A methodology for evaluating the usability of audiovisual consumer electronic products

Jiyoung Kwahka, Sung H. Han* b

a Grado Department of Industrial and Systems Engineering, Virginia Polytechnic Institute and State University, 250 Durham Hall, Blacksburg, VA 24061, USA
b Division of Mechanical and Industrial Engineering, Pohang University of Science and Technology, San 31, Hyoja-Dong, Nam-Gu, Pohang, Kyungbuk, 790-784, South Korea

Received 26 November 2001; received in revised form 15 April 2002; accepted 2 May 2002

Abstract

Usability evaluation is now considered an essential procedure in consumer product development. Many studies have been conducted to develop various techniques and methods of usability evaluation hoping to help the evaluators choose appropriate methods. However, planning and conducting usability evaluation requires considerations of a number of factors surrounding the evaluation process including the product, user, activity, and environmental characteristics. In this perspective, this study suggested a new methodology of usability evaluation through a simple, structured framework. The framework was outlined by three major components: the interface features of a product as design variables, the evaluation context consisting of user, product, activity, and environment as context variables, and the usability measures as dependent variables. Based on this framework, this study established methods to specify the product interface features, to define evaluation context, and to measure usability. The effectiveness of this methodology was demonstrated through case studies in which the usability of audiovisual products was evaluated by using the methods developed in this study. This study is expected to help the usability practitioners in consumer electronics industry in various ways. Most directly, it supports the evaluators’ plan and conduct usability evaluation sessions in a systematic and structured manner. In addition, it can be applied to other categories of consumer products (such as appliances, automobiles, communication devices, etc.) with minor modifications as necessary. © 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Usability; Audiovisual consumer electronic products; Usability evaluation framework

1. Introduction

Decades of ergonomics research changed the consumers’ criteria by which they value and choose a product. Compared with the emphasis placed on new functions, reliability, and good after-sale service in the past (Butters and Dixon, 1998), usability has become an important user requirement to improve the customer acceptance in the market. Stimulated also by legislations that require products to meet certain standards of usability (Stewart, 1991; ISO, 1993), efforts to cope with user and legal requirements made usability a business phenomenon (Rubin, 1994) or even a part of the industry’s responsibility to its customers (Jordan, 1998a). This trend is amplified especially when it comes to consumer products such as consumer electronics and telecommunication devices. Although it was originally intended to explain how efficient and effective a product was to use (Bennet, 1984; Shackel, 1984), the concept of usability is undergoing a major change as the application area is broadened (Caplan, 1994; Logan, 1994; Nagamachi, 1995; Hofmeester et al., 1996; Nielsen, 1996; Jordan, 1997). Examples include studies on new aspects such as guessability to facilitate ease of learning (Buurman, 1997), pleasure in products use (Jordan, 1998b), and emotional, experiential aspects related to appeal, aesthetics, or product image (Suri and Marsh, 2000). Common to these attempts is that usability should comment on something beyond functional concerns and be extended to consider a broader range of the subjective aspect. Although some valuable insights were provided to redefine the boundary, there

*Corresponding author. Tel.: +82-54-279-2203; fax: +82-54-279-2870.
E-mail address: shan@postech.edu (S.H. Han).
has been no widely accepted definition appropriate for consumer products yet. Only recently, Han et al. (2001) has suggested a new definition of usability for consumer electronic products by integrating the performance and impression aspects into the concept.

Regardless of the type of products, usability evaluation is now considered an essential step in the development of consumer products. Decades of efforts resulted in various usability evaluation techniques and methods ranging from discount methods to formal testing protocols (Nielsen, 1993), contemplating the beyond such as remote evaluation methods and supporting tools (Kwahk et al., 2001). Then, as the number of methods exploded, some studies focused on providing the practitioners with the guidelines to choose appropriate methods (Stanton and Young, 1998; Stanton and Baber, 1996).

In many situations, however, even a well-trained evaluator finds it difficult to perform an evaluation in an effective and efficient way. One reason is that there are too many test issues, influencing factors, evaluation methods, and analysis paradigms that the evaluator needs to consider. Planning and conducting a usability evaluation study requires far more than inventing new techniques or selecting appropriate methods. More efforts are needed to bring together various issues surrounding the usability evaluation process and to establish usability evaluation methodology well enough to support the work in the consumer product design environment.

