

Konstruktion eines Fragebogens für jugendliche Personen zur Messung der User Experience



Andreas Hinderks

RMT Soft GmbH & Co. KG
Carl-Zeiss-Str. 14
28816 Stuhr
andreas@hinderks.org

Martin Schrepp

SAP AG – User Experience
Dietmar-Hopp-Allee 16
69190 Walldorf
martin.schrepp@sap.com

Maria Rauschenberger

MSP Medien-Systempartner GmbH & Co. KG
Peterstraße 28-34
26121 Oldenburg
maria.rauschenberger@medien-systempartner.de

Siegfried Olschner

DATEV eG
90329 Nürnberg
siegfried.olschner@datev.de

Jörg Thomaschewski

Hochschule Emden/Leer
Constantiaplatz 4
26723 Emden
joerg.thomaschewski@hs-emden-leer.de

Abstract

Der User Experience Questionnaire (UEQ) ist ein etablierter Fragebogen zur Messung des Nutzungserlebens, der als semantisches Differential mit 26 Gegensatzpaaren realisiert ist. Fragebögen dieses Typs setzen ein genaues Verständnis der verwendeten Items durch die befragten Personen voraus. Beim Einsatz des UEQ zur Befragung von jugendlichen Nutzern einer Web-Seite wurden Probleme in Bezug auf die Verständlichkeit einiger Begriffe durch diese Zielgruppe beobachtet. Hauptprobleme waren Fremdwörter und Begriffe, deren konkrete Interpretation im gegebenen Befragungskontext ein höheres Maß an Abstraktionsfähigkeit voraussetzt. Um diese Probleme zu beheben, wurde eine spezielle Version des UEQ in einfacher Sprache konstruiert. Wir beschreiben in diesem Beitrag den Konstruktionsprozess und erste Evaluationsstudien dieser Version.

Keywords:

/// UEQ
/// User Experience
/// UX Evaluation
/// UX Beurteilung

1. Einleitung

User Experience fasst alle Aspekte zusammen, die für die subjektive Bewertung eines Produkts durch seine Nutzer von Bedeutung sind. Dazu zählen neben klassischen Usability-Aspekten, wie z. B. Effizienz, Steuerbarkeit, Fehlertoleranz oder Erlernbarkeit (siehe z. B. Molich & Nielsen, 1990 oder ISO 9241-110), auch Aspekte wie Joy of Use (Hatscher, 2001; Norman, 2003) oder ästhetisches Design (Kurosu & Kashimura, 1995; Tractinsky, 1997). Die User Experience ist dabei laut Hassenzahl (2010) ein Zusammenspiel aus pragmatischer und hedonischer Qualität, die vor, während oder nach der Benutzung eines Produktes erlebt wird.

Eine gute User Experience (=Nutzungserlebnis) ist ein zentraler Faktor für den Markterfolg eines Produkts. Falls hier Probleme bestehen, müssen diese deshalb schnell erkannt und behoben werden. Aus diesem Grund ist es eine wichtige Anforderung, die User Experience eines

Produkts schnell, einfach und kostengünstig messen zu können.

Da es sich beim Konzept der User Experience um eine rein subjektive Einschätzung eines Produkts durch seine Nutzer handelt, bieten sich Fragebögen als einfache und kostengünstige Messmethode an. Es existieren hierzu im deutschsprachigen Raum insbesondere die Fragebögen UEQ – User Experience Questionnaire (Laugwitz, Schrepp & Held, 2006) und AttrakDiff (Hassenzahl, Burmester & Koller, 2003).

Wir beschreiben in diesem Beitrag die Konstruktion einer Variante des UEQ für die Zielgruppe jugendlicher Nutzer. Ziel des UEQ ist eine effiziente Messung des Gesamteindrucks, den ein Benutzer in Bezug auf die Interaktion mit einem Produkt entwickelt hat. Der UEQ besteht aus 26 bipolaren Begriffen (Gegensatzpaare), die die Form eines 7-stufigen semantischen Differentials haben, z. B.:

kompliziert ○ ○ ○ ○ ○ ○ **einfach**

Die Items sind den folgenden sechs Skalen zugeordnet: **Effektivität, Durchschaubarkeit, Vorhersagbarkeit, Stimulation, Originalität** (jeweils 4 Items), sowie **Attraktivität** (6 Items).

