A Conjoint Analysis of Attributes Affecting the Likelihood of Technology Use

Anna Elisabeth Pohlmeyer[1] and Lucïenne Blessing[2]

[1] University of Technology Berlin, Center of Human-Machine Systems, Franklinstr. 28/29, FR 2-6, 10587 Berlin, Germany
[2] University of Luxembourg, Engineering Design and Methodology, Campus Limpertsberg, 162A, avenue de la Faïencerie, L-1511 Luxembourg
anna.pohlmeyer@zmms.tu-berlin.de, lucienne.blessing@uni.lu

Abstract. Products are a composition of multiple attributes and should be evaluated in full-profile also in research contexts. Research studies on the use of technological products which only assess the importance of individual attributes do not reflect real-life scenarios of multi-attribute judgments and miss out on information that only becomes apparent in relative terms. In order to study the predictive value and relative importance of six attributes (functionality, cognitive ergonomics, classical ergonomics, quality, aesthetics, and emotional involvement) with respect to the likelihood of use, the method of conjoint analysis was borrowed from consumer research. The study was conducted with 104 participants divided in two groups of low and high self-competence. Group differences were only revealed when attributes were considered jointly, but not in single ratings. An intuitive interface, easy handling, and emotional involvement were greater motivators for the low competence group. Methodological implications are discussed.

Keywords: Technology Adoption, Self-Competence, Conjoint Analysis, Multi-Attribute Rating, Preferences.

1 Introduction

The consideration of what users want is central in user-centered design approaches. Also in marketing and product development, “the voice of the customer” has become an integral part of design specifications [1]. But what should be done if users want *everything*? If users are presented with a list of attributes such as good ergonomics or durable quality and asked how important these are, they might rate them all as very important. Without the necessity of trade-offs, independent ratings might lead to ceiling effects. But in everyday life, and also in product development as a special case, trade-offs are indispensable. Moreover, with too little variance in the responses, there is not enough information to differentiate preferences and requirements for different user groups. For example, both novices and experts might agree that ease of use is very important, but might prioritize this differently in relation to other attributes such as quality. Clearly, it is possible that different groups do agree on how important different attributes are. However, multi-attribute ratings are still necessary to validate

A. Marcus (Ed.): Design, User Experience, and Usability, Pt II, HCII 2011, LNCS 6770, pp. 303–312, 2011.

the finding. The additional value and source of variance that can be derived from multi-attribute ratings is of interest here. The methodological challenge addressed is whether we gain more insight by rating products as a whole or by collecting ratings of their parts?

Concerning the definition of attributes, we follow Grunert [2, p.229]: *"An attribute can be defined as any aspect of the product itself or its use that can be used to compare product alternatives."* The attributes we address were identified in a previous study in which participants documented real-life examples of interactive technology over the course of one week and elaborated why they liked or disliked the products. In addition, product-independent reasons were given about what motivates the use of technology and what hinders it, respectively. These statements were categorized according to qualitative content analysis procedures [3]. The set of attributes included in the present study includes: *functionality, ease of use, ergonomics, quality, aesthetics, and emotional involvement* (see Table 1). Note that "*ease of use*" refers to the cognitive, information-processing side of ergonomics, while "*ergonomics*" is explicitly related to classical, physical aspects of ergonomics such as size and weight [4].

One prominent question in HCI touches the field of technology adoption: how should a system be designed to increase the likelihood of usage? In the present study we focus on usage scenarios but not on purchase situations. A purchase context with a monetary investment depends on consumer resources [5] and influences decision-making processes in a way that is not of relevance here. For this reason, 'financial issues' were excluded from the original list of attributes. Also, we tried to control for 'usefulness' by concentrating on one product, a digital camera, and on one scenario, namely hobby photography. The set of attributes was used to describe different models of digital cameras by varying the attributes within two levels (see Table 1). Participants were asked how likely they would be to use each camera for hobby photography. As in real life, all attributes were *considered jointly*.

Relative weights can serve as useful indicators for differentiating user preferences and thus setting priorities in early phases of product development as recommended by Ulrich and Eppinger [6]. In these early phases, user preferences have to be anticipated and are not derived from evaluating a final prototype. For this reason, established questionnaires such as Hassenzahl's AttrakDiff [7] that evaluate existing products and prototypes are no option for weighting user preferences at this stage.

