

Exposition und Erleben in der Mensch-Maschine-Interaktion

Marlene Vogel¹, Nina Hallier², Manfred Thüring³

Graduiertenkolleg prometei, Technische Universität Berlin, ¹

Spiegel Institut Stuttgart GmbH c/o Daimler AG ²

Kognitionspsychologie und Kognitive Ergonomie, Technische Universität Berlin ³

Zusammenfassung

Das Nutzererleben (User Experience, UX) wird als ein sich zeitlich dynamisch entwickelndes Phänomen im Umgang mit technischen Produkten als auch Service Anwendungen betrachtet. Dabei kann das Erleben in unterschiedliche Phasen unterteilt werden: vor, während und nach einer Interaktion sowie der wiederholten Nutzung (Pohlmeier, 2011). Diese Phasen unterliegen verschiedensten Einflussfaktoren und tragen zu einem spezifischen und individuellen Nutzererleben bei. Der vorliegende Beitrag untersucht den Einfluss der Expositionshäufigkeit verschiedener Varianten eines Interface auf die Wahrnehmung von pragmatischen Erlebiskomponenten und dem allgemeinen Gefallen (engl.: Liking). Dabei zeigen sich Auswirkungen auf die antizipierte Usability von User Interfaces in der *pre-use* Phase, allerdings ohne dass diese Einflüsse auf die anschließende Nutzungsphase übergreifen.

1 Einleitung

Um das ganzheitliche Nutzererleben bei der Mensch-Maschine-Interaktion zu verstehen, ist es wichtig den Charakter verschiedener Phasen (Pohlmeier, 2011) näher zu betrachten. So verändern sich die Wahrnehmungen verschiedener Produkteigenschaften (Thüring & Mahlke, 2007) und das emotionale Erleben durch Nutzer- als auch Systemeigenschaften und spezifischer Interaktionscharakteristika während des gesamten Prozesses der Interaktion. Die Wahrnehmung der hedonischen und pragmatischen Produktqualitäten (Hassenzahl, 2003) ist dabei nicht unabhängig voneinander zu sehen, außerdem kann es zu HALO-Effekten kommen (Minge, 2011). Da ein komplexes Phänomen wie das des Nutzererlebens unterschiedlichsten Faktoren und Einflüssen unterliegt, müssen für ein umfassendes Verständnis dessen, diese Faktoren näher untersucht werden. Die vorliegende Arbeit fokussiert dabei den Einfluss der Expositionshäufigkeit eines Nutzers mit einem Interface in der *pre-use* Phase und daraus resultierende Effekte auf die Bewertung nach einer anschließenden Nutzungssituation (*use* Phase).

2 Studie

Im Rahmen der vorliegenden Untersuchung wurden zwei Varianten (siehe Abbildung 1) einer ÖPNV (öffentlicher Personennahverkehr) Applikation (App) in *pre-use* und *use* Phase verglichen, die sich hinsichtlich ihrer Usability stark unterschieden. In einer Vorbefragung ($N=43$, 20 weiblich/ 23 männlich, Alter: $M=30.8$, $SD=10.3$) konnten Screenshots dieser ÖPNV-App anhand eines siebenstufig, likert-skalierten Single-Items (gebrauchstauglich: 1=gar nicht, 7=voll und ganz) als höher ($M=4.77$, $SD=1.8$) und niedriger ($M=3.28$, $SD=1.7$) gebrauchstauglich (usable) validiert werden (siehe Abbildung 1).

