Die Kürzung des User Experience Questionnaire UEQ

Catherine Alberola[1], Henning Brau[1], Götz Walter[2]

Corporate Strategic Marketing – User Experience, BSH Hausgeräte GmbH [1]
International School of Management, München[2]

Zusammenfassung

Der Fokus dieser Arbeit liegt auf der Kürzung des quantitativen Messinstruments User Experience Questionnaire (UEQ) von 26 auf etwa 13 Items. Hierdurch soll für besondere Anwendungsfälle eine noch effizientere Messung der User Experience (UX) von Produkten ermöglicht werden, unter Beibehaltung einer grundsätzlichen Vergleichbarkeit mit anderen UEQ-Studien. Die Kürzung erfolgt anhand eines Datensatzes eines Unternehmens in der Hausgeräteindustrie, der 1534 Personen umfasste. Mittels Faktor- und Reliabilitätsanalysen konnte eine reliable Kurzversion des Messinstruments entwickelt werden, die nur elf Items umfasst und alle Dimensionen des ursprünglichen Instruments abdeckt. Korrelationen zwischen den Dimensionen der UX und verschiedenen Operationalisierungen der Kundenzufriedenheit weisen auf eine hohe Kriteriumsvalidität der gekürzten Version des UEQ hin. Die neue Kurzversion kann beispielsweise bei seriellem Ausfüllen während einer Testsitzung mit mehreren Stimuli Ermüdung der Teilnehmer vorbeugen.

1 Einleitung

Das Konstrukt der „User Experience“ (UX) oder auch des „Nutzererlebnisses“ (Rauschenberger et al., 2013, S. 72) bezieht sich im Rahmen der Produktinteraktion auf „sämtliche Emotionen, Vorstellungen, Vorlieben, Wahrnehmungen, physiologische und psychologische Reaktionen, Verhaltensweisen und Leistungen, die sich vor, während und nach der Nutzung ergeben“ (DIN EN ISO 9241-210, 2011). Demnach fällt der Nutzer auf Basis seiner Erlebnisse mit interaktiven Produkten und Servicedienstleistungen ein subjektives und umfassendes Urteil (Schrepp & Müller, 2015).

Ein häufig verwendetes quantitatives Instrument zur Messung der UX ist der User Experience Questionnaire (UEQ) (Laugwitz et al., 2009). Im Rahmen von Nutzertests mit Messwiederholungen innerhalb einer Sitzung oder in Längsschnittstudien kann sich der Fragebogen jedoch trotz seines geringen Umfangs als mühsam für die Teilnehmer erweisen. Ermüdungseffekte können dann die Datenqualität beeinträchtigen.

Veröffentlicht durch die Gesellschaft für Informatik e. V. 2017 in
M. Burghardt, R. Wimmer, C. Wolff, C. Womser-Hacker (Hrsg.):
Mensch und Computer 2017 – Tagungsband, 10.–13. September 2017, Regensburg.
Copyright (C) 2017 bei den Autoren. https://doi.org/10.18420/muc2017-mci-0178

Die vorliegende Arbeit wurde in Kooperation mit einem Unternehmen der Hausgeräteindustrie verfasst, welches die Erzeugung und Validierung einer positiven UX als festen Bestandteil in den Produktentwicklungsprozess integriert hat. Das Konsumgüterunternehmen begegnet der oben geschilderten Problematik des Fragebogens regelmäßig bei der Durchführung von Nutzertests mit seriellem Einsatz des UEQ in einer Testsitzung mit mehreren Stimuli. Durch eine Kurzversion des UEQ kann eine höhere Praktikabilität und Zuverlässigkeit in der Anwendung erreicht werden, da ein schneller und mehrmaliger Einsatz des Fragebogens während eines Testdurchlaufs möglich ist. So kann einem zufälligen Ankreuzen von Antwortwerten aufgrund von Ermüdung vorgebeugt werden (Hassenzahl et al., 2008).

2 Der User Experience Questionnaire (UEQ)

Der User Experience Questionnaire (UEQ) ist 2006 von den SAP-Mitarbeitern Laugwitz, Held und Schrepp entwickelt worden und basiert auf dem Modell der hedonischen und pragmatischen Qualitätswahrnehmung (Hassenzahl et al., 2000). Die UX wird somit als ein Konstrukt definiert, das aus der durch den Nutzer wahrgenommenen pragmatischen (PQ) und hedonischen Qualität (HQ) sowie einer allgemeinen Bewertung des Produkts besteht. Beide Qualitäten werden unabhängig voneinander evaluiert (Hassenzahl, 2008). Die Beurteilung sowie Gewichtung der beiden Qualitätsaspekte fließen schließlich gemeinsam in die globale Valenzdimension Attraktivität (ATT), welche die allgemeine Zustimmung oder Ablehnung des Produkts oder der Servicedienstleistung beschreibt (Schrepp, 2015).

Insgesamt besteht der Fragebogen UEQ aus 26 Items in Form von semantischen Differentialen, wobei diese zu sechs Skalen zusammengefasst werden: Attraktivität (ATT), Effizienz (EFF), Steuerbarkeit oder Vorhersagbarkeit (VOR), Durchschaubarkeit (DUR), Stimulation (STI) und Originalität (ORI) (Laugwitz et al., 2006).

