

IM-UX – Fragebogen zu intrinsischer Motivation in der User Experience

Jette Selent und Michael Minge

1 Einleitung

Intrinsische Motivation gilt als wesentlicher Einflussfaktor dafür, dass Verhaltensweisen mit erhöhter Wahrscheinlichkeit und über längere Zeiträume hinweg stabil gezeigt werden (Ryan & Deci 2000a, Vallerand 1997). Dies lässt sich auch bei der Nutzung von Software-Produkten beobachten, deren Nutzung nicht aus Notwendigkeit sondern freiwillig erfolgt. So gilt etwa die Playfulness (Verspieltheit) von Systemen nicht nur als Quelle für intrinsische Motivation, sondern auch als einer der wichtigsten Faktoren für das Entstehen und Aufrechterhalten von Akzeptanz und der Bereitschaft zur Nutzung eines Systems (Venkatesh 2000). Hwang und Yi (2002) sowie Lee, Cheung und Chen (2005) beschreiben, dass das Erleben von Lernerfahrungen bei der Nutzung eines Systems die Wahrscheinlichkeit erhöht, dass es langfristig genutzt wird. Sowohl Spaß, als auch Wissens-, Fähigkeits- und Fertigkeitserwerb sowie spielerische Erfahrungen können demnach als Quellen intrinsischer Motivation betrachtet werden und es gibt Grund zu der Annahme, dass sich das Hervorrufen intrinsischer Motivation positiv auf die Wahrscheinlichkeit, Häufigkeit und Dauer der Nutzung eines Software-Produktes auswirkt.

Auch die *User Experience* (UX) bei der Nutzung interaktiver Produkte wird mit Faktoren intrinsischer Motivation in Verbindung gebracht. So beschreibt etwa Hassenzahl (2008) eine gute UX als Konsequenz der Befriedigung des Bedürfnisses unter anderem nach Autonomie, Kompetenz, und Verbundenheit. Diese Bedürfnisse gelten laut der *Selbstbestimmungstheorie* (Self-Determination Theory, SDT) von Deci und Ryan (1985) als Grundlage für das Entstehen intrinsischer Motivation. Kompetenz beschreiben Deci und Ryan

(2002b) hier als das Gefühl dafür, in welchem Ausmaß eine Person einen effektiven Einfluss auf ihre soziale und physische Umwelt ausüben kann. Darüber hinaus beinhaltet Kompetenz im Sinne der SDT das Erleben von Möglichkeiten, die eigenen Fähigkeiten und Fertigkeiten auszuüben und in kontrollierbarer Art und Weise weiterzuentwickeln (Deci & Ryan 2002a, White 1959). Nimmt eine Person sich selbst als Ursprung und Initiator ihrer Handlungen wahr, so wird dies als Erleben von Autonomie bezeichnet. Diese Empfindung wird dadurch begünstigt, dass die Person erlebt, die eigene Umwelt beherrschen und beeinflussen zu können (Angyal 1941). Autonomes Handeln bedeutet ebenso, dass das gezeigte Verhalten als Ausdruck des eigenen Willens verstanden, aus eigenem Interesse initiiert und im Sinne der eigenen Wertvorstellungen ausgeführt werden kann (Deci & Ryan 2002a). Als drittes Grundbedürfnis wird Geborgenheit, Einbettung in ein emotional stabiles soziales Umfeld sowie die Bindung an Bezugspersonen und eine Gesellschaft beschrieben (Ainsworth 1989, Baumeister & Leary 1995, Bowlby 1979). Das Gefühl von Verbundenheit im Sinne der SDT wird ebenfalls durch das Erleben einer Verbindung zu anderen Menschen, die auf gegenseitigem Füreinander-Sorge-Tragen beruht, gefördert (Ryan & Deci 2000b).

Für den langfristigen Erfolg von Technik und die Integration von Produkten in den Alltag von Nutzern ist die Frage entscheidend, ob es gelingt, diese intrinsischen Bedürfnisse zu befriedigen. Bestehende Fragebögen zur Erfassung intrinsischer Motivation, wie z.B. die Kurzskala intrinsische Motivation (KIM, Wilde et al. 2009) oder der Fragebogen zur aktuellen Motivation (FAM, Rheinberg et al. 2001) sind allerdings nicht für den Bereich UX validiert und bilden in ihrer Struktur nur selten etablierte theoretische Modelle ab. Daher wurde auf Grundlage der SDT ein neues Messinstrument konstruiert und anhand mobiler Applikationen validiert.

