

Developing and Validating an English Version of the meCUE Questionnaire for Measuring User Experience.

Michael Minge, Manfred Thüring, Ingmar Wagner
 Technische Universität Berlin

Department of Cognitive Psychology and Ergonomics
 {michael.minge; manfred.thuering}@tu-berlin.de, ingmar.wagner@zoho.com

Based on the Component model of User Experience (CUE), a standardized questionnaire (meCUE) was developed measuring key aspects of user experience (UX) for interactive products. The questionnaire consists of 34 items and covers four components: product perceptions (usefulness, usability, visual aesthetics, status, commitment), user emotions (positive, negative), consequences of usage (intention to use, product loyalty), and overall judgment. The modules were separately validated in a series of studies. Therefore, meCUE can be easily adapted to specific research goals by simply choosing those modules which are required. The original German questionnaire was translated into an English version that was validated in an online study. Fifty-eight native English speakers assessed a wide variety of interactive products, such as cell phones, digital cameras, PCs, laptops, tablets, software and mobile applications. Results show that the English version reliably assesses the key components of UX and that the internal consistency of its scales is high.

INTRODUCTION

According to the user-centered design approach, a particularly important aspect for the evaluation of design solutions is how users experience the interaction with a technical system [9]. Perceptions, judgments, emotions, personal values and motivational tendencies significantly influence whether a system is likely to be used and whether it will be accepted and integrated into everyday life. Positive impressions ensure that pleasant memories are socially shared [2] and that users are likely to reuse the system because they expect similar experiences in the future [“prospective user experience”, 18]. Since subjective impressions and emotions may differ significantly from usability-oriented performance data [10], designing and evaluating the user experience (UX) is the key to success for many technical devices.

As an analytic framework for defining UX, the Components model of User Experience (CUE) by Thüring and Mahlke [22] integrates a number of theories and approaches. The CUE model distinguishes between the perception of instrumental qualities (e.g., usability, usefulness, etc.) and non-instrumental qualities (e.g. aesthetics, product image, etc.). Both kinds of perceptions are influenced by interaction characteristics, which relate to the product, the user, and the context of use (see Figure 1). In particular, the model assumes that emotions play a major role for UX. Emotional responses can be characterized in terms of valence and arousal which are independent aspects of any emotional reaction [19]. According to the CUE model, emotions are influenced by the instrumental and non-instrumental product qualities that users perceive when interacting with the product. Such influences have been reported in a number of studies [e.g. 13, 20, 21, 22].

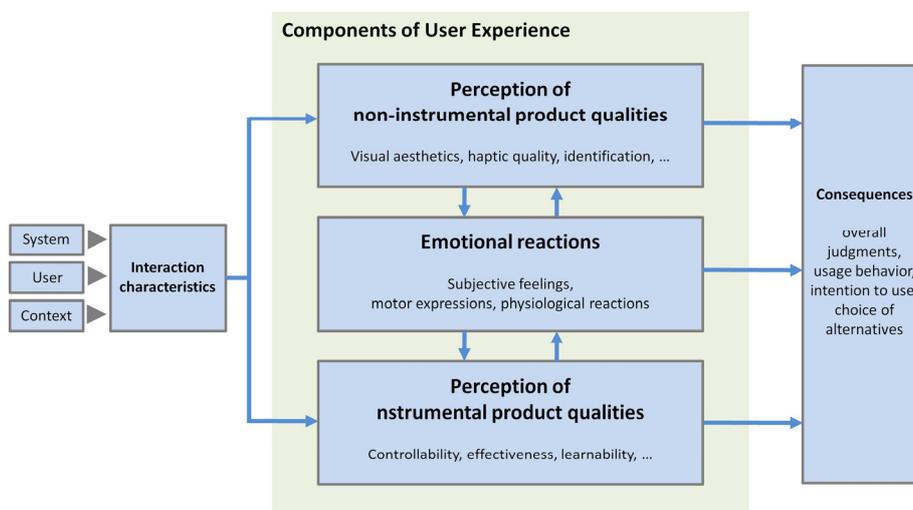


Figure 1. Components of User Experience (CUE model)

However, emotions themselves may in turn influence perception, attention and decision-making [e.g. 4] as well as judgment and thought [e.g. 5]. In the domain of user experience, a comparable relation was found by Aranyi and van Schaik [1] who adapted the CUE model to describe the interplay between product perceptions and emotions for news websites. Their study was based on structural model equations and showed that positive emotions affected the non-instrumental quality, but not the instrumental one, while negative emotions affected the instrumental quality, but not the non-instrumental one.