Jede der erwähnten Skalen umfasst vier miteinander hoch korrelierende Items – mit Ausnahme der Skala ATT, die über sechs Items verfügt. Die Anzahl der Items wird durch die Autoren mit einer gesteigerten Robustheit begründet (Rauschenberger et al., 2013). EFF, DUR und VOR zählen zur PQ oder pragmatischen bzw. aufgabenbezogenen Qualität. STI und ORI hingegen zählen zur HQ oder hedonischen bzw. nicht-aufgabenbezogenen Qualität (vgl. Abbildung 1) (Laugwitz et al., 2006).

Seit 2009 liegt eine Kurzversion des UEQ vor, der Short UEQ. Allerdings umfasst der Short UEQ noch immer 20 Items und ist daher nur unwesentlich kürzer. Im Rahmen der Kürzung wurden die Skala VOR sowie zwei Items der Skala ATT eliminiert, da diese im Bezug auf die interne Konsistenz und Faktorladungen am schwächsten abschnitten (Sarodnick & Brau, 2016). Grundsätzlich wird beim UEQ davor gewarnt, einzelne Items einer Skala zu eliminieren, da dies die Reliabilität aufgrund der Interdependenz der Skalen schwächt und eine Vergleichbarkeit mit anderen Produkten beziehungsweise Servicedienstleistungen erschwert (Schrepp, 2015).

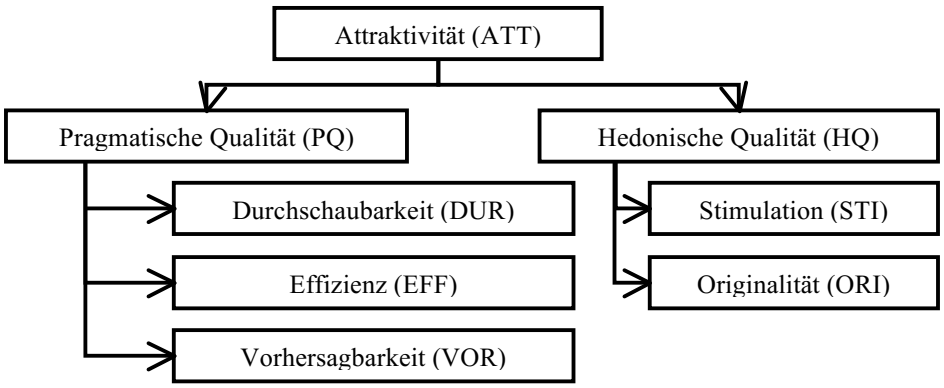

Abbildung 1: Dimensionen und Skalen des UEQ und ihre Beziehung untereinander; Quelle: Hinderks et al., 2012.

Das Ziel dieser Arbeit ist die Kürzung des UEQ um etwa die Hälfte seiner Items, ohne dabei große Einbußen in Reliabilität und Validität des Messinstruments hinzunehmen. Die Kürzung erfolgt anhand eines Datensatzes aus der Hausgeräteindustrie. Die praktische Relevanz einer gekürzten Version des UEQ erwächst aus der Forderung nach einem seriellen Einsatz bei mehreren Stimuli in einer einzigen Session unter Beibehaltung einer grundsätzlichen Vergleichbarkeit mit „normalen“ UEQ-Studien.

3 Methode

3.1 Stichprobe

Die Kürzung des UEQ wurde anhand eines Online-Panels vorgenommen, welches 1534 Teilnehmer umfasste. Alle Teilnehmer hatten innerhalb der letzten 36 Monate ein Hausgerät einer bestimmten Marke erworben und füllten den UEQ sowie einige ergänzende Skalen und Items zum gekauften Produkt aus.

Die Probanden gehörten zu gleichmäßigen Anteilen den Altersklassen 25 bis 29 Jahre, 30 bis 34 Jahre, 35 bis 39 Jahre, 40 bis 44 Jahre, 45 bis 49 Jahre, 50 bis 55 Jahre und 56 bis 65 Jahre an. Insgesamt wurden 786 Frauen und 748 Männer befragt. Die Stichprobe war in drei Länder unterteilt: 971 Teilnehmer kamen aus Deutschland (63,3 %), 366 Teilnehmer aus der Türkei (23,9 %) und 197 Teilnehmer aus China (12,8 %). Bezüglich des Nutzungszeitraums gaben 327 Teilnehmer (21.3 %) bis zu sechs Monate an, 539 Teilnehmer (35.1 %) bis zu einem Jahr, 455 Teilnehmer (29.7 %) bis zu zwei Jahre und 213 Teilnehmer (13.9 %) bis zu drei Jahre. Das Online-Panel bewertete fünf verschiedenen Gerätekategorien: Waschmaschinen (n=382; 24,9 %), Geschirrspüler (n=289; 18,8 %), Einbauherde (n=277; 18,1 %), Kühlschrank-Gefrierschrank-Kombinationen (n=390; 25,4 %) und Staubsauger (n=196; 12,8 %). Es wurden nicht in jedem Land zu jeder Kategorie Daten erhoben. Eine Präzisierung bezüglich des jeweiligen Gerätemodells erfolgte bei der Datenerhebung nicht.