Conjoint analysis is generally used on a very detailed level of product features. Here, a more abstract level of product attributes has been applied that can be transferred to other systems in future work. Although conjoint analysis is commonly applied in consumer research, purchase behavior was not of interest. Instead, the method was modified to assess anticipated technology adoption by asking for the likelihood of usage rather than acceptable pricing or purchase intent.

The most prominent model of technology acceptance was introduced by Davis in 1989 [8]. In essence, the technology acceptance model (TAM) claims that system use is best predicted by the intention to use the system. This, in turn, is influenced by the perceived ease of use and perceived usefulness of the system. The model has gone through numerous iterative modifications and extensions and dominates the field. Unfortunately, its wide acceptance does not equate to an equal amount of knowledge accumulation [9]. The approach in this study will not propose yet another TAM-modification. Instead, a methodological approach will be introduced to the field of technology acceptance. The method is not new by any means, but, to our knowledge, has not been applied at such an abstraction level and not in the context of technology adoption (with the criterion "likelihood of use"). In general, conjoint analyses are rarely seen in HCI despite their evident potential. In traditional TAM-studies only one system is being evaluated. Conducting a conjoint analysis allows the consideration of multiple system alternatives in the assessment. The resulting utility scores can predict preferences even of systems that have not been shown to the participants.

Finally, group differences with respect to self-competence (closely related to self-efficacy) with technology will be analyzed. Self-efficacy refers to an individual's belief whether they are able or competent to perform a specific task. Computer self-efficacy (ability to competently use computers) beliefs have been shown to be a valuable antecedent of technology use [10]. On the other hand, computer self-efficacy is also linked to consequences such as the affect (or liking) concerning computer use. An objective of the present study is to compare groups with self-perceived high or low technology competence with respect to relative attribute importance regarding technology use. We expect that the low competence group emphasizes attributes concerning the understanding and handling of technology (ease of use and ergonomics) more than the high competence group. It is also of interest how the outlook of positive affect (emotional involvement/ joy of use) will influence the likelihood of usage in the two groups.

2 Methods

The method of choice was conjoint analysis, and a full-profile rating in particular. A full-profile is a description of a product alternative including all attributes of investigation. Based on this information, participants rate several product alternatives. From the overall judgments, part-worth utilities are derived through regressional modeling. Thus, instead of computing a composite score of *single* attribute ratings, the path is a de-compositional calculation of attribute weights. Attention is focused on the relevance of attributes, rather than on differences between product models themselves [11]. The impact that a variation of attribute levels has on the likelihood of use will be expressed in relative importance values.

2.1 Participants

The sample consisted of 104 participants all living in Berlin, Germany. Two age groups were recruited: 52 younger adults (20-30 years, M_{young}= 25.88, SD_{young}= 2.73) and 52 older adults (65-75 years, M_{old}= 67.9, SD_{old}= 2.38). Each age group had equal numbers of men and women. The sample was well-educated; 43.3% had a university degree, and 26% had an "Abitur" secondary school qualification as their highest qualification. Younger adults were recruited through an online database of study volunteers. Older adults were additionally recruited through an advertisement in a weekly newspaper. Only volunteers who had used a digital camera previously qualified as study participants.

The sample was split into two groups regarding perceived subjective competence. Subjective competence is a five-item subscale of a German 19-item-questionnaire of technology affinity [12]. A median split conducted over the entire sample was confounded by age (more older adults felt less competent, $\chi^2\,(1) = 9.85$, $p < .05$; and Spearman's $r\,(104) = -.3$, $p = .002$). Therefore, median splits were conducted within each age group. As a result, young and older adults were divided equally between the groups. The two groups did not differ with respect to age ($t\,(101.5) = .04$, $p > .05$). However, as intended, the two groups differed regarding perceived subjective competence ($t\,(94.5) = -11.48$, $p < .001$). Also, the older subsample of the high competent group had significantly higher competence values than the young subsample in the low competent group ($M_{young\_low} = 3.3$, $SD_{young\_low} = .56$; $M_{old\_high} = 3.93$, $SD_{old\_high} = .42$; $t\,(50) = -4.61$, $p < .001$). This serves as proof that the grouping is valid and no longer biased by age effects.