Die Skalen **Durchschaubarkeit, Effizienz, und Steuerbarkeit** beschreiben aufgabenbezogene (pragmatische) Qualitätsmerkmale eines Produkts. Die Skalen **Stimulation** und **Originalität** beschreiben nicht aufgabenbezogene (hedonische) Qualitätsmerkmale. Die Skala **Attraktivität** ist eine reine Valenzdimension, die eine positive, bzw. negative Einstellung/Befindlichkeit gegenüber dem Produkt darstellt. **[Abb. 1]**

Zusätzlich zur deutschen Originalversion des UEQ sind auch Versionen in englischer, spanischer (Rauschenberger et al., 2012), französischer und italienischer Sprache verfügbar.

In mehreren Validierungs-Studien zur deutschen und englischen Version konnte

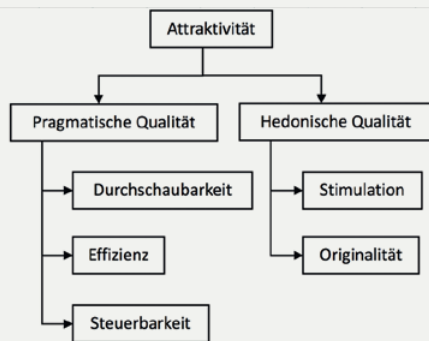


Abb. 1.
Skalen des UEQ und ihre
Beziehung untereinander.

gezeigt werden, dass die Skalen des UEQ eine hohe Reliabilität bzw. Zuverlässigkeit aufweisen. Weiterhin deuten die bisher durchgeführten Validierungs-Studien auf eine zufriedenstellende Konstruktvalidität hin (z. B. Laugwitz, Schrepp & Held, 2006 oder Laugwitz, Held & Schrepp, 2008).

Die Items des UEQ wurden über eine Faktorenanalyse aus einem großen Pool von potentiellen Items extrahiert. Details zur Konstruktion des Fragebogens sind in Laugwitz, Schrepp & Held (2006) und Laugwitz, Held & Schrepp (2008) beschrieben.

Informationen zum Einsatz des UEQ in verschiedenen Evaluationsszenarien finden sich in z. B. in Laugwitz, Schubert, Ilmberger, Tamm, Held & Schrepp (2009), Rauschenberger, Hinderks & Thoma-schewski (2011), Hartmann (2011), Dicke, Jakus, Tomazic & Sodnik (2012), Ebner, Stickel, Scerbakov & Holzinger (2009) oder Wieschnowsky & Paulheim (2011).

2. Probleme beim Einsatz des UEQ mit jugendlichen Nutzern

Eine Anwendungsvoraussetzung von semantischen Differentialen ist es, dass die Teilnehmer einer Befragung die Bedeutung der Gegensatzpaare relativ gut verstehen müssen, um ein gutes Ergebnis zu erzielen. Das erschwert den Einsatz bei Nicht-Muttersprachlern oder bei

Personengruppen mit eingeschränkterem Wortschatz. Bei Versuchen den UEQ zur Evaluation einer Web-Seite (www.stayblue.de) durch überwiegend jugendliche Nutzer einzusetzen, wurden Schwierigkeiten hinsichtlich der Verständlichkeit einiger Items festgestellt.

Die beobachteten Schwierigkeiten jugendlicher Nutzer bei der Fragebogenanwendung können dabei in zwei Kategorien eingeteilt werden:

- Das Item enthält Fremdwörter, deren exakte Bedeutung den Nutzern nicht bekannt ist. Beispiele waren die Begriffspaare **konservativ – innovativ** oder **erwartungskonform – nicht erwartungskonform**, etc.
- Die Gegensatzpaare des Items sind zwar gängig, aber die Übertragung auf den Kontext der Bewertung eines interaktiven Produkts erfordert ein hohes Maß an Abstraktion und kann daher leicht zu Fehlinterpretationen führen. Ein Beispiel war das Begriffspaar **wertvoll – minderwertig**.

