International Journal of Occupational Safety and Ergonomics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tose20

A Systematic Procedure for Modeling Usability Based on Product Design Variables: A Case Study in Audiovisual Consumer Electronic Products

Kwang-Jae Kim[a], Sung H. Han[a], Myung Hwan Yun[a] & Jiyoung Kwahk[a]
[a] Division of Mechanical and Industrial Engineering, Pohang University of Science of Technology, Korea
Published online: 08 Jan 2015.

INTERNATIONAL JOURNAL OF OCCUPATIONAL
SAFETY AND ERGONOMICS (*JOSE*) 2002, VOL. 8, NO. 3, 387–406

A Systematic Procedure for Modeling Usability Based on Product Design Variables: A Case Study in Audiovisual Consumer Electronic Products

Kwang-Jae Kim
Sung H. Han
Myung Hwan Yun
Jiyoung Kwahk

Division of Mechanical and Industrial Engineering,
Pohang University of Science of Technology, Korea

A systematic modeling approach to describing, prescribing, and predicting usability of a product has been presented. Given the evaluation results of the usability dimension (UD) and the measurement of the product's design variables, referred to as the human interface elements (HIEs), the approach enables one to systematically assess the relationship between the UD and HIEs. The assessed relationship is called a usability model. Once built, such a usability model can relate, in a quantitative manner, the HIEs directly to the UDs, and thus can serve as an effective aid to designers by evaluating and predicting the usability of an existing or hypothetical product. A usability model for elegance of audiovisual consumer electronic products has been demonstrated.

usability evaluation human interface elements
consumer electronic products

This research was supported in part by a research fund from the Ministry of Science and Technology (through STEPI) and the Ministry of Education (BK21 Project) of Korea.

Correspondence and requests for offprints should be sent to Kwang-Jae Kim, Division of Mechanical and Industrial Engineering, Pohang University of Science of Technology, San 31, Hyoja-Dong, Nam-Gu, Pohang, Kyungbuk, 790-784, Korea. E-mail: <kjk@postech.ac.kr>.

1. INTRODUCTION

Today usability is becoming increasingly more important as a factor determining the overall quality and thus the success probability of a product (Jordan, 1998a; Nagamachi, 1995; Nielsen, 1993). This is particularly true for the consumer product industry (Butters & Dixon, 1998), where the buyers are the end-users. Function-oriented products are losing their attractiveness because consumers look for products providing high usability, with basic functions taken for granted. Manufacturers ought to consider and build in usability in their new product development in order to be competitive in the market (Jordan, 1998b; Rubin, 1994).

The concept of usability in this study embraces the degree of ease of use as well as that of subjective satisfaction. Ease of use, which has traditionally been considered as the major component of usability, refers to the system performance aspect such as task completion time, error rate, and learnability. On the other hand, subjective satisfaction refers to the human feeling evoked when users see, touch, or use the product, such as pleasure (Jordan, 1997) or attractiveness (Woodson, Tillman, & Tillman, 1992). In short, it represents product image and impression perceived by the users.

Techniques for evaluating usability have mainly been concerned with measuring the system performance (or objective performance) when the users perform intended tasks. Some examples include verbal protocol in which the users think out loud while conducting a task (Brinkman, 1993), direct observation in which the evaluator observes the users' behavior (Nielsen, 1993), design review in which designers and developers look for potential usability problems (Nielsen, 1993; Williges, Williges, & Elkerton, 1987), and questionnaires or interviews (Rogers, Gilbert, & Cabrera, 1997; Stanton & Young, 1998). Various attempts have been made to evaluate the subjective aspects of usability especially in the field of software design (Chin, Diehl, & Norman, 1988; Gelderman, 1998; Kalawsky, 1999; Keinonen, 1997). However, the main focus of such work has been limited to the traditional dimensions of usability such as effectiveness, learnability, flexibility, and attitude. Consequently, development of techniques for evaluating the level of subjective satisfaction related to image and impression aspects has scarcely been addressed (Han, Yun, Kwahk, & Hong, in press), although it is perceived to be increasingly more critical in today's marketplace.

Aiming to evaluate the subjective as well as objective aspects of usability in a systematic and quantitative manner, this study employs the notion of usability dimensions (UDs; see Han et al., in press; Kwahk, 1999).

UDs represent a set of dimensions by which users evaluate the overall usability of a product. Therefore, UDs, as a group, constitute the evaluation criteria of usability, which is abstract and complex in nature.

Once UDs are defined, one can design and conduct an experiment where participants evaluate the usability of a product (or products) on each of the UDs. The evaluation scores of UDs of a product are certainly affected by the characteristics of the product's design variables, which will be referred to as human interface elements (HIEs) throughout this paper. It is postulated that there exists a statistical relationship between the UDs and the HIEs of a product. This study aims to develop an approach to building such an empirical relationship, called a usability model, for audiovisual consumer electronic products such as video cassette recorders and CD players. Once built, a usability model can relate, in a quantitative manner, the HIEs (e.g., the number and size of control knobs) directly to the UDs (e.g., degree of elegance or luxuriousness; see Figure 1).