2.2 Material and Task

Verbal product descriptions of different digital camera models served as study material. Models differed on six two-level attributes (see Table 1): *functionality, ease of use, ergonomics, quality, aesthetics, and emotional involvement*. The authors tried to define the levels in a way that firstly, made ranges between level 1 and 2 comparable across factors, and secondly, excluded knock-out criteria. Participants received thorough instructions and examples regarding attribute levels, which are not apparent in Table 1. For more elaborate description of the attribute levels see [13].

Table 1. Attributes and levels

attribute	level 1	level 2
functionality	primary functions	secondary functions
ease of use	takes getting used to	intuitive to use
ergonomics	handling requires physical effort	easy handling
quality	prone to defects; poor performance	reliable + durable; excellent performance
aesthetics	average appearance	appealing appearance
emotional involvement	not engaging, just functional	pleasurably engaging in addition to functional

Six attributes, each at two levels, allow 64 possible model combinations (2^6). Presenting all possible profiles would have been very tiring for the participants and consequently also endangering the results' reliability. Therefore, a model selection was made in order to reduce the number of models presented to a manageable amount. An orthogonal fractional factorial design was created using SPSS ORTHOPLAN. Orthogonal designs try to represent the entire set of models in a best possible way. The final design consisted of 20 model combinations (or so-called *stimulus cards*). A stimulus card is a list of the six attributes with a unique level-combination, describing a specific 'camera profile' (see Figure 1). There was no definite 'good' or 'bad' product description: each description was good with respect to some, but inferior with respect to other attributes.

Participants were asked to indicate how likely it was (on an 11-point-Likert scale ranging from 0% to 100% in steps of 10) that they would use the described camera for the purpose of hobby photography (see Figure 1).

Additionally, participants were asked to rate the importance of each attribute individually on a 10-point Likert scale (1 = unimportant; 10 = very important). This is what is referred to as *single* attribute ratings. A detailed, written elaboration of the attributes and their levels was provided and could be checked whenever necessary.

Finally, the 10-item version of the Achievement Motives Scale was assessed. An inventory that captures the two factors *hope of success* and *fear of failure* (5 items each) with 4-point Likert scales [14].

Fig. 1. Stimulus material and conjoint rating sheet

2.3 Setup and Procedure

In single sessions, each participant was instructed individually, following a standardized protocol. An entire session, including questionnaires regarding demographic information, participants' background in technology use and computer literacy, as well as a Kano analysis of user satisfaction, lasted 60-90 minutes. Participants were reimbursed with €10/hour. The data was collected using paper-and-pencil questionnaires.

The order of the product descriptions (stimulus cards) was randomized by shuffling the cards before rating. Prior to the rating of the camera models, participants completed the questionnaire on technology affinity and were asked for a baseline measure of how likely they would use a digital camera for hobby photography in general (on the same 11-point Likert scale of usage likelihood as provided for the product ratings).

2.4 Data Analysis

Metric conjoint analysis was employed to evaluate so-called part-worth utilities (analogous to weight estimates) and relative importance scores. Competence group differences were modeled as interactions within the regression. This was realized by dummy coding the groups and attributes as demonstrated by Bloch et al [15]. With this approach, between-group comparisons of regression coefficients can be tested statistically. The relative importance of each attribute is computed by calculating relative ranges (utility range of the attribute divided by the sum of the part-worth utility ranges of all six attributes). These values are percentages and consequently sum to 100. Data was analyzed using SPSS 19 software. α-level was considered at .05.

3 Results

The overall regression model was significant ($F(13, 26) = 224.43$, $p < .001$) and explained 99.1% of variance.

3.1 Predictive Value of Attributes

All attributes except functionality were significant predictors of usage likelihood. As expected, the prospect of intuitive use ($\beta = .58$, $t = 22.07$, $p < .001$), easy handling ($\beta = .70$, $t = 6.94$, $p < .001$), high quality ($\beta = .56$, $t = 21.36$, $p < .001$), an appealing appearance ($\beta = .18$, $t = 7.0$, $p < .001$), and emotional involvement ($\beta = .17$, $t = 6.33$, $p < .001$) increased likelihood of usage. The non-significance of functionality can be explained by equal numbers of participants who preferred the camera to have only primary functions integrated, who preferred the addition of secondary functions, and who were indifferent about this matter. This was revealed by a Kano analysis (see [13]).

