How to create short forms of UEQ+ based questionnaires?

Martin Schrepp
TI SAP CP Experience
SAP SE
Walldorf, Germany
martin.schrepp@sap.com

Heike Sandkühler
NOZ Digital GmbH
Osnabrück, Germany
heike.sandkuehler@noz-digital.de

Jörg Thomaschewski
Hochschule Emden/Leer
Emden, Germany
joerg.thomaschewski@
hs-emden-leer.de

ABSTRACT

The UEQ+ is a modular framework that can be used to set up UX questionnaires that fit to a research question. The UEQ+ consists of a set of UX scales that can be combined to form a concrete questionnaire. Thus, the product owner or UX researcher can define which aspects of UX are important for his or her project and should thus be measured in the questionnaire. A UEQ+ scale consists of 4 items that measure the UX aspect represented by the scale and an additional item that measures how important this UX aspect is for the overall UX impression. Some situations require that a UX questionnaire is extremely short and for such situations the standard scale format may be too time consuming. We describe how short forms of the UEQ+ can be built. The impact of shortening the scales on the results is investigated in several studies.

CCS CONCEPTS

Human-centered computing → HCI design and evaluation methods

KEYWORDS

User Experience Questionnaire; UX Measurement, UEQ, UEQ+

1 Introduction

Questionnaires are an efficient and cheap method to collect UX feedback concerning a product from larger samples of users. Thus, they are quite popular amongst UX researchers [1].

However, different research situations pose different requirements concerning the maximal length of a questionnaire. If the time to answer the questionnaire is seen as non-acceptable by potential participants, many of them will refuse to submit answers and the dropout rate will be high. Thus, there is a need to provide short forms of UX questionnaires for such situations. Such short versions exist for many established UX questionnaires (for example, [2, 3, 4]), thus the researcher can, depending on the needs of his or her current research project,

select between a more accurate long and a less accurate short form.

We describe in this paper how short forms of UEQ+ based questionnaires can be created. We describe some typical situations in which such a short form makes sense. In addition, we show with the help of some empirical examples how shortening the questionnaire impacts the results.

2 The UEQ+ framework

Standardized UX questionnaires (for example, UEQ [2, 11], SUS [5], VISAWI [6], ATTRAKDIFF2 [7], meCUE[8]) are powerful tools to measure the UX impression towards a product. Such questionnaires consist of one or several scales. Each scale measures a dedicated aspect of UX, for example, efficiency of use, ease of learning, aesthetic value of the design, novelty of the design, or trust in the product. Many of the existing UX standard questionnaires offer some support for researchers in form of a handbook giving hints how to use the questionnaire and tools to automate the data analysis.

However, standard questionnaires come with a fixed set of scales. Sometimes this is not an issue, since these scales cover what is of interest to a UX researcher in a concrete evaluation project. But in other situations, this may not be the case, i.e. there is no standard UX questionnaire that covers all the UX aspects relevant for a certain product. Sometimes it may be possible to use several UX questionnaires in parallel. But since the item formats of these questionnaires typically differ, this also has some drawbacks on the mental effort for the participants, the time required to fill all questionnaires and the ability to integrate scale values obtained from different questionnaires into a consistent picture about the UX of the evaluated product [1].

The UEQ+ [10] is not a UX questionnaire in the classical sense. It is a collection of UX scales that can be combined by a researcher to create a UX questionnaire that measures exactly those UX aspects that are relevant to answer his or her research questions. To allow such a combination, the UEQ+ uses a special scale format.

Items have the form of a semantic differential, i.e. consist of two terms that describe the opposite ends of a semantic dimension. The participant can describe his or her impression of the product concerning this semantic dimension on a 7-point Likert scale. The four items of a scale are grouped, and an introductory sentence is used to set a common context for these 4 items. In addition, there is an item that asks about the

importance of the UX aspect represented by the scale for the overall UX impression of the product. This item is used to calculate an overall UX KPI by weighting the rating of a scale with its importance. See [10] for details of the KPI calculation. As an example, we present the scale *Efficiency*.