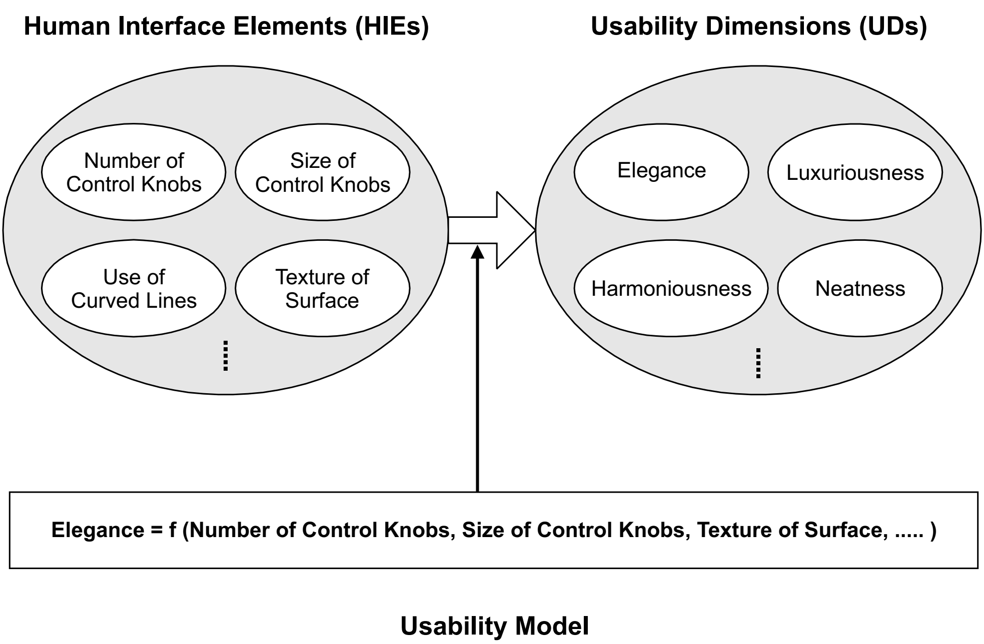


Figure 1. Concept of usability model relating human interface elements (HIEs) to usability dimensions (UDs).

A usability model serves three major purposes: (a) description, (b) prescription, and (c) prediction. For a given UD, the usability model enables one to answer such questions as: "Which HIEs have significant impact on the UD and in which direction (i.e., positive or negative) do they affect the UD?" (description); "Which HIEs should be adjusted in which direction and how much in order to change the users' evaluation score of the UD to a desired target level?'' (prescription); "How good would the evaluation score of the UD be in the market, given simple measurements of the HIE values of a product?'' (prediction) The three purposes of the usability model frequently overlap in practice.

Most existing evaluation methods can be applied only if there is a final product, or at least a prototype, when much time and money have already been spent. In short, the usability model can serve as an effective and efficient aid to designers by evaluating and predicting the usability of an existing or hypothetical product.

Whereas such a model should be applicable to either objective performance or subjective satisfaction-type UDs, its usefulness would be more significant for the latter case. This is because research on the evaluation and quantification of subjective satisfaction has received very little attention in the literature, both in the traditional human-computer interaction context and the consumer product industry.

Section 2 explains the definition, structure, and measurements of UDs and HIEs. Section 3 discusses the proposed modeling approach, which is applied to assess the elegance of audiovisual electronic products in section 4. Section 5 discusses directions for extending the proposed modeling approach. Finally, conclusions are given in section 6.

2. USABILITY DIMENSIONS AND HUMAN INTERFACE ELEMENTS

In order to derive an empirical model as proposed, its major components, namely UDs and HIEs, should be formally defined in the context of the product in question—consumer electronic products in this study. This section discusses the definition, structure, and measurement of UDs and HIEs for consumer electronic products.

2.1. UDs of Consumer Electronic Products

Product usability is not a simple and straightforward concept. It should be defined considering the unique characteristics of the product in question.

TABLE 1. UDs on Performance Dimensions

Group	Dimension	Definition
Perception--cognition	Directness	Degree of user's perception of directly controlling the objects represented by the product
	Explicitness	User's perception that the way the product looks and works is clear and accurate
	Modellessness	Capability that allows people to do whatever they want when necessary
	Observability	Ability to evaluate the internal state of the product based upon displayed information
	Responsiveness	Degree of presenting feedback information for user input in terms of speed
	Simplicity	The way the product looks and works is simple, plain, and uncomplicated
Learning--memorization	Consistency	Similarity in the way the product looks and works and the input-output behavior arising from similar situations or similar task objectives
	Familiarity	Extent to which the user's knowledge and experience in other domains or real world can be applied interacting with a new product
	Informativeness	Degree to which the product is informational and gives all the necessary information to the user in a proper manner
	Learnability	Time and effort required for the user to learn how to use a product
	Memorability	Degree to which the product is easy to remember
	Predictability	Ability of the user to determine the effect of future action based on past interaction experience
Control-action	Accessibility	Degree to which a product is easy to approach, enter, or operate
	Adaptability	Degree to which a product is changed easily to fit different users or conditions
	Controllability	Ability of the user to regulate, control, and operate the product
	Effectiveness	Accuracy and completeness with which specified users achieved specified goals in a particular environment
	Efficiency	Degree to which the product enables the tasks to be performed in a quick, effective, and economical manner or hinders performance
	Error prevention	Ability to prevent the user from making mistakes and errors
	Flexibility	Extent to which the product can accommodate changes to the tasks and environments beyond those first specified
	Helpfulness	User's perception that the product communicates in a helpful way and assists in the resolution of operational problems
	Multithreading	Ability of the product to support user interaction pertaining to more than one task at a time
	Recoverability	Ability of the user to take corrective actions once an error has been recognized
	Task conformance	Degree to which the product supports all of the tasks the user wishes to perform in such a way that the user easily understands them

Notes. Adapted from Han, Yun, Kwahk, and Hong (in press).