In this perspective, this study introduces a framework of usability evaluation that guides the evaluators to a structured evaluation of consumer products. A model procedure accompanied by some examples from case studies is introduced to demonstrate the usefulness of the framework. This study focussed on establishing the methodology for audiovisual consumer electronic products. The approach can be extended to other categories such as home appliances, automobiles, or communication devices with proper adaptations to each group.

2. A usability evaluation framework

As shown in Fig. 1, the basis of the usability evaluation framework is a concise model that captures how the usability of a product should be evaluated within the context of user, product, activity, and environment. This four-component framework has long been accepted as the principal components in a human–machine system upon which good system design depends (Shackel, 1984).

Numerous factors around the usability evaluation process can be organized within three groups of
variables: design variables, context variables, and dependent variables. The interface features of a product are considered the design variables, which elicit the primary interest in the usability evaluation process. Certain characteristics of the product, user, activity, and environment correspond to the context variables that define the situation and context in which usability evaluation takes place. Finally, the specific measures of usability are considered dependent variables.

Development of this framework resulted from academic and practical concerns based on the author’s experience of collaborating with related industry. Firstly, as Charlton and O’Brien (1996) mentioned, evaluations are not always systematically planned, but are often conducted based on the preference of the evaluator without careful considerations of various issues of usability evaluation. This is likely to result in irrelevant or useless results, and the evaluation efforts may turn out to be inefficient and unstructured. A new framework is required for the usability evaluation methodology, which is simple and useful to the practitioners at work.

Secondly, the usability evaluation of consumer products should be handled in a different way from that of software interfaces, since a consumer electronic product includes both hardware (e.g., dedicated control buttons to which functions are assigned) and software features (e.g., programmable menus and graphical interfaces). Research interests merge to how these different features interact with each other to produce good coordination that leads to better usability. So, a well-structured usability evaluation framework should cover hardware and software-oriented aspects as well as the interactions between them.

Thirdly, note that a typical usability evaluation is user-, product-, task-, and environment-specific as is obvious from the widely-accepted definition of usability (Shackel, 1984, ISO, 1998; Jordan, 1998a). It goes without saying that the usability of a consumer product is formed around the interactions among the user, product, user activity, and environment. However, as far as most consumer products are concerned, it is less likely to come up with clearly specified users, goals, or context of use (Buurman, 1997). Hence, when it comes to consumer products, it is important to consider the design features of a product with primary interest, and more flexibility is required in defining the context of usability evaluations.

Finally, a practical support is required regarding the measurement of usability. Presumably, a widely accepted definition of usability implies that it should be measured in terms of effectiveness, efficiency, and satisfaction (ISO, 1998). Then it often gets fuzzy when the practitioners need to figure out exactly what measures stand for these three aspects. Also there are other concerns such as if these three aspects properly represent the complex nature of usability, or how to apply various techniques of usability evaluation for the measurement of usability.

The following three sections describe the specific methods related to the three groups of variables showed in the framework.

3. Taxonomy of context variables for usability evaluation

The role of the taxonomy of context variables is to suggest a scheme to specify the test situation and to collect the information for a structured usability evaluation. A widely accepted four-component framework consisting of the product, user, activity, and environment is used as the basic scheme (Fig. 1). The taxonomy of context variables is shown in Table 1.

3.1. User information

A survey was conducted to identify user-related variables essential to the usability evaluation (Salvendy and Carayon, 1997; Rubin, 1994; Woodson et al., 1992; Hix and Hartson, 1993; Jordan, 1998a), in which a total of 55 variables were collected. As listed in Table 1, the results were then classified into five categories based on the functional groupings of the human characteristics: demographic/social information, physical/psychomotor/physiological characteristics, knowledge/experience characteristics, perceptual/cognitive characteristics, and emotional/psychological characteristics.

User variables are especially important when the product is targeted at the users with special needs such as seniors and people with physical or cognitive challenges. The taxonomy of user variables is expected to help the evaluator to define the characteristics of target users in a systematic way.

