

Vergleich von UX Fragebögen

Andreas Hinderks¹, Martin Schrepp², Jörg Thomaschewski³

Universidad de Sevilla, ETS Ingeniería Informática, Sevilla¹

SAP SE, User Experience²

Hochschule Emden Leer³

andreas.hinderks@iwt2.org, martin.schrepp2@sap.com,
joerg.thomaschewski@hs-emden-leer.de

Zusammenfassung

Wir untersuchen drei verbreitete Anwendungen (Whatsapp, Netflix, Moodle) mit drei im deutschsprachigen Raum häufig verwendeten UX Fragebögen (VISAWI, meCUE, UEQ). Ziel ist die Beziehung der Skalen dieser Fragebögen näher zu untersuchen. Überraschenderweise zeigte sich, dass die Korrelationen der Skalen innerhalb eines Fragebogens stets deutlich höher sind, als zwischen verschiedenen Fragebögen. Zusätzlich hat die Art des untersuchten Produkts offenbar einen Einfluss auf die Beziehungen zwischen den Skalen verschiedener Fragebögen.

1 Fragestellung

Für die Einschätzung der User Experience (UX) eines Produkts durch Nutzer*innen spielen viele verschiedene UX Aspekte eine Rolle, z.B. Effizienz, Schönheit oder Spaß bei der Nutzung (siehe Schrepp, 2018). Zusätzlich ist die Wahrnehmung dieser verschiedenen UX Qualitäten sehr subjektiv. Wie intuitiv Nutzer*innen ein User Interface einschätzen, hängt z.B. von den Vorerfahrungen mit anderen Produkten ab. Damit sind Fragebögen ein sehr geeignetes Mittel, diese subjektiven Einschätzungen zu erfassen.

Abhängig von der Art des Produkts sind jeweils unterschiedliche UX Aspekte für den Gesamteindruck wichtig. Für eine professionell häufig genutzte betriebswirtschaftliche Anwendung sind andere UX Aspekte relevant, als für ein Spiel (Winter et al., 2017). Aus diesem Grund gibt es verschiedene UX Fragebögen, die jeweils andere Kombinationen von UX Aspekten als Skalen beinhalten. Es ergibt sich allerdings die Frage, in welcher Beziehung die Skalen der unterschiedlichen Fragebögen zueinanderstehen. Das gilt insbesondere, wenn Skalen verschiedener Fragebögen verwandte Konzepte messen. Weiterhin ist es interessant, ob die Beziehung der Skalen untereinander auch vom evaluierten Produkt abhängig ist. Wir untersuchen diese Fragen in einer Studie, in der drei verschiedene Anwendungen mit Hilfe von drei im deutschsprachigen Raum verbreiteten UX Fragebögen untersucht wurden.

2 Verwendete Fragebögen

VISAWI (Visual Aesthetics of Website Inventory)

Der VISAWI erlaubt die Messung der visuellen Ästhetik von Web-Seiten (Mooshagen & Thielsch, 2010) mit den Skalen *Vielfalt* (Wirkt das Layout originell und interessant?), *Einfachheit* (Ist das Layout übersichtlich?), *Farbigkeit* (Sind die Farben schön und aufeinander abgestimmt?) und *Kunstfertigkeit* (Wirkt die Gestaltung professionell und zeitgemäß?). Mehr Informationen finden sich auf www.meinald.de/forschung/visawi.

meCUE

Der meCUE (Minge & Riedel, 2013) besteht aus drei Modulen. Das Modul *Produktwahrnehmung* enthält die Skalen *Nützlichkeit* (Hilft mir das Produkt meine Ziele zu erreichen?), *Benutzbarkeit* (Lässt sich das Produkt einfach erlernen und benutzen?), *Visuelle Ästhetik* (Ist das Design attraktiv?), *Status* (Verschafft mir die Nutzung des Produkts einen höheren sozialen Status?) und *Bindung* (Ist das Produkt wichtig für mich?). Das Modul *Nutzeremotionen* enthält Skalen zur Erfassung von *positiven* und *negativen Emotionen* zum Produkt. Das Modul *Konsequenzen* erfasst mit den Skalen *Produktloyalität* (Man wird das Produkt nicht gegen ein anderes eintauschen) und *Nutzungsintention* (Man will das Produkt so oft wie möglich nutzen) die Verhaltensweisen, die sich aus den Reaktionen auf das Produkt ableiten. Mehr Informationen finden sich auf www.mecue.de.