While the model proposed by Thüring and Mahlke [22] proposes uni-directional relationships between product perceptions and emotions, the current version of the CUE model postulates bi-directional relationships between perceived product qualities and emotions. Perceived qualities and emotions together determine the consequences of use, such as the overall product judgment, acceptance and intentions of future usage (see Figure 1).

To ascertain how users experience the interaction with a system, all UX components matter and a number of questionnaires can be employed to assess some of them. For example, AttrakDiff [7] and the User Experience Questionnaire (UEQ) [12] measure product perceptions and acceptance, while SAM [3], PrEmo [6] and LEM-Tool [8] capture emotional reactions. However, there has been no instrument so far which addresses all components in a unified way. Instead, questionnaires with different scales, formats and instructions must be used to get a holistic and comprehensive picture of user experience. Such a “test battery” may be confusing for test persons and require additional effort on behalf of the researcher who must find the best combination of tests and aggregate data from different scales.

In order to provide an alternative to the existing instruments, the meCUE¹ questionnaire addresses all key components of UX together. It is structured into four modules which correspond to the components and sub-concepts that are specified by the CUE model. Each module contains items in a consistent Likert-scale format and can be used separately or in combination with the other modules. The final version of the questionnaire consists of 34 items which are categorized into 10 scales (see Figure 2). One item, the overall evaluation, is captured on a single-item scale.

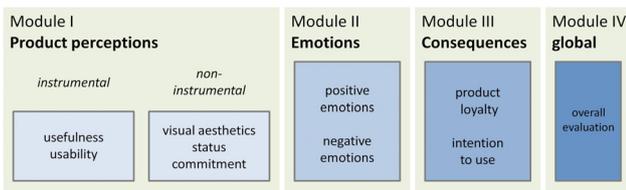


Figure 2. Final structure of the meCUE questionnaire

¹ meCUE: modular evaluation of key Components of User Experience

The development of meCUE started with a German version, for which 67 items were generated. Item selection and validation of the questionnaire were based on five data collections [11, 14, 15, 16, 17]. Three surveys were conducted online ($n_1 = 238$, $n_2 = 238$, $n_3 = 237$) and two in a laboratory setting ($n_4 = 67$, $n_5 = 24$). In all studies, participants rated a wide range of different interactive products (e.g., mobile devices, apps and software, home appliances).

The first two online studies focused on determining those items which loaded high on the scales of the questionnaire. Data was analyzed using principle component analyses and resulted in the selection of 33 items measuring 9 different scales (plus the single-item for the overall rating). The constructed scales showed a high internal consistency (Cronbach's Alpha values between .69 and .83 [14, 17]). Proportions of explained variance were acceptable for all modules (module 1: 69.9 %, module 2: 57.4 %, module 3: 63.5 %).

The final set of items and the structure of the questionnaire were confirmed under laboratory conditions (Cronbach's Alpha values of the scales between .76 and .94, proportions of explained variance for module 1: 81.1 %, for module 2: 74.3%, and for module 3: 74.1 %) [14, 17].

The validity of the final version of the meCUE questionnaire was tested using three different approaches. First, the correlations between the dimensions of meCUE and the dimensions of other questionnaires measuring similar constructs were examined (convergent validity). Here, it was found that meCUE consistently led to comparable values [11, 14, 17]. For example, significant correlations were observed between the meCUE single-item “overall rating” (module IV) and the dimension “attraction” of AttrakDiff ($r = .559$) [16, 17], AttrakDiff mini ($r = .919$) and UEQ ($r = .887$) [11]. Second, the relationships between the dimensions of the questionnaire and total task times were analyzed (criterion-related validity). In this experimental usability test scenario, sixty-seven participants worked with two versions of text-editing software. With respect to product perceptions, it could be shown that total task times as objective performance measures were highly correlated with the subjective ratings of instrumental qualities (usefulness and usability), whereas no significant correlations were observed between performance and non-instrumental quality ratings (visual aesthetics, status, commitment) [14, 17]. Third, it was systematically investigated whether the questionnaire was able to detect the same differences previously identified by an expert evaluation (discriminative validity). It was found that the dimensions of the meCUE questionnaire nicely replicated the results of the expert evaluation and even minimal variations were consistently represented [11, 15, 17].