Combined with the method for usability measurement, which will be explained in a later section, the taxonomy of user variables is also useful when the evaluators need to determine the logistics of usability evaluation. It supports the evaluators by suggesting which user variables should be considered in determining the measurement procedure. For example, different dimensions and criteria of usability should be considered in evaluating a product for audio enthusiasts and for ordinary intermittent users.

3.2. Product information

Aside from the interface features, which will be discussed as design variables later in this paper, some information about the product is important to understand the context of evaluation. Ten representative items were identified as listed in Table 1 based on a survey on consumer product evaluation issues related to product
characteristics (McClelland, 1990; Consumer Report, 1997).

The first item is concerned with the form or status in which a product can be evaluated such as paper-based descriptions, mockups, computer-based prototypes, or finished products (McClelland, 1990). It is closely related to the stage of product development, and affects the fidelity of the evaluation and selection of evaluation apparatus (e.g., a finished product versus a working prototype on a computer). It also determines the scope of interface features that can be evaluated at the moment. If a product is to be evaluated in a paper-and-pencil form (i.e., an initial, rough sketch of the product design concept), the interface features can only be evaluated in much less detail compared with a finished product.

The second item is the level of a product line, which is usually related to the perceived value of the brand name (Consumer Reports, 1997). Most consumer electronic products can be categorized into three major groups based on the pricing strategy: the prestige (high-end), mass-market (mid-range), or the value (low-end) lines. Different product groups may use different evaluation criteria.

The third one is the product type (or product subcategory). A survey was conducted on more than 20 representative consumer electronics companies to examine how the products were classified. Most audiovisual products can be classified into (1) home theatre or full-size systems, (2) shelf, mini- or micro-systems, component systems, personal systems/portables, and mobile products), brand name, company, country, after-sale service quality (or warrantee information), manufactured date, network connectivity, and availability of peripheral devices such as a remote control, earphone, connector cables, microphones, etc.

3.3. User activity information

The word ‘activity’ was used instead of ‘task’ because the activity that the users do with an audiovisual
product is not confined to performing a specific task. Representative user activity information essential to the usability evaluation is summarized in Table 1.

Depending on the user activity characteristics, different criteria should be applied in usability evaluation. For example, if the use of a product is repetitive, simple, and momentary (e.g., loading/unloading a compact disc to/from the loading mechanism), such criteria as responsiveness, accessibility, efficiency, or error prevention would be of primary concerns. On the contrary, for infrequent and complex activities (e.g., programming a recording function of a videocassette recorder), learnability, familiarity, or informativeness would be more important than the others.

Also some characteristics influence the choice of evaluation techniques and evaluation procedure. For example, when the consequence of erroneous handling is expected to be harmful or to damage the product, a possible technique would be to bring the participants to a laboratory-testing environment and conduct an evaluation under control.

3.4. Environment information

The environment information includes both the physical ambient and psychological (or social) conditions of home or office. Some representative items are listed in Table 1 (Woodson et al., 1992; Meister, 1989).

The environment in which a product operates has been considered important in the traditional ergonomics studies. This is partly because the user performance is known to deteriorate in such abnormal environmental conditions as dark, cold, or noisy ones. As far as the usability of audiovisual products is concerned, however, it can be assumed that the environment is under a normal condition.

4. Method to specify the interface features of a product

In this study, interface features are defined as the product attributes that are presumed to affect the usability of a product. Some popular examples include number, size, and shape of buttons used, form and density of information contents presented on a display, size of a product body, to name only a few. A systematic method is required to define the interface features as design variables, which eventually correspond to ‘what should be evaluated’ (i.e., things or objects to be evaluated). The method, if successful, would be very useful as a practical guide to specify what the product under evaluation is like (i.e., how it looks and works). Sometimes it is not obvious to differentiate the interface features of a product from its functions. Without a systematic approach to identify design variables, the evaluators may end up with evaluating the usefulness of the functions that the product has, rather than systematically evaluating the usability of relevant interface features. It is also more likely that critical evaluation objects are omitted from the evaluation or the evaluation concentrates on a few salient features of a product.