TABLE 2. UDs on Image-Impression Dimensions

Group	Dimension	Definition
Basic sense	Shape	Feeling about the shape of a product developed by integrated characteristics (ratio, length, area, etc.) of its components such as line and curvature
	Color	The conceptual image of a product developed by its color (e.g., warm, cool)
	Brightness	The image of a product developed by its brightness (e.g., dark, bright)
	Texture	The image of a product developed by its texture or touch (e.g., soft, coarse)
	Translucency	The image of a product developed by its translucency (e.g., opaque, translucent, transparent)
	Balance	Feeling that a product looks properly balanced or unbalanced
	Heaviness	Feeling that a product looks light or heavy
	Volume	Feeling that a product looks voluminous or slim
Description of image	Metaphoric design image	Image of a product expressed by the user using a simile or metaphor
	Elegance	Degree to which a product is graceful or stylish
	Granularity	Degree to which a product is worked out with great care and in fine detail
	Harmoniousness	Feeling that the components of a product are well-matched or in harmony
	Luxuriousness	Feeling that a product looks flashy, splendid, or extravagant
	Magnificence	Feeling that a product looks grand and spectacular
	Neatness	Feeling that a product looks clean, tidy, simple, and well-arranged
	Rigidity	Feeling that a product looks stout, stable, and secure
	Salience	Degree to which a product is outstanding, prominent, and catches one's eye
	Dynamicity	Feeling that a product looks dynamic or steady
Evaluative feeling	Acceptability	Degree to which the user feels a product is agreeable or acceptable
	Comfort	Degree to which the user feels easy and comfortable with a product
	Convenience	Feeling that a product is handy and suitable
	Reliability	Feeling that a product is dependable, fit to be trusted, or inspires confidence
	Attractiveness	Degree to which a product is pleasing, charming, and arouses interest
	Preference	Degree to which the user likes or dislikes a product over another one
	Satisfaction	Degree to which a product gives contentment or makes the user satisfied

Notes. Adapted from Han, Yun, Kwahk, and Hong (in press).

The concept of usability traditionally accepted in the human computer interaction area (Bennet, 1984; Shackel, 1984) is not equally applicable to other types of products like, for example, consumer electronic products.

This study adopts the concept of usability as consisting of two different types of dimensions, namely, performance dimensions and image-impression dimensions. Performance dimensions, which measure user performance, are further classified into three categories: perception-cognition, learning-memorization, and control-action. The perception-cognition category dimensions are used to examine how easy it is for users to perceive and interpret the interface of a product. The learning-memorization category dimensions explain how fast users can get used to the product and how well they can remember its functions. The control-action category dimensions explain users' control activity and its results. All together a total of 23 performance dimensions have been defined (Han et al., in press), and listed in Table 1.

Similarly, the image-impression dimensions are further classified into three categories: basic sense, description of image, and evaluative feeling-attitude. The basic sense category dimensions are related to the primitive image and impression of the product. The description category dimensions explain the image and impression of a product that the users would describe based on their experience with the product. The evaluative feeling-attitude category dimensions explain the attitude or judgmental feeling about the product. A total of 25 image-impression dimensions have been defined (Han et al., in press), and listed in Table 2.

2.2. HIEs of Consumer Electronic Products

As mentioned earlier, the HIEs refer to the characteristics of a product's design components with which users communicate when they see, touch, or operate the product (Kwahk, 1999). The HIEs of a typical audiovisual electronic product can be analyzed in four different categories: body-chassis-connection, control, display, and loading mechanism. These four categories have been further decomposed into 88 specific HIEs, which can be measured; 16 HIEs on body-chassis-connection, 29 HIEs on control, 23 HIEs on display, and 20 HIEs on loading mechanism. Four different types of scales were used to measure the HIEs in this study: rating (selecting a proper value among the rating scales provided); measurement (measuring the physical dimension of an HIE); category (selecting a proper category for an HIE); and binary choice (selecting *yes* or *no*). For a detailed description on the decomposition of HIEs, see Han, Yun, Kim, and Cho (1998). To measure HIEs in an efficient and systematic way, a measurement checklist covering all the 88 HIEs has been developed.

2.3. Experiments for Measuring UDs and HIEs

Thirty-six units of audiovisual (AV) electronic products were used in the experiment. Some examples include video cassette recorders, CD players, amplifiers, and boom boxes. The HIEs of each product were measured using the aforementioned measurement checklist. Five different individuals independently measured the products and the measurement results were cross-checked for consistency. It is notable that there were only negligible measurement variations even for the rating-type HIEs.