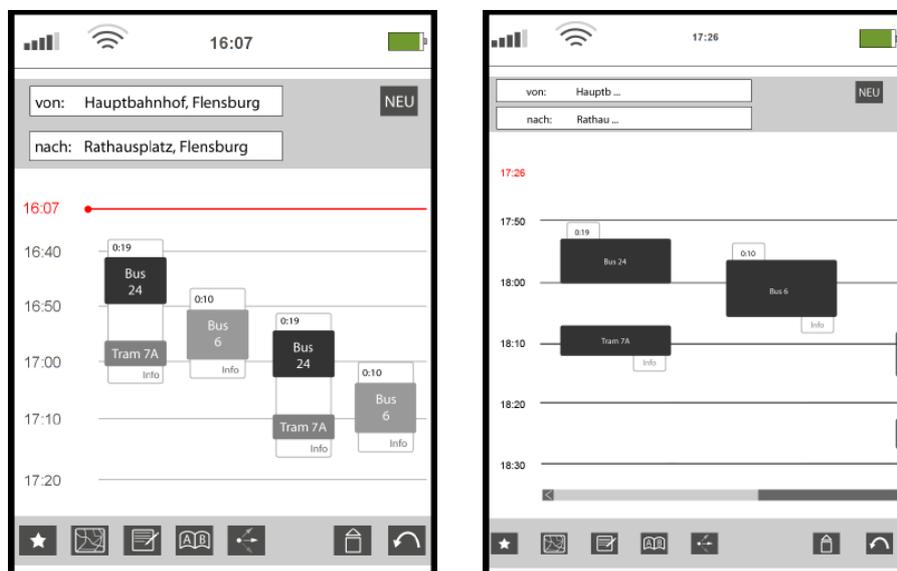


Abbildung 1: Interfacevarianten mit hoher (li.) und niedriger Usability (re.)

Bei einem Vergleich von diesen Gestaltungsvarianten über die Zeit und zunehmender Exposition hinweg, wird von einem scherenhaften Expositionseffekt ausgegangen. Dieser sollte sich in einer Verbesserung der Bewertung der ‚subjektiv wahrgenommenen Gebrauchstauglichkeit‘ hinsichtlich der hohen Usability-Variante und einer Verschlechterung dieser Bewertung für die niedrigere Usability-Variante zeigen. Ein derartiger Effekt wurde bereits für die Bewertung der visuellen Attraktivität von Interfaces in der *pre-use* Phase von Vogel (2013) festgestellt. Weiterführend stellt sich die Frage, ob ein derartiger Expositionseffekt in der *pre-use* Phase, Einfluss auf darauffolgende Phasen und das Nutzererleben während und nach der Interaktion hat.

2.1 Versuchsdesign

Im Rahmen der an die Vorbefragung anschließenden Laboruntersuchung wurde der Faktor ‚Expositionshäufigkeit‘ als Zwischensubjektfaktor variiert ($f = 1$ vs. 30 Darbietungen), dabei wurden die Probanden zufällig auf die Treatment- und Kontrollgruppe verteilt, wobei auf die gleichmäßige Verteilung von Männern und Frauen geachtet wurde. Der Faktor zur Gestaltung der ‚Usabilityvariante‘ (hoch vs. niedrig) ging als Innersubjektfaktor in das Versuchsdesign mit ein. Als abhängige Maße wurden die ‚subjektiv wahrgenommene Gebrauchstauglichkeit‘ und das ‚allgemeine Gefallen‘ (engl.: Liking) mittels siebenstufig, likert-skalierten Single-Items (gefällt mir: 1=gar nicht, 7= voll und ganz) gemessen. Den Probanden wurde vor Abgabe der ersten Bewertung für die ‚subjektiv wahrgenommene Gebrauchstauglichkeit‘ eine einheitliche Definition zur Usability präsentiert, die sich an der ISO 9241-210 (2010) und Attributen des AttrakDiffs (Hassenzahl, Burmester & Koller, 2003) orientierte. Als objektives abhängiges Maß wurde in der *pre-use* Phase die Reaktionszeit (RT), welches die Zeit vom Beginn der Reizdarbietung (SOA) bis zur Abgabe der Bewertung für die ‚subjektiv wahrgenommene Gebrauchstauglichkeit‘ auf der Tastatur darstellte, gemessen. In der *use* Phase wurde die Gesamtbearbeitungszeit (Total Task Time, TTT) für die zwei Aufgaben eines Aufgabenblocks gemittelt (siehe Abbildung 2).