To achieve my goals, I consider the product as

slow	o	o	o	o	o	o	o	fast
inefficient	o	o	o	o	o	o	o	efficient
impractical	o	o	o	o	o	o	o	practical
cluttered	o	o	o	o	o	o	o	organized

I consider the product property described by these terms as

completely irrelevant	o	o	o	o	o	o	o	very important

The UEQ+ is an extension of the UEQ [11]. It contains the 6 UEQ scales and some additional scales provided by different authors, see [12,13,14,15]. All scales and the available materials to use the UEQ+ can be found on the UEQ+ homepage www.ueqplus.ueq-research.com.

Thus, the application of the UEQ+ is more powerful but also more difficult than using a typical standard questionnaire. It requires on the side of the UX researcher a pretty good understanding of the UX aspects relevant for his or her research questions. But this is most likely an important step to understand how to improve a product in terms of UX.

In principle, there are two reasons to select a UEQ+ scale. First, the scales that represent UX aspects which are highly relevant for the typical users should be selected. Second, it is often required to measure UX aspects not very important from the point of view of the users, but important for the product placement because of marketing reasons.

If just a single product should be evaluated to get a rough idea about its UX quality it is much easier to use a standard questionnaire, for example the UEQ [11]. If it is planned to establish a repeated measurement of a product to check the development of the UX quality over time or if UX measurement should be established as a part of the quality process, then the additional efforts required to set up a UEQ+ based questionnaire will pay off. In addition, the effort to set up a UEQ+ for a project can also create a deeper common understanding of the UX requirements and their importance in the project team and can therefore create some additional value [16]. See the UEQ+ handbook [17] for a more detailed description of the recommended usage scenarios.

3 Typical usage scenarios requiring a short form?

Questionnaires are a quite efficient method to collect user feedback to an interactive product. If the questionnaire is handed over as part of a usability test or if the participants are highly motivated to give feedback, the length of the questionnaire and thus the time to fill it out are not so important. But in other typical application scenarios for a UX questionnaire it is important to keep the questionnaire short to ensure a

sufficiently high response rate. We provide in this section some typical examples of such situations.

As a first example, assume that you want to collect feedback when a user leaves a web-shop. The display of the questionnaire will be triggered after the user submitted an order. Here it is quite important that the user has the impression that filling out the questionnaire will not require much time and effort. Otherwise he or she will simply ignore the questionnaire and close the web page. This is also true for other situations in which a questionnaire is triggered by some user actions in a system, i.e. where the user does not click intentionally on an element to start the questionnaire.

Often product experience questionnaires are sent out by marketing departments some weeks after customers purchased a product. For example, if you buy a new car you typically receive a questionnaire that asks about your satisfaction with the buying experience, in addition to some demographic questions. Such questionnaires try to cover the complete product experience and tend to be quite lengthy, i.e. contain already many questions that are important for marketing purposes. But for UX researchers such questionnaires are often a good chance to collect also some UX feedback. Since companies of course do not want to annoy their customers with too many surveys it is in such situations typically not possible to send out a second questionnaire concentrating on UX aspects. Thus, including a short UX section in the customer experience survey is often a good chance to collect feedback.

A third example are situations in which a participant should provide feedback to several products or variants of a product. This situation occurs, for example, in experimental settings that compare several variants of a new product design. Participants in such studies try out the different variants in a random order and provide feedback in the form of a UX questionnaire. Since the same questionnaire must be filled out several times it is important to keep it short. Otherwise the participant will get bored and stressed and the quality of the answers will decrease over time.

Another example are situations, in which the same target group is asked in fixed time intervals (for example, every 6 month) to give feedback about the UX of a product. For example, a media house that offers a payed online journal may ask the readers every 6 month about their UX impression of the product. In such situations it is quite interesting to see how the ratings of a reader evolve over time. If the questionnaire requires much time to fill it out, then readers will refuse to participate a second time.