3.2 Datenerhebung

Die Datenerhebung erfolgte im Jahr 2015. Die Umfrage bestand aus einem Online-Fragebogen mit fünf Teilen. Zunächst wurden allgemeine Daten zur Nutzung von Hausgeräten sowie anschließend zum spezifischen Gerät abgefragt. Des Weiteren sollen die Teilnehmer Fragen zur Zufriedenheit mit dem Gerät beantworten sowie biographische Angaben machen. Abschließend erfolgt die Bewertung der UX des Geräts mithilfe des UEQ in der jeweiligen Muttersprache. Der Fragebogen enthielt insgesamt 46 Items. Den Teilnehmern wurde eine anonyme Auswertung der Daten garantiert.

3.3 Analysemethoden

Die Kürzung des UEQ erfolgte im ersten Schritt mittels eines faktorenanalytischen Vorgehens. In Anlehnung an den Konstruktionsprozess des UEQ wurden die Daten in zwei Gruppen geteilt – eine Gruppe mit den sechs Items der Skala ATT und eine mit den übrigen 20 Items. In jeder der zwei Gruppen wurden im Folgenden explorative Faktorenanalysen gerechnet. Die Faktorenextraktion erfolgte mittels Hauptkomponentenanalyse und der Varimax-Rotation (Moosbrugger & Schermelleh-Engel, 2012). Die Anzahl der Faktoren wurde basierend auf den drei Dimensionen des UEQ (ATT, PQ, HQ) für die erste Itemgruppe auf eins und für die zweite Gruppe auf zwei festgelegt.

Im nächsten Schritt wurden sukzessive die Items mit der geringsten Kommunalität und der geringsten Itemladung aus den Faktorenanalysen entfernt. In allen Schritten der Faktorenanalysen lag das KMO-Kriterium über .70 und der Bartlett-Test erreichte Signifikanz ($p < .05$), sodass davon auszugehen ist, dass die Daten die Voraussetzungen für das Rechnen einer explorativen Faktorenanalyse erfüllen. Ziel war es, jeweils eine Faktorlösung zu identifizieren, (1) die mindestens 60 % der Varianz der Items erklären kann, (2) deren Items Kommunalitäten und Faktorladungen von > .60 aufweisen und eindeutig auf einen Faktor laden (Field, 2009), und (3) die in Summe aus maximal 13 Items besteht. Für die zweite Gruppe von Items war ein weiteres Kriterium, dass (4) jede der fünf Skalen des UEQ (EFF, VOR, DUR, STI, ORI) mit mindestens einem Item in der gekürzten Version vertreten ist, um das Konstrukt UX ganzheitlich abzubilden.

In Anschluss an die Faktorenanalysen wurde zusätzlich je identifiziertem Faktor (ATT, PQ, HQ) eine Reliabilitätsanalyse durchgeführt. Das Cronbachs Alpha (α) sollte jeweils mindestens einen Wert von .70 erreichen (George & Mallery, 2003). Die Reliabilitätsanalysen wurden außerdem ergänzend für die Items gerechnet, die jeweils eine Dimension in der ursprünglichen Version des UEQ bilden, um einen Vergleich der internen Konsistenz der kurzen und langen Version des UEQ zu ermöglichen. Zu guter Letzt wurde über alle Dimensionen des gekürzten UEQ eine konfirmatorische Faktorenanalyse gerechnet, um eine Einschätzung der Güte des Messmodells zu erhalten.

Die Daten des Onlinepanels erlaubten außerdem eine Einschätzung der Kriteriumsvalidität sowohl des originären UEQ (Langversion) als auch des gekürzten UEQ. Die Validität wird durch die Korrelation der drei UEQ-Dimensionen mit sechs Items zur Produktzufriedenheit überprüft. Drei Items erfassen hierbei die allgemeine Zufriedenheit mit der Produktinterak-

tion (Nutzerzufriedenheit, Weiterempfehlung, erneuter Kauf) und drei weitere Items erfassen die Zufriedenheit mit speziellen Produktaspekten (Nutzerfreundlichkeit, Design, innovative Features). Alle Items wurden über eine fünfstufige Likert-Skala abgefragt. Ein hoher Zusammenhang zwischen diesen Items und den UEQ-Dimensionen gilt als Kriterium für ein positives beziehungsweise negatives Nutzererlebnis (Moosbrugger & Kelava, 2012). Eine mittlere Kriteriumsvalidität ist gegeben, wenn die Korrelation einen Wert von mindestens r = +/- .30 annimmt (Bühner & Ziegler, 2009).