The method suggested in this study was established based on the idea that the interface features of a product could be represented in terms of structural and descriptive aspects. The structural aspect can be explained by identifying the components that a product consists of, while the descriptive aspect needs descriptions about how each component looks and works. The interface feature specification method defines human interface elements and their properties for the structural and the descriptive aspect, respectively (See Fig. 1).

4.1. Human interface elements

Human interface elements are defined as the components of a product through which the users communicate (e.g., see, touch, or operate) with the product. Note again that this method is targeted at audiovisual consumer electronic products such as videocassette recorders (VCR), compact disk players (CDP), digital video-disk (DVD) players, and so on.

An extensive survey was conducted to collect the human interface elements relevant to audiovisual products. Ergonomics design guidelines (Bennet et al., 1997; Bullinger et al., 1997; Hix and Hartson, 1993; Shneiderman, 1992), style guides (Apple Computer, 1992; OSF, 1993; IBM, 1989), standards (ANSI, 1988; ISO, 1993), design handbooks and engineering data compendiums (Woodson et al., 1992) were reviewed in this survey. To find out human interface elements that do not appear in the literature (e.g., loading mechanism, connector, etc.), more than 500 existing audiovisual products were examined as well. Photographs of the products obtained from the product catalogs and homepages of consumer electronics companies were used as well as sampled products.

Summarized in Table 2 is the resulting list of 41 human interface elements. Note that they are grouped into two categories: hardware and software elements. Hardware elements lead to the physical configuration of the user interface of a product. They are classified into six subgroups: control, display, loading mechanism, connector, control panel/connector panel, and body/chassis. Software elements include programmed or programmable interface elements that typically appear on a display panel. They were classified into six subcategories: displayed item, menu, form, feedback, message, and help. As consumer electronic products become more integrated and more intelligent, the proportion of software elements to hardware elements seems to increase.
4.2. Properties of human interface elements

The descriptive aspect of interface features (i.e., the way a product looks and works) can be represented by the properties of each human interface element. An extensive literature survey was conducted on design handbooks, guidelines, and style guides to define them. In addition, non-academic publications such as brochures, buying guides, or review magazines for the audiovisual products were also examined (Consumer Reports, 1997; Gallea, 1996).

The surveyed properties were first classified into individual and inter-element properties. Individual properties constitute the characteristics of an individual human interface element, while the inter-element properties stem from the interrelationship between two or more interface elements. The individual properties were then classified into two distinct groups: static and dynamic properties. The static properties are the attributes of an interface feature that do not change over time, while the dynamic properties are transient in nature or change over time. Some examples of static properties are size, shape, color, texture, material, and polishing, while noise, speed, displacement, and resistance are typical dynamic properties. Also, in terms of sensory modality, all the visual, tactile, and auditory aspects were considered.

The class–subclass relationship of the object-oriented modeling technique (Hong et al., 1998; Kim, 1990) was used in this study for effective representation of the individual properties. An example of the class hierarchy is shown in Fig. 2 (See Fig. 2 with reference to Table 3). For example, the ‘rotary control’ interface element is a subclass of the ‘control’ interface element. This means that the ‘rotary control’ class possesses all the properties of the ‘control’ class (e.g., material, color, feedback, etc.) as well as its own properties (e.g., size in diameter, degree of protrusion, dead space, etc.). Likewise, with the ‘knob’ interface element, which is a subclass of the ‘rotary control’ interface element, it is not necessary to list all the common properties of the ‘rotary control’ interface element by noting that it is a subclass of the ‘rotary control’ class. In this way, a total of 295 distinct properties for the 41 human interface elements were identified. See Kwahk (1999) for the complete list.

