The Development of a Formative and a Reflective Scale for the Assessment of On-line Store Usability

Timo CHRISTOPHERSEN
Department of Business, University of Kiel, Germany

and

Udo KONRADT
Department of Psychology, University of Kiel, Germany

ABSTRACT

In usability research, difference between formative and reflective measurement models for the assessment of latent variables has been ignored largely. As a consequence, many usability scales are misspecified. This might result in reduced scale validity because of the elimination of important usability facets within the procedure of scale development. The aim of the current study was to develop a questionnaire for the evaluation of On-line store usability (UFOS-V2) that includes both a formative and a reflective scale. 378 subjects participated in a laboratory experimental study. Each participant visited two out of 35 On-line stores. The usability and intention to buy was assessed for both stores. In addition, actual purchase behaviour was observed by combining the subjects' reward with the decision to buy. In a two-construct PLS structural equation model the formative usability scale was used as a predictor for the reflective usability measure. Results indicate that the formative usability scale UFOS-V2f forms a valid set of items for the user-based assessment of online store usability. The reflective usability scale shows high internal consistency. Positive relationships to intention and decision to buy confirm high scale validity.

Keywords: Usability, E-Commerce, On-line store, Scale Development, Evaluation Questionnaire, Formative and Reflective Measurement Models

1. INTRODUCTION

In part 11 of the international norm ISO 9241 [1], the usability of a software product is defined by the three criteria efficiency, effectiveness, and satisfaction. Usually, scholars also draw on the principles of dialogue design in part 110 of the same norm [2] which are suitability for the task, self-descriptiveness, controllability, conformity with user expectations, error tolerance, suitability for individualization, and suitability for learning. As stated in the norm, evaluation of usability always has to be referred to the individual user, task, and context of usage.

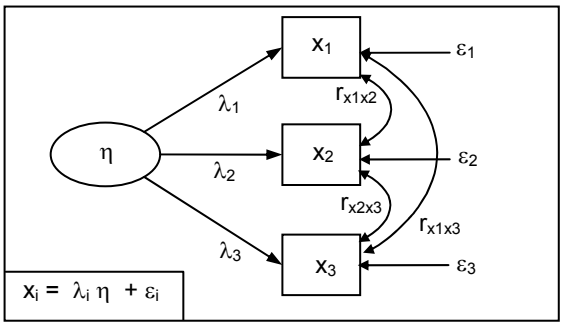
In general, investments in the improvements of the user interface will give a good return [3]. Several empirical studies suggest that usability is a critical success driver of E-commerce sites [4]. Usability problems on sites may be so serve that the customers are simply not able to accomplish the On-line purchase although they originally intended to [5]. Also, less important design flaws might have a negative impact on success because the On-line transaction becomes more inconvenient and more time-consuming [6]. The *intention to buy* turns out to be positively influenced by usability and related constructs *perceived ease of use* and *perceived usefulness* [4, 7, 8]. Other studies show positive relationships between the usability and the *intention to revisit* the On-line store [9]. Few authors have investigated the impact of usability on actual purchases [4, 9, 10].

2. THEORETICAL BACKGROUND

Formative and reflective measurement models

For the assessment of latent constructs, two kinds of measurement models might be applied which differ in the underlying assumption of causal relationship between the latent variable (LV) and its manifest indicators [11].

Figure 1: Reflective Measurement Model of a Latent Variable with three Indicators

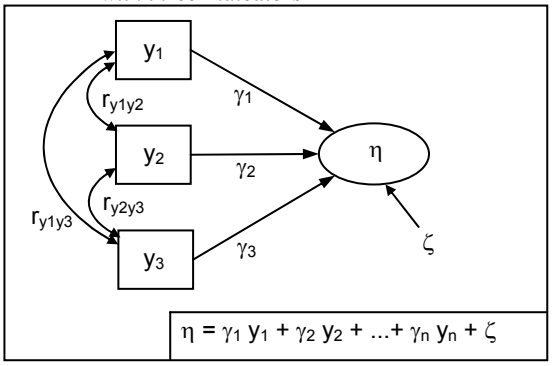


η: latent variable; λ: loading; x: reflective indicator; ε: measurement error on level of indicators; r: correlation between indicators

Traditional scale development usually draws on reflective measurement models where the observed indicators are assumed to be caused by the LV [12]. Figure 1 shows a LV that is assessed by three reflective indicators. According to the causal relationship, value changes in the LV result in changing values of all reflective indicators [13]. Hence, high correlations between the indicators are expected and can be interpreted as a criterion for high internal consistency. In the sense of domain sampling, reflective indicators are interchangeable. Thus, leaving out one specific indicator of a reflective scale won't result in alternations in its sense of content [14].