Aus bisherigen Veröffentlichungen zur UX und dem UEQ lassen sich weitere Kriterien bezüglich eines gekürzten UEQ ableiten. Dies führt zu sechs Postulaten, deren empirische Überprüfung eine Einschätzung der Reliabilität und Validität des gekürzten Messinstruments ermöglicht: (1) Die Valenzdimension ATT sollte am stärksten von allen Dimensionen mit den drei allgemeinen Zufriedenheits-Items positiv korrelieren. Für die Gruppe der Hausgeräte sind besonders die pragmatischen Skalen wichtig (EFF, DUR, VOR), ebenso wie das Erreichen der Usability-Ziele während der Nutzung des Produkts (Koller & Andert, 2004). Demnach wird postuliert, dass (2) die positive Korrelation zwischen PQ und den allgemeinen Zufriedenheits-Items größer sein soll als die Korrelation zwischen HQ und diesen Items. Weiterhin sollte gelten, dass (3) PQ und HQ nicht oder nur sehr niedrig miteinander korrelieren und gleichzeitig hoch positiv mit der Dimension ATT (Schrepp et al., 2013). Bezüglich der Zufriedenheit mit spezifischen Produktaspekten wird postuliert, dass (4) das Item „Nutzerfreundlichkeit“ positiv mit der PQ korrelieren soll, (5) „Design“ positiv mit der HQ korrelieren soll und (6) das Item „innovative Features“ positiv mit der HQ korrelieren soll, da dort der Aspekt Originalität abgefragt wird.

4 Ergebnisse

Durch das beschriebene Vorgehen konnte eine gekürzte Version des UEQ mit elf Items erstellt werden, die alle postulierten Anforderungen erfüllt und somit als reliabel und valide angesehen werden kann. Tabelle 1 gibt einen Überblick über den gekürzten UEQ, inklusive einer Zuordnung der elf Items auf die drei Dimensionen ATT, PQ und HQ.

Die explorativen Faktorenanalysen zu den sechs Items der Dimension ATT führen zu der Löschung von zwei Items („attraktiv-unattraktiv“ sowie „sympathisch-unsympathisch“), sodass diese Skala nach der Kürzung noch vier Items enthält. Die Kommunalitäten aller verbleibenden Items der Skala ATT sind größer als .70 und der extrahierte Faktor erklärt 74.5 % der Varianz der vier Items. Alle Itemladungen sind höher als .80.

Die explorativen Faktorenanalysen zu den 20 Items der Dimensionen PQ und HQ führten zu einer Eliminierung von 13 Items („uninteressant-interessant“, „unberechenbar-voraussagbar“, „schnell-langsam“, „herkömmlich-neuartig“, „sicher-unsicher“, „erwartungskonform-nicht erwartungskonform“, „unpragmatisch-pragmatisch“, „aufgeräumt-überladen“, „konservativ-innovativ“, „leicht zu lernen-schwer zu lernen“, „übersichtlich-verwirrend“ und „langweilig-spannend“). Von den sieben verbleibenden Items laden vier auf den Faktor 1 (PQ) und drei auf den Faktor 2 (HQ). Die Kommunalitäten aller verbleibenden Items sind

größer als .65 und die extrahierten Faktoren erklären 73.4 % der Varianz der sieben Items. Alle Items laden auf ihren jeweiligen Faktor mit mindestens .80, mit der Ausnahme des Items „ineffizient – effizient“, welches über eine ebenfalls sehr hohe Faktorladung von .72 verfügt. Die Faktorladungen der Items mit dem jeweils anderen Faktor liegen allesamt unter .31, womit eine eindeutige Zuordnung von Items auf Faktoren gegeben ist.

ItemNR in Originalversion	Item	Faktor
1	unerfreulich - erfreulich	ATT
2	unverständlich -verständlich	PQ
3	kreativ - phantasielos	HQ
5	wertvoll - minderwertig	HQ
11	behindernd - unterstützend	PQ
12	gut – schlecht	ATT
13	kompliziert - einfach	PQ
14	abstoßend - anziehend	ATT
16	unangenehm - angenehm	ATT
18	aktivierend - einschläfernd	HQ
20	ineffizient - effizient	PQ

Tabelle 1: Auflistung der Items der gekürzten UEQ-Version sowie des jeweiligen Faktors.

Die Cronbach-Alphas der Dimensionen ATT, PQ und HQ in der gekürzten Version des UEQ liegen allesamt über dem Wert von .80 und weisen auf eine hohe Reliabilität des Messinstruments hin. Ein Vergleich mit den internen Konsistenzen unter Beachtung aller 26 Items auf Dimensionsebene zeigt, dass die interne Konsistenz der Kurzversion etwas abgenommen hat, aber immer noch in einem sehr guten Bereich liegt (vgl. Tabelle 2).

Dimension	Cronbachs Alpha (gekürzte Version)	Cronbachs Alpha (originale Version)
ATT	.89	.91
PQ	.86	.92
HQ	.85	.88

Tabelle 2: Gegenüberstellung der Werte der internen Konsistenz beider Fragebogenversionen auf Dimensionsebene.

Die konfirmatorische Faktorenanalyse über die drei Dimensionen des gekürzten UEQ weist ebenfalls sehr gute Ergebnisse vor (vgl. Abbildung 2). Die standardisierten Koeffizienten und R^2 der Items liegen allesamt über .60 (mit Ausnahme des R^2 von Item 13, das einen Wert

von .49 erreicht). Alle Fit-Indizes des Modells liegen im guten bis akzeptablen Bereich (RMSEA = .065, CFI = .976, SRMR = .031).