4 How to create short forms?

The typical approach to create a short form of an existing UX questionnaire is to select a small subset of the items in the original form. Obviously, this subset will not cover as much information as the original form. Thus, typically the measurement results are less fine granular. In addition, the more items are used to measure UX the smaller is the impact of random response errors or misunderstandings in some items.

Thus, a reduction in the number of items will also decrease the reliability of measurement.

For example, the UEQ consists of 26 items grouped to 6 UX scales. The short form consists of just 8 of these items that are grouped into the two meta-scales *pragmatic quality* and *hedonic quality*. Thus, the short form just measures two high level sub-aspects of UX instead of 6 more detailed aspects. The VISAWI consists of 18 items, measuring 4 distinct aspects of visual beauty. The short form consists of just 4 items and only measures an overall impression, i.e. just one value indicating overall impression concerning visual beauty.

For a modular framework like the UEQ+ this approach cannot be followed. The basic idea of the UEQ+ is that the researcher picks the scales he or she considers as important in a concrete project. It is thus not possible to simply measure meta-scales. Thus, we must follow a different path to shorten such questionnaires.

4.1　Save time by reusing measurements for the importance question

A UEQ+ scale consists of 4 items and a question concerning the importance of the scale for the overall UX impression. Assume that a product is measured every 6 months with a questionnaire derived from the UEQ+ and that the scale *Efficiency* is included. Is it plausible that the importance of this scale differs between two measurements? Maybe the scale will show different values if the efficiency of the product gets better or worse. But how important this aspect is for the overall product satisfaction will typically not change. This depends mainly on the type of product and the typical use case [1,9].

Thus, if a UEQ+ questionnaire is used multiple times to measure the same product, the question concerning the importance can be used the first time and then be dropped for the other studies. This is also true if several products of the same type are measured. Use the importance question in one study and reuse the measured values in the other studies.

However, there is one point that should be considered if you reuse importance values from older studies. If the nature of the product changes dramatically by introducing new functionality that allow completely new interactions, then the importance of the scales may change. For example, assume that a completely task-oriented product to manage customer relations (a CRM product) is enhanced with some new functionality to support internal communication between employees including features of a social network. This will create completely new usage patterns and thus the importance of UX aspects may change. It may also be required to include some new scales into the questionnaire in such a case.

We provide now some data from first studies with the UEQ+ that help to understand the impact that a reuse of importance ratings will have on the results.

As a first example we look at the importance ratings of 4 digital newspapers. The four newspaper apps are similar in their technology and deviate mainly concerning their content. Details of this study are described in [19].

Table 1 shows the importance ratings and standard deviations for the four apps. The value range of the importance ratings is between -3 to +3.

Scales	Studies (Participants)			
	A (242)	**B** (1125)	**C** (1770)	**D** (190)
Content Qual.	2.24 (1.09)	2.25 (0.99)	2.31 (0.93)	2.27 (0.89)
Efficiency	2.06 (1.01)	2.07 (1.00)	2.10 (1.00)	2.08 (0.94)
Value	1.40 (1.74)	1.46 (1.33)	1.55 (1.27)	1.71 (1.17)
Perspicuity	2.34 (0.85)	2.20 (0.96)	2.29 (0.92)	2.27 (0.97)
Usefulness	2.02 (1.12)	2.02 (1.03)	2.07 (1.03)	2.08 (1.03)
Attractiveness	1.85 (1.17)	1.76 (1.24)	1.91 (1.13)	1.89 (1.22)

Table 1: Importance ratings and standard deviations (in brackets) for four newspaper apps. The numbers in brackets in the column headers show the number of participants per study.

The following differences are significant on the 5% level (t-test): *Value* (between A, C and B, D), *Perspicuity* (between A, B and A, D) and *Attractiveness* (between A, D). But the actual differences are quite small.

The KPIs for the four products were 1,66 (A), 1,58 (B), 1,82 (C) and 1,85 (D). Assume now that we would calculate the KPI of A, B, and C not using the ratings from the corresponding study but based on the importance ratings obtained in the smallest study (i.e. for product D).