A total of 30 participants participated voluntarily in the experiment. The participants, students in Korea, consisted of 15 males and 15 females and ranged from 20 to 29 years in age. A rating scale based on the modified free-modulus method of the magnitude estimation technique (Han, Jung, Chung, Kwahk, & Park, 1998; Han, Song, & Kwahk, 1999) was used to indicate the degree of usability felt by the participants.

Our experimental design was a within-participant design, where each participant evaluated the usability of each product. Six participants performed the evaluation tasks simultaneously for the sake of efficiency but they were not allowed to interact. Upon completion of one UD, the next UD was assigned to each participant. The order of presentation of the products and the UDs was completely randomized to prevent any learning or transfer effect. As a result, each participant had a completely different order of evaluations.

3. USABILITY MODEL BUILDING PROCEDURE

As noted before, it is postulated that there exists a relationship between the usability scores and the HIE measures of a product, referred to as a usability model in this study. We seek to assess usability models using multivariate statistical techniques. This section discusses the usability model building approach. The major components of the approach are the following: (a) preprocessing the HIE data set, (b) evaluating the degree of relevancy of HIEs to the given UD, (c) choosing the form of the model to be fitted, and (d) fitting the usability model. Figure 2 briefly illustrates the model building procedure.

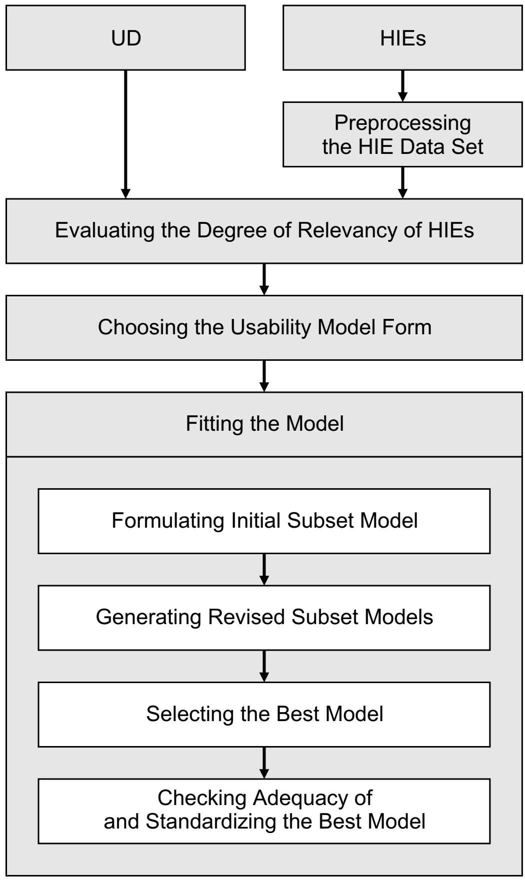


Figure 2. Usability model building framework. *Notes.* HIEs—human interface elements, UDs—usability dimensions.

3.1. Preprocessing the HIE Data Set

Some HIEs may have multiple measures or missing values depending upon the product and the characteristics of the HIEs. For example, for HIE indicating the size of controls, there may be multiple measures if there are multiple control buttons on a product. Because HIEs are to serve as the independent variables of the usability model, it is required that each HIE be represented by a single value. In this view, for the multiple measure case, the most proper one among the following four statistics—average, mode, maximum, or minimum—was selected in this study.

In the same token, the missing value of an HIE has to be replaced with some non-null value. In this study, either zero or the average HIE value of the other products was used when the frequency of missing values was low (e.g., less than half of the number of the products measured). When the frequency of missing values was high, the HIE was excluded in the following modeling procedure. In addition, the HIEs with a constant value (i.e., the measured values of certain HIEs were found to be identical across all the products used in the experiment) were excluded from the analysis. And finally, the values of HIEs were standardized (i.e., rescaled and normalized) for the sake of computational efficiency and precision.

3.2. Evaluating the Degree of Relevancy of HIEs

It is not necessary to include all the HIEs in the model because all of them are not equally important in explaining a specific UD. To reduce data handling effort and increase the efficiency of the model building procedure, a reduced set of HIEs for each UD needs to be determined. Typically, brainstorming or focus group discussion (or a combination of both) should be useful for this purpose. In this study, each HIE was evaluated and rated as one of the four levels—strongly relevant, moderately relevant, weakly relevant, or irrelevant—depending upon its degree of relevancy to the given UD. Irrelevant HIEs can be excluded in the following modeling procedure.

3.3. Choosing the Form of the Model

Given the major HIEs identified in the previous step, it is necessary to determine the order of the HIE terms and the interaction terms to be included in the usability modeling. In order to avoid undue complexity and derive meaningful insights, a second-order model with only the interaction items that are expected to have significant influence on the given UD is suggested. An interaction table (Coleman & Montgomery, 1993) or a priori interaction plots (Barton, 1997) may be used to select interaction items among HIEs that might affect the given UD. A traditional matrix analysis approach may also be employed, where each pair of HIEs is evaluated as to its potential degree of relevance to the UD.

3.4. Fitting Usability Model

The model fitting procedure consists of the following three steps: formulating initial subset model, generating revised subset models, and selecting the best model.