Bezüglich der Kriteriumsvalidität erreicht der gekürzte UEQ gute Werte, die allerdings nicht an das Niveau der originären Langversion heranreichen. Die Korrelationen r der UEQ-Dimensionen der gekürzten Fragebogenversion mit den drei allgemeinen Zufriedenheits-Items liegen zwischen .34 und .51 ($p < .01$). Im Vergleich dazu sind die Korrelationswerte der originalen UEQ-Version etwas höher und reichen von r = .41 bis .53 ($p <.01$) (vgl. Tabelle 3).

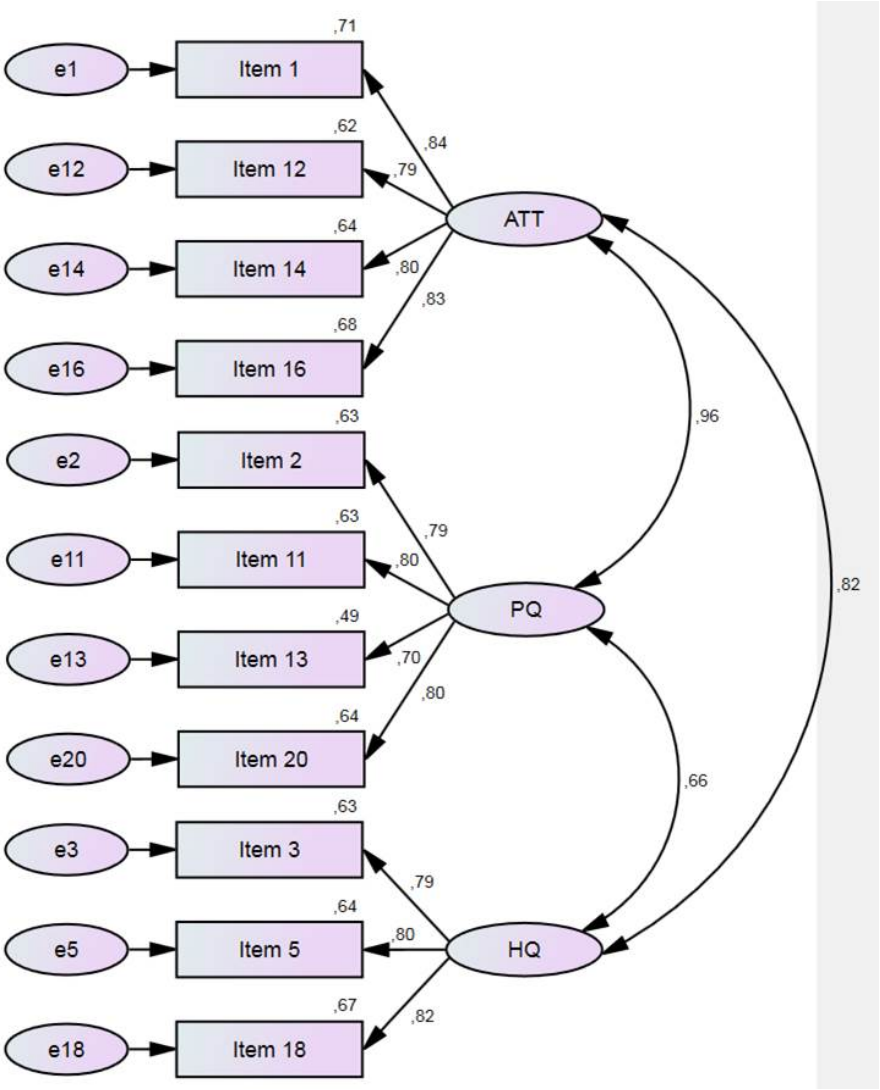

Abbildung 2: Ergebnisse einer konfirmatorischen Faktorenanalyse mit den drei Dimensionen des gekürzten UEQ.

Auch bezüglich der weiteren Postulate zeigt der gekürzte UEQ gute Ergebnisse: Die Dimension ATT korreliert am höchsten von allen drei Dimensionen mit den sechs Items zur allgemeinen Zufriedenheit. Postulat (1) ist somit erfüllt. Die Dimension HQ erzielt in beiden Tests bessere Werte als die Dimension PQ, was zur Ablehnung des zweiten Postulats (2) führt. Das dritte Postulat (3) ist nur teilweise erfüllt: Die beiden Dimensionen PQ und HQ korrelieren zwar in beiden Fragebogenversionen hoch mit der Dimension ATT, doch korrelieren die Dimensionen PQ und HQ ebenfalls hoch miteinander, wobei dieser Zusammenhang in der gekürzten Version abnimmt (vgl. Tabelle 3).

Korrelation nach Pearson	**Nutzerzufriedenheit**	**Weiterempfehlung**	**erneuter Kauf**
ATT-MW (lang)	.46**	.53**	.51**
PQ-MW (lang)	.41**	.47**	.46**
HQ-MW (lang)	.46**	.51**	.49**
ATT-MW (kurz)	.45**	.51**	.50**
PQ-MW (kurz)	.34**	.40**	.41**
HQ-MW (kurz)	.43**	.48**	.46**

***. Die Korrelation ist auf dem Niveau von .01 (2-seitig) signifikant.*
Nutzerzufriedenheit: n= 1534; Weiterempfehlung: n= 1525; erneuter Kauf: n= 1528;

Tabelle 3: Vergleich der Korrelationswerte der Dimensionen der langen und der gekürzten UEQ-Version mit den allgemeinen Zufriedenheits-Items.

