



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE DELLA RICERCA

Alma Mater Studiorum Università di Bologna Archivio istituzionale della ricerca

A method for systematic usability evaluation of interactive product interfaces

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

Mengoni M, Ceccacci S, PERUZZINI, M., Germani M. (2011). A method for systematic usability evaluation of interactive product interfaces. Padova : Libreria Internazionale Cortina Padova.

Availability:

This version is available at: <https://hdl.handle.net/11585/956645> since: 2024-02-11

Published:

DOI: <http://doi.org/>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)



A method for systematic usability evaluation of interactive product interfaces

M. Mengoni ^(a), S. Ceccacci ^(a), M. Peruzzini ^(a), M. Germani ^(a)

^(a) Polytechnic University of Marche, Department of Mechanical Engineering, Italy

Article Information

Keywords:

K1, Interactive design
K2, User-centred Design
K2, Virtual Prototyping
K4, Augmented Reality
K5, Usability

Abstract

Purpose:

The paper aims at defining a user-centred design method to develop interactive product interfaces that allows to stretch time to market and to reduce prototyping costs.

Method:

It consists in an experimental protocol to evaluate product interface usability and to correlate the achieved performance with its design features and attributes. It exploits low cost high-fidelity prototypes based on Augmented Reality techniques that allow users to interact with the Graphic User Interface and at the same time touch the object, explore its shape, the material it is realized to stimulate both cognitive and emotional response.

Result:

Experimental sessions have been carried out on a real industrial case that is the design of an interactive remote control device for bathtubs and showers. Time for the development of the final design solution halves because main usability errors have been detected at the early design stages and no iterative testing sessions have been carried out on the final physical prototype.

Discussion & Conclusion:

Conclusion provides some discussions of experimental results to highlight the effectiveness of the method to detect usability problems and support the definition of what to improve. Finally two research lines have been drawn to foster improvements in virtual prototyping for interactive design.

1 Introduction

Many consumer products on the market have not been fully successful due to the limitations of traditional design processes, where decision-making is only assigned to designer and to producing company. Nowadays when speaking about product innovation we refer to the realization of objects whose level of quality depends not only on the functions they implement, but mainly on their technological and aesthetic performance and on the cultural, social and symbolic meaning they embody and convey to user. It is quite evident that such a product innovation cannot prescind from an innovation concerning primarily design methods. This can be reached by defining a new design philosophy within which methods capable of considering and elaborating whatever, material or immaterial, is related to the product itself must be developed. In this scenario, it is very important to place user at the centre of design process to reach a high level of perceived quality and a high level of product usability.

The user-centered design method develops according to an iterative process based on the knowledge of the context in which the interaction user- product takes place and of nature of the interaction. A similar approach implies the awareness that human subjectivity and behavioural differences shown by the same person in different contexts and activities cannot be described a priori and that the final evaluation parameter of every design proposal is the user, who is called to interact with the chosen solution. Furthermore, knowledge of the

environment of use allows the definition of a design solution which is not directed to the object itself, but which is optimal with respect to the reference context. Evaluation represents the key stage of the creation of a user- sized product. It allows not only to point out design errors, but above all to measure the quality perceived by user and the resulting product experience, so that the design changes can be better finalized.

The introduction of computer-based technologies has lead to an increase of the degree of interactivity of consumer products, which can be remotely programmed, customized according to user's needs and used not only as utilities but also for entertainment. In this context, product usability represents one of the strategic elements for the product's success on the market, since it guarantees the level of provided interactivity, the degree of contents availability, the psycho-physical wellness reached while using the product and, more in general, the perceived quality. Evaluation of product usability is therefore fundamental for user-centred design. It is carried out through usability tests, which allow to acquire detailed information on the product experience and the way consumers use it. To fully evaluate product usability it is necessary to develop an protocol analysis which allows to correlate user's response to the specific product features and to consider both emotional, affective and cognitive aspects of the user-product interaction. In the case of handled devices the evaluation of these aspects can be carried out only by adopting a high fidelity prototype capable of simulating the interaction user- product and reacting correctly to the inputs from user.

Currently usability tests are generally carried out by final users when a reliable physical prototype of the product has been realized. As a consequence potential interaction problems can be detected only at an advanced stage of the product development cycle with an increase of design costs and time. The capability of evaluating usability even before the physical prototype is realized could bring a considerable competitive advantage to companies. This can be reached by using virtual prototypes able to simulate interaction between product and user from the early design stages. However production of interactive high-fidelity virtual prototypes in immersive environments is very complex and expensive.

The aim of this study is therefore to define a method for usability evaluation which can be adopted by designers to involve end users of a product right from the first stages of design, with the advantage of making it possible to check more design proposals. This method uses low cost virtual prototypes which allow users to interact with a product in a similar way as with high-fidelity prototypes. The method implies:

- 1) definition of an experimental protocol for usability analysis allowing to evaluate users' responses to stimuli from product. Both emotional and cognitive reactions are assessed;
- 2) use of Mixed Reality (MR) techniques based on Augmented Reality (AR) technologies for the creation of low cost interactive prototypes.

The method is valid enough to measure users' response to numerous consumer products provided with an interactive computer-based interface for the selection of operating programs and the visualization of relative information (e.g. washing machines and automotive dashboards, mobile phones, audio-visual media, etc.).

The present work aims at illustrating the method proposed, the technologies developed to support its application and the experimental results reached in the case of handheld remote control devices for bathtubs and showers.