Diese Überlegungen zeigen, dass der im UEQ verwendete Sprachgebrauch für die intendierte Zielgruppe problematisch ist. Es besteht daher die Gefahr, dass das Ergebnis einer Evaluation durch diese Probleme mit der Verständlichkeit der Items zu stark beeinflusst wird.

3. Erzeugen einer UEQ-Version in vereinfachter Sprache

Das Ziel ist die Konstruktion einer Variante des UEQ, der ohne Verständnisprobleme für die Zielgruppe jugendlicher Nutzer eingesetzt werden kann. In dieser Variante sollen nur diejenigen Items, die für die Zielgruppe potentiell schwer verständlich sind, durch einfacher zu verstehende Items ersetzt werden. Die neuen Items sollen dabei die gleichen Dimensionen messen, wie die ersetzten Items.

Analyse des bestehenden Fragebogens

In einem ersten Schritt wurden alle Items des UEQ (siehe Appendix) einer

gründlichen inhaltlichen Analyse durch die Autoren und einem Lehrer unterzogen. Hierbei wurde in mehreren Diskussionen eine Reihe von potentiell problematischen Items identifiziert, d. h. Items bei denen vermutet wurde, dass diese Probleme für Personen der intendierten Zielgruppe bereiten können.

Insgesamt wurden für die Skala Attraktivität eines von 6 Items als potentiell problematisch identifiziert (attraktiv/unattraktiv). Für die 4 Items der anderen Skalen wurden jeweils folgende Anzahlen von Items als problematisch identifiziert: Effizienz 3 (unpragmatisch/pragmatisch, ineffizient/effizient, aufgeräumt/überladen), Steuerbarkeit 3 (erwartungskonform/nicht erwartungskonform, sicher/unsicher, behindernd/unterstützend), Durchschaubarkeit 0, Stimulation 2 (wertvoll/minderwertig, aktivierend/einschläfernd) und Originalität 3 (originell/konventionell, konservativ/innovativ, herkömmlich/neuartig).

Alternative Items erzeugen

Im zweiten Schritt wurden in mehreren Diskussionsrunden der Autoren Alternativen zu den problematischen Items erzeugt. Bei der Konstruktion der alternativen Items wurde versucht, die oben genannten Probleme (Fremdwörter, schwieriger Transfer) zu vermeiden.

Insgesamt wurden 27 neue, zusätzliche Items erzeugt. Aus diesen 27 neuen Items und den vorhandenen alten 26 Items wurde eine initiale Version eines Fragebogens erstellt. Die Reihenfolge der Items wurde randomisiert. Pro Item wurde zusätzlich die Reihenfolge des Gegensatzpaares randomisiert.

Dieser initiale Fragebogen wurde in mehreren Erhebungen zur Bewertung interaktiver Produkte verwendet.

Datenerhebung

Die Datenerhebung des initialen Fragebogens wurde mit einem Online-Tool innerhalb von zwei Monaten durchgeführt. Zusätzlich zu den Items wurden



Bewertetes Produkt	Gesamtanzahl	Verwertbar
Eclipse	8	7
EMP Onlineshop	10	10
KO.PAS	8	8
MS Word	35	29
Visual Studio	13	10
Keine Angaben	11	2
	85	66

Tab. 1.
Anzahl der ausgefüllten Fragebögen

soziologische Daten, wie Alter, Geschlecht und Ausbildung abgefragt, die in diesem Artikel keine weitere Verwendung finden. Innerhalb des Fragebogens wurde explizit darauf hingewiesen, dass unverständliche Items bzw. nicht intuitiv anwendbare Items übergangen werden sollten.

Der Fragebogen wurde mit 85 Probanden an 5 verschiedenen Produkten angewendet. Von den insgesamt 85 erhobenen Fragebögen waren 66 vollständig bearbeitet und konnten für die Datenanalyse verwendet werden, wie in Tabelle 1 dargestellt. [Tab. 1]

Datenanalyse und Erzeugung einer neuen UEQ Version

Die 66 erhobenen Datensätze wurden in einer mehrstufigen Vorgehensweise

analysiert, um die am besten passenden neuen Items zu identifizieren.