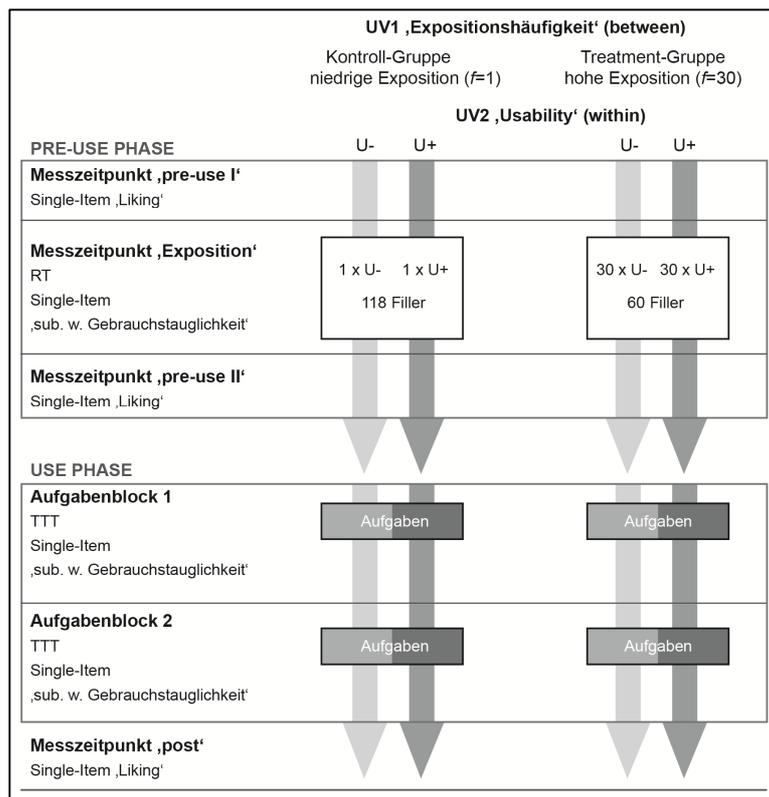


Abbildung 2: Versuchsdesign und Ablauf (RT=Reaktionszeit, TTT= Total Task Time)

Da im Rahmen des ContinUE Modells (Pohlmeier, 2011) die Vorerfahrung als wichtiger Einflussfaktor für das Nutzererleben in der *pre-use* Phase benannt wird, wurde diese als dichotomer Zwischensubjektfaktor (vorhanden vs. nicht vorhanden) im Umgang mit Apps für den ÖPNV in das Design mit einbezogen.

2.2 Versuchsablauf & Stichprobe

Der Versuch unterteilte sich in die *pre-use* und *use* Phase (siehe Pohlmeier, 2011), wobei die Manipulation der ‚Expositions Häufigkeit‘ nur in der *pre-use* Phase stattfand. In der Treatment-Gruppe (TG) wurden das System mit ‚niedriger Usability‘ (U-) und mit ‚hoher Usability‘ (U+) jeweils 30 × präsentiert. In der Kontroll-Gruppe (KG) wurden diese Zielstimuli jeweils einmal dargeboten. Zusätzlich wurden „Filler-Bilder“ (FB) eingefügt, die Ansichten anderer Funktionen (u.a. Kartenkauf) der ÖPNV-App abbildeten. Die Zielstimuli wurden mit diesen FBern in sechs Blöcken randomisiert dargeboten. Die Zielstimuli traten demzufolge in der TG maximal zwei Mal hintereinander auf. Die Bewertung der ‚subjektiv wahrgenommenen Gebrauchstauglichkeit‘ erfolgte sowohl nach einem Zielstimulus als auch nach einem FB. Es mussten insgesamt 120 Bewertungen in der Expositionsphase abgegeben werden (siehe Abbildung 2).