2 State of the art

User-centered design (UCD) is a design philosophy and a process which extensively addresses needs, wants, and limitations of end users of a product at each stage of the design process. According to ISO 13407 standard [1], UCD process is structured in the following iterative phases: a) identification of users' needs and establishment of requirements for product; b) development of alternative designs to meet such needs; c) building of interactive prototypes which can be communicated and assessed; d) evaluation of what is being built throughout the process and of the user experience it offers. In this context, ergonomics, intended as "the discipline concerned with the understanding of interactions among humans and other elements of a system" [2], has a central role. It refers to both the physical and psychological elements which determine the quality of interaction with a product in order to assure its safety and usability.

Product usability is a well-know concept in UCD. It refers to the investigation of product performance in terms of efficiency, effectiveness and users' satisfaction as ISO 9241-11 guidelines synthesize [3]. This definition points out two important fields of interest. The first one relates to the qualities of product commonly known as pragmatic attributes [4] (product functionality, dimensional aspects, safety, performances) which concern Physical Ergonomics [5] and which can be evaluated in terms of

effectiveness and efficiency. The second area, instead, relates to user's wellness and is characterized by his/her satisfaction in relation to the product (emotional reactions, level of comfort perceived, sensations felt), that is hedonic attributes [6]. These aspects concern Cognitive Ergonomics [7]. So that, a product may be perceived as pragmatic because it provides effective and efficient means to manipulate the environment. A product may be perceived as hedonic because it provides stimulation, identification or provokes memoris [8].

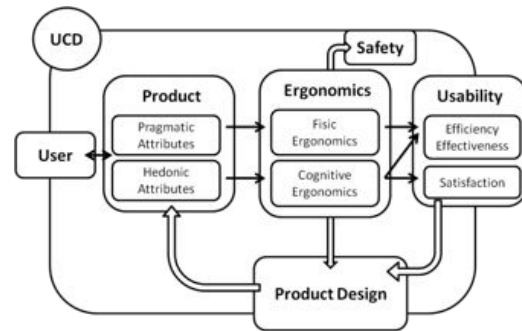


Figure 1. UCD process: relations among product attributes, ergonomic analyses and usability dimensions.

Properties of the object such as weight, shape and material affect product perception (product experience) through two mechanisms; on one side the object is evaluated on the basis of what it allows to do, on the other it is given a symbolic and emotional value which determines the level of psychological wellness. Invitations to action from the object are known as affordances [9]. The second communicative aspect, instead, falls within the definition of synaesthesia [10]. A proper ergonomic analysis must face both affordance and synaesthesia properties which determine respectively usability, utility and aesthetics.

When a user interacts with a product, all the properties perceived influence his/her judgement and consequent actions. Users' response can be considered as the final step of a complex communication process based on theories of perception [11]. In particular, behavioral and cognitive aspects characterized user response [12,13]. Affordance properties qualify the behavioural response, while synaesthesia properties express the cognitive response [14]. Behavioral response refers to the way in which user behaves in front of product, how he/she acts and how he/she can reach his/her goals, while cognitive response refers to the judgment that user makes about product on the basis of the information perceived through the sensorial modalities.

Although subjective satisfaction is included in the original usability definition, most studies have been concerned only with functional and behavioural performance (e.g. efficiency and accuracy in task completion) [15]. Affective and emotional aspects have been neglected as they are difficult to be objectively measured. However, the improvement of product interface performance does not necessarily mean that consumers are satisfied during use. Although in the case of either aesthetic artefacts (e.g. some furniture artefacts) or specific functional products (e.g. manufacturing machines, mechanical components) a separation between subjective and objective assessments is appropriate, this is not acceptable in the case of personal interactive devices where function has progressively moved from the

exclusively practical use also to symbolisms and aesthetics.

From the whole analysis of the state of the art of current UCD methods and of the interaction process between users and products a proper protocol for usability assessment should allow both subjective and objective measurements of behavioural and cognitive response to hedonic and pragmatic product attributes.

Another challenging issue in UCD application consists in the construction of interactive prototypes able to support a proper product ergonomic evaluation in shorter time [16]. Two main prototyping techniques can be classified: low-fidelity prototyping (e.g. paper sketches, cardboard mock-up) and high-fidelity prototyping (e.g. software-based and VR-based prototypes, physical mock-up) [17]. Low-fidelity prototypes are good for testing aspects such as the layout of controls and displays, but not for evaluating the effects of tactile, auditory and visual feedback. High-fidelity prototypes are able to make users realistically appraise product aesthetic attributes and functionalities [18], but they are costly and can be built only at the end of the design process.

Some studies demonstrate how Virtual Reality (VR) - based prototypes can be used to rapidly carry out usability testing, to reduce evaluation time and costs and to involve end-users from the earliest stages of the design process without having to build costly physical mock-ups [19], [20]. These works showed two main technological limitations of VR-based environment for usability testing:

- Interaction is mainly based on vision. Physical interaction is poor;
- The achieved sense of immersion in the virtual environment and the necessary presence in the real product context are difficult to be achieved simultaneously.