In einem ersten Schritt wurden alle Skalen einzeln über eine Faktorenanalyse betrachtet.

Für die Skalen **Attraktivität**, **Durchschaubarkeit** und **Originalität** ergab eine Analyse mit dem Kaiser-Guttman Kriterium, dass ein einzelner Faktor die Daten bereits ausreichend gut erklärt. In diesem Fall wurden diejenigen neuen Items als Ersatz für die als problematisch markierten Items ausgewählt, die auf diesem Faktor die höchsten Ladungen zeigten. Abbildung 2 zeigt als Beispiel den entsprechenden Scree-Plot der Dimension Attraktivität. [Abb. 2]

Bei den **Skalen Effizienz**, **Steuerbarkeit** und **Stimulation** ergab die Analyse mit

dem Kaiser-Guttman Kriterium, dass durch die zusätzlichen Items ein neuer Aspekt in die Skala kam, so dass eine Lösung mit zwei Faktoren besser passte. Abbildung 3 zeigt als Beispiel den Scree-Plot der Skala [Abb. 3]

In diesem Fall wurde geprüft, auf welcher dieser beiden Dimensionen die Items aus der alten UEQ Skala die höchsten Ladungen zeigten, d.h. dieser Faktor wurde als Repräsentant der alten Dimension betrachtet. Danach wurden die neuen Items, die auf diesem Faktor die höchsten Ladungen aufwiesen, als Ersatz für die zu streichenden Items gewählt.

Im zweiten Schritt wurde untersucht, welche Items sehr hohe Korrelationen mit anderen Skalen aufwiesen, d.h. nicht einen abgrenzbaren Faktor repräsentieren, sondern eher unspezifisch auf mehreren Qualitätsaspekten laden. Dies war bei zwei der neuen Items der Fall, die daraufhin durch andere alternative Items aus dem ersten Schritt ersetzt wurden.

Als Resultat der Analyse wurden folgende Items der Originalversion ersetzt:

- **Skala Attraktivität:**
 - attraktiv/unattraktiv ersetzt durch schön/hässlich
- **Skala Effizienz:**
 - ineffizient/effizient ersetzt durch zeitraubend/zeitsparend
 - unpragmatisch/pragmatisch ersetzt durch stockend/flüssig

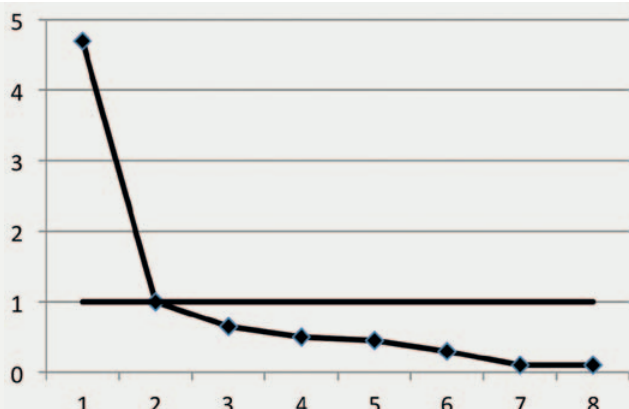


Abb. 2.
Scree-Plot für die Items der Dimension Attraktivität (6 Items des UEQ und 2 alternative Items)

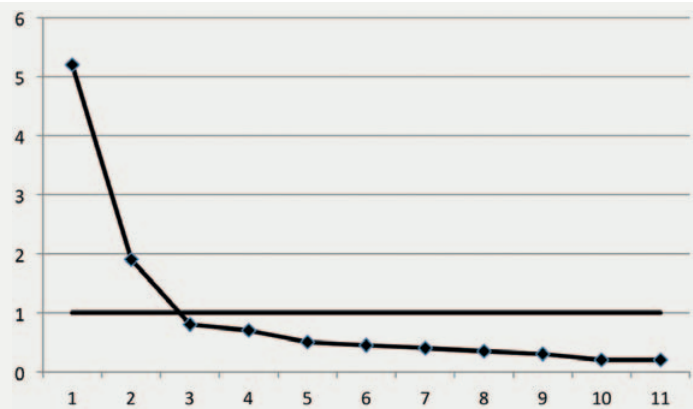


Abb. 3.
Scree-Plots für die Items der Dimension Steuerbarkeit (4 Items des UEQ und 7 alternative Items).