3.4.1. Formulating initial subset model

A multiple regression model is run with the evaluation scores of the given UD as the dependent variable and the relevant HIEs and the chosen interaction terms as the independent variables. Some of the model items that are not significant enough can be eliminated using the stepwise or backward elimination procedures. The result from the elimination procedure would contain only significant effects.

At this point, it is recommended to check the existence of multicollinearity. When there is a serious multicollinearity problem, it is known that the fitted model would show a poor prediction performance because of a high prediction variance. The variance inflation factor (VIF) can be used to detect the existence of the multicollinearity. When the highest VIF is greater than 10, the model is generally considered to have a multicollinearity problem (Myers, 1990). When this happens, the model item with the highest VIF is eliminated from the model and the model can be fitted again. This procedure is repeated until all of the VIFs fall below the threshold value of 10. The resulting model, which has the highest VIF of less than 10, is called the initial subset model.

3.4.2. Generating revised subset models

As the number of model items increases, the error degree of freedom decreases and, as a result, the model built results in a poor power in testing. A regression model is generally accepted as a reasonable one when the error degree of freedom is greater than that of the model items. Models with a smaller error degree of freedom may go through another elimination procedure in which the model item with the highest value of VIF is eliminated and the model is fitted again. This procedure can be repeated until some predetermined stopping rules are met. The stopping rules used in this study, as an example, are that all of the VIFs are less than 2 and the error degree of freedom is greater than that of the model items. There is no theoretical justification as to why such rules should be employed, yet those are the empirical rules that were found to work quite well in our case study (to be presented in section 4).

3.4.3. Selecting the best model

The revised subset models generated in the previous step are the candidate models, from which the best model is to be selected. Three different criteria are suggested to select the best model—the highest value of R^2, the lowest value of prediction sum of squares (PRESS), and the lowest p value. A model with the lowest p value tends to be the one satisfying the other two criteria as well. When the three criteria are not in agreement, the model that shows the best compromise with respect to the three criteria should be selected as the best model.

3.4.4. Checking adequacy of and standardizing the best model

The model selected as the best model should be checked for the model adequacy including the residual normality check. Finally, the best model can be standardized so that the model parameter estimates represent the importance-impact of the model items on a fair basis.

4. CASE STUDY: ELEGANCE OF AV SYSTEMS

This section discusses, in a step-by-step format, the usability model for elegance, one of the empirical models built in this study. Elegance is defined as the degree to which a product looks graceful or stylish.

4.1. Preprocessing the HIE Data Set

The measurement data of HIEs were preprocessed as discussed in section 3.1. For example, for products with no hidden controls, the “number of hidden controls” was set to zero; for products with no loading mechanisms, the “rounding and curved surface of loading mechanisms” was set at the average value of the others. In addition, “representation format of analogue displays” and “body shape” were excluded from the modeling process as the former had too many missing data and the latter was of a constant value.

4.2. Evaluating the Degree of Relevancy of HIEs

Through a brainstorming session, those HIEs that are believed to be relevant to each of the UDs were selected. For example, for the UD of “elegance of AV systems,” 52 HIEs were selected among the 88 HIEs, including 15 measurement type-HIEs (such as "size of the product body"), 23 rating type-HIEs (such as "degree of rounding and curved surface of the product body"), and 14 category type-HIEs (such as "color of the product body").

4.3. Choosing the Form of the Model

A second-order model was chosen as the form of the model to be fitted. Main effects, pure quadratic effects, and some selected interaction effects were included in the model. For the sake of maintaining interpretability and complexity of the model at a reasonable level, only two-way interaction terms of important HIEs were considered. Here, important HIEs refer to those HIEs that were deemed to be strongly relevant to elegance (see section 4.2).

4.4. Fitting Usability Model

The average of the 30 usability evaluation scores (resulting from 30 different participants) was used as the dependent measure of each product, whereas the measured HIE values were employed as the independent variable values in fitting the model.

4.4.1. Generating subset models

The subset models, that is, the initial subset model and the revised subset models, were generated following the procedure described in sections 3.4.1 and 3.4.2. The revised subset models are summarized in Table 3.

TABLE 3. Usability Model for Elegance: Revised Subset Models

Model	RSM 1[+]	RSM 2	RSM 3	RSM 4
R^2	**.9935**	.9899	.9872	.9861
Adjusted R^2	**.9880**	.9824	.9786	.9779
p value	**.0001**	.0001	.0001	.0001
Maximum VIF	**5.80**	3.82	3.22	1.95
PRESS	**127.6**	175.5	184.4	185.8
Number of parameters	**16**	15	14	13

Notes. RSM—revised subset model, +—RSM 1 is identical to the initial subset model, VIF—variance inflation factor, PRESS—prediction sum of squares.