3.2 Self-perceived Competence Group Differences

The groups did not differ with respect to baseline likelihood of using a digital camera for hobby photography ($M_{low} = 51.15$, $SD_{low} = 4.31$; $M_{high} = 59.23$, $SD_{high} = 3.84$; $t(102) = -1.40$, $p > .05$), which allows the straightforward interpretation of competence group differences.

There were no significant competence group main effects regarding *single* attribute ratings ($F(1, 102) = .23$, $p > .05$), nor group x attribute interaction effects ($F(3.6, 364.4) = .64$, $p > .05$; degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity). In other words, evaluating each attribute individually led to comparable ratings of the high and low competence groups (see Figure 2 and Table 2).

Part-Worth Utilities. Significant interactions of attributes and competence grouping could be observed with respect to ease of use ($\beta = -.23$, $t = -7.34$, $p < .001$), ergonomics ($\beta = -.11$, $t = -3.58$, $p < .05$), and emotional involvement ($\beta = -.12$, $t = -3.65$, $p < .05$). The low competence group demonstrated higher part-worth utilities of these three attributes (see Table 2).

Relative Importance Rating. The range of the part-worth utility values for each attribute is divided by the sum of utility ranges of all attributes. This provides a measure of how important the attribute is to overall preference. The size of the *relative importance* for each attribute indicates the impact that a level variation of this attribute has on the likelihood of use.

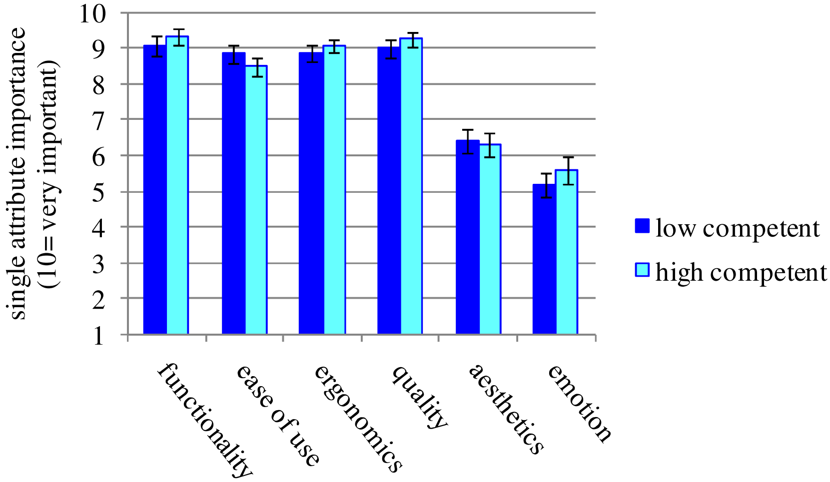

Fig. 2. Mean single attribute ratings (± *SE* mean) for low and high self-competence groups

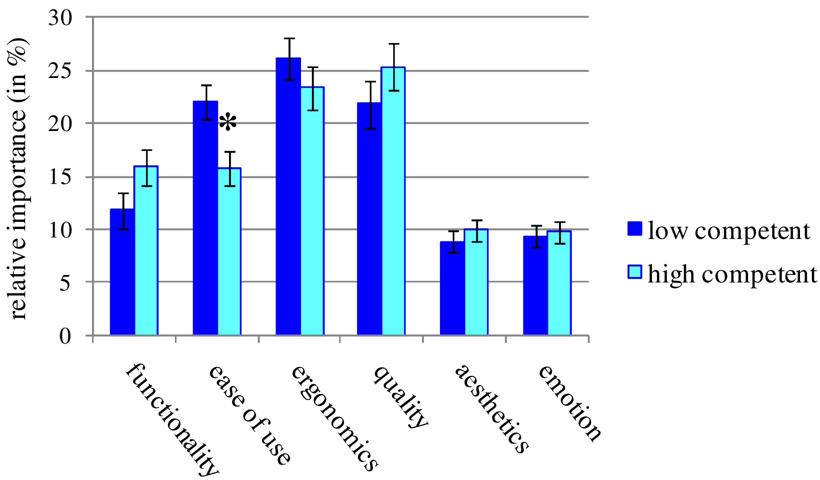

Fig. 3. Mean relative importance (± *SE* mean) for low and high self-competence groups

For the low competence group, (physical) ergonomics was the most important attribute while the high competence group's rating was mostly influenced by the degree of quality (see Figure 3 and Table 2).