For a *formative* measurement model, the opposite direction of causal relationship between the LV and the manifest indicators is assumed. In this case, the indicators cause the LV (see figure 2) [15]. High correlations between formative indicators might occur, but are not generally expected. Hence, factor analysis and Cronbachs Alpha are inappropriate methods of evaluating a formative measure [16]. Any elimination of relevant formative indicators will result in reductions of scale validity.

Figure 2: Formative Measurement Model of a Latent Variable with three Indicators



η: latent variable; γ: weight; y: formative indicator;
ζ: measurement error on level of the latent variable;
r: correlation between indicators

As an example for a formative construct, consider social status [15]. This LV arises from several different individual characteristics of a person like occupation and ethnic group membership, religion, gender, voluntary associations, hobbies, and several other formative facets [17]. Status inconsistency exists when correlations between status criteria are low, e.g. in case of an unemployed academic. Hence, high correlations between formative indicators are possible but not generally expected. In the context of human computer interaction research it has to be noted that the term *formative* does not refer to the difference between summative and formative evaluation methods.

There are two major problems of dealing with formative LVs statistically. Firstly, the formative measurement approach does not allow estimating the parameters of a formative model within a structural equation model without linking the LV to at least one other LV [15]. Secondly, the estimates are biased if a critical degree of multicollinearity between the formative indicators exists [18]. The best way to handle this problem is to aggregate the indicators that show high multicollinearity into an index [19]. Furthermore, it is important to notice that some variables may be assessed by both a reflective and a formative measurement model [20]. For example, a scale on customer satisfaction may include a certain number of formative indicators, each covering a different aspect of the construct (e.g. satisfaction with service hot lines, service personal, product quality etc.) [21]. On the other hand, a reflective scale would include items on customers' global judgement of satisfaction. For constructs like this, it is possible to validate the formative scale by means of a two-construct model [15]. In this kind of model, the formative LV is used as a predictor for the reflective LV [20].

Studies show that in case of many scales, the measurement model unintentionally has been specified as being reflective although the construct affords a formative model [22]. As a potential consequence of this kind of misspecification, scale validity might be reduced. Invalid scales entail the danger of false interpretations of statistical results. In addition, results of structural equation models (SEM) may be strongly biased if the measurement model of one LV is misspecified [22, 23].

Misspecifications of Usability Scales

To our knowledge, no published study so far has attempted to develop a usability scale with respect to the difference between formative and reflective measurement models for the assessment of LVs. Consequently, it may be suspected that many scales are misspecified in terms of the measurement model.

Exemplarily, we illustrate the misspecification of an usability questionnaire by means of the UFOS-V1 [4] which is a questionnaire for the usability assessment of On-line stores from the customer's point of view. This instrument includes several items which have to be considered as being formative usability indicators, e.g. 'I can easily access the search functions of this store' and 'The shopping cart is clearly arranged'. Obviously, high correlations between both indicators *may* occur but are not generally expected. For example it may be given that the shopping cart of a store is arranged in a clear way, but at the same time the accessibility of its search functions is limited. Thus, correlations between these two indicators would turn out to be low. UFOS-V1 [4] also includes reflective usability indicators, e.g. 'The purchase can be performed quickly'. This item reflects the overall level of subjective usability and can be considered as a global judgement on the part of the customer.

Although both formative and reflective indicators were included in the item pool of UFOS-V1 [4], exploratory factor analysis was applied in order to identify different dimensions of On-line store usability. According to the general procedure of scale purification [12] several indicators were eliminated in order to optimize Cronbachs Alpha. For example the highly relevant formative indicator 'The screen text is hardly readable' was

eliminated. As outlined above, the application of this procedure of scale development which is based on classical test theory is not appropriate for formative measures. As a consequence of this misspecification, several aspects of On-line store usability have been excluded from the scale. This might have caused a reduction of content validity of UFOS-V1 [4].
A thorough analysis of the adequate measurement model reveals that almost all usability scales lack of an adequate specification of the underlying measurement model. As in case of UFOS-V1 [4], many usability scales include both formative and reflective indicators without considering this methodologically. Rare examples for adequate specifications of the measurement model are the SUS questionnaire [24], and the first subscale of the QUIS [25].