In the last years Mixed Reality (MR) environments have gained a great attention in the field of user-centered design because they are able to overcome the above-mentioned problems by combining real and virtual worlds in various proportions and presenting them as a unified whole [21]. Within the MR framework, Augmented Reality (AR) technique is one of the most adopted due to the low cost of the technologies and to its ability to enhance the real scene with computer graphics and emerging tactile and sound rendering displays [22]. It generally uses mobile and wearable systems (e.g. see-through glasses, wearable cameras, Head Mounted Displays, etc.) to superimpose the virtual interfaces designed with real control panels [23]. Many different solutions have been proposed with the intent of providing devices capable of interacting with the AR environment in a more intuitive way [24,25]. The main identifiable problems concern with difficulty of systems integration, high complexity of systems' interfaces, low realism of virtual scenes, the sometimes-unnatural manipulation, non-intuitiveness of the devices adopted and of system control for non- expert users. The developed toolkits provide only a small number of predefined interaction techniques that are expected to be used regardless of context. Moreover, most works mainly address usability in terms of efficiency and effectiveness, and they do not correlate user response with product features and related attributes to support UCD application.

3 The method

Interactive consumer products, such as handheld remote controls and mobile devices (e.g. mobile phones, TV remote controls, PDAs, music players, recorders,

cameras, etc.), are characterized by some physical interface elements, such as display panel, control buttons and knobs, body, which are used to interact with the product (e.g. external case, display, control buttons, etc.) and by some logical elements (e.g. icons, menu structure, semantics) which are used to access and make the product functions work.

Although the present research focuses on the analysis of the logical elements, which are embodied in the Graphical User Interface (GUI), it is assumed that both physical and logical features affect the achieved usability performance. As a consequence, prototypes should be realized to stimulate both behavioral and cognitive response as well as to make the user interact with the GUI in a natural manner.

Different interface design features characterize each interface element. For example the physical elements affecting usability are location and size of control buttons, operating conditions, body shape and weight, materials, etc. On the other side, the design variables of GUI features are: structure of menu, data representation method, size and mutual position of text and icons, icons position within the display panel, etc. These variables influence both pragmatic and hedonic product qualities that are respectively information navigability and quality, comfort in use and aesthetic impression. Navigability refers to easiness of moving within the screen and finding information. It is determined by the level of clarity and adequacy of information, by the way information are organized in the interface and by the level of affordance. Comfort refers to the degree of ease experienced by user while using a device, both in physical and psychological terms. Aesthetics, at last, is a very important characteristic since it is the first to come into play when user interacts with the object and it continues to have a significant role throughout all its use. Such characteristic affects cognitive response at first at a visceral level, determining user's behavioural response, then at a reflective level, influencing user's satisfaction during interaction with product. The method for product interface usability assessment consists of the following steps:

- definition of different types of analysis able to assess each dimension of the product interface usability;
- setting of metrics to measure both cognitive and physical ergonomics of alternative design solutions. The metrics are specific for the evaluation of the GUI elements, but they also take into consideration the influence of the physical elements on user response. For each metric some usability criteria are defined to objectify user response;
- identification of proper observation techniques to capture data during prototypes interaction;
- definition of a method for data elaboration to compare subjective and objective responses and to qualify the GUI;
- development of high-fidelity AR-based virtual prototypes operating similarly to the final GUI and enabling user to interact with the physical elements to trigger out emotional and affective response;
- application of the experimental protocol to assess usability of different GUIs and identification of which interface features allow to achieve the highest performance.

3.1.1 The experimental protocol

The experimental protocol is based on traditional task analysis and on four types of ergonomic investigations that can be used to assess GUI usability by also considering the influence of physical elements on the product's global performance. Task analysis is based on the definition of a set of tasks to be carried out by sample users to verify usability problems. Each task is the course of action the user goes through in order to achieve a specific state. It can be divided into sub-tasks till elementary units to better understand user behavior and product performance.

According to previous studies [14], we considered four types of ergonomic investigations:

- 1.Touch & feel;
- 2.Mental workload;
- 3.Emotional analysis;
- 4.Physical stress.

Each analysis allows to investigate the different dimensions of usability that are effectiveness, efficiency and satisfaction. Analyses are carried out by a set of metrics that practically enable experts to objectify user response during interaction. For each metric some evaluation criteria are defined to objectify the analysis (e.g. completion time, error occurrences, number of facial expression and gestures that reveal stress or pain [26], etc.).

Touch&feel analysis investigates the communicative aspects of artifacts, which determine how the product is able to lead user to a correct use by transmitting its functions (affordance) and the tasks they implement. Such aspects concern the usability dimensions typical of physical ergonomics: effectiveness and efficiency. Three evaluation metrics are defined to support Touch&Feel analysis: information legibility, adequacy and arrangement. They respectively refer to the degree of difficulty in information reading, the level of information intelligibility and adequacy and the degree of difficulty in information searching to execute a specific task.

Mental load analysis, as specified by ISO 10075 standard, aims at determining possible criticalities in a particular context of use in order to evaluate adequacy of

workload. Usually the nature of criticalities is subjective, even if they objectively affect task completion, since they influence the efficiency dimension. The measurement of the ease to use and friendliness of the product interface, the workload adequacy in respect with task complexity and the ability of the GUI to avoid mental workload and repetitive actions (absence of monotony) is used to carry out mental load analysis.

Emotional analysis investigates the sensations arising during interaction at cognitive level. These sensations affect human behaviour by determining the way user relates to product. The main element emotional analysis intends to evaluate is perception of product aesthetics. This analysis allows to obtain subjective data concerning satisfaction. Three metrics are defined to evaluate emotional aspects: pleasure in use that refers to the pleasure perceived by users during interaction, the sense of order, that measures the perceived interface organization, and aesthetic appreciation that regards with the perceived sense of beauty and aesthetic liking.

Physical stress analysis focuses on the degree of comfort user experiences at a physical level during interaction and fully expresses the concept of physical ergonomics analysis. It can be evaluated by measuring the pleasure in handling that is the perceived comfort in product handling (e. g. easiness in manipulation, touch sensation, etc.)

