools and a more effective way of getting to understand the workflow

18

between the tools. There were significant negative sentiments towards the workflow of VerSEd-GUI, with many designers calling for laborious tasks, such as the image event-mapping (Task 2.4), to be automated as much as possible. Furthermore, there was significant confusion about different modes, such as the difference between the modalities in VerSim-GUI and how to change between them, and small non-intuitive aspects of the user interface, such as having to enable the drag and drop functionality for placing hot areas in VerSEd-GUI. For example: "A much more integrated tool chain, with software module to allow the designer to open only the simulator, [is needed]"; "Found it a little difficult to remember the order of tasks for certain workflows. It would be very straightforward for an experienced user but a novice might require more support"; "I like the idea, but the workflow is not clear, and the tool does not support me in understanding the workflow (no help, no sequence)".

5.3. Requirements and Expected Features

The requirements and expected features theme that emerged represents the sentiment that designers felt towards the ability of the tools to satisfy their needs. In particular, this can be further decomposed into several key sub-themes: (i) the capabilities of the tools; (ii) features that were expected to be in the tools; and (iii) the ability of the tool to support new users. Generally, designers were positive about the capabilities of the tools, however some designers wanted to create more sophisticated user models to reflect different disabilities (i.e., left and right differences, cognitive impairments, reaction times, etc.). Furthermore, many expected a more integrated tool-chain which would allow designers to simulate models as they were creating them. A range of features that designers expected to be in the tools were not present. These ranged from simple utilities that had not been enabled by default in the prototypes such as auto-saving, to more sophisticated editing options such as a console window to manually edit XML files, and facilities (e.g., wizards) to enable the batch processing simulations for different user models and scenarios. The most prominent problems that were identified with respect to supporting new users related to the workflow of the tools and the lack of help features. For example, designers noted that there were no tooltips in the simulation editor: "The installation and set-up of the other tools (jack, etc.) was very time-consuming [...] because [there was] no tutorial [or] integrated menu"; "A detailed report for the hearing simulation should be created".

5.4. System Feedback

The feedback theme that emerged represents the sentiments designers held towards the types of feedback presented by the tool and their understanding of that feedback. Predominantly, comments about feedback were accompanied by codes that formed part of the comprehensibility theme; however, there were four sub-themes that could be considered independent of this: (i) assurances that the tools were being used correctly; (ii) the level guidance provided by error messages; (iii) the types of general feedback that the tools provided; and (iv) the final reporting. Generally, designers were able to successfully guide themselves through the tool, however some designers expressed a desire to have confirmation dialogues and other assurance-based feedback to let them know that models and scenarios they generated were valid and that they were configuring the tools correctly. This was often accompanied by sentiments about the lack of guidance provided by error messages; typically, neither being easy to understand, nor being clear how to resolve. There were a small number of requests for additional types of feedback, which included estimates of the usefulness of a model (i.e., its resemblance to real-world disabilities and estimates of the number of people affected by the disability modelled) and improvements to the live feedback presented during a simulation. There were also a number whom believed the content of the reports generated by VerSim-GUI could be improved. This could be achieved by, for example: incorporating qualitative and criterion-based analyses; producing reports in immersive simulation mode; and compare new results against benchmarks or previous results: "A greater feedback of the results and a better simulation of physical disabilities and immersive simulations"; "I would have appreciated a preview to show how each parameter affects the VUM and the simulation (i.e., an avatar for the motor impairment, the visual filter for the glaucoma, etc.)"; "The feedback and error messages could be improved - they are useless if only those who created the tools understand what they're trying to say".

Table 3: Codes with more than three occurances in the dataset

Tool	Comprehensibility for Designers	Simulation Workflow	Requirements & Features	System Feedback
VerGen	small graphs; button meanings; differences between persona and anthropometrics; regression representations; meaning of parameter values; units of measurement; severity of impairments; lack of clinical expertise.	differences to similar tools; button order; meaning of buttons (initialize / start); toggle functions; manipulating parameters; manipulating fields (% / tot%).	more sophisticated characteristics and impairments; integrated framework for tool-chain; preview of effects in simulation; graphic preview of changes to model; left and right impairments; parameters for cognitive impairment.	accuracy of parameter estimations; usefulness of model (% pop affected); feedback from manipulating model parameters; preview of simulation effects.
VerSEd-GUI	feedback.	differences to similar tools; effort required; left-to-right flow; automation of positioning recording area; automation of mappings; unclear reasoning for task being manual; how to enable features (i.e., drag and drop).	integrated tool-chain; highlighting cursor in hot area; reaction times; auto-save; undo; different hot area shapes; save button in hot area; console to modify xml.	step-by-step guide; confusing error messages; poor warning messages; missing assurances.
VerSim-GUI	immersive-mode button; quality of simulation; meaning of values; goals; audio settings; not enough detail.	how to change modalities; how to simulate a different disability.	optimization for different modalities and combinations of impairments; batch processing of models; model validator; recommendations.	report in immersive mode; dynamic; live display movement; audio setup; more detail needed.