Die Postulate (4), (5) und (6) sind wiederum erfüllt: Sowohl in der Kurz- als auch der Langversion des UEQ korreliert die Dimension HQ stärker mit der Bewertung von Design und innovativen Features als die Dimension PQ, während für die Bewertung der Nutzerfreundlichkeit das Gegenteil gilt (vgl. Tabelle 4).

	Zufriedenheit mit		
Korrelation nach Pearson	**Design**	**Nutzerfreundlichkeit**	**innovativen Features**
ATT-MW (lang)	.50**	.53**	.52**
PQ-MW (lang)	.44**	.53**	.46**
HQ-MW (lang)	.50**	.48**	.58**
ATT-MW (kurz)	.47**	.51**	.49**
PQ-MW (kurz)	.38**	.50**	.38**
HQ-MW (kurz)	.47**	.44**	.56**

***. Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.*
Design: n= 1528; Nutzerfreundlichkeit: n= 1533; innovative Features: n= 1474;

Tabelle 4: Vergleich der Korrelationswerte der Dimensionen der langen und der gekürzten UEQ-Version mit den Items der Zufriedenheit mit einzelnen Produktaspekten.

5 Diskussion

Mit dem gekürzten UEQ liegt ein reliables und valides Messinstrument vor. Das Ziel dieser Studie wurde somit erreicht. Die Anzahl der Items ist um mehr als 50 % gekürzt worden, sodass ein effizienter und dennoch effektiver Einsatz in der Praxis möglich ist. Trotz geringerer Itemanzahl ist die interne Konsistenz weiterhin sehr hoch. Die Reliabilität der gekürzten Version weicht nur geringfügig von der Originalversion ab. Die angeführte Problematik der Eliminierung einzelner Items nach Schrepp (2015) konnte somit nicht bestätigt werden.

Die Kriteriumsvalidität des gekürzten UEQ ist ebenfalls als befriedigend zu bewerten. Wie erwartet, sind die Korrelationen der drei Dimensionen des gekürzten UEQ mit den drei Items „Nutzerzufriedenheit“, „Weiterempfehlung“ und „erneuter Kauf“ allesamt signifikant und im mittleren bis hohen Bereich angesiedelt, allerdings etwas geringer als die gleichen Korrelationen in der Originalfassung des UEQ. Von sechs zusätzlichen Postulaten, die den Zusammenhang der drei Dimensionen untereinander und mit allgemeinen und spezifischen Zufriedenheitsitems betreffen, finden vier Postulate ihre Bestätigung in den Daten. Entgegen der Erwartungen erzielt die Dimension HQ bessere Werte als die Dimension PQ, und zwar sowohl in der Kurzversion des UEQ als auch in der Originalfassung. Dies könnte eventuell mit der Zwei-Faktoren-Theorie nach Herzberg et al. (1959) begründet werden. Demnach bewahrt eine gute PQ vor der Unzufriedenheit des Nutzers. Wenn aber ein Grundwert an PQ erreicht worden ist, kann die Steigerung dieser nicht zu weiterer Zufriedenheit beim Nutzer führen. Die HQ ist demnach der Motivator. Sind die Werte positiv, ist auch die Zufriedenheit hoch. Sind die HQ-Werte allerdings eher niedrig, bedeutet dies nicht automatisch, dass der Nutzer unzufrieden ist (Hassenzahl et al., 2008). Weiterhin wurde entgegen Laugwitz et al. (2006) eine starke Korrelation zwischen den beiden Dimensionen PQ und HQ in den Daten vorgefunden – sowohl in der langen, als auch in der gekürzten UEQ-Version.

Es wird empfohlen, die Reihenfolge der Items der gekürzten UEQ-Version in Anlehnung an die Vollversion beizubehalten, um eine zufällige Reihenfolge zu gewährleisten, und die Anwendung des gekürzten UEQ auf die Dimensionen ATT, PQ und HQ zu beschränken. Die Skalen umfassen teilweise nur noch ein Item (ORI, EFF und VOR) und werden daher nur unvollständig abgebildet. Das Modell, dem der UEQ zugrunde liegt, wird weiterhin durch die Struktur der Kurzversion abgebildet, allerdings nur auf Dimensions- und nicht auf Skalenebene.

Es ist eine ausreichende Kriteriumsvalidität nachgewiesen worden. In einem weiteren Schritt wäre es zudem sinnvoll, auch die Konstruktvalidität zu prüfen, indem die Korrelation mit einem konstruktverwandten Test berechnet wird (Bühner, 2011). Denkbar wäre dies beispielsweise durch eine Korrelation mit dem AttrakDiff2, da dieses Instrument als Vorlage des UEQ gilt und bereits zu dessen Validierung verwendet worden ist. Es wäre ebenfalls sinnvoll, im Rahmen der Kürzung das UEQ-Analyse-Tool entsprechend anzupassen, sodass eine schnelle Auswertung nach Nutzertests möglich ist.