The results would be 1,65 (A), 1,57 (B) and 1,81 (C). Thus, the results would be nearly the same if the user groups are similar.

As a second example we reanalyze some data sets obtained from the first validation studies of a set of new UEQ+ scales. In two investigations students judged the UX of the streaming platforms *Amazon Prime* and *Netflix* and the web-shops *otto.de* and *zalando.de*. Details of the study are described in [10]. Here we have the situation that we have two pairs of quite similar products.

Table 2 shows the importance ratings for the two streaming platforms.

UEQ+ Scale	Amaz. Prime (57)		Netflix (73)	
	Mean	**STD**	**Mean**	**STD**
Attractiveness	1.61	0.99	1.45	1.25
Perspicuity	2.27	0.83	2.22	0.80
Intuitive Use	1.86	1.03	1.96	1.05
Visual Aesthetics	1.11	1.36	1.10	1.39
Quality of Content	1.63	1.12	1.79	1.08
Trustworthiness of Content	1.49	1.13	1.08	1.47
Trust	1.91	1.00	1.85	1.06

Table 2: Importance ratings for 7 UEQ+ scales for *Amazon Prime* and *Netflix*.

The differences in the importance ratings for the scales of both products are statistically not significant on the 0.05 level (t-test).

The measured KPI for *Amazon Prime* was 1,35 and the KPI for *Netflix* was 1.73. If we would use the importance ratings

obtained from *Netflix* to calculate the KPI for *Amazon Prime* the result would be 1.33. If we conversely use the importance ratings obtained from *Amazon Prime* to calculate the KPI for *Netflix* the result would be 1.72. Thus, in both cases the result of the KPI would be quite similar if we had used the importance ratings obtained from another product of the same type.

Table 3 shows the importance ratings for the two web shops.

UEQ+ Scale	otto.de (42)		zalando.de (46)	
	Mean	STD	Mean	STD
Attractiveness	1.14	1.30	0.93	1.37
Dependability	1.59	1.01	1.89	0.89
Intuitive Use	1.86	1.10	2.17	0.96
Visual Aesthetics	0.95	1.46	1.35	1.43
Quality of Content	1.71	1.04	1.87	1.03
Trustworthiness of Content	1.81	0.98	2.04	0.95
Trust	2.02	1.02	2.22	0.88
Value	0.67	1.21	1.27	1.31

Table 3: Importance ratings for 8 UEQ+ scales for *otto.de* and *zalando.de.*

None of the scales shows a statistically significant difference (t<0.05 t-test).

The measured KPI for *otto.de* was 1,27 and the KPI for *zalando.de* was 1.70. Again, if we use the importance ratings from the other product to calculate the KPI's we would get 1,31 for *otto.de* and 1.75 for *zalando.de*. As in the previous example, the differences are small.

This example supports the assumption that the importance of a UX aspect depends mainly on the type of product. Thus, it seems even possible to reuse importance ratings obtained from a similar product that supports the same usage scenarios if the full form of the scales cannot be used.

Our suggestion would be to measure the importance of the scales in a first study and to reuse these values for other studies as long as there is no indication that changes in the product will allow quite different usage scenarios and thus require different importance ratings. In such cases the importance should be measured again.

4.2 Reducing the number of items per scale

The UEQ+ scales measure an UX aspect with 4 items that are set into a common context by an introductory sentence.

Compared to the original UEQ that presents the items in a random sequence, this format reduces potential misunderstandings of items. Thus, the reliability of the UEQ+ scales (for example measured by their Cronbach Alpha values) is quite high. Thus, another potential way to shorten the questionnaire is to reduce the number of items per scale to 3 or even 2.

Of course, this will have an impact concerning reliability. But we will show with an example that this loss may be acceptable for practical purposes.

We reanalyze a data set (see [19] for details of the study) where 1125 participants rated a payed digital newspaper concerning 6 UEQ+ scales.