- **Skala Steuerbarkeit:**
 - behindernd/unterstützend ersetzt durch unbedienbar/bedienbar
 - sicher/unsicher ersetzt durch vorhersagbar/unvorhersagbar
 - erwartungskonform/nicht erwartungskonform ersetzt durch zuverlässig/unzuverlässig
- **Skala Stimulation:**
 - wertvoll/minderwertig ersetzt durch erfrischend/einschläfernd
 - aktivierend/einschläfernd ersetzt durch abwechslungsreich/eintönig
- **Skala Originalität:**
 - originell/konventionell ersetzt durch neu/alt
 - herkömmlich/neuartig ersetzt durch veraltet/modern
 - konservativ/innovativ ersetzt durch unauffällig/auffällig

Das Item aufgeräumt/überladen der Skala Effizienz konnte nicht ersetzt werden, da nur zwei der neuen Items für diese Skala geeignete Ladungen auf dem entsprechenden Faktor zeigten.

4. Items der Version in vereinfachter Sprache

Nach diesen Analysen ergab sich folgende neue Version des UEQ für den Einsatz in der Zielgruppe jugendlicher Nutzer interaktiver Produkte. Die gegenüber der Originalversion veränderten Items sind durch fette Formatierung hervorgehoben. [Tab. 2]

5. Validierung

Die Verständlichkeit der Items wurde durch einen Experten geprüft. Dieser kam wie erwartet zu der Ansicht, dass die neuen Items wesentlich einfacher zu verstehen sind, was natürlich aufgrund der bei der Konstruktion der Items zugrundeliegenden Grundidee auch zu erwarten war.

	1	2	3	4	5	6	7		
unerfreulich	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	erfreulich	1
unverständlich	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	verständlich	2
kreativ	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	phantasielos	3
leicht zu lernen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	schwer zu lernen	4
erfrischend	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	einschläfernd	5
langweilig	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	spannend	6
uninteressant	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	interessant	7
unberechenbar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	voraussagbar	8
schnell	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	langsam	9
neu	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	alt	10
unbedienbar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	bedienbar	11
gut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	schlecht	12
kompliziert	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	einfach	13
abstoßend	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	anziehend	14
veraltet	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	modern	15
unangenehm	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	angenehm	16
vorhersagbar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	unvorhersagbar	17
abwechslungsreich	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	eintönig	18
zuverlässig	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	unzuverlässig	19
zeitraubend	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	zeitsparend	20
übersichtlich	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	verwirrend	21
stockend	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	flüssig	22
aufgeräumt	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	überladen	23
schön	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	hässlich	24
sympathisch	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	unsympathisch	25
unauffällig	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	auffällig	26

Tab. 2.

Studie 1

In einer ersten kleinen Studie beurteilten 44 Studenten (Studienfach Informatik, Erstsemester) ein soziales Netzwerk (Facebook) bezüglich seiner User Experience. 25 Studenten benutzten die Originalversion des UEQ, während 19 Studenten den neu erstellten Fragebogen nutzten. Wegen der geringen Größe der

Stichprobe, kann diese erste Untersuchung natürlich nur grobe Hinweise bzgl. der Qualität der neuen Version liefern. [Abb. 4]

Abbildung 4 zeigt die Skalenmittelwerte der alten und neuen Version des UEQ. Mit Ausnahme der Dimension Steuerbarkeit waren die Unterschiede der Mittelwerte auf dem 5% Niveau nicht signifikant.