Metrics related to satisfaction are evaluated only on the basis of subjective data (subjective satisfaction). Those relating to effectiveness and efficiency are evaluated on the basis of both objective (number of errors, task completion time, facial expressions of discomfort) and subjective data. Subjective data are collected using a Likert 1-5 scale.

The product attributes evaluated in all analyses are the ones mainly influencing user response in the case of handled devices, in other words: aesthetics, navigability and comfort. Table 1 sums-up the identified evaluation criteria (metrics) for each type of analysis and correlates them with the main usability dimensions and product interface attributes.

Table 1 Protocol analysis: metrics description, usability dimensions and data collection

PRODUCT ATTRIBUTES	ERGONOMIC ANALYSIS	USABILITY DIMENSION	EVALUATION METRICS	OBJECTIVE USABILITY METRICS		SUBJECTIVE USABILITY METRICS		INVESTIGATION TECHNIQUES
				Description	Unit	Description	Unit	
NAVIGABILITY	Touch & feel	Effectiveness	Information legibility	Frequency of some user behaviour such as squinting, moving close the display	occurrence number	The dimensions of icons and characters is really appropriate, so that reading the information is very simple.	1-5 Likert Scale (1=strongly disagree, 5=strongly agree)	<u>Objective data collection methods:</u> VIA, direct observation
			Information adequacy	Assistances occurrence: user does not associate icon to respective command	occurrence number	The informations provided by texts and icons are very intuitive		
			Information arrangement	Errors number: user does not discern a specific command in a menu	occurrence number	It is easy to find the information I need		
COMFORT	Mental workload	Efficiency	Easy to use	Time to complete task	seconds	Overall, this product is easy to use		
			Workload adequacy	Frequency of some user behaviour such as asking clarification questions	occurrence number	It is easy to learn to use this product		
COMFORT	Physical stress	Effectiveness	Absence of monotony	-----	----	It is not annoying at all to use the product		
			Pleasure in handling	Frequency of some user behaviour such as touching the physical elements, facial expressions and gestures of discomfort (e.g. stress or pain)	occurrence number	I feel comfortable using this product		
			Pleasure in use	-----	----	It is very pleasant to use the product		
AESTHETICS	Emotional	Satisfaction	Sense of order	-----	----	The organization of the information on the menus is very clear		
			Aesthetics appreciation	-----	----	This product looks very nice		

3.1.2 Data capture

In order to record data and monitor users during product interaction, three investigation techniques are combined: Video Interaction Analysis (VIA) [27], interviews and direct observation. VIA technique consists in video-recording the testing sessions. It allows to capture interaction by analyzing both users' behaviours, words, gesture and facial expressions. Recorded data are divided into segments according to the main undertaken activities. For each segment of verbal communication, completion times, metric occurrences, user behaviours and frequency of actions are transcribed to support the objective measurement of some evaluation criteria. Interviews consist in a pre-defined list of questions that experts ask users while they are evaluating the product prototype. The aim is to assess the protocol metrics in a more friendly way in respect to traditional post-hoc questionnaires. According to the question users express their judgement according to 1-5 Lickert scale. Direct observation allows to have an immediate feedback from users' actions and to better investigate testing sessions. It is carried out by an expert that monitors the testing session and fulfils a Diary Study MSWord format. It contains both data about users' and experts' impressions and sketches of user actions and behaviours. Diary Study is a valid support to VIA as it makes explicit some moment-to-moment impressions and captures what cameras are not able to record.

3.1.3 Data elaboration: how to compare different product alternatives

To compare the usability performance from different analyses it is necessary to have uniform scales, which are obtainable through normalization. For objective data the normalized value is calculated in three steps: 1) by considering the maximum average value in the whole task, 2) by assuming that it is the condition limit for the task acceptability, 3) by normalizing the average numbers in respects to it. The metric "time to complete task" is an exception, in fact it is calculated by the ratio between experts' completion times and user times. Different products can be ranked according to the averaged values of completion time for all tasks. Subjective average judgments are normalized in respects to 5, that is the upper limit of a 5-point Likert scale.

In order to obtain a single score in each metric it is possible to sum average subjective and objective data, once assumed that they have equal weights. In the same way it is possible to calculate a single value for each usability dimension by working out the average of the values obtained for the relative metrics.

It can be argued that we have not considered the possible correlations between objective and subjective collected data and consequentially that some collected data may be redundant. Of course it is true, but this is not very crucial in a qualitative analyse.

3.2 How to create interactive virtual prototypes

The MR prototyping technique here described exploits the advantages of two traditional prototyping methods for usability testing: rapid prototyping and paper prototyping. The MR prototypes obtained are therefore characterized by a tangible interaction in AR environment and by a functional simulation of the GUI behaviour. The tangible interaction is produced thanks to a physical prototype of

the product to which a projection of the virtual prototype in real scale is overlapped. The behaviour of the device is managed by an operator who, by changing the markers according to users' options, interactively modifies the screenshots visualized on the physical prototype. A human-scale environment is used to visualize the interactive AR-based prototype. The user hand is displayed as a pointer moving synchronously with the user in the real scene.

The elaboration of the prototypes was carried out with the support of the following software tools:

- CATIA V5 (by Dessault Systemes) to represent all design alternatives through 2D drawings and 3D models of components and as-semblies.