UEQ (User Experience Questionnaire)

Der UEQ umfasst 6 Skalen (Laugwitz et al., 2006): *Attraktivität* (Gesamteindruck zum Produkt), *Effizienz* (Kann man seine Aufgaben ohne unnötigen Aufwand abarbeiten?), *Durchschaubarkeit* (Ist es leicht sich in die Bedienung des Produkts einzuarbeiten?), *Steuerbarkeit* (Hat der Nutzer die Kontrolle über die Interaktion mit dem Produkt?), *Stimulation* (Ist die Beschäftigung mit dem Produkt spannend und motivierend?), und *Originalität* (Ist das Produkt innovativ und kreativ gestaltet?). Mehr Informationen zum UEQ finden sich auf www.ueq-online.org.

3 Studien

Mit *Whatsapp*, *Netflix* und *Moodle* wurden drei bekannte Anwendungen mit stark unterschiedlichem Nutzungskontext ausgewählt und von studentischen Teilnehmern mit dem VISAWI, dem meCUE und dem UEQ bewertet, d.h. jeder Teilnehmer füllte alle drei Fragebögen zu einem Produkt aus. Die Bewertung der Anwendungen erfolgte über einen Online-Fragebogen, der an Teilnehmer von Lehrveranstaltungen der Hochschule Emden/Leer verschickt wurde. Insgesamt wurden die Online-Fragebögen 259-mal begonnen. Daten von Teilnehmenden, die nicht alle Fragen beantworteten, wurden aus der Auswertung eliminiert. Insgesamt verblieben 137 vollständige Datensätze (Durchschnittsalter 28,6 Jahre, 36% weiblich und 64% männlich) in allen drei Teilstudien (31 Whatsapp, 63 Netflix, 43 Moodle).

Wir fokussieren uns in diesem Beitrag auf die Beziehung der Skalen (detaillierte Auswertungen, insbesondere die numerischen Korrelationen, finden sich in Hinderks et al., 2018). Abbildung 1 visualisiert die Korrelationen zwischen den Skalen für *Netflix* und *Whatsapp* mit einer multi-dimensionalen Skalierung (Torgerson, 1958), d.h. je näher zwei Dimensionen zusammenliegen, desto höher ist die Korrelation. Hierbei können nur die Distanzen interpretieren werden, nicht die Lage eines Punktes in der Graphik.

Es ist deutlich, dass die Korrelationen zwischen den Skalen, d.h. die Beziehung zwischen den von diesen Skalen gemessenen Aspekten von UX, zwischen den Produkten stark variieren. Zum Beispiel bestehen für *Whatsapp* offenbar hohe Korrelationen (s. Abb. 1 rechts) zwischen den hedonischen UEQ Skalen und den Skalen der VISAWI, was inhaltlich auch zu erwarten ist. Dieser Zusammenhang ist für *Netflix* jedoch offenbar nicht vorhanden (s. Abb. 1 links). Das gleiche gilt für Moodle (ohne Abbildung). Ähnliche Abweichungen sind auch für andere Skalen zu beobachten.



Abbildung 1: Korrelation der Skalen von meCUE (grüne Punkte), VISAWI (orangene Punkte) und UEQ (blaue Punkte) für *Netflix* (links) und *Whatsapp* (rechts).

Items werden immer im Kontext einer Befragungssituation interpretiert, d.h. im Kontext des evaluierten Produkts. Damit können natürlich Bedeutungsunterschiede auftreten, die die Korrelationen zwischen den Skalen beeinflussen. Den oben gezeigten Unterschied zwischen *Netflix* und *Whatsapp* könnte z.B. dadurch erklärt werden, dass bei *Netflix* der visuelle Eindruck vor allem durch die Inhalte entsteht, während bei *Whatsapp* vor allem die Gestaltung der App selbst für diesen Eindruck verantwortlich ist.

Eine weitere Beobachtung (dies ist sehr deutlich bei *Netflix* zu sehen, aber auch in den hier nicht visualisierten Daten von Moodle) ist, dass die Skalen innerhalb eines Fragebogens deutlich höher korrelieren, als zwischen Fragebögen. Das ist zunächst einmal irritierend. Von der Semantik der Skalen wäre z.B. zu erwarten, dass die Ästhetik Skala des meCUE stärker mit den Skalen des VISAWI korreliert, als mit den anderen Skalen des meCUE. Das ist aber nicht der Fall. Ähnliches gilt für die pragmatischen Skalen des UEQ und die Skalen Benutzbarkeit und Nützlichkeit des meCUE.

4 Schlussfolgerungen

Ein überraschendes Ergebnis ist, dass Skalen innerhalb eines Fragebogens höher korrelieren, als mit Skalen in einem anderen Fragebogen, auch wenn diese semantisch eigentlich verwandt sein sollten. Eine mögliche Erklärung (die genauer untersucht werden muss) ist, dass die Item-Formate (z.B. semantisches Differential im UEQ und Statements im meCUE und VISAWI) und die konkrete Art Items zu formulieren, hier einen Einfluss hat, der die semantische Nähe überdeckt. Ein weiteres Ergebnis ist, dass die Art des Produkts offenbar einen Einfluss auf die Korrelationen zwischen Skalen hat.