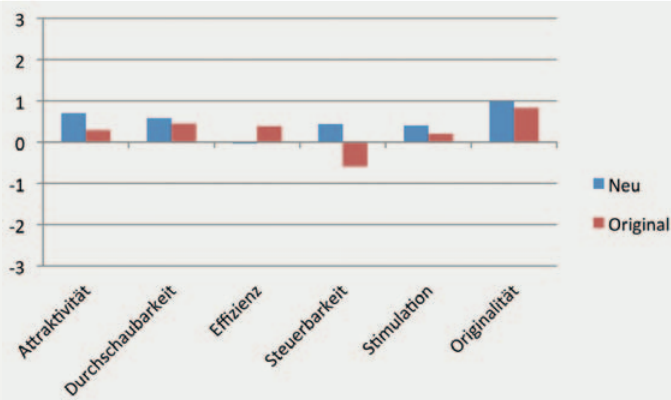


Abb. 4.
Skalenmittelwerte der alten und neuen Version des UEQ anhand von Facebook.

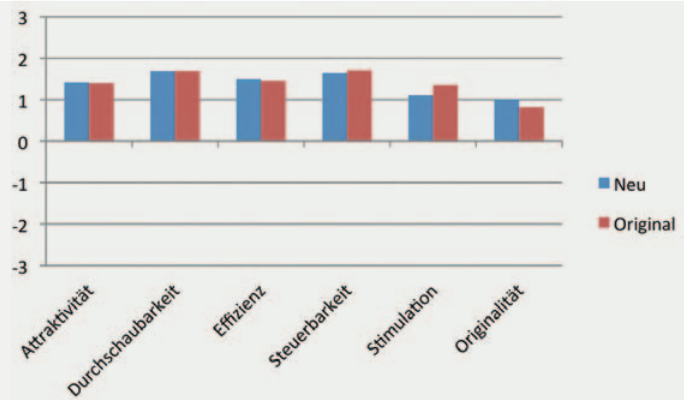


Abb. 5.
Skalenmittelwerte der alten und neuen Version des UEQ anhand einer Buchhaltungs-Software.

In Bezug auf die Konsistenz der Skalen wurden folgende Werte für Cronbachs Alpha-Koeffizient beobachtet (originale Version/neue Version):

- Attraktivität: 0,94 / 0,93
- Durchschaubarkeit: 0,80 / 0,77
- Effizienz: 0,73 / 0,67
- Steuerbarkeit: 0,68 / 0,83
- Stimulation: 0,77 / 0,79
- Originalität: 0,51 / 0,58

Hier unterscheiden sich die beiden Versionen also nicht gravierend. Die Konsistenz der Skalen ist zufriedenstellend, wobei man hier die geringe Stichprobengröße berücksichtigen muss. Einzige Ausnahme ist die Skala **Originalität**, die bei beiden Versionen einen relativ geringen Wert des Alpha- Koeffizient zeigt. Dies könnte daran liegen, dass dieser Faktor für die Teilnehmer in Bezug auf das bewertete soziale Netzwerk (Facebook) nicht wirklich relevant erschien.

Die Unterschiede in der Skala **Steuerbarkeit** resultierten im Wesentlichen aus einer stark negativen Bewertung des alten Items **sicher/unsicher** im Vergleich zu den anderen Items der Skala. Hier liegt der Schluss nahe, dass dies im Kontext von sozialen Netzwerken anders interpretiert wurde als intendiert. Dies wurde in Nachgesprächen auch von einigen Teilnehmern explizit angesprochen.

Ein weiteres Problem ergab sich bei dem Item **zeitraubend/zeitsparend** der neuen Version. Dieses unterschied sich bzgl. seines Mittelwerts (-1,3) gravierend von allen anderen Items der Skala **Effizienz**. Auch hier liegt nahe, dass dieses Item im Kontext sozialer Netzwerke anders als gedacht (z. B. als soziale Netzwerke rauben mir meine Zeit, da ich mich dort zu oft aufhalte) interpretiert wurde. Dies wurde ebenfalls von einigen Teilnehmern nach der Untersuchung explizit erwähnt. Als Konsequenz aus dieser Beobachtung wurde dieses Item durch das alte Item **ineffizient/effizient** aus dem Originalfragebogen ersetzt.