Es gilt anzumerken, dass die genutzten Zufriedenheit-Items der Studie nur einzelne, sehr globale Aussagen darstellen und das Konstrukt Zufriedenheit nicht vollständig abbilden. Eine weitere Validierungsstudie mit einer detaillierteren Erhebung der Nutzerzufriedenheit

wäre wünschenswert. Eine Verallgemeinerung der Studienergebnisse auf andere Produkte und Services ist nur begrenzt möglich, da es sich um einen Datensatz zu spezifischen Produkten, nämlich Hausgeräte, handelt. Ein weiterer möglicher Kritikpunkt am Datensatz ist, dass es sich um Produkte handelt, welche die Tester zum Teil schon seit längerer Zeit benutzen. Zudem mussten die Probanden im Rahmen des Tests keine Aufgabe mit den Hausgeräten erledigen, sondern das Erlebnis aus ihren Erinnerungen bewerten. Problematisch ist hierbei, dass sich Nutzer vor allem an Erlebnisse erinnern, die sie erst vor kurzem durchlebt haben. Es kann daher gegebenenfalls zu einer verzerrten Beurteilung des Nutzererlebnisses gekommen sein (Hassenzahl & Sandweg, 2004).

6 Fazit

Im Rahmen dieser Arbeit wurde eine mögliche Kurzversion des UEQ vorgestellt. Es kann in diesem Bereich eine Forschungslücke geschlossen werden: Es gibt zwar bislang eine etwas kürzere Version des Fragebogens, den „Short UEQ“, dieser unterscheidet sich allerdings in der Länge nur unwesentlich von der Originalversion. Zudem wurde bei dieser Version eine Skala vollständig ausgeschlossen. Durch die in dieser Arbeit entwickelte Kurzversion erfolgt eine umfassende Abbildung des Konstrukts der UX, indem noch immer alle Skalen der Originalversion durch Items abgebildet werden. Zudem ist das abgewandelte Messinstrument um 50 % kürzer als das Original. Der praktische Nutzen dieser Kürzung ist die Ermöglichung einer effizienteren Durchführung von Nutzertests. Eine mehrfache Erhebung von Daten einer Testperson innerhalb eines Tests ist ebenfalls möglich. Durch diesen Fragebogen besteht nun ein noch schlankeres aber dennoch valides und reliables Instrument zur Messung der UX. Es wird allerdings empfohlen, die Anwendung des gekürzten UEQ auf die Dimensionen ATT, PQ und HQ zu beschränken, und zusätzliche Validierungsstudien durchzuführen, insbesondere um die Verallgemeinerung der Studienergebnisse zu überprüfen.

Literaturverzeichnis

Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion*. 3. Aufl. München: Pearson Studium.

Bühner, M.& Ziegler, M. (2009). *Statistik für Psychologen und Sozialwissenschaftler*. München: Pearson Studium.

DIN EN ISO 9241-210 (2011). *Ergonomie der Mensch-System-Interaktion – Teil 210: Prozess zur Gestaltung gebrauchstauglicher interaktiver Systeme*. Berlin: Beuth.

Field, A. (2009). *Discovering Statistics using SPSS*. Los Angeles: Sage.

George, D. & Mallery, P. (2003). *SPSS for Windows step by step. A simple guide and reference*. 11.0 Update. 4. Aufl. Boston, MA: Allyn & Bacon.

Hassenzahl, M. (2008). User Experience (UX). Towards an experiential perspective on product quality. In Brangier, E., Michel, G., Bastien, J. M. C. & Carbonell, N. (Hrsg.): *Proceedings of the 20th international conference on Association Francophone d'Interaction Homme-Machine*. Metz: Vieweg & Teubner, S. 11-15.

Hassenzahl, M., Burmester, M. & Koller, F. (2008). Der User Experience (UX) auf der Spur. Zum Einsatz von www.attrakdiff.de. In Brau, H., Diefenbach, S., Hassenzahl, M., Koller, F., Peissner, M. & Röse, K. (Hrsg.): *Usability Professionals 2008*. Stuttgart: German UPA e.V., S. 78-82.

Hassenzahl, M., Platz, A., Burmester, M. & Lehner, K. (2000). Hedonic and Ergonomic Quality Aspects Determine a Software's Appeal. *Proceedings of the CHI 2000*, S. 201-208.

Hassenzahl, M. & Sandweg, N. (2004). From Mental Effort to Perceived Usability. Transforming Experiences into Summary Assessments. *Proceedings of the CHI 2004*, S. 1283-1286.

Herzberg, F., Mausner, B. & Snyderman, B. (1959). *The Motivation to Work*. New York, NY: Wiley.

Hinderks, A., Schrepp, M., Rauschenberger, M., Olschner, S. & Thomaschewski, J. (2012). Konstruktion eines Fragebogens für jugendliche Personen zur Messung der User Experience. In Brau, H., Lehmann, A., Petrovic, K. & Schroeder, M. C. (Hrsg.): *Usability Professionals 2012.* Stuttgart: German UPA e.V., S. 78-83.

Koller, F. & Andert, M. (2004). Total User Experience für Hausgeräte. Ganzheitlich konzipieren und testen. In Hassenzahl, M. & Peissner, M. (Hrsg.): *Usability Professionals 2004*. Stuttgart: German UPA e.V., S. 70-73.