Die *use* Phase bestand aus der Bearbeitung von zwei Aufgabenblöcken für jede Gestaltungsvariante in ausbalancierter Reihenfolge. Jeder Block bestand aus zwei Aufgaben: 1. Anzeige aller möglichen Abfahrten an einem Standort zu einer bestimmten Uhrzeit, 2. Anzeige einer selbstausgewählten Verbindung zu einer bestimmten Uhrzeit. Die Messung des ‚Likings‘ erfolgte für jeden Zielstimulus zu drei Messzeitpunkten: 1. zu Beginn der *pre-use* Phase, 2. nach der Expositionsphase (Ende der *pre-use* Phase) und 3. am Ende der *use* Phase. Die ‚subjektiv wahrgenommene Gebrauchstauglichkeit‘ wurde nach jedem Aufgabenblock (jeweils zwei Messzeitpunkte für jede Systemvariante) erhoben (siehe Abbildung 2).

Zum Abschluss wurden demografische Angaben (Alter, Geschlecht, Bildungsstand) und die Vorerfahrung im Umgang mit ähnlichen Applikationen für den ÖPNV erhoben.

An dem Versuch nahmen 40 Personen freiwillig teil, davon konnten 32 Probanden ($N=17$ weiblich / 15 männlich, Alter: $M=25.8$, $SD=4.2$) in die Auswertung miteinbezogen werden. Acht Personen mussten aufgrund technischer Probleme ausgeschlossen werden, da die Daten nicht zuverlässig aufgezeichnet wurden. Der TG sowie KG wurden jeweils 16 Personen zufällig zugeordnet. Die Versuchspersonen wurden monetär (EUR 10,00) oder mittels für das Studium nötige Versuchspersonenstunden (1 Std.) entlohnt.

3 Ergebnisse

Aufgrund der unterschiedlichen Messzeitpunkte und Phasen wurden mehrere, voneinander unabhängige Varianzanalysen für die abhängigen Maße ‚Liking‘ und ‚subjektiv wahrgenommene Gebrauchstauglichkeit‘ berechnet, wobei die Phasen (*pre-use*, *use*) ebenfalls einzeln voneinander analysiert wurden.

3.1 Pre-Use Phase

Um einen Expositionseffekt für die Treatment-Gruppe identifizieren zu können, wurde eine $2 \times 2 \times 30$ Varianzanalyse mit den Innersubjektfaktoren ‚Usabilityvariante‘ (hoch/niedrig) und der ‚Exposition‘ (30 Messzeitpunkte), sowie dem Zwischensubjektfaktor ‚Vorerfahrung‘ (hoch/niedrig) berechnet. Als abhängige Maße wurden die ‚subjektiv wahrgenommene Gebrauchstauglichkeit‘ und die ‚Reaktionszeit‘ (RT) in einer Varianzanalyse untersucht.

Es konnten ein signifikanter Haupteffekt des Faktors ‚Usabilityvariante‘ [$F(1,14)=5.25$, $p=.04$, $\eta^2_{PART}=.28$] und ‚Vorerfahrung‘ [$F(1,14)=10.12$, $p=.007$, $\eta^2_{PART}=.42$] auf die Bewertung der ‚subjektiv wahrgenommenen Usability‘ gezeigt werden. Dabei wurde das Interface mit ‚hoher Usability‘ ($M=4.75$, $SD=0.3$) erwartungsgemäß besser bewertet als die Variante mit ‚niedriger Usability‘ ($M=3.92$, $SD=0.3$). Die Probanden mit Vorerfahrung gaben dabei generell eine höhere Bewertung ab ($M=5.12$, $SE=.30$) als Probanden ohne Vorerfahrung ($M=3.56$, $SE=.39$). Zusätzlich zeigt sich eine signifikante Interaktion zwischen den Faktoren ‚Usabilityvariante‘ und ‚Exposition‘ [$F(29,406)=1.86$, $p=.005$, $\eta^2_{PART}=.12$], welches den vermuteten scherenhaften Expositionseffekt bestätigt (siehe Abbildung 3).