- LinceoVR Software (by Seac02) to render 3D models and to create animations of the virtual prototypes in Camera Matching (AR) modes. This software also supports active stereoscopic rendering to perceive 3D depth using proper visualization displays (e.g. AR dedicated see-through glasses and large volume displays).

- ZPrint™ software to import the polygonal mesh model, to repair and optimize it, to analyze parts in order to check whether the whole surface is closed and whether all facets normal vectors are pointed outwards, and finally to prepare the RP process (e.g. slicing the model into cross-sectional slices, positioning the parts in the platform built, setting-up the process parameters, performing exports in the format required by the adopted RP machine).



Figure 2. Example of a MR prototype used during the tests.

LinceoVR software allows to work with Augmented Reality. The three-dimensional models of the remote control device were imported, the materials chosen during the design phase were applied to them and the relations between the screenshots to be analyzed and the markers were created. A screenshot of the proposal examined is associated to each marker. A separate marker is set to be associated to a virtual pointer simulating the finger the user utilizes to interact with the model. The physical models were constructed by Rapid Prototyping (RP) techniques. They were produced by 3D ZPrinter® 450 by Z-CORPORATION, which creates a 3D physical model directly from digital data, layer by layer. The CAD file is imported in .stl format and sliced into cross-sectional slices and the printer creates the model one layer at a time by spreading a layer of powder and inkjet printing a binder in the cross-section of the part. A special glue (cyanoacrylate) is used to strengthen the prints. The physical prototypes were subsequently elaborated to

obtain a realistic surface finishing. The MR prototype is therefore obtained by overlapping in real time the image of the AR prototype with the physical one.

4 Experimental test case

Experimental tests have been carried out on remote handheld control devices for bathtubs and hydro-spa in the health and wellness sector. These highly interactive devices merge aesthetics, functions, ergonomics and technical performance. Four different design solutions have been developed. For each of them a MR prototype was created. The experimental procedure was organized in the following phases:

1. Conception of several design alternatives and description of their features referring to protocol elements (e.g. navigability, comfort and aesthetics);
2. Construction of the experimental set-up exploiting MR techniques to create interactive virtual prototypes implementing different GUIs;
3. Definition of protocol tasks according to the product functions;
4. Selection of a proper sample of participants in relation to the market target;
5. Application of the experimental protocol to assess the product usability of the different design alternatives;
6. Data evaluation, collection and elaboration.

4.1 Experimental set up

The lab arranged for experimentations is equipped with:

- an active 3D stereoscopic projector F10 AS3D ZOOM by ProjectionDesign able to implement the patented as "dual head", an IR emitter system and APG6000 active glasses by NuVision;
- a front-projected flexible display 150" DIAMOND (300x225 cm);
- A HW workstation with PCI-E Nvidia Quadro FX4800 1,5GB 2Xdvi;
- 3D Connexion Space Navigator Professional with 6 DOF to manipulate the prototype;
- a webcam C910 Logitech with Zeiss® optics, autofocus and 30 frames per second video to video capture the user during prototype interaction.

4.2 Product design and prototyping

The four remote handheld control devices analyzed differ for:

- 1) the way the information are organized: in a matrix or in a list;
- 2) the number of icons per screenshot: 4 or 9;
- 3) the use of descriptive text for the several functions;
- 4) the size of text and icons.

In all alternative design solutions the graphics used for the icons is the same. Moreover, all the icons have a square shaped body made of a material which can be assimilated to smooth and shiny perspex and implement the touch-screen with iconic interfaces.

The analyzed four proposals consist of (figure 3):

1. an entirely iconic interface with nine macro-categories with no description;
2. an iconic interface with four macro-categories and description of the function below each icon;
3. an iconic interface with nine macro-categories and description below each icon;
4. a list interface with four macro- categories, each icon positioned on the left with a fixed description aside.

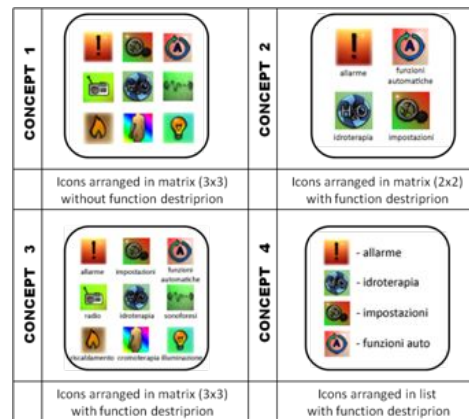


Figure 3. Different design solutions

4.3 Experimental sessions

Product usability was evaluated by task analysis and by applying the above-mentioned four types of investigations. The testing users were chosen using a subjective sampling method based on the field studies provided by the producer company and on two selection criteria: age and gender of the target market. Five women, mobile technology unskilled, were involved in the tests. This number is sufficient to reveal about 80% of all usability problems, as supported by different studies [28]. Concerning age distribution, 2 of the women were 45-50 years old, 2 were 35-45 and 1 was over 50 years old.

For each interface proposal users were asked to carry out tasks, which had been planned in advance. The tasks had been defined to evaluate the quality of the interaction with every product function according to the above-mentioned metrics of measurement. The procedure adopted during the test was very similar to that of a paper prototypes analysis. The tests were carried out with the support of two operators. The first one had to simulate the behavior of the GUI by replacing the pieces where markers are printed according to user options.