Als Antwortformat wurde das semantische Differenzial (Osgood, 1952) gewählt, da Friberg, Martinussen und Rosenvinge (2006) zeigen konnten, dass dies bei der Erfassung positiver psychologischer Konstrukte, zu denen auch intrinsische Motivation gezählt werden kann, Likert-Skalen vorzuziehen ist. Grund hierfür sei die geringere Verzerrung durch kontraintuitive Formulierungen, die bei Likert-Skalen durch die Negation positiver Konstrukte oder Begriffe entstehen können. Darüber hinaus haben sich semantische Diffe-

renziale bei der Messung von Einstellungen, Emotionen und bewertungsbezogenen Wahrnehmungen als adäquat erwiesen (Heise 1970). Außerdem sind sie bei der Bewertung von Software-Produkten und Designs, wie etwa bei Hassenzahl, Platz, Burmester und Lehner (2000), bei Hsu, Chuang und Chang (2000) bei Huang, Chen und Khoo (2012), bei Khalid und Helander (2004) sowie bei Lin (1992), weit verbreitet und akzeptiert.

Das Ziel war demnach, mit Hilfe semantischer Differenziale ein kompaktes Messinstrument für die Betrachtung der Befriedigung der drei Grundbedürfnisse nach der SDT bei der Nutzung interaktiver Softwareprodukte zu entwickeln.

2 Konstruktion des Fragebogens

Die intendierte Struktur des zu entwickelnden Messinstruments zielte darauf ab, die drei Grundbedürfnisse der SDT, *Kompetenz*, *Autonomie* und *Verbundenheit*, sowie zwei zusätzliche Komponenten zu erheben. *Stimulation* als Eigenschaft, Neugierde zu wecken und Spaß zu bereiten, wurde als wichtiger Einflussfaktor für intrinsische Motivation betrachtet (z.B. Vallerand 1997). Mit der *Akzeptanz* sollte außerdem ein Gesamturteil erfasst werden können. Nach Recherche bestehender Fragebögen zu intrinsischer Motivation und zum Nutzererleben sowie mehreren Brainstorming-Sessions wurden insgesamt 17 Items als Gegensatzpaare formuliert, die den fünf Dimensionen zugeordnet wurden (siehe Tabelle 1). Als Antwortformat wurde in Anlehnung an Hassenzahl, Burmester und Koller (2003) ein siebenfach gestuftes semantisches Differenzial gewählt (Osgood, 1952), da dies zur Erfassung positiver psychologischer Konstrukte likertskalierten Antwortformaten vorzuziehen ist und es darin unterstützt, spontane Urteile abzugeben (z.B. Fribourg et al. 2006).

2.1 Erste Datenerhebung

Methode

Die Überprüfung der dimensional Struktur erfolgte auf Basis einer Online-Datenerhebung, bei der $n=119$ Probanden (78 weiblich, 41 männlich) eine ihnen bekannte mobile Applikation bewerteten. Die Auswertung der Daten

erfolgte mittels einer varianzmaximierenden Hauptkomponentenanalyse unter Anwendung des Minimum Average Partial (MAP-)Tests nach Velicer (1976).

Ergebnisse

Die Eignung des Datensatzes zur Berechnung einer Hauptkomponentenanalyse ist durch ein Kaiser-Meyer-Olkin(KMO)-Kriterium in Höhe von .788 sichergestellt. Im Ergebnis entsprach das Ladungsmuster der Items auf den extrahierten fünf Komponenten nicht zufriedenstellend der intendierten inhaltlichen Struktur des Fragebogens. Teilweise wiesen die Itempaare deutliche Schwächen auf. Insbesondere für die geplante Skala Kompetenz ist es nicht gelungen, eine in sich konsistente und gegenüber anderen Komponenten trennscharfe Skala zu konstruieren. Im Bereich der Akzeptanz fiel das Itempaar „kritisch-unkritisch“ deutlich heraus. In Folge wurde der Itempool überarbeitet und dort, wo bereits möglich, reduziert. Außerdem wurden neue Itempaare ergänzt. Durch diesen Schritt entstand eine neue Rohversion mit 14 Items.

Kompetenz: (behindernd - unterstützend) / (harmlos - herausfordernd) / (unterfordernd - überfordernd)
Autonomie: (bestimmt mich - berücksichtigt mich) / (starr - flexibel) / (einseitig - vielseitig) /
Verbundenheit: (entfernt mich von anderen - bringt mich anderen näher) / (ausgrenzend - einbeziehend) / (isolierend - verbindend)
Stimulation: (langweilig - unterhaltsam) / (herkömmlich - neuartig) / (belanglos - interessant) / (lästig - spaßig)
Akzeptanz: (schlecht - gut) / (unangenehm - angenehm) / (kritisch - unkritisch) / (abstoßend - ansprechend)

Tabelle 1: Liste der Itempaare in der Rohversion des Fragebogens für die erste Validierung.