The results reported so far were obtained with the German version of meCUE. The study, which is described in the following, empirically validated a version of the questionnaire, which had been translated into English.

METHOD

To obtain an English version of meCUE, three native speakers were engaged who had been working as professional translators or language teachers for several years. In a first step, two of them translated the items independently from each other into English. For 25 items out of 34, their translations coincided. Next, all English items, including the inconsistent ones, were retranslated by the third speaker into German. This resulted into 28 correct retranslations. To find a solution for the remaining six ambiguous items, two of the speakers thoroughly discussed different alternatives and finally agreed on a wording. Finally, the answer form of the questionnaire was translated. All native speakers were involved in finding adequate English expressions for the seven points of the Likert-scale (see Figure 3).

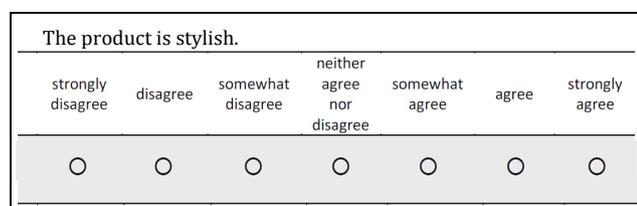


Figure 3. Seven-point Likert-scale answer form.

Note that the single-item of module IV (overall rating) is a semantic differential, ranging from -5 to 5 with an increment of .5 respectively (see Figure 4).

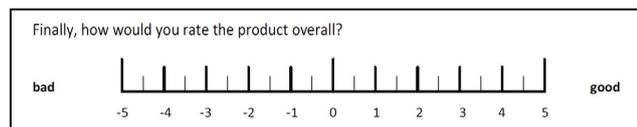


Figure 4. Single-item of module IV (global assessment).

To test the factorial structure of the English version of the questionnaire, it was deployed in an online survey. Fifty-eight native speakers from the United Kingdom and the United States (ages between 23 and 56, $M=32.6$ years) rated their experience with an interactive technical product from their personal environment (e.g. mobile devices, laptop, TV, software, mobile application, household appliances). Participants were free to choose which product they wanted to rate. On average, they had owned the selected device for 9.4 months.

For each module of the questionnaire, a principle component analysis based on the Minimum Average Partial (MAP-) Test [23] was calculated. In order to determine the reliability of the instrument, factor loadings were checked. The internal consistency coefficients of the scales were analyzed using Cronbach's Alpha.

RESULTS

The pattern of factor loadings shows that the five-dimensional structure of module I is adequately represented in the English version. Loadings higher than .4 correspond to the expected dimensions (see Table 1: product perceptions).

Item:	usefulness	usability	visual aesthetics	status	commitment
The product is creatively designed.			.848		
The design looks attractive.			.784		
The product is stylish.			.745		
The product is easy to use.		.864			
The operating procedures of the product are simple to understand.		.842			
It is quickly apparent how to use the product.		.797			
By using the product, I would be perceived differently.				.886	
The product would enhance my standing among peers.				.861	
My friends could be quietly envious of this product.				.716	
I could not live without this product.					.849
The product is like a friend to me.					.789
If I ever lost the product, I would be devastated.			.498		.517
With the help of this product, I will achieve my goals.	.875				
I consider the product extremely useful.	.629		.533		
The functions of the product are exactly right for my goals.	.584	.531			

Table 1. Module I (product perceptions): Factor loadings > .4

Items and factor loadings of module II (user emotions) are displayed in Table 2. The analysis revealed two independent factors for the assessment of emotions in human-computer technology. According to the corresponding items, these dimensions differ in the quality of emotional reactions, namely negative and positive.

Item:	negative emotions	positive emotions
The product annoys me.	.864	
The product angers me.	.860	
The product frustrates me.	.857	
When using this product I feel exhausted.	.814	
The product makes me tired.	.759	
The product makes me feel passive.	.560	
The product relaxes me.		.766
The product exhilarates me.		.748
The product makes me happy.		.732
The product makes me feel euphoric.		.723
The product calms me.		.716
When using this product, I feel cheerful.		.623

Table 2. Module II (user emotions): Factor loadings > .4

Finally, a two-dimensional structure was found for module III (consequences of use). The pattern of factor loadings is equivalent to the German version with the two dimensions “intention to use” and “product loyalty” (Table 3).