Laugwitz, B., Schrepp, M. & Held, T. (2006). Konstruktion eines Fragebogens zur Messung der User Experience von Softwareprodukten. In Heinecke, A. M. & Paul, H. (Hrsg.): *Mensch & Computer 2006. München*: De Gruyter, S. 125-134.

Laugwitz, B., Schubert, U., Ilmberger, W., Tamm, N., Held, T. & Schrepp, M. (2009). Subjektive Benutzerzufriedenheit quantitativ erfassen. Erfahrungen mit dem User Experience Questionnaire UEQ. In Brau, H., Diefenbach, S., Hassenzahl, M., Kohler, K., Koller, F., Peissner, M., Petrovic, K., Thielsch, M., Ullrich, D. & Zimmermann, D. (Hrsg.): *Usability Professionals 2009*. Stuttgart: German UPA e.V., S. 220-225.

Moosbrugger, H. & Kelava, A. (2012). Qualitätsanforderungen an einen psychologischen Test (Testgütekriterien). In Moosbrugger, H. & Kelava, A. (Hrsg.): *Testtheorie und Fragebogenkonstruktion*. 2. Aufl. Heidelberg: Springer, S. 7-26.

Moosbrugger, H. & Schermelleh-Engel, K. (2012). Exploratorische (EFA) und Konfirmatorische Faktorenanalyse (CFA). In Moosbrugger, H. & Kelava, A. (Hrsg.): *Testtheorie und Fragebogenkonstruktion*. 2. Aufl. Heidelberg: Springer, S. 325-343.

Rauschenberger, M., Thomaschewski, J. & Schrepp, M. (2013). User Experience mit Fragebögen messen, Durchführung und Auswertung am Beispiel des UEQ. In Brau, H., Lehmann, A., Petrovic, K. & Schroeder, M. (Hrsg.): *Usability Professionals 2013*. Stuttgart: German UPA e.V., S. 72-76.

Sarodnick, F. & Brau, H. (2016). *Methoden der Usability Evalution. Wissenschaftliche Grundlagen und praktische Anwendung*. 3. Aufl. Bern: Hogrefe.

Schrepp, M. (2015). User Experience Questionnaire Handbook. All you need to know to apply the UEQ successfully in your projects. http://www.ueq-online.org/ (abgerufen am 30.05.2016).

Schrepp, M. & Müller, K. E. (2015). Übersichtlichkeit als Mediator zwischen Ästhetik und Usability? In Diefenbach, S., Henze, N. & Pielot, M. (Hrsg.): *Mensch & Computer 201*5. Stuttgart: De Gruyter, S. 73-82.

Schrepp, M., Olschner, S. & Schubert, U. (2013). User Experience Questionnaire Benchmark. Praxiserfahrungen zum Einsatz im Business-Umfeld. In Brau, H., Lehmann, A., Petrovic, K. & Schroeder, M. (Hrsg.): *Usability Professionals 2013*. Stuttgart: German UPA e.V., S. 348-353.

Autoren

Alberola, Catherine

Catherine Alberola studierte Wirtschaftspsychologie an der International School of Management in München. Sie hat diese Studie im Rahmen ihrer Masterarbeit zum Thema "Erweiterung des User Experience Questionnaire zur Anwendung auf Produkte und Services in der Hausgeräteindustrie" in Kooperation mit der UX-Abteilung eines Hausgeräteunternehmens durchgeführt. In diesem Zusammenhang konnte sie erste praktische Erfahrungen im Bereich der UX sammeln.

Brau, Henning

Henning Brau studierte Psychologie an der TU Berlin. Danach folgte freiberufliche Tätigkeit als Marktforscher und Usability-Consultant. Er war ab Mai 2003 zunächst in der Forschung, später im zentralen IT Management der Daimler AG tätig. Dort war er bis 2010 verantwortlich für das Themenfeld User-Centered Technologies. Ab 2010 leitete er als 'Director of User Experience Design' den Münchener Standort der User Interface Design GmbH („UID“). Seit 2015 ist er als Manager Corporate UX verantwortlich für die Entwicklung des UX Reifegrads aller Geschäftsbereiche der BSH Hausgeräte GmbH in München sowie für die Einführung von Design Thinking. Von 2007 bis 2014 war er im Vorstand der German UPA tätig.

Walter, Götz

Prof. Dr. Götz Walter ist Hochschullehrer für Psychology & Management an der International School of Management (ISM) in München. Er studierte Diplom-Psychologie an der Universität Regensburg und als DAAD-Jahresstipendiat an der University of Melbourne (Australien). Nach Studienabschluss war Prof. Dr. Götz Walter sieben Jahre als Unternehmensberater für die Energiewirtschaft bei der Schweizer Managementberatung The Advisory House tätig. Er promovierte berufsbegleitend am Lehrstuhl Sozialpsychologie der Universität Zürich zum Thema „Erneuerbare-Energie-Kraftwerksprojekte: Dimensionen von Akzeptanz und Opposition“. Seine Lehr- und Forschungsschwerpunkte sind Change Management, Consulting und Konsumentenpsychologie.