2.2 Zweite Datenerhebung

Methode

Die Überprüfung der dimensional Struktur dieser zweiten Rohversion und die Itemselektion erfolgten auf Basis einer erneuten Datenerhebung, bei der $n=79$ Probanden (46 weiblich, 31 männlich, 2 keine Angabe) analog zu 2.1 eine selbst gewählte mobile Applikation online bewerteten. Die Auswertung erfolgte durch Hauptkomponentenanalyse mit Varimax-Rotation.

Ergebnisse

Die Eignung des Datensatzes zur Berechnung einer Hauptkomponentenanalyse ist durch ein Kaiser-Meyer-Olkin(KMO)-Kriterium von .733 sichergestellt. Die Ladungsverteilung der selektierten Items zeigt in diesem Fall ein hoch trennscharfes Muster mit inhaltlich intendierten Zusammenhängen auf den erwarteten extrahierten fünf Komponenten (siehe Tabelle 2).

Diese Komponenten erklären 79,567 % der Varianz (siehe Tabelle 3). Die interne Konsistenz der Skalen wurde mittels Reliabilitätsanalyse geprüft. Cronbach's Alpha-Werte sind, außer für Autonomie, zufriedenstellend bis gut. Jedoch kann auch ein Alpha von .5 für beginnende Forschung mit einem neuen Messinstrument als ausreichend betrachtet werden (Streiner, 2003).

Itempaar	I	II	III	IV	V
hinderlich - förderlich	.907				
bremst mich aus - bringt mich weiter	.828				
starr - flexibel		.827			
bestimmt mich - berücksichtigt mich		.681			
entfernt mich von anderen - bringt mich anderen näher			.860		
isolierend - verbindend			.816		
langweilig - unterhaltsam				.924	
lästig - spaßig				.735	.507
schlecht - gut					.845
abstoßend - ansprechend					.768

Tabelle 2: Varimax-rotierte Ergebnisse der Hauptkomponentenanalyse (dargestellt sind alle Faktorladungen $>.4$)

Komponente	Bezeichnung	Varianzaufklärung in Prozent	Cronbrachs Alpha
I	Kompetenz	17,081	.764
II	Autonomie	6,693	.506
III	Verbundenheit	14,270	.653
IV	Stimulation	8,670	.741
V	Akzeptanz	32,916	.767
Summe:		79,567	--

Tabelle 3: *Varianzaufklärung durch die Komponenten sowie Ergebnisse der Reliabilitätsanalyse der Skalen.*

3 Diskussion

Es konnte ein hinsichtlich seiner faktoriellen Struktur und seiner internen Konsistenz überzeugender Kurzfragebogen zur Erfassung der wahrgenommenen Bedürfnisbefriedigung bei Interaktion mit technischen Systemen entwickelt werden, der zentrale Komponenten intrinsischer Motivation erfasst.

Einschränkend muss erwähnt werden, dass die zwei in den Studien akquirierten Stichprobenumfänge für den Zweck der Itemselektion mittels Hauptkomponentenanalyse mit $n=119$ bzw. $n=79$ Probanden als vergleichbar klein und als lediglich ausreichend akzeptabel eingestuft werden müssen. Die mit .788 bzw. .733 deutlich über den geforderten Mindestwerten von .5 (Cleff, 2015 & Field, 2013) bzw. .6 (Möhring & Schlütz 2013) liegenden Kaiser-Meyer-Olkin-Kriterien lassen die Durchführung einer Hauptkomponentenanalyse und die Interpretation der Ergebnisse allerdings zweifelsfrei zu. Außerdem unterstützt die stabile Replikation der faktoriellen Struktur bei explorativem Vorgehen in zwei unabhängigen Stichproben die Annahme, dass die extrahierten Dimensionen von substanzieller Bedeutung und Trennschärfe sind.

Neben der in den vorgestellten Datenerhebungen fokussierten Sicherstellung der Reliabilität ist das Gütekriterium der Messgültigkeit (Validität) bislang noch nicht untersucht. Die Überprüfung der inhaltlichen Validität ist daher ein wichtiger weiterer Schritt. Neben der Überprüfung konvergenter und divergenter Validität, beispielsweise mittels Korrelation zu Fragebögen, die inhaltsnahe bzw. -ferne Konstrukte erfassen, ist außerdem wünschenswert, den Zusammenhang zu externen Außenkriterien zu ermitteln (z.B. Häufigkeit

freiwilliger Nutzung, Nutzungsintensität und –dauer oder Bevorzugung bestimmter Produkte oder Services gegenüber möglichen Alternativlösungen).