Significant group differences were found for ease of use with a higher importance value for the low competence group (t (102) = 2.73, p< .05). However, no differences were found with respect to ergonomics (t (102) = .99, p> .05) nor emotional involvement (t (102) = -.30, p> .05)

Relative importance values can be interpreted in direct comparison because they are ratio scaled. For example, the variation of an intuitive usable interface in contrast to one that necessitates learning (ease of use) is twice as important as the variation of functionality (whether a device has only primary or additionally secondary features) for the low competence group (22.0 : 11.8). In contrast, the high competence group regards both attributes to be equally important (15.7 : 15.9).

Table 2. Overview of Single Importance Ratings, Part-Worth Utilities, and Relative Importance Values. Filled boxes indicate significant group differences.

attribute	group	single rating	part-worth utility	relative importance
functionality	low high	8.56 8.81	.927 .865	11.835 15.862
ease of use	low high	8.35 7.98	16.888 8.942	22.005 15.738
ergonomics	low high	8.35 8.56	20.619 16.750	26.104 23.329
quality	low high	8.50 8.75	16.350 17.442	21.831 25.300
aesthetics	low high	5.92 5.81	5.350 3.327	8.856 9.961
emotional involvement	low high	4.69 5.10	4.850 .904	9.368 9.809

Achievement Motives Scale. Participants in the low competence group showed significantly lower *hope of success* (M_{low}= 3.15, SD_{low}= .48; M_{high}= 3.33, SD_{high}= .56; t (101) = -2.13, p < .05) and a significantly *higher fear of failure* (M_{low}= 2.42, SD_{low}= .58; M_{high}= 2.12, SD_{high}= .60; t (101) = 2.62, p < .05).

4 Discussion

The likelihood of technology use, in this case a digital camera, depends on the combination of the product's attributes. There was consensus in the sample that an intuitive interface (cognitive ergonomics), easy handling of the device (classical ergonomics), high quality, an appealing appearance, and a pleasurable engagement increase the likelihood of usage. This list of attributes extends the two factors in the original TAM model [8]. Surprisingly, functionality did not contribute significantly to the prediction of technology adoption despite high *single* attribute importance scores. This was due to the operalization of functionality: the two levels (primary vs. secondary functions) were equally attractive.

Nonetheless, the model explained a near to perfect 99.1% of the variance. However, this should not be over-interpreted as the objective of this study was not to test a model, but to *compare* relevant attributes. Conjoint analysis is a method that decomposes overall ratings, in contrast to frequently used techniques of generating a composite score (e.g. sum) from independent ratings. The relative weights and importance values are of interest and not the overall rating itself.

As expected, the low competence group put more emphasis than their high competence counterparts on aspects of cognitive and classical ergonomics as these directly affect the probability of success or failure of an interaction. In addition, the prospect of an enjoyable interaction also increased the likelihood of usage to a higher extent in the low competence group. This should be considered as a promising motivator when designing for a hesitant user group such as a group with low self efficacy beliefs. As seen in the Achievement Motives scores, the low competence group is more likely to avoid anything that might lead to failures and has a lower tendency to aim for situations that might lead to success. Thus, any attribute that can encourage this group should be taken seriously. Emotional involvement is a hedonic attribute that is not directly linked to product performance in a pragmatic sense [16], but apparently to user motivation which is a prerequisite of interaction. In contrast, to get the high competence group interested, primarily high quality standards need to be met.

Group differences were not found in *single* attribute ratings, supporting our notion that additional information is gained when presenting products in full-profile rather than by independent attribute ratings.

One limitation of the study that should be mentioned is that the dependent variable was only a theoretically stated likelihood of usage. The link between such a predicted usage probability and actual behavior remains to be verified. On the other hand, this criterion creates the possibility to prospectively test expectations of user groups at very early stages of product development.