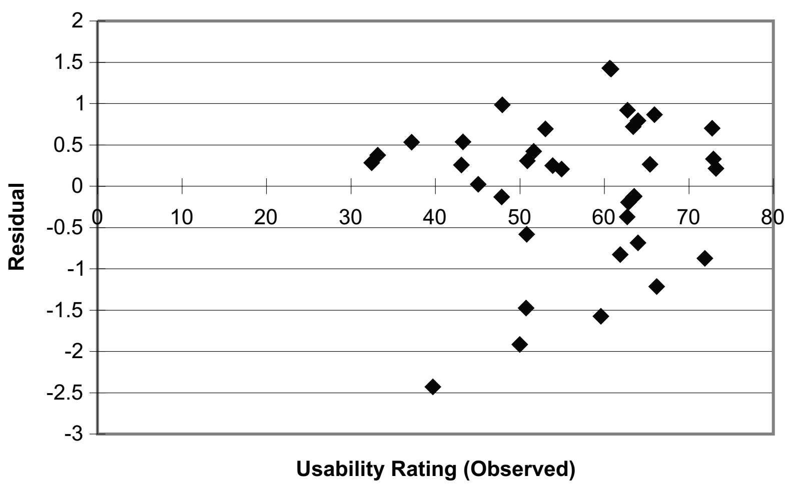
4.4.2. Selecting the best model

Among the four candidate models listed in Table 3, RSM 1 was selected as the best model because it has the highest R^2 (and adjusted R^2), the lowest p value, and the lowest PRESS value. It was a rather straightforward decision because all three criteria were in agreement, although such a situation does not always happen.

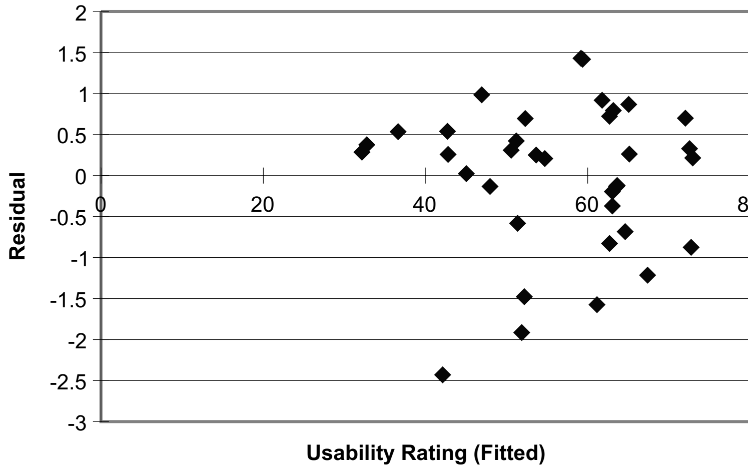
4.4.3. Checking adequacy of the best model

A model adequacy test was performed to check if the assumptions underlying the multiple regression are adequately satisfied by the chosen best model. Ideally, the modeling error should be a random error with mean zero and a constant variance. Additionally, it is usually assumed to follow a normal distribution for hypothesis testing and confidence interval estimation.

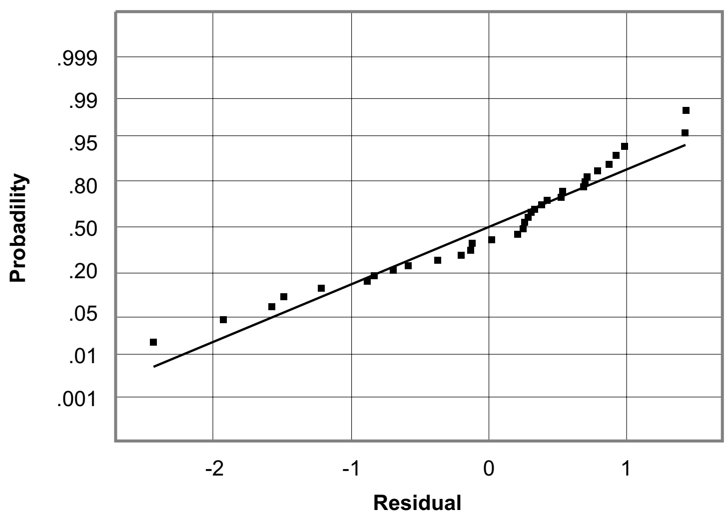
A check of the constant variance assumption was made by plotting the residuals. The residuals were plotted against the observed elegance evaluation scores (Figure 3a) and the fitted elegance evaluation scores (Figure 3b). The plots do not exhibit any systematic pattern as to the size of the variance of the residuals.



(a) Residuals Versus Observed Values



(b) Residuals Versus Fitted Values



(c) Normal Probability Plot of Residuals

Average: -0.0002778
SD: 0.901236
N: 36

W test for normality
R: .9673
p value (approximate): .0451

Figure 3. Usability model for elegance: Plot of residuals.

A normal probability plot of the residuals was constructed and displayed in Figure 3c. The general impression from examining this display is that the error distribution may be slightly skewed, with the left tail being longer than the right. This plot is not grossly non-normal, however, as manifested by the low *p* value (.0445) of the Shapiro-Wilk test (Montgomery, 1997). Overall, the chosen best model is considered to successfully pass the model adequacy test. (The standardization of the model parameters is mentioned next in section 4.5.)

4.5. Best Model: Result and Interpretation

The chosen best model for elegance is given in Table 4. Estimated model parameters, both non-standardized and standardized, are given. Model parameters are listed in the order of their standardized parameter estimates (in absolute values). The texture of surface is considered the most significant factor. More specifically, it has a positive quadratic effect on the elegance dimension. This implies that as the texture of surface changes from rubbery to plastic and to metallic, the elegance rating improves rapidly in a quadratic fashion. The quadratic effect of the use of curved lines and the interaction effect between glossiness of controls and brightness of surface color are identified as the next most significant factors. The best model was discussed with a group of design practitioners of a consumer electronic company, which provided some of the products evaluated in this study.