3. RESEARCH QUESTIONS AND HYPOTHESES

The aim of the study was to further develop the usability questionnaire for On-line stores UFOS-V1 [4] by constructing both a formative (UFOS-V2f) and a reflective scale (UFOS-V2r). Our hypotheses draw on the value criteria of both scales including their relationships to the *intention to buy* and the *decision to buy*.

4. METHOD

Scale Development and Measures

Initial point for the development of both new UFOS-V2 scales was the item pool of the UFOS-V1 questionnaire [4]. Based on a pre-test with a group of four experts in scale development, it was determined whether a formative or a reflective measurement model applied for each item. In several cases, item formulations were improved. Several items from the pool were eliminated because of high degrees of redundancy to other items or too high amounts of missing data. In order to assure completeness of the item pool, recent literature on aspects of E-Commerce user experience was reviewed [amongst others 26-28]. Based on this review eight additional items were added reflecting usability aspects concerning the ordering process, product lists, and the availability of relevant information.

The resulting questionnaire consisted of 9 reflective items for the scale UFOS-V2r (e.g., item ur1 ‘The purchase can be completed quickly’), and 58 items for the formative scale UFOS-V2f (e.g., item uf25 ‘I can easily access the search functions of this On-line store’). A 7-point Likert-type answering scale ranging from 1 (‘fully disagree) to 7 (‘fully agree’) was used. In addition, the option ‘not applicable’ was offered for each item on order to avoid potential uniformed response errors [29].

The *intention to buy* was measured with three items (e.g., ‘It is likely that I will purchase from this store again within the next three month’) and answered on a 7-point Likert-type scale.

Sample

378 users participated in a laboratory experimental study. Participants were recruited via placards on notice-boards and flyers in public facilities, through local internet portals, and advertisements in the local paper. The sample included 44.3 % females, and 55.7 % males with an average age of 33.3 years (*SD* = 12.8; *Min* = 15; *Max* = 72). On-line shopping experience in terms of the number of past purchases strongly varied across the sample with an average of 38.9 (*SD* = 61.1).

Procedures

Prior to the experiment, subjects were asked to select one product of their own choice out of five product groups (CDs, DVDs, books, printer cartridges, or concert tickets). For each of these five groups of products, seven different On-line stores were included in the study, resulting in a total number of 35 On-line stores. Depending on the product choice, the subjects where assigned to an experimental condition which was determined by two of the seven stores. The experiment was conducted by using an On-line hypertext environment that included all instructions for the participants. During the whole procedure, an experimenter was present to help out in rarely occurring cases of questions or technical problems. The experiment was carried out either in sessions with single subjects or in groups with at most five participants.

In the first part of the experiment, information on the goals and procedure were given. The subjects were asked to fill out a questionnaire that included both demographic items (age, income, sex) and control variables (On-line shopping experience, speed of internet connection, product involvement). Then, the participants were told to visit the first store and to interact with the store by accomplishing three tasks. These tasks can be considered as typical scenarios of On-line store usage [30]. Thereafter, a questionnaire for store evaluation was presented which included the usability scales and items on additional aspects of the customer experience. Then, the same procedure of interaction by means of the three tasks and the identical questionnaire were applied for the second store.

After the evaluation of the second store, the shopping carts of both stores included one exemplar of the product that the subjects had chosen in advance. The participants had to decide whether they wanted to order this particular product from the first or second store or alternatively chose a voucher amounting to 15 Euro for a local Off-line store chain. This *decision to buy* was recorded automatically. After the subjects had placed the order, the experimenter rewarded them for their attendance by either handing out 15 Euro in cash or the voucher. Finally, additional control variables were assessed by means of a post-test questionnaire. In average, the whole experimental procedure took 65.5 minutes (SD = 17.9).

5. RESULTS

Ten cases that showed more than 20 % missing data were eliminated from the data set, resulting in a reduced sample size of $N = 368$. The multiple imputation technique was applied to impute the remaining missing data. For the following steps of analysis, each case of the data set was split into two store evaluations, resulting in a doubled sample size of $N = 736$.