6. Discussion

This evaluation with professional designers and developers has provided a valuable opportunity to discover new dimensions of the VERITAS framework and to learn how to maximize the expected results. In general, the VERITAS concept was well-received across the five application areas, whilst also being seen as an innovative effort to fill-in-the gap in virtual user modelling for various disability groups. Most designers welcomed the insight the tools provided to see for various disability groups. Of the 1620 sentiments coded using the common themes identified in the data, 63% were of a positive or neutral nature. Typically, such comments endorsed: the clarity of the goal of the tools; the potential outcomes that can be achieved with the help of the tools; and the ease of common tasks, such as loading or saving different scenarios. However, a substantial amount of feedback was provided that highlighted opportunities to enhance the toolset. These fell into four key themes: (i) comprehensibility within the tools, largely focusing on how designers interpreted the user models and the reports generated from the simulations, particularly as many participants did not have clinical expertise or previous experience using simulation tools; (ii) the workflow, as many designers were unclear on how to complete tasks and there were many calls to automate laborious tasks such as setting up to the hot areas; (iii) requirements and expected features, as many expected features such as undo and auto save were missing, while different application areas held different concepts of how to use the tool in their context; and (iv) the quality of the feedback, as many error messages could not be understood by the designers.

In addition to these core themes, there was a diverse range of priorities and usability concerns expressed by professional designers across each of the tools. Most notably, different designers had varying levels of clinical expertise and experience with simulations. These represent a further number of potential barriers that should be addressed. As examples, many designers called for step-by-step guides, video documentation, and increased automation. It is, therefore, important that the use of the toolset should be accompanied with appropriate guidance and the different usability concerns should be addressed, where it is feasible to do so.

Although a Virtual Learning Environment was created for designers, as well as structured learning processes in developer workshops, the depth of knowledge was not equally and uniformly checked across all participants. It is, then, possible that this issue was caused by some participants, or even

22

groups of participants from different application domains, having different experiences or reactions to the training. Nevertheless, the complexity of the tools and the necessary training are vital components of future application. Therefore, in future similar studies, training programs and naturalistic methodologies should be taken into consideration. For example, professional designers might need to actually use the tools in their real professional environments for at least a month within an extended user-centered design formative evaluation framework.

Based on the findings and the authors' own experience with the tools, the following recommendations could be incorporated in future virtual user simulation tools aimed at designers:

- Balance the complexity associated with the generation of user models and their subsequent use with the flexibility needed to drive key improvements;

- Refine user model parameters for specific tasks;

- Create different models or variation of models for different application areas;

- Extend application areas;

- Incorporate other disabilities.

As mentioned previously, the thematic analysis also highlights several key areas where usability could be improved and additional features could be incorporated to better meet the knowledge and needs of designers across different fields. Taking the general concerns about feedback from each tool and the usability of user interface aside, several themes that occurred regularly point towards potential improvements to:

- The presentation of the units of measurement for each parameter in the user model, possibly introducing natural language indicators that could be used to help designers without clinical expertise understand what severity a value represents;

- The help available, potentially improving step-by-step guides and video tutorials to assist new users to familiarize themselves with the terminology used throughout the system (such as immersive vs non-immersive,

persona vs anthroprometrics, etc.), how to interpret a model and simulation results, and learn the general workflow of the entire toolset;

- Batch processing and logging features to enable designers to compare multiple user models, designs and scenarios;

- The sophistication of the user models, for example, enabling designers to define left and right differences;

- Automating as many of the processes as possible, specifically aiming to reduce the amount of time required to manipulate images and tasks in the simulation editor;

- and incorporating more general support tools, such as a user database to better facilitate how designers understand disabilities in the contexts of legal frameworks and commercial markets.

These recommendations complement the findings from the quantitative analyses, suggesting that acceptance and usability could be further enhanced and consequently, ease of use increased. This can be addressed through making the graphical interface clearer and more refined in order to increase efficiency and, therefore, productivity.

It is important to acknowledge that further work is required to broaden the evaluation of VERITAS. Notably, this study has focused on the evaluation of visual interfaces used in ICT products and services. However, the VERITAS project overall seeks to address a much broader range of usecases. Some examples that include physical interfaces are: designing car interiors; designing domestic appliances; and even designing collaborative tabletop games[4]. Consequently, a broader range of tools are currently undergoing development and evaluation. Another concern is that an over-emphasis on the medical aspects of disability may undermine its social aspects and how these correspond with system interaction. As such, future studies could incorporate a social view of disability to determine potential impact in this area.