The inter-element properties consist of integration and interaction properties. The integration properties are the properties generated by two or more human interface elements grouped together such as grouping, arrangement, proximity, sequence, spacing, density, etc. On the other hand, the interaction properties are issued when the dynamic reciprocal relationship (i.e., the interaction) between two or more human interface elements occurs. It is classified into three groups based

<table>
<thead>
<tr>
<th>Group</th>
<th>Subgroup/human interface element</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hardware</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>Switch (toggle switch, rocker switch)</td>
</tr>
<tr>
<td></td>
<td>Button (push, tactile)</td>
</tr>
<tr>
<td></td>
<td>Slider (continuous control slider, discrete slider)</td>
</tr>
<tr>
<td></td>
<td>Knob (continuous control knob, selector knob, ganged knob, jog shuttle)</td>
</tr>
<tr>
<td></td>
<td>Thumb wheel</td>
</tr>
<tr>
<td></td>
<td>Display panel</td>
</tr>
<tr>
<td></td>
<td>Status indicator</td>
</tr>
<tr>
<td></td>
<td>Analog level meter</td>
</tr>
<tr>
<td></td>
<td>Direct reading counter</td>
</tr>
<tr>
<td>Display</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Single-media loading type (open-front, open-top, cartridge-tray, slot-in)</td>
</tr>
<tr>
<td></td>
<td>Multiple-media loading type (carousel tray, file type)</td>
</tr>
<tr>
<td></td>
<td>Grouped loading mechanism</td>
</tr>
<tr>
<td>Loading mechanism</td>
<td></td>
</tr>
<tr>
<td>Connector</td>
<td></td>
</tr>
<tr>
<td>Control panel/connector panel</td>
<td></td>
</tr>
<tr>
<td>Body/chassis</td>
<td></td>
</tr>
<tr>
<td>Software</td>
<td></td>
</tr>
<tr>
<td>Displayed items</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Digitized-bar level meter</td>
</tr>
<tr>
<td></td>
<td>Numerical item</td>
</tr>
<tr>
<td></td>
<td>Verbal item (text, animated text, spoken)</td>
</tr>
<tr>
<td></td>
<td>Nonverbal item (sound, icon, animated icon, image, animation)</td>
</tr>
<tr>
<td></td>
<td>Built-in display menu</td>
</tr>
<tr>
<td></td>
<td>On-screen display menu</td>
</tr>
<tr>
<td>Menu</td>
<td></td>
</tr>
<tr>
<td>System message</td>
<td></td>
</tr>
<tr>
<td>Feedback</td>
<td></td>
</tr>
<tr>
<td>Form</td>
<td></td>
</tr>
<tr>
<td>Help</td>
<td></td>
</tr>
</tbody>
</table>
on the type of interaction involved: order-and-response type, interlock type, and concurrency type. Representative properties of order-and-response type interaction are movement compatibility, spatial compatibility, control/display ratio, and response delay time. Properties for the interlock type interaction include action sequence, interlocking delay time, and the number of interlocking steps. Representatives for the concurrency type interaction are latency and the number of concurrent elements.
Through repeated trial runs, a drawback of the interface feature specification method came to the author's attention. It was that the process could be time-consuming because of the amount of work and size of data without a proper support. A database application tool was developed as shown in Fig. 3 to support this procedure.

In summary, the human interface element specification method provides useful information in determining what to evaluate and specifying the interface features of a product in a structured manner. A successful application of the specification method was reported by Han et al. (2000). They adopted this method to discover the functional relationship between the design features of a product and its usability by developing empirical models. Besides, the usefulness of this method can be further increased when it is used by the designers who want to know the design features that may affect the usability of their design alternatives.

Fig. 3. A database application tool to support the interface feature specification procedure.
5. Method to measure usability

Apparently, the concept of usability is very complex and multidimensional in nature. Also, the selection of good measures and evaluation techniques determines the quality of evaluation studies. In fact, some questions are frequently raised by the practitioners who read through some books on usability evaluation methods: ‘So, how many different aspects should I measure before I can conclude my product is very usable?’; ‘What are good measures that can tell me my product is usable?’; and ‘Which evaluation technique should I choose to efficiently solve my problem?’ Indeed, the measurement of usability requires careful considerations of defining specific usability dimensions, selecting effective measures, and choosing appropriate evaluation techniques as outlined in Fig. 1.