The reflective usability scale UFOS-V2r showed a high internal consistency with a Cronbachs Alpha $\alpha = 0.93$. By eliminating one of the nine items, this value was slightly optimized to $\alpha = 0.94$ ($M = 3.63$; $SD = 1.54$). An ANOVA was performed to address the question whether the reflective scale UFOS-V2r was able to detect differences in the subjective usability between the 35 stores that were considered in this study. A significant result indicated sensitivity of the reflective usability scale ($F = 3.84$; $df = 34$; $p < .001$).

Eight out of the 58 formative usability items showed a high amount of more than 10 % missing data. These indicators were eliminated from the final UFOS-V2f scale because it was claimed that all items could be easily answered. In addition, it was analyzed whether any semantic redundancies could be identified within the scale. To avoid multicollinearity, four items were eliminated based on an inspection of the correlation coefficients and semantic similarities for each pair of formative indicators. After these eliminations, 46 items remained in the formative scale UFOS-V2f.

In order to investigate whether the set of 46 formative items adequately covered the overall impression of On-line store usability, a two-construct model was estimated using Partial Least Squares (PLS) path modelling. The sample size of $N = 736$ was sufficient, because the required number of cases for this PLS analysis is only 10 times the number of indicators in the formative LV [20]. Even after the elimination of four semantic similar indicators, a high degree of multicollinearity between the formative indicators was observed as indicated by low values of tolerance (< 0.4), high VIF-values (> 2), and a condition index of 50.96. In order to overcome the problem of multicollinearity indices were computed [19]. Therefore, we followed a methodological approach which is similar to the principal of main component regression [31]. A main component analysis with orthogonal Varimax rotation was computed. Based on the Eigenwert criterion, nine components were extracted. The factor scores of the0 nine components were used as formative indicators within the PLS two-construct model (ufindex1 to ufindex2). In this model which is illustrated in figure 4, the LV *formative usability* (scale UFOS-V2f) functions as a predictor for the dependent LV *reflective usability* (scale UFOS-V2r). The model was estimated using the software Smart-PLS (www.smartpls.de). A bootstrapping procedure was run in order to determine significance of the model estimates.

Table 1 shows the specification of the outer model. Eight of the formative indices showed both weights $\gamma > 0.1$ and significant t-values ($t > 1.66$). Only the weight for the ninth component ufindex9 was rather low with $\gamma = 0.03$ and insignificant ($t = 0.51$). All loadings of the eight reflective usability indicators were above the critical value of $\lambda > 0.707$ and showed significant t-values. The composite reliability of the reflective scale UFOS-V2r reached a high value of 0.95. The Average Variance Extracted was $AVE = 0.7$, which is above the critical value of 0.5. Overall, the measurement model can be evaluated as being excellent (cf. [20]), which is a precondition for the interpretation of the structural model.