Was bedeuten diese Ergebnisse das für die Anwendung von UX-Fragebögen? Bei Arbeiten, die versuchen Zusammenhänge zwischen UX Aspekten zu untersuchen, sollte stark darauf geachtet werden, dass diese mit mehreren möglichst verschiedenen Produkten nachgewiesen werden. Ansonsten besteht die Gefahr, dass die Beziehung nur für eine spezielle Art von Produkten existiert und man sich dessen nicht bewusst ist.

Generell ist es offenbar wichtig, genauer zu untersuchen, wie sich die Skalen verschiedener UX Fragebögen zueinander verhalten. Auch wenn Skalen zweier UX Fragebögen semantisch ähnliche Konzepte realisieren, können die gemessenen Qualitäten offenbar stark abweichen. Einschränkend muss angemerkt werden, dass die Datensätze unserer Studie recht klein sind und nur drei verschiedene Produkte untersucht wurden. Es ist geplant, die gefundenen Effekte in weiteren Studien genauer zu untersuchen.

Literatur

- Hinderks, A.; Schrepp, M. & Thomaschewski, J. (2018). Vergleich von UX Fragebögen: Evaluationsbericht von Whatsapp, Netflix und Moodle mit meCUE, Visawi und UEQ. DOI: 10.13140/RG.2.2.36124.18560.
- Laugwitz, B.; Schrepp, M. & Held, T. (2006). Konstruktion eines Fragebogens zur Messung der User Experience von Softwareprodukten. A.M. Heinecke & H. Paul (Eds.): Mensch & Computer 2006 - Mensch und Computer im Strukturwandel. Oldenbourg Verlag, S. 125 – 134
- Minge, M. & Riedel, L. (2013). meCUE – Ein modularer Fragebogen zur Erfassung des Nutzungserlebens. In: S. Boll, S. Maaß & R. Malaka (Hrsg.): Mensch und Computer 2013: Interaktive Vielfalt, S. 89-98. München, Oldenbourg Verlag.
- Moshagen, M. & Thielsch, M. T. (2010). Facets of visual aesthetics. *International Journal of Human-Computer Studies*, 68, S. 689-709.
- Schrepp, M. (2018). *User Experience mit Fragebögen messen*. ISBN: 9781986843768. Printed by CreateSpace (Amazon).
- Torgerson, W. S. (1958): *Theory & Methods of Scaling*. New York: Wiley.
- Winter, D., Hinderks, A., Schrepp, M. & Thomaschewski, J., (2017). Welche UX Faktoren sind für mein Produkt wichtig? In: Hess, S. & Fischer, H. (Hrsg.), *Mensch und Computer 2017 - Usability Professionals*. Regensburg: Gesellschaft für Informatik e.V. (S. 191 – 200).

Autoren



Hinderks, Andreas

Andreas Hinderks ist Diplom-Informatiker und Master of Science in Medieninformatik der University of Applied Science Emden/Leer. Von 2001 bis 2016 arbeitete er in verschiedenen Führungsrollen als Business Analyst und Programmierer. Sein Schwerpunkt lag auf der Entwicklung anwenderfreundlicher Unternehmenssoftware. Derzeit ist er freiberuflich als Scrum Master, Business Analyst und Senior UX Architekt tätig. Außerdem ist er Doktorand an der Universität Sevilla. Seit 2011 ist er an Forschungsaktivitäten beteiligt, die sich mit UX-Fragebögen, Prozessoptimierung, Informationsarchitektur und User Experience beschäftigen.



Schrepp, Martin

Dr. Martin Schrepp studierte Mathematik und Psychologie an der Universität Heidelberg. 1990 Abschluss als Diplom-Mathematiker. 1990 – 1993 Promotion in Psychologie. Seit 1994 bei der SAP AG tätig. Bisherige Tätigkeitsfelder waren hier die Konzeption technischer Dokumentation, Software-Entwicklung, User Interface Design und Barrierefreiheit. Hauptinteressen sind die Anwendung kognitionswissenschaftlicher Erkenntnisse auf das Design interaktiver Anwendungen, Barrierefreiheit und die Entwicklung von Methoden zur Evaluation und Datenanalyse.



Thomaschewski, Jörg

Dr. Jörg Thomaschewski ist Professor an der Hochschule Emden/Leer und Leiter der „Research Group for Agile Software Development and User Experience“ mit den Lehr- und Forschungsschwerpunkten Usability und User Experience, Human Computer Interaction, Requirement Engineering, Agile Software Development, Kanban, Scrum, Internet-Programming und E-Learning. Er ist Autor verschiedener Online-Module der Virtuellen Hochschule (VFH) und verfügt über umfangreiche Erfahrungen in IT-Analysen und Beratungen.