5.1. Usability dimensions

Defining usability dimensions requires us to start from a lengthy discussion about what defines the usability applicable to audiovisual consumer electronic products. Han et al. (2001) has recently suggested a viable definition that states ‘satisfying the user with respect to both performance and image/impression aspects’. This definition was derived based on the idea that the usability of audiovisual products should consider both the traditional performance-oriented aspect and the consumers’ emotional or attitudinal aspect equally important. The performance aspect means “how efficient and effective a product is for a user to perform a task to achieve some intended goals”. On the other hand, the impression aspects is concerned with “how favorable or unfavorable a product is in terms of the images, impressions, or subjective feelings about the product”.

The study by Han et al. (2001) is considered to have made a valuable breakthrough by suggesting the specific dimensions of usability based on this broadened definition. The authors came up with a total of 48 dimensions (23 for performance and 25 for image/impression) through an extensive literature and user survey.

Dimensions for the performance aspect of usability were classified into three categories borrowing the human information processing and response production model: dimensions related to human perception process (seven dimensions such as directness, responsiveness, and simplicity), those about learning/memorization (six dimensions such as learnability, memorability, and helpfulness), and those related to action (10 dimensions such as effectiveness, task conformance, and recoverability).

The authors recall that they had more difficulty in defining the dimensions for the image/impression aspect. This could be because they had to understand users’ feelings toward or against the product, which are more ill-structured in nature and difficult to clearly identify. Another reason is that less resource was available for the subjective aspect since it has been treated less important than the performance. The image/impression dimensions were classified into three categories based on the degree of subjectivity: basic sense (eight dimensions such as shape, color, and texture), description of image (10 dimensions such as metaphoric design image, elegance, and neatness), and evaluative feeling (seven dimensions such as preference, satisfaction, and attractiveness). The basic idea for this classification was that the image/impression felt by the user could be organized in a hierarchical manner depending on the degree to which subjective factors such as human experience, judgement, and feeling are involved. More details on the procedure and the definition of each dimension can be found in Han et al. (1998), Kwahk (1999), and Han et al. (2001).

5.2. Usability measures

A total of 54 measures of usability were found from the literature (Hix and Hartson, 1993; Dix et al., 1993; Meister, 1985; Gechieder, 1985). The number was reduced to 36 after eliminating redundant, irrelevant, or infeasible ones. They were then classified into six groups: (1) temporal (six measures such as task completion time and time spent in errors), (2) frequency (three measures such as frequency of errors and number of tasks completed per unit time), (3) complexity (five measures such as number of user inputs requested and number of functions used to perform a task), (4) psychophysical (eight measures such as magnitude estimation and simple rating scale), (5) descriptive (two measures such as user comments and verbal reports), and (6) physiological (12 measures such as electromyogram and electroencephalogram). The complete list of the usability measures can be found in Kwahk (1999).

5.3. Evaluation techniques

More than 100 evaluation techniques were surveyed from the literature (Dix et al., 1993; Nielsen, 1993; Meister, 1985; Chapanis, 1991; Sinclair, 1990; Christie and Gardiner, 1990; McClelland, 1990; Polson et al., 1992; Monk et al., 1993; Bainbridge, 1990; Card et al., 1983; Meister, 1990). The techniques surveyed were analyzed in terms of redundancy, relevancy, similarity, and efficiency of measurement, and the number of evaluation techniques was reduced to 35. They were then classified into five groups: (1) observation/inquiry (nine techniques such as time line analysis and focus groups), (2) empirical testing (three techniques such as benchmark testing and user trials), (3) introspection
(four techniques such as cognitive walkthrough and thinking aloud protocol), (4) inspection (four techniques such as feature inspection and heuristic evaluation), (5) modeling/simulation (two techniques such as model-based evaluation and computer-aided simulation). More details on the evaluation techniques are presented in Kwahk (1999).

5.4. Selection of measures and evaluation techniques

Combining the usability dimensions, usability measures, and evaluation techniques, the process of selecting proper usability measures and evaluation techniques can be explained as the following three-step procedure: (1) to list up the usability measures applicable to the usability dimensions of interest, (2) to identify the evaluation techniques applicable to the candidate measures, and (3) to choose the most relevant measures and evaluation techniques based on selection criteria. Kwahk (1999) suggested the applicability of measures and evaluation techniques for given dimensions by using the Affinity Diagram Technique (Holtzblatt and Beyer, 1993). Various criteria for selecting measures and evaluation techniques are also summarized in Kwahk (1999). When several criteria are in conflict, relative importance of each criterion could be used. The analytic hierarchy process (AHP) is a good tool to help the investigator make decisions in a structured manner when conflicts exist (Saaty and Kearns, 1985).