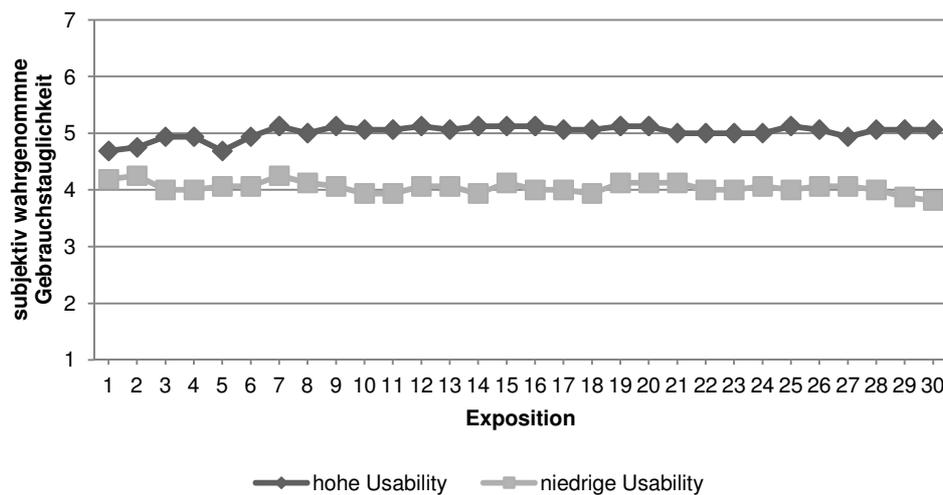


Abbildung 3: Interaktion und Einfluss von ‚Exposition‘ und ‚Usabilityvariante‘ auf die Bewertung der ‚subjektiv wahrgenommenen Gebrauchstauglichkeit‘.

Bei der Analyse der Reaktionszeit (RT) konnte ein signifikanter Haupteffekt der ‚Exposition‘ [$F(29,406)=8.25$, $p<.001$, $\eta^2_{PART}=.37$] ermittelt werden, wobei die RT mit zunehmender Exposition für beide Systemvarianten in gleicher Weise abnimmt.

3.2 Use Phase

Für die Analyse der *use* Phase wurde eine vier-faktorielle MANOVA mit den Innersubjekt-faktoren ‚Usabilityvariante‘ (hoch/niedrig), ‚Aufgabenblock‘ (1/2), den Zwischensubjekt-faktoren ‚Expositionshäufigkeit‘ (hoch/niedrig) und ‚Vorerfahrung‘ (hoch/niedrig) und den abhängigen Maßen ‚subjektiv wahrgenommene Gebrauchstauglichkeit‘ und ‚Total-Task-Time‘ (TTT in ms) berechnet.

In dieser Phase wurde ein signifikanter Haupteffekt für den Faktor ‚Usabilityvariante‘ und dessen Einfluss auf die ‚subjektiv wahrgenommene Gebrauchstauglichkeit‘ (siehe Tabelle 1) sowie der ‚TTT‘ (siehe Tabelle 1) ermittelt. Dieser Effekt ist erwartungskonform ausgerichtet, d.h. es kommt zu einer besseren Bewertung und schnelleren Erfüllung der Aufgaben für die Variante mit ‚hoher Usability‘ als bei der Variante mit ‚niedriger Usability‘.