Mit lediglich 10 Items ist der vorgestellte Fragebogen IM-UX ein äußerst ökonomisches Messinstrument, dessen Beantwortung in aller Regel weniger als eine Minute benötigt. Es erlaubt eine standardisierte Messung und damit eine hohe Vergleichbarkeit zwischen unterschiedlichen Bewertungsobjekten (z.B. Produktalternative A und B), zwischen verschiedenen Nutzergruppen (z.B. Experten und Novizen) bzw. zu verschiedenen Messzeitpunkten (z.B. in frühen oder späteren Nutzungsphasen bei ein- und denselben Personen). Die Interpretation von Werten bei einmaliger Messung erfordert eine vorherige Definition von Ziel- bzw. Akzeptanzkriterien. Es liegen bislang keine Vergleichswerte vor, die einen branchen- oder produktspezifischen Benchmark vorgeben, oder einen standardisierten relativen Vergleich von Messwerten ermöglichen.

Die vorliegende Validierung des Messinstruments erfolgte unter Abfrage des Erlebens motivationaler Aspekte bei Interaktion mit mobilen Anwendungen. Aufgrund des Einsatzes generisch formulierter Adjektivpaare in der Form eines semantischen Differenzials erscheint der Einsatz auch für weitere Technologie- und Anwendungsbereiche (wie z.B. Websites, betriebliche Software, Fahrzeuge und Fahrzeuginterfaces, physische Produkte, Gebrauchsgegenstände, Haushaltsgeräte, etc.) prinzipiell möglich, wenn diese mittels erlebbarer Oberfläche für Anwender in Erscheinung treten. Interessiert das Erleben von Dienstleistungen oder Services sowie die UX spezieller Anwendungsbereiche, sollte im Einzelfall auf Basis der Augenscheinvalidität vorab geprüft werden, ob die Items für den Anwendungsbereich und die jeweilige Fragestellung der Erhebung sowohl inhaltlich passend als auch eindeutig formuliert sind.

Innerhalb der Phasen der menschenzentrierten Gestaltung (ISO 9241, Teil 210), kann der Fragebogen IM-UX bereits früh bei der Evaluation von Produktideen, Wireframes oder Klick-Prototypen eingesetzt werden, ebenso wie für bereits implementierte und marktgängige Produkte, mit dem Ziel, Optimierungsbedarfe für eine stärkere Kundenbindung und –begeisterung zu identifizieren. So lassen sich auf Basis bislang unterberücksichtigter Aspekte innovative Ideen für Erweiterungen und neue Funktionalitäten begründen und entwickeln, die das Potenzial, dass das Produkt durch Nutzer im Alltag als

stärker integriert erlebt wird, strategisch ausschöpfen. Im Rahmen der Evaluation bietet sich der Einsatz sowohl bei der Überprüfung von zuvor gesetzten Anforderungen in Laborstudien und User-Experience-Testings (Verifizierung) an, als auch bei der Validierung von Produkten, also der Überprüfung, ob sich Gestaltungslösungen im Alltag von Anwendern bewähren und für Sie einen wahrgenommenen Mehrwert generieren. Der Einsatz als Paper-Pencil-Version oder als Online-Fragebogen ist gleichfalls denkbar.

4 Ausblick

Neben der Erfassung von Emotionen ist die Erhebung motivationaler Aspekte eine wichtige Aktivität der Berücksichtigung einer guten User Experience in der Gestaltung und Evaluation von Technik. Der hier präsentierte Kurzfragebogen bietet eine Möglichkeit, diese Aspekte früh in der Entwicklung und kontinuierlich während der Implementierung zu erfassen. Die Überprüfung der Validität, z.B. mittels Korrelation zu inhaltlich ähnlich aufgebauten Fragebögen (Konstruktvalidität) sowie die Übereinstimmung mit Verhaltenskriterien (externe Validität) sind zentrale nächste Schritte. Ebenso wird momentan eine englischsprachige Version des Messinstruments erarbeitet und validiert. Der Fragebogen ist frei einsetzbar und kann von Interessierten für eigene Forschungs- und Anwendungszwecke verwendet werden.