Conjoint analysis has been introduced as an effective approach to address the intention of technology use in HCI by considering attributes jointly. Attribute utility values from full-profile ratings and relative importance values showed group differences while independent importance ratings were not able to differentiate between users with low and high technology self-competence, respectively. Researchers should be cautious when relying on independent attribute importance ratings as these might yield to ceiling effects. An additional consideration of relative importance values is recommended. The derived preferences for different user groups can be translated into actionable priorities in a design process.

Acknowledgements. This research was supported by the German Research Foundation DFG [no. 1013/2 Prospective Design of Human-Technology Interaction] and by the National Research Fund Luxembourg [AFR].

References

1. Bradley, N.: Marketing Research. Tools and Techniques. Oxford University Press, Oxford (2007)
2. Grunert, K.G.: Attributes, Attribute values and their characteristics: a unifying approach and an example involving a complex household investment. Journal of Economic Psychology 10, 229–251 (1989)
3. Pohlmeyer, A.E., Blessing, L., Wandke, H., Maue, J.: The Value of Answers Without Question[s]: A Qualitative Approach to User Experience and Aging. In: Kurosu, M. (ed.) HCII 2009. LNCS, vol. 5619, pp. 894–903. Springer, Heidelberg (2009)

4. Hollnagel, E.: Cognitive Ergonomics: It's All in the Mind. Ergonomics 40(10), 1170–1182 (1997)
5. Blackwell, R.D., Miniard, P.W., Engel, J.F.: Consumer Behavior. Thomson/South-Western, Mason (2006)
6. Ulrich, K.T., Eppinger, S.D.: Product Design and Development. McGraw-Hill, New York (2007)
7. Hassenzahl, M., Burmester, M., Koller, F.: AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: Ziegler, J., Szwillus, G. (eds.) Mensch & Computer 2003. Interaktion in Bewegung, pp. 187–196. Teubner, Stuttgart (2003)
8. Davis, F.D.: Perceived Usefulness, Perceived Ease of Use and User Acceptance of Information Technology. MIS Quarterly 13(3), 319–339 (1989)
9. Benbasat, I., Barki, H.: Quo vadis, TAM? Journal of the Association for Information Systems 8(4), 212–218 (2007)
10. Compeau, D.R., Higgins, C.A.: Computer Self-Efficacy: Development of a Measure and Initial Test. MIS Quarterly 19(2), 189–211 (1995)
11. Huber, J.: What We Have Learned from 20 Years of Conjoint Research: When to Use Self-Explicated, Graded Pairs, Full Profiles or Choice Experiments. Sawtooth Software, Research Paper Series (1997)
12. Karrer, K., Glaser, C., Clemens, C., Bruder, C.: Technikaffinität erfassen - der Fragebogen TA-EG. In: Lichtenstein, A., Stößel, C., Clemens, C. (eds.) Der Mensch als Mittelpunkt technischer Systeme. Berliner Werkstatt Mensch-Maschine-Systeme, vol. 8, pp. 196–201. VDI Verlag GmbH, Düsseldor (2009)
13. Pohlmeyer, A.E., Machens, F., Blessing, L.: Attractive or Not – What's the Difference? Inter- and Intra-Group Comparisons in the Kano Model. In: Marjanovic, D., Storga, M., Pavkovic, N., Bojcetic, N. (eds.) 11th International Design Conference - DESIGN 2010, pp. 413–422. University of Zagreb, Croatia (2010)
14. Lang, J.W.B., Fries, S.: A Revised 10-Item Version of the Achievement Motives Scale - Psychometric Properties in German-Speaking Samples. European Journal of Psychological Assessment 22(3), 216–224 (2006)
15. Bloch, P.H., Brunel, F.F., Arnold, T.J.: Individual Differences in the Centrality of Visual Product Aesthetics: Concept and Measurement. Journal of Consumer Research 29, 551–565 (2003)
16. Hassenzahl, M.: The thing and I: understanding the relationship between user and product. In: Blythe, M., Overbeeke, C., Monk, A.F., Wright, P.C. (eds.) Funology: From Usability to Enjoyment, pp. 31–42. Kluwer Academic Publishers, Dodrecht (2004)