---

[4]Please refer to `http://veritas-project.eu/deliverables/index.html` for further detail and additional use-cases

7. Conclusion

The VERITAS framework received an encouraging evaluation, which will hopefully pave the way for a radical change in how accessibility concerns are incorporated into the design of ICT products and services. However, a number of concerns were raised as a result of the evaluation process: that some aspects of the tool were difficult for designers to comprehend; that the workflow, particularly when defining the simulation, was not as clear as it could be; some expected features were not included, such as a recommendation system for improving user interface designs; and the feedback from the system was not clear enough for non-technical audiences. Consequently, it is recommended that additional features be incorporated to better meet the knowledge and needs of designers. Namely, these targets hiding technical details that designers do not want by, perhaps, automating processes and implementing alternative interactions styles (e.g., natural language options). Alternatively, where this is not feasible, appropriate resources to support these processes, such as guidence from tooltips and wizards, should be provided.

References

Abascal, J., Arrue, M., Fajardo, I., Garay, N., Tomás, J., 2004. The use of guidelines to automatically verify web accessibility. Universal Access in the Information Society 3 (1), 71–79.

Barbotte, E., Guillemin, F., Chau, N., 2001. Prevalence of impairments, disabilities, handicaps and quality of life in the general population: a review of recent literature. Bulletin of the World Health Organization 79 (11), 1047–1055.

Braun, V., Clarke, V., 2006. Using thematic analysis in psychology. Qualitative research in psychology 3 (2), 77–101.

Chisholm, W., Vanderheiden, G., Jacobs, I., Jul. 2001. Web content accessibility guidelines 1.0. interactions 8 (4), 35–54.
URL http://doi.acm.org/10.1145/379537.379550

Choi, Y. S., Yi, J. S., Law, C. M., Jacko, J. A., 2006. Are "universal design resources" designed for designers? In: Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility. Assets '06. ACM, New York, NY, USA, pp. 87–94.
URL http://doi.acm.org/10.1145/1168987.1169003

Cooper, A., 1999. The Inmates Are Running the Asylum. Macmillan Publishing Co., Inc., Indianapolis, IN, USA.

Crabtree, B. F., Miller, W. F., 1992. A template approach to text analysis: Developing and using codebooks. In: Miller, W. L. (Ed.), Doing qualitative research. Sage Publications, Thousand Oaks, CA, USA, pp. 93–109.

Eriksson, J., Johansson, G., Akselsson, K., Sep 1995. A planning tool prototype for environment adaptations. Rehabilitation Engineering, IEEE Transactions on 3 (3), 283–287.

Fereday, J., Muir-Cochrane, E., 2008. Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. International journal of qualitative methods 5 (1), 80–92.

Hafkin, N. J., Huyer, S., 2007. Women and gender in ict statistics and indicators for development. Information Technologies & International Development 4 (2), pp–25.

Hull, L., Mar. 2004. Accessibility: It's not just for disabilities any more. interactions 11 (2), 36–41.
URL http://doi.acm.org/10.1145/971258.971270

Huyer, S., Hafkin, N., Ertl, H., Dryburgh, H., 2005. Women in the information society. In: Sciadas, G. (Ed.), From the digital divide to digital opportunities: Measuring infostates for development. Orbicom, Ottawa, Canada, pp. 135–196.

Janarthanan, V., April 2012. Serious video games: Games for education and health. In: Information Technology: New Generations (ITNG), 2012 Ninth International Conference on. pp. 875–878.

Kasday, L. R., 2000. A tool to evaluate universal web accessibility. In: Proceedings on the 2000 Conference on Universal Usability. CUU '00. ACM, New York, NY, USA, pp. 161–162.
URL http://doi.acm.org/10.1145/355460.355559

Keates, S., Clarkson, P. J., Harrison, L.-A., Robinson, P., 2000. Towards a practical inclusive design approach. In: Proceedings on the 2000 Conference on Universal Usability. CUU '00. ACM, New York, NY, USA, pp. 45–52.
URL http://doi.acm.org/10.1145/355460.355471

Lange, B., Chang, C.-Y., Suma, E., Newman, B., Rizzo, A., Bolas, M., Aug 2011. Development and evaluation of low cost game-based balance rehabilitation tool using the microsoft kinect sensor. In: Engineering in Medicine and Biology Society,EMBC, 2011 Annual International Conference of the IEEE. pp. 1831–1834.

Law, C. M., Yi, J. S., Choi, Y. S., Jacko, J. A., 2006. Are disability-access guidelines designed for designers?: Do they need to be? In: Proceedings of the 18th Australia Conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments. OZCHI '06. ACM, New York, NY, USA, pp. 357–360.
URL http://doi.acm.org/10.1145/1228175.1228244

Leder, R., Murillo, N., Ibarra, C., Gushiken, H., Anaya, G., Escalona, C., Nunez, I., Alvarez, A., Tyler, M., Bach-y Rita, P., 2001. Computer game motivating rehabilitation with objective measures of improvement in motor function. In: Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE. Vol. 2. pp. 1388–1390 vol.2.