Table 2 Tasks description and expert performance data

	Description	EXPERT TIME [s]			
		Concept 1	Concept 2	Concept 3	Concept 4
TASK 1	Switch on the radio and set the station FM 104.5.	20.74	31.98	14.17	20.77
TASK 2	Start the Hydrosonic treatment for a period of 20 mins.	42.34	18.35	25.86	20.59
TASK 3	Start the Sonophoresis treatment and select "Uniform massage" option.	22.56	25.71	17.15	18.73
TASK 4	Turn on the spotlight and start Cromotherapy in fixed-green modality.	19.60	28.74	11.76	25.77
TASK 5	Start Sanification as automatic function. Then immediately stop it and go back to the main menu.	43.92	26.25	30.14	23.79

In the meantime the second operator watched and collected data according to Direct Diary Study method. Once the interaction was over, users were asked some questions to obtain the subjective data relating to the metrics. In the end users were asked to give their global judgment on the four alternatives on a Likert 1- 5 scale and to arrange them from the least to the most satisfying.

Before starting the test, a calibration phase was necessary to set the MR prototype, in order to assure a perfect superimposition between the user finger and the virtual pointer: in order to set the position of the pointer using LinceoVR dedicated commands, each user was asked to put his index fingers on the display area, then was asked to move it slowly in order to check the goodness of the superimposition.

The tasks the users were asked to complete are described in the following summarizing chart (table 2) within which, for completeness sake, also the times taken by an expert user are given. The tasks proposed are the same for each structure, in order to limit the variables to be considered, and every user completed them for each proposal according to the order defined above. The expert user belonged to the industrial partner of research. He was trained for two hours to use the virtual prototyping technology in order to achieve times that are independent from the adopted interface but that are affected only by the designed GUI.

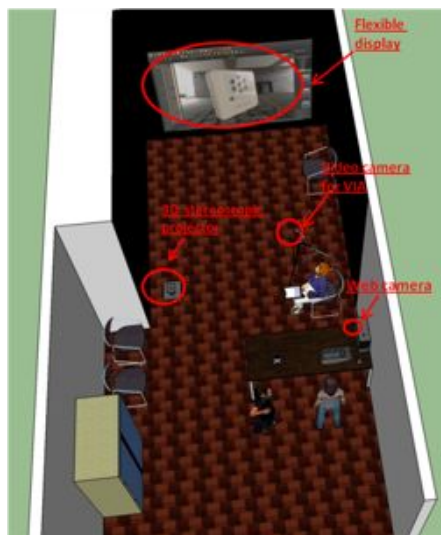


Figure 4. Experimental set up

4.4 Experimental results

Collected data were normalized using the method described above. Proposals 3 and 4 turned out to be winning in respect to proposals 1 and 2 (table 3). As far as concept 3 is concerned, the result is not surprising. In fact, this way of structuring information is similar to the one currently used by the well-known iPhones (by Apple). This result is unexpected: the users that were involved in the tests had not any familiarity with this type of interface, indeed.

Several are the reasons for this success. Users recognise a known metaphor, perceive a menu structure which is similar to well-known commonly used products, appreciate the social and cultural values transferred by the type of high-performance technology and finally find the product pleasant from an aesthetic point of view. All the icons in the structure are represented in the main screenshot with a visualization of the information which can be defined “horizontal”, in the sense that user, right from starting the device, is potentially capable of launching the majority of the functions, since there is no need to look for information in the sub menu. This makes user’s task undoubtedly easier and more immediate. The down side of this structure could be that in this way there is less space available for each icon, with a negative influence on legibility. However this is compensated by the presence of the descriptive text. Proposal 4 is less widespread. In this structure the information are arranged “vertically”, in other words the icons are structured on more levels of details, nested in more and more specific sub menus. The main advantages of this design proposal are given by the bigger size of the icons and of the texts, which allow user to quickly find the metaphor and select it with precision, since the sweet spot to activate the control is considerable. Furthermore, the visualization of the information in columns makes it easier to read on the screen: in fact user can scroll the display with no problems. Interface 2 adopts the same vertical structure of menu as interface 4, but the icons are arranged in a matrix instead of in a list. Its failure is probably due to the fact that, to the contrary of proposal 4, the down side of not having all the information immediately available is not compensated by the bigger size of the text. Moreover, in this case the description is situated below the icon and the character is smaller because the space available is less. With regards to proposal 1, there is no difference with proposal 3 concerning the way the information are arranged.

Table 3. Results: performances of the design alternatives in terms of usability.

USABILITY DIMENSIONS	EVALUATION METRICS	Concept 1 Collecting data				Concept 2 Collecting data				Concept 3 Collecting data				Concept 4 Collecting data			
		Objective	Subjective	Global	TOTAL	Objective	Subjective	Global	TOTAL	Objective	Subjective	Global	TOTAL	Objective	Subjective	Global	TOTAL
EFFECTIVENESS	Information legibility	0,42	0,68	0,55	0,61	0,46	0,72	0,59	0,68	0,70	0,88	0,79	0,87	0,74	0,76	0,75	0,85
	Information adequacy	0,43	0,84	0,64		0,82	0,88	0,85		0,92	0,96	0,94		0,94	0,88	0,91	
	Information arrangement	0,37	0,88	0,63		0,34	0,68	0,51		0,87	0,88	0,88		0,84	0,88	0,86	
	Pleasure in handling	0,40	0,84	0,62		0,70	0,84	0,77		0,85	0,88	0,87		0,85	0,92	0,89	
EFFICIENCY	Workload adequacy	0,56	0,92	0,74	0,75	0,92	0,92	0,92	0,83	0,96	0,92	0,94	0,90	0,96	0,92	0,94	0,89
	Absence of monotony	---	0,88	0,88		---	0,80	0,80		---	0,88	0,88		---	0,84	0,84	
	Easy to use	0,46	0,84	0,62		0,60	0,88	0,77		0,73	0,96	0,87		0,74	0,96	0,89	
SATISFACTION	Pleasure in use	---	0,90	0,90	0,89	---	0,80	0,80	0,84	---	0,96	0,96	0,96	---	0,96	0,96	0,92
	Sense of order	---	0,96	0,96		---	0,96	0,96		---	0,96	0,96		---	0,92	0,96	
	Aesthetics appreciation	---	0,88	0,88		---	0,88	0,88		---	0,96	0,96		---	0,88	0,88	