Zusätzlich wurden ein marginal signifikanter Haupteffekt für den Faktor ‚Aufgabenblock‘ auf die Bewertung der ‚subjektiv wahrgenommenen Gebrauchstauglichkeit‘ (siehe Tabelle 1) und ein signifikanter Einfluss auf die ‚TTT‘ (siehe Tabelle 1) identifiziert.

| Quelle der Varianz | df | F-Werte und Effektstärke (η^2_{PART}) | |
|------------------------------|----|---|---------------------------|
| | | Single-Item „subjektiv wahrgenommene Gebrauchstauglichkeit“ | Total Task Time |
| ‚Usabilityvariante‘ (UV) | 1 | 91.20*** (0.77) | 35.35*** (0.56) |
| ‚Expositionshäufigkeit‘ (E) | 1 | 0.19 (0.01) | 2.33 (0.08) |
| ‚Aufgabenblock‘ (AB) | 1 | 3.36(*) (0.11) | 41.34*** (0.60) |
| UV × E | 1 | 0.05 (0.002) | 0.50 (0.02) |
| UV × AB | 1 | 0.57 (0.02) | 11.48** (0.29) |
| Fehler innerhalb der Gruppen | 28 | [1.85] | [159306.42] |

Tabelle 1: Varianzanalyse für das Nutzererleben in der use Phase. Werte in runden Klammern repräsentieren die zugehörige Effektstärke (η^2_{PART}). Werte in eckigen Klammern repräsentieren die mittleren Quadratfehler.
Anmerkung: (*) $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Es zeigt sich zudem eine signifikante Interaktion zwischen den Faktoren ‚Usabilityvariante‘ und ‚Aufgabenblock‘ für die ‚TTT‘ (siehe Tabelle 1 und Abbildung 4).

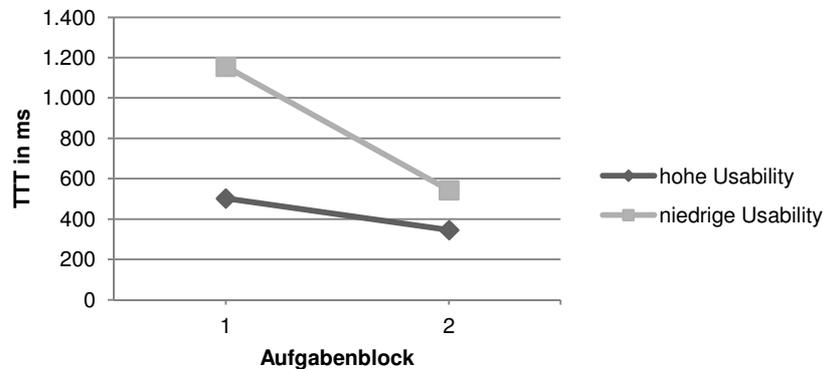


Abbildung 4: Interaktion von ‚Usabilityvariante‘ und ‚Aufgabenblock‘ auf die ‚Total-Task-Time‘ (TTT in ms) in der use Phase.

Jedoch konnte kein Einfluss der ‚Expositionshäufigkeit‘ in der use Phase beim Vergleich der Treatment- und Kontroll-Gruppe hinsichtlich der ‚subjektiv wahrgenommenen Gebrauchstauglichkeit‘ (siehe Tabelle 1) beobachtet werden.

3.3 Liking

Hinsichtlich des ‚Likings‘ konnte kein Effekt der ‚Expositionshäufigkeit‘ [$F(1,29)=0.55$, $p=.46$, $\eta^2_{PART}=.02$] und des ‚Messzeitpunktes‘ (vor/ nach Expositionsphase) [$F(1,28)=2.44$, $p=.13$, $\eta^2_{PART}=.08$] in der pre-use Phase festgestellt werden. Erst in der use Phase hat die ‚Usabilityvariante‘ [$F(1,28)=24.68$, $p<.001$, $\eta^2_{PART}=.47$] einen signifikanten Einfluss auf die Bewertung des ‚Likings‘ (siehe Abbildung 5).