Table 1: Specifications of the Outer Model for the Estimated PLS-Model

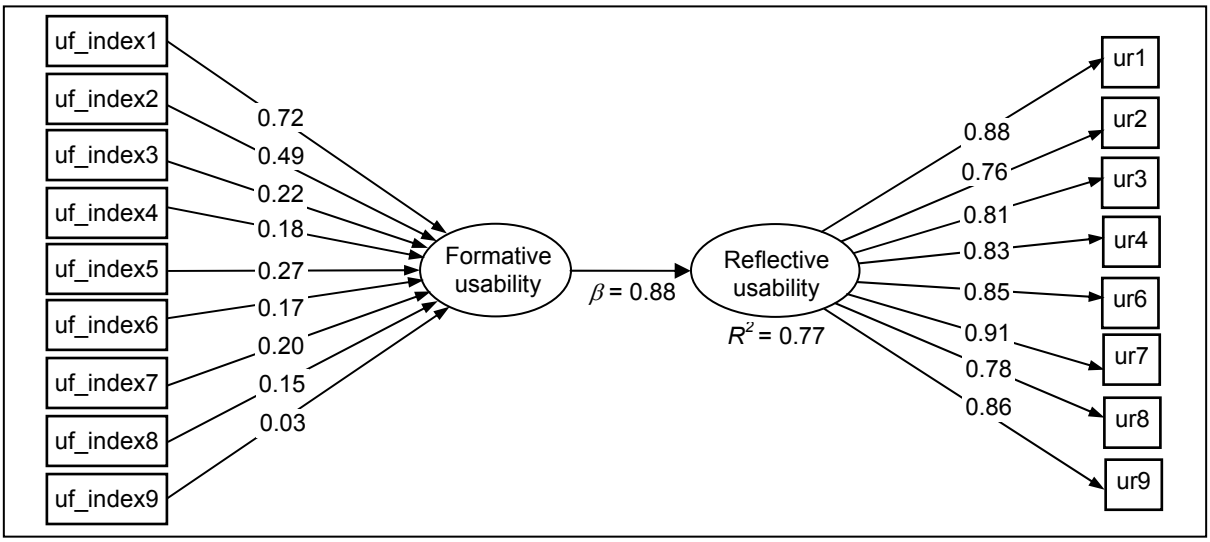
Latent variable	*Indicator*	*Weight/ loading*	*SE*	*t-Statistic*
Formative usability (UFOS-V2f scale)	ufindex1	0.72	0.06	11.51
	ufindex2	0.46	0.07	6.61
	ufindex3	0.22	0.05	4.16
	ufindex4	0.18	0.08	2.63
	ufindex5	0.27	0.08	3.51
	ufindex6	0.17	0.08	2.31
	ufindex7	0.20	0.06	3.17
	ufindex8	0.15	0.07	2.15
	ufindex9	0.03	0.06	0.51
Reflective usability (UFOS-V2r scale)	ur1	0.88	0.03	27.33
	ur2	0.76	0.06	13.02
	ur3	0.81	0.04	19.86
	ur4	0.83	0.03	26.92
	ur6	0.85	0.03	26.08
	ur7	0.91	0.02	39.83
	ur8	0.78	0.04	18.37
	ur9	0.86	0.03	32.50

Figure 3 illustrates the structural model of the PLS analysis. The path coefficient turned out to be very high ($\beta = 0.88$). Also, the determination coefficient showed a very high value of $R^2 = 0.77$. Results revealed a strong positive relationship between both newly developed usability scales. This leads to the conclusion that the set of 46 formative usability indicators (scale UFOS-V2f) very well covers the global judgement of usability which is assessed by means of the reflective scale UFOS-V2r.

In order to evaluate validity of both UFOS-V2 scales, correlations to the *intention to buy* ($\alpha = 0.91$) were observed. Coefficients with this dependent variable of $r = 0.62$ for the reflective scale UFOS-V2r and $r = 0.59$ for the formative sale UFOS-V2f indicate high concurrent criterion validity.

Furthermore, a sequential multinomial logistic regression (MLR) analysis [32] was applied to predict whether the subjects decided to purchase in the first store, in the second store or to choose a voucher. With a total number of 284 (77.17 %) most of

Figure 3: Two-Construct-Model for the Validation of the Formative Scale UFOS-V2f



the participants decided to place an order, while 84 (22.83 %) subjects chose the voucher. The MLR analysis was based on $N = 352$ cases, after 16 subject cases were identified and eliminated as being outliers. In the first step of analysis, the variables age, income, sex, On-line shopping experience, duration of the experiment, speed of internet connection, product involvement, and product group (CD, DVD, books etc.) were used as predictors. This model showed a Nagelkerke $R^2 = 0.21$ and a percentage of correctly classified cases $CCR = 59.1$ %. In the second step of the MLR, the latent variable scores of the reflective usability scale UFOS-V2r for both stores were added to the model. A Nagelkerke $R^2 = 0.49$ and $CCR = 67.64$ indicate high predictive validity of the newly developed. In an alter-native second step of analysis, the scores of the formative scale UFOS-V2f were added as predictors of the MLR instead of the UFOS-V2r scores. According to the results of this analysis, the predictive power of the formative usability scale is only slightly lower than for the reflective one (Nagelkerke $R^2 = 0.46$, $CCR = 64.72$ %).

6. DISCUSSION

The goal of our study was to develop and validate both a formative and a reflective scale for the assessment of perceived usability of On-line stores. Both scales demonstrated good value criteria. The reflective scale UFOS-V2r shows high internal consistency and is able to discriminate across different stores. A high postive relationship between the formative scale UFOS-V2f and the reflective scale UFOS-V2r indicated high content validity. The set of formative items very well covers the customer's global judgement on store usability. Furthermore, high concurrent and predictive criterion validity of both scales were demonstrated by positive relationships to the intention to buy and the decision to buy respectively.