TABLE 4. Usability Model for Elegance: Parameter Estimates

	Parameter Estimates	
Model Item	**Standardized**	**Non-Standardized**
(Texture of Surface)^2	0.672	3.058
(Use of Curved Lines)^2	–0.395	–2.281
(Glossiness of Controls) × (Brightness of Surface Color)	0.294	3.268
(Sound of Loading Mechanism) × (Density of Display Items)	–0.210	–2.978
(Sound of Loading Mechanism) × (Use of Curved Line)	0.174	1.764
(Sound of Loading Mechanism) × (Glossiness of Controls)	–0.164	–3.169
(Saturation of Surface Color)^2	0.155	2.403
(Loading Time)^2	0.149	0.118
(Speed of Loading Mechanism) × (Texture of Surface)	–0.144	–1.489
Brightness of Surfaces Color	–0.143	–1.161
Texture of Controls	0.131	1.063
(Number of Indicator Controls)^2	0.116	0.135
(Saturation of Surface Color) × (Number of Controls)	–0.111	–0.187
Stiffness of Control Touch	–0.087	–1.734
(Sound of Controls) × (Density of Display Items)	0.085	1.083
(Brightness of Surfaces Color)^2	–0.082	–0.565
Intercept	0.000	50.616

They agreed that the modeling result conformed fairly well to their experience and insights. The real validity of the model, however, would have to be justified eventually using market data.

4.6. Notes on Modeling Performance

A total of 33 usability models were built to describe the relationships between the usability evaluation scores and the HIEs for 33 UDs (out of the 48 UDs mentioned in section 2.1). The remaining 15 UDs were not considered in the model building process because they were not appropriate to model with the evaluation data collected in this study. The R^2 values of the fitted models ranged from .40 to .99 (average .84 with a standard deviation of .12), which are considered remarkably successful results. The precision of the interval estimates was also satisfactory. As an example, the average width of the 95% confidence intervals for the 36 observations (i.e., the 36 AV products tested) of the best model for elegance discussed earlier is only 6.20 (on a scale of 100). This is another strong indication that the model building procedure worked very well in our case study.

5. DISCUSSION

It should be noted that socioeconomic factors such as cultural, ethnic, and educational background would have a significant impact on the perception or subjective satisfaction of a given UD (e.g., elegance in this study). In principle, the derived usability model is valid only for the market segment having the same socioeconomic background of the participants of the experiment. The applicability of the usability model should be determined accordingly.

The modeling approach proposed in this study can be extended along the following directions. First, our modeling approach requires that for a given UD, irrelevant HIEs be identified and excluded from the modeling procedure. Such a subjective evaluation necessitates a significant involvement of usability specialists with expertise in the product domain, which might be hard to come by in practice. A possible solution to alleviate this difficulty is to develop a new variable screening method that may be performed without the help of usability experts. Multivariate statistical techniques (such as partial least squares, principal component analysis, and factor analysis) may be considered for this purpose.

Secondly, among the four types of HIEs (i.e., measurement, rating, category, and binary choice), category type HIEs were not considered in the modeling procedure proposed in this study. This is mainly due to the increased data requirement as each category type HIE generates a number of dummy variables. Thus it would be of value to develop a method that can handle category type HIEs in the model without significantly increasing the data requirement or complexity of the modeling procedure. Neural networks combined with a genetic algorithm would be a promising avenue.

Thirdly, the purpose of the models built in this study is to predict the average score of the UDs. However, in order to investigate the diversity of the users' attitude toward a product, one should consider the models to predict the degree of variation of the UD scores as well as the "average models." Furthermore, a combination of these two types of models can be used in suggesting optimal design alternatives to achieve the designer's aspiration levels on usability.

6. CONCLUSIONS

An innovative and systematic modeling approach to describing, prescribing, and predicting product usability has been presented. The usability evaluation approach developed and illustrated may be applied to a wide variety of product designs or evaluations with only minor modifications to a specific product domain. Currently, the suggested modeling approach is being applied to other product domains including office chairs (Yun, Han, Ryu, & Yoo, 2001) and mobile phones (Han, Yun, Kim, Hong, & Kim, 2001). The empirical models built in this study can help identify important design variables and also predict the level of usability of a given AV system.

The contribution of this research is quite significant from the methodological as well as pragmatic perspectives. The usability modeling procedure can serve as an aid for new product development. The approach suggested in this study provides a different point of view in usability engineering. It attempts to model the relationship between the design variables and the level of product usability, whereas traditional ones have focused on finding usability problems. That is, it is an initial attempt to understand the underlying process of human perception of product designs.

Usability models can be used to predict the usability of new products as well as to evaluate that of existing products. Moreover, the models, if combined with a proper optimization scheme, can be employed to generate design features that would yield given usability aspirations. In sum, usability models can accelerate the adoption of the usability concept and method by various organizations in practice.