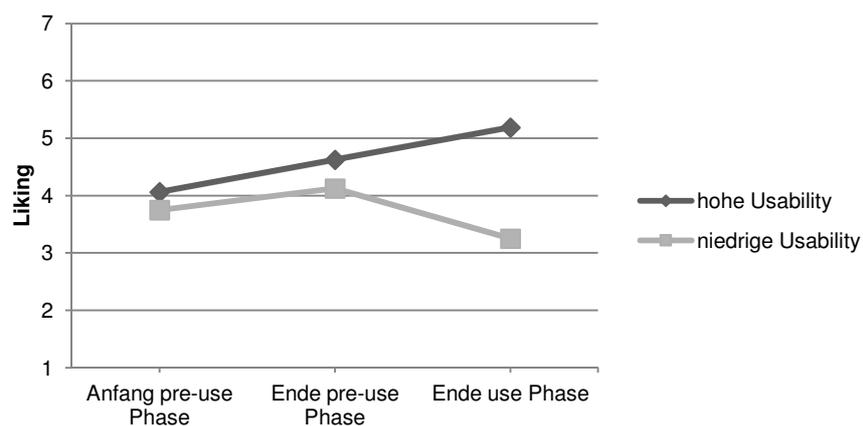


Abbildung 5: Einfluss von ‚Usabilityvariante‘ auf die Bewertung des ‚Likings‘ in der pre-use und use Phase (1=gefällt mir gar nicht, 7=gefällt mir voll und ganz).

Allerdings konnte in der *pre-use* Phase ein signifikanter Einfluss der Vorerfahrung [$F(1,29)=5.67, p=.02, \eta^2_{PART}=.17$] im Umgang mit ÖPNV-Apps und eine signifikante Interaktion der ‚Usabilityvarianten‘ mit dieser ‚Vorerfahrung‘ [$F(1,28)=23.03, p<.001, \eta^2_{PART}=.45$] auf das ‚Liking‘ gezeigt werden (siehe Abbildung 6).

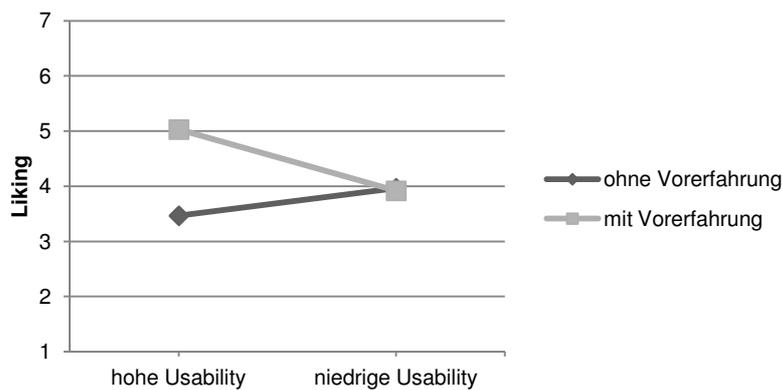


Abbildung 6: Einfluss und Interaktion der ‚Vorerfahrung‘ und ‚Usabilityvariante‘ auf die Bewertung des ‚Likings‘ (1=gefällt mir gar nicht, 7=gefällt mir voll und ganz) in der *pre-use* Phase.

4 Diskussion & Ausblick

Im Rahmen der zuvor beschriebenen Auswertung und Analyse konnte ein Einfluss der ‚Expositionshäufigkeit‘ auf die Bewertung der ‚subjektiv wahrgenommenen Gebrauchstauglichkeit‘ in der Treatment-Gruppe identifiziert werden. Bei einem derartigen Effekt wird auch von einem *Mere Exposure Effekt* (Zajonc, 1968) gesprochen. Im vorliegenden Experiment wird eine scherenhafte Ausprägung erkennbar, ähnlich wie bei den Versuchen von Brickman, Redfield, Harrison und Crandall (1972) und Grush (1976). Es kommt dabei zu einer Verbesserung der Bewertung der höher gebrauchstauglichen Systemvariante und zu einer Verschlechterung der Bewertung für die geringer gebrauchstaugliche Variante. Im weiteren Verlauf des Nutzererlebenszykluses wird aber erkennbar, dass dieser Effekt der Expositionshäufigkeit keinen Einfluss auf die Bewertung der ‚subjektiv wahrgenommenen Gebrauchstauglichkeit‘ in der anschließenden *use* Phase hat. Die Bewertung der Gebrauchstauglichkeit orientiert sich sowohl bei der Treatment- als auch Kontroll-Gruppe an der tatsächlichen Ausprägung dieser pragmatischen Produktqualität. Die Erfahrung während der tatsächlichen Interaktion ist in dieser Phase ausschlaggebend für die Bewertung.