As far as we know, all existing usability questionnaires to date have been developed without any reference to formative and reflective measurement models. Most usability scales include both formative and reflective indicators. Because of these misspecifications it seems unclear whether sufficient content validity of these scales is assured. In comparison to other usability questionnaire, the difference between both kinds of measurement models has been considered for the development of UFOS-V2. The reflective scale UFOS-V2r is comparatively short with only eight items. It offers the opportunity to economically assess the global usability impression. Hence, it is very well suited for purposes of summative evaluation like benchmark projects or rapid-prototyping procedures. The 46 items of the formative scale UFOS-V2f include all relevant single aspects of On-line store usability. This offers the possibility of screening a store for potential design flaws. By building item profiles, different store interfaces may be compared. Those aspects that are identified as being problematic may be addressed more thoroughly in further evaluation approaches like usability tests in order to deduce design recommendations for interface improvement.

The distinction between formative and reflective measurement has important implications for human-computer-interaction research. A theoretical foundation of usability seems inappropriate that is based upon a factor analytical breakdown of the construct. Amongst others, the dialogue design criteria in part 110 of the DIN 9241 [2] go back to an empirical study in which factor analysis was applied for the purpose of classifying different usability aspects [32].

7. REFERENCES

[1] DIN EN ISO 9241-11, **Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) - Part 11: Guidance on Usability**, Berlin: Beuth, 1999.

[2] DIN EN ISO 9241-110, **Ergonomics of Human-System Interaction - Part 110: Dialogue Principles**, Berlin: Beuth, 2006.

[3] R.G. Bias & D.J. Mayhew, **Cost Justifying Usability. An Update for the Internet Age**, San Francisco, CA: Morgan Kaufmann, 2005.

[4] U. Konradt, H. Wandke, B. Balazs & T. Christophersen, "Usability in Online Shops: Scale Construction and the Influence on the Buyers' Intention and Decision". **Behaviour & Information Technology**, Vol. 22, No. 3, 2003, pp. 165-174.

[5] J. Nielsen, R. Molich, C. Snyder & S. Farrell, **E-Commerce User Experience**, Fremont: Nielsen Norman Group, 2001.

[6] I. Ceaparu, J. Lazar, K. Bessiere, J. Robinson & B. Shneiderman, "Determining Causes and Severity of End-User Frustration", **International Journal of Human-Computer Interaction**, Vol. 17, No. 3, 2004, pp. 333-356.

[7] D. Gefen & D. Straub, "The Relative Importance of Perceived Ease of Use in IS Adoption: A study of E-Commerce Adoption", **Journal of the Association for Information Systems**, Vol. 1, No. 8, 2000, pp. 1-30.

[8] A. Muthitacharoen, P.C. Palvia, L.D. Brooks, B.C. Krishnan, R.F. Otondo & D. Retzlaff-Robert, "Reexamining Technology Acceptance in Online Task Behaviours", **Electronic Markets**, Vol. 16, No. 1, 2006, pp. 4-15.

[9] T.P. Liang & H.-J. Lai, "Effect of Store Design on Consumer Purchases: an Empirical Study of On-line Bookstores", **Information & Management**, Vol. 39, 2002, pp. 431-444.

[10] R. Henderson & M.J. Divett, "Perceived Usefulness, Ease of Use and Electronic Supermarket Use", **International Journal of Human Computer Studies**, Vol. 59, 2003, pp. 383-395.

[11] K.A. Bollen, **Structural Equations with Latent Variables**, New York: Wiley, 1989.

[12] G.A. Churchill Jr., "A Paradigm for Developing Better Measures of Marketing Constructs", **Journal of Marketing Research**, Vol. 16, 1979, pp. 64-73.

[13] C. Fornell & F.L. Bookstein, "Two Structural Equation Models: LISREL and PLS Applied to Consumer Exit-Voice Theory", **Journal of Marketing Research**, Vol. 19, 1982, pp. 440-452.

[14] J.C. Nunnally & I.H. Bernstein, **Psychometric Theory** (3rd ed.), New York: McGraw-Hill, 1994.