Item:	intention to use	product loyalty
I would not swap this product for any other.		.885
In comparison to this product, no others come close.		.865
I would get exactly this product for myself (again) at anytime.		.826
I can hardly wait to use the product again.	.848	
If I could, I would use the product daily.	.808	
When using this product, I lose track of time.	.757	

Table 3. Module III (consequences): Factor loadings > .4

As a measure of internal consistency, Cronbach’s Alpha values were determined for each scale (see Table 4). All values are between .76 (acceptable) and .91 (excellent). Table 4 also shows the proportion of explained variance for each scale and the cumulative proportions for each module.

Modules and scales:	% variance	Cronbach’s alpha
Product perceptions (module I)		
usefulness	12.3	0.78
usability	19.2	0.90
visual aesthetics	19.8	0.91
social identity: status	15.5	0.84
social identity: commitment	12.8	0.76
cumulative %	79.5	
User emotions (module II)		
positive emotions	27.2	0.82
negative emotions	32.1	0.88
cumulative %	59.3	
Consequences of usage (module III)		
intention to use	39.2	0.87
product loyalty	35.3	0.78
cumulative %	74.5	

Table 4. Internal consistency of the meCUE questionnaire.

In the present study, the overall evaluation captured by the single-item of module IV indicates that participants preferred to choose products which were associated with a positive attitude ($M=1.98, s=2.18$).

CONCLUSION

To summarize, meCUE is a flexible, adaptable, lean and validated questionnaire for measuring user experience based on the CUE model. The German as well as the English version have a good internal consistency and reliably assess the key components of user experience. In contrast to instruments which capture single aspects or a

subset of them, meCUE addresses all UX components together – including emotions – in a unified format using a Likert-scale. Since the questionnaire consists of only 34 items, it is efficient to use, requiring only between two and five minutes on average to be filled in.

As demonstrated in our studies, meCUE can be applied in UX surveys on all kinds of interactive systems. The use of such a standardized tool is particularly suitable for comparing different products or design options and for detecting changes of experience in the course of long-term usage.

The English version of meCUE is freely available under the following link: www.mecue.de/english. The website also provides an Excel file that supports data collection and analysis.

ACKNOWLEDGMENTS

This research was supported by the German Research Association (DFG) as part of the Research Training Group "Prospective Engineering of Human-Technology Interaction" (GrK 1013). We would like to thank Laura Riedel and Carina Kuhr for their work on the studies.

REFERENCES

- Gabor Aranyi, Paul van Schaik. 2015. Testing a model of user-experience with news websites. *Journal of the Association for Information Science and Technology*. DOI: 10.1002/asi.23462.
- Katja Battarbee. 2004. *Co-Experience. Understanding User Experiences in social interaction*. Ph.D. dissertation at the University of Art and Design Helsinki.
- Margaret Bradley, Peter Lang. 1994. Measuring emotion: The Self-Assessment Manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1), 49-49.
- Tobias Brosch, Klaus R. Scherer, Didier Grandjean & David Sander. 2013. The impact of emotion on perception, attention, memory, and decision-making. *Swiss Medical Weekly*. 2013;143:w13786.
- Gerald L. Clore, Jeffrey R. Huntsinger. 2007. How emotions inform judgment and regulate thought. *Trends in Cognitive Science*, 11, 393-399.
- Peter Desmet. 2003. Measuring Emotion: Development and application of an instrument to measure emotional responses to products. In M.A. Blythe, A.F. Monk, K. Overbeeke, P.C. Wright (Eds.). *Funology: From Usability to Enjoyment* (pp. 111-123). Kluwer Academic Publishers.
- Marc Hassenzahl, Michael Burmester, Franz Koller. 2008. Der User Experience (UX) auf der Spur: Zum Einsatz von *www.attrak.diff*. In H. Brau, S.