6. Usability evaluation procedure and examples from case studies

After they had practiced the suggested methodology in various ways, the authors developed a model of usability evaluation procedure that can be generally applied to the usability evaluation of audiovisual consumer electronic products. Fig. 4 outlines the suggested procedure, and explains how the three methods suggested in this study should be applied in planning a usability evaluation. For better understanding, some examples from two different case studies are given in explaining each step of the procedure.

The usability evaluation procedure starts from a preparatory step 0, where the objective of usability evaluation is stated. It is very important to include, in defining the objective, whether the evaluation requires involvement of users or it had better be done by usability experts. The decision can be made by various factors such as confidentiality constraints (i.e., can the design be released to the public before the product is

![Fig. 4. The suggested usability evaluation procedure.](image-url)
introduced to the market?) or availability of volunteering participants, to name only a few. For example, the objective of the first case study was 'to find out the functional relationship between the design features of typical audiovisual consumer electronic products and various dimensions of usability by developing empirical predictive models' (Han et al., 2000), and the second case study was intended 'to report usability problems of a DVD player' (Kwahk, 1999). Aiming to the development of empirical predictive models, the first case study required participation of users in the evaluation. On the other hand, the second case study recruited usability evaluation experts, aiming to a 'quick and thorough' evaluation to find out any usability problems before the final release of the product.

Step 1 is where the evaluators review the taxonomy of context variables to define the test situation. With a list of context variables that are considered important, the evaluators need to set each of them as a constant, controlled, or random variable. In the first case study, for example, the evaluators agreed that target users' age, sex, and experience of using audiovisual products were expected to affect the underlying relationship between the design features and the usability of a product. The evaluators recruited participants in 20's (i.e., age as constant variable) whose experience of using audiovisual products varied randomly with an average of 17.8 h per week (i.e., experience as random variable). Then they decided to develop empirical models for male and female groups separately (i.e., sex as controlled variable). About the activity-related aspect, the evaluators were interested in the first impression formed while the users were playing around with the product at the store, which was considered in developing testing scenario and the apparatus. So the experiment was conducted in an exhibition room, where the products were displayed on a table. And the participants were allowed to see, touch, or operate the products during the evaluation (See Han et al. (2000) for more details). Even when the evaluation does not aim at any particular test situation, as is the case with the second case study, it is important to go through the taxonomy of context variables and define each of them as either constant or random. It is because the evaluation results should not be extrapolated to other situations in case the variables are set as constant variables. In the second case study, a computer-based fill-in form was developed based on the taxonomy of context variables in advance. The evaluators used the form to define about 35 context variables as constant such as target users' visual and hearing ability (normal), product usage experience (inexperienced), frequency of use (infrequent to intermittent), or illumination and noise (average office level), and the rest of the context variables were assumed to be random (See Kwahk (1999) for the complete list).

Step 2 provides detailed information about the interface features of the products to be evaluated. The supporting tool shown in Fig. 3 can be used to facilitate this step. First, human interface elements of the target products are analyzed. Second, for each human interface element, related individual and inter-element properties are listed up. Then each product is measured with respect to the selected properties of human interface elements. To return to our case study 1, a total of 88 properties related to Control (e.g., number of controls, consistency of control size, color coding), Display (e.g., information density, identifiability of status indicator, shape of primary display), Loading Mechanism (e.g., loading time, noise of loading/unloading, brightness of tray), and Body (e.g., size of product, ratio of horizontal and vertical length, number of colors used) human interface elements were selected as a result of this step. A product measurement form was developed based on the selected list of properties, which was used by the evaluators in measuring a total of 36 randomly sampled audiovisual products. In case study 2, a checklist was developed based on the human interface elements, through which 17 human interface elements were identified as the components of the evaluation object. The evaluators used the checked items as the features to be examined in discovering any usability-related problem.