REFERENCES

Barton, R. (1997). Pre-experiment planning for designed experiments: Graphical methods. *Journal of Quality Technology, 29*(3), 307–316.

Bennet, J. (1984). Managing to meet usability requirements: Establishing and meeting software development goals. In J. Bennet, D. Case, J. Sandelin, & M. Smith (Eds.), *Visual display terminals* (pp. 161–184). Englewood Cliffs, NJ, USA: Prentice-Hall.

Brinkman, J.A. (1993). Verbal protocol accuracy in fault diagnosis. *Ergonomics, 36*(11), 1381–1397.

Butters, L., & Dixon, R.T. (1998). Ergonomics in consumer product evaluation: An evolving process. *Applied Ergonomics, 29*(1), 55–58.

Chin, J.P., Diehl, V.A., & Norman, K.L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. In *Proceedings of CHI '88* (pp. 213–218). New York, NY, USA: ACM Press.

Coleman, D., & Montgomery, D. (1993). A systematic approach to planning for a designed experiment. *Technometrics, 35*, 1–27.

Gelderman, M. (1998). The relation between user satisfaction on information systems and performance. *Information and Management, 34*, 11–18.

Han, S.H., Jung, E.S. Chung, M., Kwahk, J., & Park, S. (1998). Psychophysical methods and passenger preferences of interior designs. *Applied Ergonomics, 29*(6), 499–506.

Han, S.H., Song, M., & Kwahk, J. (1999). A systematic method for analyzing magnitude estimation data. *International Journal of Industrial Ergonomics, 23*, 513–524.

Han, S.H., Yun, M.H., Kim, K., & Cho S. (1998). *Development of a usability evaluation method.* Unpublished technical report No. 97-G-17-01-A-15, Ministry of Science and Technology, Korea.

Han, S.H., Yun, M.H., Kim, K., Hong, S., & Kim, J. (2001). *Identifying mobile phone design features critical to user satisfaction.* Manuscript submitted for publication.

Han, S.H., Yun, M.H., Kwahk, J., & Hong, S.W. (in press). Usability of consumer electronic products. *International Journal of Industrial Ergonomics.*

Jordan, P.W. (1997). The four pleasures: Taking human factors beyond usability. In P. Seppälä, T. Luopajärvi, C.-H. Nygård, & M. Mattila (Eds.), *Proceedings of the 13th Triennial Congress of the International Ergonomics Association, Tampere, Finland, 1997* (Vol. 2, pp. 633–635). Helsinki, Finland: Finnish Institute of Occupational Health.

Jordan, P.W. (1998a). *An introduction to usability*. London, UK: Taylor & Francis.

Jordan, P.W. (1998b). Human factors for pleasure in product use. *Applied Ergonomics, 29*(1), 25–33.

Kalawsky, R.S. (1999). VRUSE—A computerized diagnostic tool for usability evaluation of virtual/synthetic environment systems. *Applied Ergonomics, 30*, 11–25.

Keinonen, T. (1997). Expected usability and product preference. In *Proceedings of the*

Conference on Designing Interactive Systems (pp. 197–204). New York, NY, USA: ACM Press.

Kwahk, J. (1999). *A methodology for evaluating the usability of audiovisual consumer electronic products.* Unpublished doctoral dissertation, Pohang University of Science and Technology, Pohang, Republic of Korea.

Montgomery, D. (1997). *Design and analysis of experiments* (4th ed.). New York, NY, USA: Wiley.

Myers, R.H. (1990). *Classical and modern regression with application* (2nd ed.). Boston, MA, USA: PWS-KENT.

Nagamachi, M. (1995). Kansei engineering: A new ergonomic consumer-oriented technology for product development. *International Journal of Industrial Ergonomics, 15*, 3–11.

Nielsen, J. (1993). *Usability engineering*. New York, NY, USA: AP Professional.

Rogers, W., Gilbert, D., & Cabrera, E. (1997). An analysis of automatic teller machine usage by older adults: A structured interview approach. *Applied Ergonomics, 28*(3), 173–180.

Rubin, J. (1994). *Handbook of usability testing.* New York, NY, USA: Wiley.

Shackel, B. (1984). The concept of usability. In J. Bennet, D. Case, J. Sandelin, & M. Smith (Eds.), *Visual display terminals* (pp. 45–87). Englewood Cliffs, NJ, USA: Prentice-Hall.

Stanton, N., & Young, M. (1998). Is utility in the mind of the beholder?—A study of ergonomics methods. *Applied Ergonomics*, *29*(1), 41–54.

Williges, R.C., Williges, B.H., & Elkerton, J. (1987). Software interface design. In G. Salvendy (Ed.), *Handbook of human factors* (pp. 1416–1449). New York, NY, USA: Wiley.

Woodson, B.T., Tillman, B., & Tillman, P. (1992). *Human factors design handbook: Information and guidelines for the design of systems, facilities, equipment, and products for human use* (2nd ed.). New York, NY, USA: McGraw-Hill.

Yun, M.H., Han, S.H., Ryu, T., & Yoo, K. (2001). Determination of critical design variables based on the characteristics of product image/impression: Case study of office chair design. In *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting* (pp. 712–716) Santa Monica, CA, USA: Human Factors and Ergonomics Society.