Studie 2

In einer zweiten Studie beurteilten 31 Nutzer einer Buchhaltungs-Software für Altenpflegeeinrichtungen die Anwendung mit einem Fragebogen, der die Items der alten und der neuen UEQ Version enthielt.

Abbildung 5 zeigt die Skalenmittelwerte der alten und neuen Version des UEQ, d. h. für die Items des kombinierten Fragebogens wurden zwei Auswertungen erstellt (Item der alten Version / Items der neuen Version). [Abb. 5]

Die Unterschiede sind auf dem 5% Niveau nicht signifikant.

In Bezug auf die Konsistenz der Skalen wurden folgende Werte für Cronbachs Alpha-Koeffizient beobachtet (originale Version/neue Version):

- Attraktivität: 0,83 / 0,82
- Durchschaubarkeit: 0,82 / 0,82
- Effizienz: 0,28 / 0,43
- Steuerbarkeit: 0,73 / 0,67
- Stimulation: 0,82 / 0,83
- Originalität: 0,86 / 0,77

Bis auf die Skala Effizienz sind die beobachteten Werte durchaus zufriedenstellend. Woraus die geringe Reliabilität der Effizienz-Skala in beiden Versionen resultiert, war nicht schlüssig zu klären.

6. Zusammenfassung

Aufgrund beobachteter Verständnisprobleme jugendlicher Nutzer wurde eine Version des UEQ in einfacher Sprache erstellt. Hierzu wurde für jedes der als schwer verständlich identifizierten Items des UEQ eine Reihe von alternativen Items erstellt. Diese neuen Items wurden zusammen mit den originalen Items einer Stichprobe von Personen zur Bewertung verschiedener interaktiver Produkte vorgegeben. Die Ergebnisse dieser Studie wurden verwendet, um die als problematisch erkannten Items durch die am besten passenden alternativen Items zu ersetzen.

Einschätzungen durch Experten weisen darauf hin, dass der neue Fragebogen für die angestrebte Zielgruppe besser verständlich ist. Dies muss aber zusätzlich noch durch weitere empirische Untersuchungen, die den neuen Fragebogen in der Zielgruppe jugendlicher Nutzer einsetzen, verifiziert werden.

Generell entsteht durch die Vereinfachung der Begriffe ein Verlust an sprachlicher Trennschärfe, d. h. hier ist theoretisch immer die Gefahr gegeben, dass sich beide Versionen bzgl. einzelner Skalen unterschiedlich verhalten.

Erste empirische Validierungen mit studentischen Teilnehmern und Nutzern von betriebswirtschaftlicher Software (d. h. außerhalb der intendierten Zielgruppe) zeigen, dass sich der neue Fragebogen ähnlich verhält, wie das Original. In weiteren Validierungsstudien mit größeren Stichproben, muss dies zukünftig noch detaillierter untersucht werden. Falls sich hier zeigt, dass beide Versionen weitgehend parallel sind, kann in Zukunft evtl. die neue Version generell anstatt der Originalversion verwendet werden. Es sind weitere Studien, auch mit jugendlichen Nutzern geplant, um die Verständlichkeit der Items für die Zielgruppe empirisch zu belegen.

Bis solche Ergebnisse vorliegen, sollte der neue Fragebogen nur in der Zielgruppe jugendlicher Nutzer eingesetzt werden bzw. in Situationen, in denen man davon ausgehen muss, dass viele Nutzer nur über einen eingeschränkten Wortschatz verfügen. In anderen Zielgruppen sollte weiterhin die Originalversion verwendet werden, da im Moment noch nicht hinreichend belegt ist, ob durch die Vereinfachung der Begriffe evtl. die Bedeutung der einzelnen Skalen verändert wurde.