Milne, S., Dickinson, A., Carmichael, A., Sloan, D., Eisma, R., Gregor, P., 2005. Are guidelines enough? an introduction to designing web sites accessible to older people. IBM Systems Journal 44 (3), 557–571.

Navarro, A. M., Mocholí, J. B., Naranjo, J. C., 2012. Veritas approach for parameterization of psychological and behavioral models. In: Proceedings of the 13th International Conference on Computers Helping People with

Special Needs - Volume Part I. ICCHP'12. Springer-Verlag, Berlin, Heidelberg, pp. 303–310.
URL http://dx.doi.org/10.1007/978-3-642-31522-0_45

Nielsen, J., Molich, R., 1990. Heuristic evaluation of user interfaces. In: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, pp. 249–256.

Palmon, O., Oxman, R., Shahar, M., Weiss, P. L., 2004. Virtual environments as an aid to the design and evaluation of home and work settings for people with physical disabilities. In: Proc. 5th Intl Conf. Disability, Virtual Reality & Assoc. Tech., Oxford, UK. pp. 119–124.

Papastergiou, M., 2009. Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation. Computers & Education 52 (1), 1–12.

Picking, R., Robinet, A., Grout, V., McGinn, J., Roy, A., Ellis, S., Oram, D., 2010. A case study using a methodological approach to developing user interfaces for elderly and disabled people. The Computer Journal 53 (6), 842–859.

Pruitt, J., Grudin, J., 2003. Personas: Practice and theory. In: Proceedings of the 2003 Conference on Designing for User Experiences. DUX '03. ACM, New York, NY, USA, pp. 1–15.
URL http://doi.acm.org/10.1145/997078.997089

Rubin, J., Chisnell, D., 2008. Handbook of usability testing: howto plan, design, and conduct effective tests. John Wiley & Sons.

Santos, J. R. A., Mitchell, D., Pope, P., 1999. Are open-ended questions tying you in knots? Journal of Extension 37 (4), n4.

Schmidt, A., Dey, A. K., Kun, A. L., Spiessl, W., 2010. Automotive user interfaces: Human computer interaction in the car. In: CHI '10 Extended Abstracts on Human Factors in Computing Systems. CHI EA '10. ACM, New York, NY, USA, pp. 3177–3180.
URL http://doi.acm.org/10.1145/1753846.1753949

Scott, M., Spyridonis, F., Ghinea, G., 2015. Designing accessible games with the veritas framework: Lessons learned from game designers.

Scott, M. J., Ghinea, G., 2013a. Integrating fantasy role-play into the programming lab: exploring the'projective identity'hypothesis. In: Proceeding of the 44th ACM technical symposium on Computer science education. ACM, pp. 119–122.

Scott, M. J., Ghinea, G., 2013b. Promoting game accessibility: Experiencing an induction on inclusive design practice at the global games jam. In: Proceedings of the Inaurgural Workshop on the Global Games Jam. pp. 17–20.

Smith, K., 2011. Innovations in accessibility: Designing for digital outcasts. Society for Technical Communication, 117–127.

Spyridonis, F., Moschonas, P., Touliou, K., Tsakiris, A., Ghinea, G., 2014. Designing accessible ict products and services: The veritas accessibility testing platform. In: Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces. AVI '14. ACM, New York, NY, USA, pp. 113–116.
URL http://doi.acm.org/10.1145/2598153.2598191

Stephanidis, C., Akoumianakis, D., 2001. Universal design: Towards universal access in the information society. In: CHI '01 Extended Abstracts on Human Factors in Computing Systems. CHI EA '01. ACM, New York, NY, USA, pp. 499–500.
URL http://doi.acm.org/10.1145/634067.634352

UNESCO, 2007. Science, Technology and Gender: An International Report. EDC, Walthan, MA, USA.

Union, E., 2004. Report of the Inclusive Communications (INCOM) Subgroup of the Communications Committee (COCOM).

Van Eck, N., Waltman, L., 2011. Text mining and visualisation using vosviewer. ISSI Newsletter 7 (3), 50–54.

Westin, T., 2004. Game accessibility case study: Terraformers–a real-time 3d graphic game. In: Proceedings of the 5th International Conference on Disability, Virtual Reality and Associated Technologies, ICDVRAT. pp. 95–100.

Westin, T., Bierre, K., Gramenos, D., Hinn, M., 2011. Advances in game accessibility from 2005 to 2010. In: Universal access in human-computer interaction. Users diversity. Springer, pp. 400–409.

Yuan, B., Folmer, E., Harris Jr, F. C., 2011. Game accessibility: a survey. Universal Access in the Information Society 10 (1), 81–100.