Step 3 is where the usability dimensions, measures, and evaluation techniques are determined. The step can be done in parallel with Step 2 if necessary. In case study 1, the evaluators selected 33 usability dimensions based on their interests. A subjective rating scale of 0~100 was selected as the usability measure for these dimensions because of the following reasons: it was a measure commonly applicable to both performance and image/impression dimensions; a subjective rating scale best describes the feelings of the users in the first acquaintance situation; and they needed to collect quantitative data for empirical model building. Hands-on experiment (or user trial) was found to be the evaluation technique that best fits their interest for several reasons. Methods from empirical testing were required because they wanted to assess the user’s feelings about the products when they were playing around with them. However, it was not feasible to apply benchmark tasking where each participant is asked to conduct a series of structured task scenarios before evaluating 36 products with respect to 33 usability dimensions; or controlled experiments where the design variables as well as context variables should be formulated as a series of small experiments. In case study 2, the evaluators were interested in discovering usability problems related to 22 usability dimensions (12 performance and 10 image/impression dimensions). Methods in the inspection group were considered alternatives, and the feature inspection method (Nielsen and Mack, 1994), which is
one of the discount usability evaluation techniques, was selected as the evaluation technique. The evaluators used the following criteria and the best method was selected based on the result of AHP technique: intended focus/purpose of the evaluation and ability to address evaluation objectives, participant availability, reliability of usability findings, and the effectiveness in generating recommendations for change.

The rest of the suggested procedure is quite straightforward once the previous steps were followed as suggested. Step 4 is where the evaluators set up specific evaluation scenarios and prepare personnel, apparatus, materials, etc. Then the evaluation is conducted in step 5 according to the scenario developed. Step 6 is a very important step in which the evaluation results are stored in a database for an efficient data management, or various analyses are done on the results such as analysis of variance, regression, or post-hoc comparisons. Also the evaluators finally come up with meaningful findings or design modifications and report the evaluation results. The results of the two case studies are reported in Han et al. (2000) and Kwahk (1999), respectively.

7. Conclusion

To recall the author’s early experience in various usability evaluation projects, we usability practitioners are often frustrated over the complex situations we encounter and overwhelmed by the number of factors to consider. What makes matters worse is the lack of appropriate references and technical supports when we are in need of them. This becomes more serious especially when we have to make critical decisions in planning an evaluation such as recruiting participants or determining what to measure. This paper is all about providing a simple new methodology of usability evaluation that supports the usability practitioners in the consumer electronics industry. This approach is expected to help them make decisions related to the usability evaluation of their products in a right, reasonable and systematic way.

The authors would like to mention a few points as the contribution of this study. Firstly, there are few studies that show what aspects of the product should be considered in a usability evaluation, as far as the audiovisual products are concerned, except the author’s earlier work reported in Hong et al. (1998). This study is one of the first breakthroughs to provide an analytic solution to this issue.

Secondly, this study can be used by the usability practitioners whenever they need information about the logistics of usability evaluation especially for audiovisual consumer electronic products. Considering it is difficult to find a model procedure that we can follow in evaluating the usability of consumer electronic products, this study can play a role as a handy guide to usability evaluation practices.

Recently, consumer electronic products are becoming more complicated and intelligent, and the change occurs faster than ever. With the trend, the importance of incorporating usability into a product is increasingly reported. Many usability specialists have anticipated that usability will become one of the most important success factors for the consumer electronic products. In this perspective, as one of the leading efforts to accommodate usability into consumer electronic products, this study is expected to draw interests of professionals and practitioners who are involved in the design and evaluation of product usability.

Those who have less experience in usability evaluations, however, may find it somewhat difficult to follow the procedure suggested in this study. It is because a large number of minute things should further be considered in the evaluation process. The author’s future efforts could be dedicated to developing an intelligent usability evaluation aiding system with a rule base that supports each step of the evaluator’s decision. It is also desired that the evaluation method be extended to other categories of consumer products such as home appliances, office-automation equipment, personal communication devices, automobile interior design, furniture design, and so on.

References