Table 4 shows the scale means based on all 4 items and the corresponding means obtained from all possible combinations of 3 items.

UEQ+ Scale	Mean	Combination of 3 items			
		1,2,3	1,2,4	1,3,4	2,3,4
Content quality	1.72	1.64	1.77	1.73	1.71
Efficiency	1.20	1.14	1.21	1.17	1.26
Value	1.48	1.63	1.47	1.39	1.44
Perspicuity	1.71	1.74	1.70	1.66	1.73
Usefulness	1.75	1.77	1.79	1.75	1.71
Attractiveness	1.53	1.54	1.51	1.51	1.56

Table 4: Scale mean and all possible means over 3 item combinations for a digital newspaper app (N=1125).

Table 5 shows the scale means based on all 4 items and the corresponding means obtained from all possible combinations of 2 items.

UEQ+ Scale	Mean	Combination of 2 items					
		1,2	1,3	1,4	2,3	2,4	3,4
Content Q.	1.72	1.69	1.63	1.83	1.60	1.80	1.74
Efficiency	1.20	1.13	1.08	1.18	1.21	1.32	1.26
Value	1.48	1.69	1.56	1.33	1.63	1.40	1.27
Perspicuity	1.71	1.74	1.68	1.62	1.79	1.73	1.67
Usefulness	1.75	1.82	1.77	1.80	1.70	1.74	1.68
Attractiveness	1.53	1.52	1.51	1.48	1.58	1.55	1.54

Table 5: Scale mean and all possible means over combinations of 2 items for a digital newspaper app (N=1125).

The results show that there is of course an impact of removing items on the resulting scale means. But the loss of accuracy in measurement is in this example in a reasonable range for practical applications. In addition, we can see that some combinations of items estimate the full-scale better than others. Thus, it is a good idea to have a pilot study or a first study with the full scales to decide which item combinations should be used for the short form. This will depend on the product that is measured, i.e. it is not possible to give a general suggestion which item combinations perform best

If we would build a short form with just two items for the example shown in Table 4 we would select the following item combinations: *Content quality* (3 and 4), *Efficiency* (2 and 3), *Value* (1 and 3 or 2 and 4), *Perspicuity* (2 and 4), *Usefulness* (2 and 4) and *Attractiveness* (1 and 2 or 3 and 4).

5 A first practical application

In a first practical application, a short form which contains just two items per scale (the importance question was still included) was applied to an ePaper App and a News Portal. To have a comparison the full scales of the UEQ+ were applied to similar

products (these products use the same technology and UI design to present the articles but are independent offerings with slightly different content). But we can assume that they will not differ much in the perception of their UX.

The ePaper App runs on Smart Phones and Tablets, while the Portal is targeted both at mobile devices and larger screens.. Both products offer different functionalities to consume the content. Thus, different scales were used for the measurement. For the ePaper Apps the UEQ+ scales *Content Quality*, *Efficiency*, *Usefulness*, *Value* and *Perspicuity* were used, while the portals were evaluated with the scales *Content Quality*, *Efficiency*, *Adaptability*, *Visual Aesthetics* and *Intuitive Use.*

The 2 items used in the short version were selected based on previous measurements with the full scales. Per scale the two items which approximated the full scale best in these previous measurements were selected.

The questionnaires were launched inside the products. The following table shows the mean and standard deviations of the scales and the measured values for the KPI. The data are based on the responses of (A) N=191, (B) N=230, (C) N=147 and (D) N=229 participants.

UEQ+ Scale	A ePaper 4 items	B ePaper 2 items	C Portal 4 items	D Portal 2 items
Content Qual.	1.68 (1.11)	1.56 (1.27)	0.86 (1.54)	0.89 (1.32)
Efficiency	1.40 (1.28)	1.20 (1.46)	0.30 (1.69)	-0.10 (1.56)
Usefulness	1.90 (1.12)	1.98 (1.16)	-	-
Value	1.33 (1.26)	1.47 (1.22)	-	-
Perspicuity	1.85 (1.20)	1.78 (1.38)	-	-
Adaptability	-	-	-0.36 (1.65)	0,14 (1.38)
Aesthetics	-	-	0.18 (1.65)	0.21 (1.53)
Intuitive Use	-	-	0.57 (1.62)	0.43 (1.66)
KPI	1,66 (0.91)	1.60 (1.13)	0.37 (1.28)	0.33 (1.22)

Table 6: Scale values and their standard deviations of all 4 studies.