The difference in performance is exclusively due to the absence in proposal 1 of descriptive texts. This has determined a considerable difference in terms of effectiveness and efficiency between concept 1 and the other proposals in which the descriptive text is used. However this does not completely justify the difference in the results. In this case a good explanation lies in the order in which the tests were carried out. In fact it is obvious that at the beginning the user needed to become familiar with the test and its elements, for this reason the majority of the request for clarifications were concentrated in the first stage of the experimentation during which the first proposal was tested. The same can be said for the physical prototype: the effort of the hands and fingers to reach the controls or the weariness for holding the prototype were felt especially in this first phase. The user gradually adapted and for this reason afterwards he did not visually express any sign of unease or give a negative answer to the questions concerning physical stress. The same happened with the icons. At the beginning it took some time to understand them, afterwards, since the icons are the same in all the proposals, it was possible to proceed faster and without further requests for clarifications. The fact that the users

were in any case satisfied with the product means that the results relating to this proposal are underestimated by the effectiveness and efficiency analysis.

Concept n.3 that gained the highest score, was then engineered and a final physical prototype was realized. It integrates a water resistive touch screen based on APR (Acoustic Pulse Recognition) technology. The relative GUI has been implemented by using a cross platform application based on Linux kernel 2.6 and a modular QT C++ class library. The experimental protocol was applied to the final product and the same number and composition of sample users were submitted to tests. Results confirmed the ones achieved with the AR-based prototype. Differences mainly relate to completion times of those tasks carried out by both users and experts, that are lower than in case of virtual prototypes. No additional testing sessions were repeated to improve the designed handheld device. By comparing this process with a traditional one based on the application of usability testing at the end of the design development, the overall time decreases about -50%, due to an essential reduction of process iterations and to an increase of detected usability problems at the early stages of design process.

Table 4. Results: evaluation of interaction qualities of the three alternatives in terms of pragmatic and hedonic attributes.

PRODUCT ATTRIBUTES	EVALUATION METRICS	Concept 1 Collecting data				Concept 2 Collecting data				Concept 3 Collecting data				Concept 4 Collecting data			
		Objective	Subjective	Global	TOTAL	Objective	Subjective	Global	TOTAL	Objective	Subjective	Global	TOTAL	Objective	Subjective	Global	TOTAL
NAVIGABILITY	Information legibility	0,42	0,68	0,55	0,62	0,46	0,72	0,59	0,67	0,70	0,88	0,79	0,86	0,74	0,76	0,75	0,84
	Information adequacy	0,43	0,84	0,64		0,82	0,88	0,85		0,92	0,96	0,94		0,94	0,88	0,91	
	Information arrangement	0,37	0,88	0,63		0,34	0,68	0,51		0,87	0,88	0,88		0,84	0,88	0,86	
	Easy to use	0,46	0,84	0,65		0,60	0,88	0,74		0,73	0,96	0,85		0,74	0,96	0,85	
COMFORT	Workload adequacy	0,56	0,92	0,74	0,79	0,92	0,92	0,92	0,82	0,96	0,92	0,94	0,91	0,96	0,92	0,94	0,91
	Absence of monotony	---	0,88	0,88		---	0,80	0,80		---	0,88	0,88		---	0,84	0,84	
	Pleasure in handling	0,40	0,84	0,62		0,70	0,84	0,77		0,85	0,88	0,87		0,85	0,92	0,89	
	Pleasure in use	---	0,90	0,90		---	0,80	0,80		---	0,96	0,96		---	0,96	0,96	
AESTHETICS	Sense of order	---	0,96	0,96	0,92	---	0,96	0,96	0,92	---	0,96	0,96	0,96	---	0,92	0,96	0,92
	Aesthetics appreciation	---	0,88	0,88		---	0,88	0,88		---	0,96	0,96		---	0,88	0,88	

5 Conclusions

A user-centered design method for interactive products has been presented. It consists of an experimental protocol to assess product interface usability and to correlate the achieved performance with product features and of a low cost high-fidelity prototype based on AR techniques. The aim of the research is to reduce time to foster the implementation of user-centered design approaches into real industrial design contexts. By analyzing the results from different users, it is possible to infer that the adopted protocol has proved to be appropriate to evaluate the GUI and the mixed reality prototype can be useful to investigate both cognitive and physical ergonomics due to its dual nature (i.e. virtual interactive graphic interface and physical interaction with the product body). Although adopted metrics mostly depend on the particular context of analysis, the proposed methodology is valid in general and can be used for evaluating product usability even in the early stages of design process. However, it still need to be improved: as shown before, some results was influenced by the order

in which the tests were carried out. This can be overcome by changing the order of evaluation for each user each time.