Literaturverzeichnis

- Ainsworth, M.S. (1989). Attachments beyond infancy. *American psychologist*, 44 (4), 709
- Angyal, A. (1941). *Foundations for a science of personality*
- Bowlby, J. (1979). *The making and breaking of affective bonds*. London, England: Tavistock Publications
- Cleff, T. (2015). *Deskriptive Statistik und Explorative Datenanalyse: Eine computergestützte Einführung mit Excel, SPSS und STATA (3., überarbeitete und erweiterte Auflage)*. Wiesbaden: Gabler Verlag
- Deci, E.L. & Ryan, R.M. (2002a). *Handbook of self-determination research*. University Rochester Press
- Deci, E.L., Ryan, R.M. (1985). *Intrinsic motivation and self-determination*. New York: Plenum
- Deci, E.L. & Ryan, R. M. (2002b). Overview of self-determination theory: An organismic dialectical perspective. *Handbook of self-determination research*, 3–33
- Field, A.P. (2013). *Discovering statistics using IBM SPSS statistics: And sex and drugs and rock 'n' roll (4th edition)*. Los Angeles: Sage

- Fribourg, O., Martinussen, M., Rosenvinge, J.H. (2006). Likert-based vs. Semantic differential-based scorings of positive psychological constructs: A psychometric comparison of two versions of a scale measuring resilience. *Personality and Individual Differences*, 40 (5), 873-884
- Hassenzahl, M. (2008). User experience (UX): towards an experiential perspective on product quality. In *IHM*, 8, 11-15
- 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* (S. 187–196). Stuttgart, Leipzig: B.G. Teubner
- Hassenzahl, M., Platz, A., Burmester, M., Lehner, K. (2000). Hedonic and ergonomic quality aspects determine a software's appeal. In *Proceedings of the sigchi conference on human factors in computing systems*, 201–208
- Heise, D.R. (1970). The semantic differential and attitude research. *Attitude measurement*, 235–253
- Hsu, S.H., Chuang, M.C., Chang, C.C. (2000). A semantic differential study of designers' and users' product form perception. *International Journal of Industrial Ergonomics*, 25 (4), 375–391
- Huang, Y., Chen, C.-H., Khoo, L. P. (2012). Products classification in emotional design using a basic-emotion based semantic differential method. *International Journal of Industrial Ergonomics*, 42 (6), 569–580
- Hwang, Y. & Yi, M. (2002). Predicting the use of web-based information systems: intrinsic motivation and self-efficacy. In *Proceedings of the 2002 - eighth americas conference on information systems*, 1076–1081
- Khalid, H.M. & Helander, M.G. (2004). A framework for affective customer needs in product design. *Theoretical Issues in Ergonomics Science*, 5 (1), 27–42
- Lee, M. K., Cheung, C. M. & Chen, Z. (2005). Acceptance of internet-based learning medium: the role of extrinsic and intrinsic motivation. *Information & Management*, 42 (8), 1095–1104
- Möhring, W., Schlütz, D. (2013). *Handbuch standardisierte Erhebungsverfahren in der Kommunikationswissenschaft*. Wiesbaden: Springer VS
- Osgood, C.E. (1952). The nature and measurement of meaning. *Psychological Bulletin*, 49 (3), 197-237
- Rheinberg, F. Vollmeyer, R., Burns, B.D. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen. *Diagnostica*, 47, 57-66
- Ryan, R.M., Deci, E.L. (2000a). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, 25 (1), 54–67
- Ryan, R.M., Deci, E.L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, 55 (1), 68
- Streiner, D.L. (2003). Starting at the beginning: an introduction to coefficient alpha and internal consistency. *Journal of Personal Assessment*, 80 (1), 99-103

- Velicer, W. (1976). Determining the Number of Components from the Matrix of Partial Correlations. *Psychometrika* 41 (3), 321-327
- Vallerand, R.J. (1997). Toward a hierarchical model of intrinsic and extrinsic motivation. *Advances in Experimental Social Psychology*, 29, 271–360
- Venkatesh, V. (2000). Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology acceptance model. *Information Systems Research*, 11 (4), 342–365
- White, R.W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, 66 (5), 297
- Wilde, M, Bätz, K., Kovaleva, A., Urhahne, D. (2009). Überprüfung einer Kurzsкала intrinsische Motivation (KIM). *Zeitschrift für Didaktik der Naturwissenschaften*, 15, 31-45

Kontakt

Jette Selent, M.Sc. Human Factors
468 Sierra Vista Ave Apt 9
94043 Mountain View, CA, USA

Dr.-Ing. Dipl.-Psych. Michael Minge
Technische Universität Berlin
Marchstraße 23, Sekr. MAR 3-2
10585 Berlin
www.tu-berlin.de/kke