The mean times required to fill the questionnaire (it does not contain only the UEQ+ scales but in addition some demographic questions) were:

- A: 6:27 minutes (Median 6:03, Std.Dev. 2:13)
- B: 5:51 minutes (Median 5:26, Std.Dev. 2:26)
- C: 6:01 minutes (Median 5:21, Std.Dev. 2:05)
- D: 5:04 minutes (Median 4:31, Std.Dev. 1:59)

Thus, removing 2 items resulted in a reasonable saving of response time (36 seconds for the pair A/B and 57 seconds for the pair A/C).

As we can see from Table 5 the reduction of the number of items had an impact on the measured scale values. But this impact seems even on the level of single scales acceptable for practical applications. For the KPI, which is based on all 5 scales, the impact of the reduction is extremely small. In addition, we see that the questionnaires derived from the UEQ+ are able to distinguish quite well between products of different UX quality.

We see that the ePaper App is clearly perceived as a product with a high level of UX, while the results for the Portal are clearly worse.

6　Summary and outlook

The UEQ+ offers a set of currently 20 UX scales (available in more than 20 languages) that can be combined to create a UX questionnaire fitting to the needs of a concrete research project. This offers the UX researcher an incredible amount of flexibility to set up a questionnaire.

But this flexibility does not come for free. It requires a good understanding of UX in general and of the aspects relevant for the products that should be evaluated to select the relevant scales. In addition, because of the huge number of possible scale combinations, it is not possible to offer a benchmark or other supporting materials in the same quality as for a questionnaire with static scales, for example the UEQ. Thus, the UEQ+ is not intended for a quick and easy evaluation of a product or prototype (use the UEQ for such cases), but for situations in which a product should be evaluated multiple times (to see how UX quality develops) or where a suite of similar products should be evaluated. For such use cases the advantages of the modular approach clearly outperform the disadvantages. In addition, the need to work out which scales are important can help to develop a deeper common understanding in product teams, so it is maybe worth to use this in bigger development projects [16].

Following this basic idea of the UEQ+ we described two methods to shorten a questionnaire based on the UEQ+ framework and methods to check the impact of these actions on the overall result. Of course, this is not an out of the box approach but more a recipe that can be applied in concrete projects and requires some effort.

We demonstrated the impact of shortening the scales with the help of some empirical examples. In these examples the reuse of importance ratings from other studies and the reduction of the number of items per scale had only a limited and for practical applications acceptable effect of the accuracy of the measurements. However, this depends always on the scales and the investigated product. Thus, it cannot be granted that this always holds. Further investigations with other big data sets will help to clarify this further.

In addition, if you use the UEQ+ framework to set up a modular UX questionnaire we recommend that you use in the first round of measurement the full version (or if this is not possible plan a pilot study with a limited number of participants) and use these data to decide which items can be removed and how big the impact of this action is on the scale value (simply follow the methods shown above). For the importance ratings it is theoretically plausible that they will not change drastically in a shorter period. Thus, it is typically sufficient to measure them once and to reuse these values for other studies. If the product changes drastically concerning the typical use cases it may be required to measure them again, to confirm that the new usage scenarios did not change the relative importance of UX aspects massively.

REFERENCES

[1] Schrepp, M. (2021). User Experience Questionnaires: How to use questionnaires to measure the user experience of your products? KDP, ISBN-13: 979-8736459766.

[2] Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). IJIMAI, 4(6), 103-108.

[3] Moshagen, M., & Thielsch, M. (2013). A short version of the visual aesthetics of websites inventory. Behaviour & Information Technology, 32(12), 1305-1311.