Hinsichtlich des ‚Allgemeinen Gefallens‘ (Liking) konnte kein Unterschied zwischen hoher und niedriger Expositionshäufigkeit identifiziert werden und damit auch kein *Mere Exposure Effekt* auf dieser für die klassischen *Mere Exposure* Experimente (Bornstein, 1989) typischen Bewertungsdimension. Wie in einem Versuch von Carbon und Leder (2005) gezeigt wurde, kann eine wiederholte Evaluation auf unterschiedlichen Dimensionen zu einer verbesserten Bewertung und höheren Attraktivität führen. Diese Evaluationstechnik wird als Repeated-

Evaluation-Technique (RET) bezeichnet. Es stellt sich die Frage, ob demzufolge die wiederholte und bewusste Evaluation der Gebrauchstauglichkeit zu einem Expositionseffekt geführt hat und ob dieser auch auftreten würde, wenn nur vor und nach der Expositionsphase nach der Bewertung dieser Eigenschaft gefragt werden würde. Inwiefern kann also von einem *Mere* (,reinen‘) Expositionseffekt ausgegangen werden. Dies sollte in zukünftigen Untersuchungen fokussiert werden. Zudem konnte festgestellt werden, dass die Vorerfahrung nur in der *pre-use* Phase einen Einfluss auf die Bewertung der ‚Gebrauchstauglichkeit‘ hat und somit die Annahmen des Modells ContinUE gestützt werden können.

Literaturverzeichnis

- Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological Bulletin*, 106(2), 265–289.
- Brickman, P., Redfield, J., Harrison, A.A. & Crandall, R. (1972). Drive and predisposition as factors in the attitudinal effects of mere exposure. *Journal of Experimental Social Psychology*, 8(1), 31-44.
- Carbon, C.-C., & Leder, H. (2005). The Repeated Evaluation Technique (RET). A method to capture dynamic effects of innovativeness and attractiveness. *Applied Cognitive Psychology*, 19(5), 587–601.
- Grush, J.E. (1976): Attitude Formation and Mere Exposure Phenomena: A Nonartifactual Explanation of Empirical Findings. *Journal of Personality and Social Psychology*, 33(3), 281-290.
- Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In G. Szwillus & J. Ziegler (Hrsg.), *Mensch und Computer 2003: Interaktion in Bewegung*, 187-196. Stuttgart: B.G. Teubner.
- ISO 9241-210 (2010). *Ergonomics of human-computer interaction – Part 210: Human Centered design process for interactive systems*. Geneva: International Standardization Organization (ISO).
- Pohlmeyer, A.E. *Identifying Attribute Importance in Early Product Development*. Technische Universität Berlin, Ph.D. thesis (2011).
- Vogel, M. (2013). Temporal Evaluation of Aesthetics of User Interfaces as one Component of User Experience. In Smith, R.T. and Wünsche, B.C. (eds.) Proceedings of the Fourteenth Australasian User Interface Conference (AUIC2013), Adelaide, Australia. *Conferences in Research and Practice in Information Technology Series*, 139, 131-132, Australien Computer Society Inc.
- Zajonc, R.B. (1968). Attitudinal Effects of Mere Exposure. *Journal of Personality and Social Psychology*, 9(2), 1-27.

Kontaktinformationen

Dipl.-Des., M.Sc. Marlene Vogel, marlene.vogel@zmms.tu-berlin.de