- Diefenbach, M. Hassenzahl, F. Koller, M. Peissner, K. Rose. (Eds.): *Usability Professionals 2008*, 78-82.
8. Gijs Huisman, Marco van Hout, Elisabeth van Dijk, Thea van der Geest, Dirk Heylen. 2013. *LEMtool - Measuring emotions in visual interfaces*. In SIGCHI Conference on Human Factors in Computing Systems, CHI 2013, 27 April - 2 May 2013, Paris, France (pp. 351-360).
 9. ISO CD 9241-210. 2010. Ergonomics of human-system interaction. Part 210: Human-centered design process for interactive systems. Geneva: International Standardization Organisation (ISO).
 10. Masaaki Kurosu, Kaori Kashimura. 1995. Apparent usability vs. inherent usability: experimental analysis of the determinants of the apparent usability. *Proceedings of CHI '95. Conference Companion on Human Factors in Computing Systems* (pp. 292-293). New York: ACM Press.
 11. Carina Kuhr, Michael Minge, Manfred Thüring. 2014. Measuring the User Experience of Mobile Applications – an Empirical Validation of a Quantitative Method. In A.C. Schütz, K. Drewing, K.G. Gegenfurtner (Eds.). *Abstracts of the 56th Conference of Experimental Psychologists, TeaP* (p. 155). Lengerich: Pabst Science Publishers.
 12. Bettina Laugwitz, Martin Schrepp, Theo Held. 2006. Konstruktion eines Fragebogens zur Messung der User Experience von Softwareprodukten. In A. M. Heinecke, H. Paul (Eds.), *Mensch und Computer 2006: Mensch und Computer im Strukturwandel* (pp. 125-134). Munich: Oldenbourg.
 13. Sascha Mahlke, Michael Minge. 2008. Consideration of multiple components of emotions in human-technology interaction. In C. Peter & R. Beale (Eds.): *Affect and emotion in human-computer interaction* (pp. 51-62). Berlin: Springer.
 14. Michael Minge, Laura Riedel. 2013. meCUE – Ein modularer Fragebogen zur Erfassung des Nutzungserlebens. In S. Boll, S. Maaß, R. Malaka (Eds.). *Mensch und Computer 2013: Interaktive Vielfalt* (pp. 89-98). Munich: Oldenbourg.
 15. Michael Minge, Laura Riedel, Manfred Thüring. 2013. Und ob du wirklich richtig stehst... Zur diskriminativen Validität des User Experience Fragebogens „meCUE“. In S. Boll, S. Maaß, R. Malaka (Eds.). *Mensch und Computer 2013 – Workshopband* (pp. 137-144). Munich: Oldenbourg.
 16. Michael Minge, Laura Riedel, Manfred Thüring. 2013. Modulare Evaluation von Technik. Entwicklung und Validierung des meCUE Fragebogens zur Messung der User Experience. In E. Brandenburg, L. Doria, A. Gross, T. Güntzler, H. Smieszek (Eds.). *Berliner Werkstatt für Mensch-Maschine-Systeme 2013: Grundlagen und Anwendungen der Mensch-Technik-Interaktion* (pp. 28-36). Berlin: Universitätsverlag TU Berlin.
 17. Michael Minge, Manfred Thüring, Ingmar Wagner & Carina V. Kuhr. Accepted. The meCUE Questionnaire. A Modular Evaluation Tool for Measuring User Experience. *Applied Human Factors and Ergonomics Conference (AHFE)*, 2016.
 18. Anna E. Pohlmeier, Martin Hecht, Lucienne Blessing. 2009. User Experience Lifecycle Model ContinUE [Continuous User Experience]. In A. Lichtenstein, C. Stöbel, C. Clemens (Eds.). *Berliner Werkstatt für Mensch-Maschine-Systeme 2009: Der Mensch im Mittelpunkt technischer Systeme* (pp. 314-317). Düsseldorf: VDI Verlag.
 19. James A. Russell. 2009. Core Affect and the Psychological Construction of Emotion. *Psychological Review*, 110(1), 145-172.
 20. Pertti Saariluoma, Jussi P.P. Jokinen. 2014. Emotional dimensions of user experience: A user psychological analysis. *International Journal of Human-Computer Interaction*, 30(4), 303-320.
 21. Kwang-Kyu Seo, Sangwon Lee, Byung Do Chung, Changsoo Park. 2015. Users' Emotional Valence, Arousal, and Engagement Based on Perceived Usability and Aesthetics for Web Sites. *International Journal of Human-Computer Interaction*, 31(1), 72-87.
 22. Manfred Thüring, Sascha Mahlke. 2007. Usability, aesthetics, and emotions in human-technology interaction. *International Journal of Psychology*, 42(4), 253-264.
 23. Wayne Velicer. 1976. Determining the Number of Components from the Matrix of Partial Correlations. *Psychometrika*, 41(3), 321-327.