[4] Thielsch, M. T. & Hirschfeld, G., (2020). In der Kürze liegt die Würze: Der Web-CLIC short zur Bewertung digitaler Inhalte. In: Hansen, C., Nürnberger, A. & Preim, B. (Hrsg.), Mensch und Computer 2020 - Workshopband. Bonn: Gesellschaft für Informatik e.V.

[5] Brooke, J. (1996). SUS - A quick and dirty usability scale. Usability evaluation in industry, 189, pp. 4-7.

[6] Moshagen, M., & Thielsch, M. T. (2010). Facets of visual aesthetics. International journal of human-computer studies, 68(10), 689-709.

[7] Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In Mensch & computer 2003 (pp. 187-196). Vieweg+ Teubner Verlag.

[8] Minge, M., & Riedel, L. (2013). meCUE-Ein modularer fragebogen zur erfassung des nutzungserlebens. Mensch & Computer 2013: Interaktive Vielfalt.

[9] Schrepp, M., (2020). A Comparison of UX Questionnaires - What is their underlying concept of user experience? In: Hansen, C., Nürnberger, A. & Preim, B. (Hrsg.), Mensch und Computer 2020 - Workshopband. Bonn: Gesellschaft für Informatik e.V.

[10] Schrepp, M., & Thomaschewski, J. (2019). Design and Validation of a Framework for the Creation of User Experience Questionnaires. International Journal of Interactive Multimedia & Artificial Intelligence, 5(7).

[11] Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. In Symposium of the Austrian HCI and usability engineering group (pp. 63-76). Springer, Berlin, Heidelberg.

[12] Hinderks, A. (2016). Modifikation des User Experience Questionnaire (UEQ) zur Verbesserung der Reliabilität und Validität [Modification of the user experience (UEQ) for the improvement of reliability and validity]. Unpublished Master Thesis, University of Applied Sciences Emden/Leer.

[13] Boos, B. & Brau, H. (2017). Erweiterung des UEQ um die Dimensionen Akustik und Haptik [Extension of the UEQ by the dimensions acoustics and haptics]. In: Hess, S. & Fischer, H. (Eds.), Mensch und Computer [Humans and computers] 2017 – Usability Professionals, Regensburg: Gesellschaft für Informatik e.V., pp. 321-327.

[14] Otten, R., Schrepp, M. & Thomaschewski, J., (2020). Visual clarity as mediator between usability and aesthetics. In: Alt, F., Schneegass, S. & Hornecker, E. (Hrsg.), Mensch und Computer 2020 - Tagungsband. New York: ACM. (S. 11–15).

[15] Klein, A. M., Hinderks, A., Schrepp, M. & Thomaschewski, J., (2020). Construction of UEQ+ scales for voice quality: measuring user experience quality of voice interaction. In: Alt, F., Schneegass, S. & Hornecker, E. (Hrsg.), Mensch und Computer 2020 - Tagungsband. New York: ACM. (S. 1–5).

[16] Hinderks, A., Winter, D., Dominguez Mayo, F.J., Escalona, M.J. & Thomaschewski, J. (in press). UX Poker: Estimating the Influence of User Stories on User Experience in Early Stage of Agile Development. Journal of Interactive Multimedia & Artificial Intelligence.

[17] Schrepp, M. & Thomaschewski, J. (2019). Handbook for the modular extension of the User Experience Questionnaire. - All you need to know to apply the UEQ+ to create your own UX questionnaire. DOI: 10.13140/RG.2.2.15485.20966.

[18] Winter, D., Hinderks, A., Schrepp, M., & Thomaschewski, J. (2017). Welche UX Faktoren sind für mein Produkt wichtig?. Mensch und Computer 2017-Usability Professionals.

[19] Sandkühler, H., Schrepp, M., & Thomaschewski, J. (2020). UX Messung mithilfe des UEQ+ Frameworks. Mensch und Computer 2020-Workshopband.