[15] A. Diamantopoulos & H.M. Winklhofer, "Index Construction with Formative Indicators: An Alternative to Scale Development", **Journal of Marketing Research**, Vol. 38, 2001, pp. 269-277.

[16] J.R. Rossiter, "The C-OAR-SE Procedure for Scale Development in Marketing", **International Journal of Research in Marketing**, Vol. 19, 2002, pp. 305-335.

[17] Hornung, C.A., "Social Status, Status Inconsistency and Psychological Stress", **American Sociological Review**, Vol. 42, No. 4, 1977, pp. 623-638.

[18] K.A. Bollen & R. Lennox, "Conventional Wisdom on Measurement: A Structural Equation Perspective", **Psychological Bulletin**, Vol. 110, 1991, pp. 305-314.

[19] S. Albers & L. Hildebrandt, "Methodische Probleme bei der Erfolgsfaktorenforschung: Messfehler, formative versus reflektive Indikatoren und die Wahl des Strukturgleichungs-Modells", **Zeitschrift für betriebswirtschaftliche Forschung**, Vol. 58, 2006, pp. 2-33.

[20] W.W. Chin, "The Partial Least Squares Approach for Structural Equation Modelling", in G.A. Marcoulides (Ed.), **Modern Methods for Business Research**, pp. 295-336, Hillsdale, NJ: Lawrence Erlbaum Associates, 1998.

[21] C. Fornell, M.D. Johnson, E.W. Anderson, J. Cha & B.E. Bryant, "The American Customer Satisfaction Index: Nature, Purpose, and Findings", **Journal of Marketing**, Vol. 60, No. 4, 1996, pp. 7-18.

[22] C.B. Jarvis, S.B. MacKenzie & P.M. Podsakoff, "A Critical Review of Construct Indicators and Measurement Model Misspecifications in Marketing and Consumer Research", **Journal of Consumer Research**, Vol. 30, 2003, pp. 199-218.

[23] S.B. MacKenzie, P.M. Podsakoff & C.B. Jarvis, "The Problem of Measurement Model Misspeci-fication in Behavioral and Organizational Research and Some Recommended Solutions", **Journal of Applied Psychology**, Vol. 90, 2005, pp. 710-730.

[24] J. Brooke, "SUS – a Quick and Dirty Usability Scale", in P.W. Jordan, B. Thomas, B.A. Weerdemeester & I.L. McClelland (Eds.), **Usability Evaluation in Industry**, pp. 189-194, London: Taylor & Francis, 1996.

[25] J.P. Chin, V.A. Diehl & K.L. Norman, "Development of an Instrument Measuring User Satisfaction of the Human-Computer Interface", in E. Soloway, D. Frye & S.B. Sheppard (Eds.), **CHI´88 Proceedings. Human Factors in Computing Systems**, pp. 213-218, New York: ACM Press, 1988.

[26] R. Kohli, S. Devaraj & A. Mahmood, "Understanding Determinants of Online Consumer Satisfaction: A Decision Process Perspective", **Journal of Management Information Systems**, Vol. 21, 2004, pp. 115-135.

[27] T. Lavie & N. Tractinsky, "Assessing Dimensions of Perceived Visual Aesthetics of Web Sites", **International Journal of Human-Computer Studies**, Vol. 60, 2004, pp. 269-298.

[28] M. Wolfinbarger & M.C. Gilly, "eTailQ: Dimensionalizing, Measuring and Predicting etail Quality", **Journal of Retailing**, Vol. 79, 2003, pp. 183-198.

[29] D.I. Hawkins & K.A. Coney, "Uniformed Response Error in Survey Research", **Journal of Marketing Research**, Vol. 18, No. 3, 1981, pp. 370-374.

[30] B. Detlor, S. Sproule & C. Gupta, "Pre-Purchase Online Information Seeking: Search versus Browse", **Journal of Electronic Commerce Research**, Vol. 4, No. 2, 2003, pp. 72-84.

[31] J. Cohen, P. Cohen, S.G. West & L.S. Aiken, **Applied Multiple Regression/Correlation Analysis for the Behavorial Sciences**, 3rd ed., Mahwah: Lawrence Erlbaum, 2003.

[32] W. Dzida, S. Herda & W.D. Itzfeldt, "User Perceived Quality of Interactive Systems", **IEEE Transactions on Software Engineering**, Vol. 4, No. 4, 1978, pp. 270-276.