Literatur

1. Dicke, C., Jakus, G., Tomazic, S. & Sodnik, J. (2012). On the Evaluation of Auditory and Head-up Displays while Driving. *ACHI* 2012.
2. DIN EN ISO 9241-11 (1999): Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten. Teil 11: Anforderungen an die Gebrauchstauglichkeit – Leitsätze. Berlin: Beuth Verlag.
3. Ebner, M., Stickel, C., Scerbakov, N. & Holzinger, A. (2009). A study on the Compatibility of Ubiquitous Learning (u-Learning) Systems at University Level. In: Shumaker, R. (Ed.): *Virtual and Mixed Reality. Lecture Notes in Computer Science* 5622, S. 34-43, Springer-Verlag.
4. Hassenzahl, M., Burmester, M., Koller, F. (2003): AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: J.Ziegler, G. Szwillus (Hrsg.): *Mensch & Computer* 2003. Interaktion in Bewegung. Stuttgart: Teubner. S. 187-196.
5. Hassenzahl, M. (2010): Experience design. Technology for all the right reasons. San Rafael: Morgan & Claypool.
6. Hartmann, J. (2011). User Experience Monitoring: Über die Notwendigkeit geschäftskritische Online-Prozesse permanent zu überwachen. *iCom*, Vol. 10, No. 3, S. 59-62.
7. Hatscher, M. (2001). Joy of use – Determinanten der Freude bei der Software-Nutzung. In: Oberquelle, H., Oppermann, R. & Krause, J. (Hrsg.): *Mensch & Computer* 2001: 1 Fachübergreifende Konferenz. Stuttgart: B.G. Teubner, S. 445-446.
8. Kurosu, M. & Kashimura, K. (1995): Apparent usability vs. inherent usability: experimental analysis of the determinants of the apparent usability. Denver, Colorado: Conference Companion of human factors in computing systems, S. 292-293.
9. 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 – Mensch und Computer im Strukturwandel. Oldenbourg Verlag, S. 125 – 134.
10. Laugwitz, B., Held, T. & Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. In: Holzinger, A. (Hrsg.): *USAB* 2008, *Lecture Notes in Computer Science* 5298, S. 63-76, Springer-Verlag.
11. Laugwitz, B., Schubert, U., Illmberger, 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, F., Peissner, M., Petrovic, K., Thielsch, M., Ulrich, D. & Zimmermann, D. (Eds.), *Usability Professionals* 2009, S. 220 – 225, Fraunhofer Verlag.
12. Molich, R. & Nielsen, J. (1990): Improving a human-computer dialog: What Designers know about traditional interface design. *Communications of the ACM*, 33, S. 338-348.
13. Norman, D. (2003): *Emotional Design: Why We Love (Or Hate) Everyday Things*. Boulder Colorado: Basic Books.
14. Rauschenberger, M., Hinderks, A. & Thomaschewski, J. (2011). Benutzererlebnis bei Unternehmenssoftware. In: Brau, H., Lehmann, A., Petrovic, K. & Schroeder, C. (Eds.), *Usability Professionals* 2011, S. 158 – 164, German UPA e.V., Stuttgart.
15. Rauschenberger, M., Schrepp, M., Olschner, S., Thomaschewski, J. & Cota, M.P. (2012). Measurement of User Experience. A Spanish Language Version of the User Experience Questionnaire (UEQ). In: Rocha, A., Calvo-Manzano, J.A., Reis, L.P. & Cota, M.P. (Eds.), *Sistemas y Tecnologías de Información – Actas de la 7a Conferencia Ibérica de Sistemas y Tecnologías de Información – Madrid, Espana, 20 al 23 de Junio de 2012*. ISBN 978-989-96247-6-4.
16. Tractinsky, N. (1997): Aesthetics and Apparent Usability: Empirical Assessing Cultural and Methodological Issues. *CHI'97 Electronic Publications*, Available URL <http://www.acm.org/sigchi/chi97/proceedings/paper/nt.htm>.
17. Wieschnowsky, T. & Heiko Paulheim, H. (2011). A Visual Tool for Supporting Developers in Ontology-based Application Integration, in: *7th International Workshop on Semantic Web Enabled Software Engineering*, 2011.

Danksagung: Wir danken Herrn Marcel Lautenbach für die Durchführung der zweiten Studie.