The problems faced during experimentation are mainly due to the test set-up that results to be quite new both for company expert and for final users. They should be trained for about 15 minutes before starting with task analysis. On the contrary, these problems provide hints to improve an analysis which is already of a good level. In fact, it is necessary for the analysis to become even more natural and intuitive, so that user can feel the sensation of using a perfectly working prototype even if the object is still at the conceiving stage. Two lines of improvement are suggested. The first one concerns visualization of augmented prototype: by adopting see-through glasses or Head-Mounted Displays it could be possible to visualize everything at the same time with an ever higher degree of realism. The second improvement concerns automation of product interactive response by means of a finite state machine that is under development.

Acknowledgement

The authors wish to thank Teuco Guzzini S.p.a for their support in product prototypes development and experimentation.

References

- [1] **ISO 13407:1999** - Human-centred design processes for interactive systems.
- [2] **IEA 2011** website. <http://www.iea.cc/> accessed 03 jan 2011
- [3] **ISO 9241-11:1998**. Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability
- [4] **N. Bevan**. Extending Quality in Use to Provide a Framework for Usability Measurement. Proceedings of HCI International 2009, San Diego, California, USA.
- [5] **Gilad, I., 1999**, "A methodology for functional ergonomics in repetitive work", International Journal of Industrial Ergonomics, Vol. 23 (3), pp. 205-221.
- [6] **D. A. Norman**. Emotional design: why we love (or hate) everyday things. Basic Books, New York, 2004.
- [7] **De Waard, D., Brookhuis, K.A., Moraal, J. and Toffetti, A., 2001**, Human Factors in Transportation, Communication, Health, and the Workplace, Shaker Publishing, Maastricht, the Netherlands.
- [8] **M. Hassenzahl**. Funology (chap. 3). Springer 2005.
- [9] **J. McGrenere and W. Ho**. Affordances: clarifying and evolving a concept. Proceedings of Graphics Interfaces. Montreal, pp. 176-186.
- [10] **H. You, K. Chen**. Applications of affordance and semantics in product design. Elsevier, 2006.
- [11] **Floch, J.M., 2000**, Visual Identities, Continuum, London
- [12] **Crilly, N., Moultrie, J. and Clarkson, P.J., 2004**, "Seeing things: consumer response to the visual domain in product design", Design Studies, Vol. 25, pp. 547-577.
- [13] **Bagozzi, R.P., Gurhan-Canli, Z. and Priester J.R., 2002**, The social psychology of consumer behaviour: applying social psychology, Sutton S. Eds, Open University Press, Philadelphia.
- [14] **M. Mengoni, M. Peruzzini, F. Mandorli, M. Bordegoni, G. Caruso**. Performing ergonomic analysis in virtual environments: a structured protocol to assess humans interaction. In: Proceedings of the ASME 2008 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference IDETC/CIE 2008, August 3-8, 2008, Brooklyn, New York, USA.
- [15] **Hornbæk, F., 2005**. Current practice in measuring usability: Challenges to usability studies and research. International Journal of Human-Computer Studies, 64, 79-102.
- [16] **Bullinger, H., Bauer, W. Wenzel, G. and Blach R., 2010**, Toward user-centred design (UCD) in architecture based on immersive virtual environments, Computer in Industry Vol. 61(4), pp. 372-379.
- [17] **Hall R. R. (2001)** Prototyping for usability of new technology. International Journal of Human-Computer Studies 55, 485-501.
- [18] **Sauer, J., Sonderegger, A., 2009**. The influence of prototype fidelity and aesthetics of design in usability tests: Effects on user behaviour, subjective evaluation and emotion. Applied Ergonomics, 40, 670-677.
- [19] **Kuutti, K., Battarbee, K., Sade, S., Mattelmaki, T., Keinonen, T., Teirikko, T., Tornberg, A. M., 2001**. Virtual prototypes in usability testing. In: Proceedings of the 34th Hawaii International Conference on System Sciences.
- [20] **Park, H.J., Son, J.S. and Lee, K.H., 2008**, Design Evaluation of Digital Consumer Products Using VR-Based Functional Behaviors Simulation," Journal of Engineering Design, vol. 19, no. 4, pp. 359-375
- [21] **Bruno F., Muzzupappa M. (2010)** Product interface design: A participatory approach based on virtual reality. International Journal of Human-Computer Studies 65, 254-269.
- [22] **Zhou, Z., Cheok, A. D., Yang, X., Qiu, Y., 2004**. An experimental study on the role of 3D sound in augmented reality environment. Interacting with Computers, 16, 1043-1068.
- [23] **Park H., Moon H.C., and Lee J.Y., 2008**. Tangible augmented prototyping of digital handheld products. In Computers in Industry, 60: 114-125, 2008.
- [24] **Seichter H., and Kvan T., 2004**. Tangible interfaces in design computing, Virtual Environment, Vol. 2, 2004.
- [25] **Bordegoni M., Caruso G., and Ferrise F.** Mixed-Reality Environment based on Haptic Control System for a tractor cabin design review. Proceedings of CIRP Design Conference, Enschede, 7-9 April, 2008.
- [26] **Nilsson T., Nelson T.M., Carlson D.** Development of fatigue symptoms during simulated driving. Accidental Analysis and Prevision, vol.29, n. 4, 479-488, Elsevier, 1997.
- [27] **Jordan, B., Henderson, A., 1995**. Interaction Analysis: Foundations and Practice, Journal of the Learning Sciences, 4: 39-103, 1995
- [28] **Landauer, T.K., Nielsen, J., 1993**. A mathematical model of the finding of usability problems. In: Proceedings of the ACM INTERCHI'93 Conference, Amsterdam, 24-29 April, pp. 206-213.