

# Kann man Website-Usability schnell valide mit Screenshots erfassen?

Meinald T. Thielsch<sup>1</sup>, Ronja Engel<sup>1</sup>, Gerrit Hirschfeld<sup>2</sup>

Institut für Psychologie, Westfälische Wilhelms-Universität Münster<sup>1</sup>  
Deutsches Kinderschmerzszentrum, Vestische Kinder- und Jugendklinik Datteln<sup>2</sup>

## Zusammenfassung

Um die Frage zu beantworten ob sich Website-Usability schnell valide anhand von Screenshots beurteilen lässt, wurde die wahrgenommene Usability mit objektiven Maßen der tatsächlichen Leistung verglichen. Insgesamt beurteilten 57 Probanden die Usability nach der unterschiedlich langen Darbietung von Screenshots (50 ms, 500 ms und 10 s) sowie nach der interaktiven Nutzung von Testwebsites. Es zeigte sich, dass die Beurteilungen anhand von Screenshots für alle Darbietungszeiten weder mit den Urteilen nach der Interaktion noch mit den objektiven Maßen der Benutzerfreundlichkeit signifikant korrelierten. Stattdessen hängen die Usability-Beurteilungen anhand von Screenshots sehr stark mit der Ästhetik-Bewertung zusammen. Nur die Usability-Bewertungen nach einer Interaktion mit der Website korrelieren hoch signifikant mit objektiven Maßen wie Bearbeitungsdauer oder Aufgabenlösung und zeigen keine signifikante Korrelation zu Ästhetik-Einschätzungen. Usability-Bewertungen, die allein auf Basis von Screenshots gemacht werden, sind durch die wahrgenommene Ästhetik beeinflusst und als alleinige Basis für Usability-Tests kritisch zu sehen.

## 1 Theoretischer Hintergrund

Usability ist ein Kernkonstrukt der Mensch-Maschine-Interaktion (Hornbaek, 2006) und definiert sich über die Effektivität, Effizienz und Zufriedenheit mit der Benutzer mit einem interaktivem System vorgegebene Ziele erreichen können (ISO, 2010). Website-Usability ist demnach nicht alleine die Eigenschaft einer Website, sondern entsteht durch die Interaktion mit dieser in einem bestimmten Kontext. Es existieren unterschiedliche Ansätze diese zu erfassen von formal-analytischen Verfahren, bei denen Experten anhand festgelegter Kriterien die Usability bewerten, bis hin zu Fragebogeninstrumenten und „Think-aloud“-Usability-Tests (vgl. Gediga & Hamborg, 2002; Salaschek et al., 2007). In verschiedenen Studien, insbesondere bei der Erfassung des Ersteindrucks, wird Website-Usability anhand von Screenshots untersucht. Hierbei wurde festgestellt, dass Urteile über die Usability nach 50 ms Darbietungszeit eines Screenshots sehr hoch mit denen nach 500 s korrelieren (Lindgaard et al., 2011). Aufgrund der fehlenden Interaktionsmöglichkeit stellt sich jedoch

die Frage worauf die Probanden ihr Urteil basierten. So ist anzunehmen, dass verwandte Konstrukte wie Ästhetik (vgl. Moshagen & Thielsch, 2010) diese Ersteindrucksurteile beeinflussen (vgl. Tractinsky et al., 2000).

Zentrale Frage unserer Studie ist, ob pre-use Usability-Urteile auf Basis von Screenshots mit denen nach einer interaktiven Auseinandersetzung mit der Webseite vergleichbar sind. Außerdem soll geprüft werden, inwieweit die wahrgenommene Usability mit objektiven Maßen zusammenhängt und durch die wahrgenommene Ästhetik beeinflusst wird.

## 2 Methodisches Vorgehen

### 2.1 Stichprobe

An der Studie nahmen 57 Personen im Alter von 19 bis 32 Jahren ( $M = 23.09$ ;  $SD = 2.79$ ) teil, 70.2 % ( $n = 40$ ) waren weiblich. Die Teilnahme war freiwillig und anonym.

### 2.2 Versuchsmaterial

Die Stimuli bestanden aus 50 Screenshots von Webseiten aus zehn Website-Kategorien, die auf Basis bestehender Testsets zusammengestellt wurden. Um Einflüsse der Bekanntheit auf die Ergebnisse zu vermeiden, wurden gezielt unbekannte Websites ausgewählt.

### 2.3 Versuchsablauf und erfasste Variablen

Die Datenerhebung fand in kleinen Gruppen von bis zu acht Personen statt. Nach einer einführenden schriftlichen Instruktion folgten die Bewertungen der Websites als Screenshots in drei verschiedenen Zeitbedingungen und als vierter Block die Interaktion.

Die Usability wurde jeweils anhand einer sechsstufigen Likert-Skala von „gar nicht benutzerfreundlich“ bis „sehr benutzerfreundlich“ gegeben. Jede Bewertung wurde abgegeben nachdem ein Screenshot für eine fest definierte Zeitspanne (50, 500 oder 10.000 ms) präsentiert wurde. Jeder Block bestand aus 50 Trials, sodass jede der 50 Webseiten in randomisierter Reihenfolge einmal pro Block präsentiert wurde.

Anschließend sollten sich die Teilnehmer interaktiv mit einer Webseite auseinandersetzen, bevor sie erneut Urteile über die Usability abgaben. Für die interaktiven Aufgaben wurde aus jeder Kategorie eine der fünf Webseiten des Gesamtsets ausgewählt (eine Darbietung des gesamten Versuchssets hätte eine zu große zeitliche Belastung für die Probanden dargestellt). Bei der Formulierung der Fragen wurden möglichst mittelschwere Aufgaben angestrebt, so dass alle Aufgaben mit drei bis vier Klicks zu lösen waren. Als Antwort wurden jeweils fünf Alternativen im Multiple-Choice Format vorgegeben von denen immer eine richtig war und drei als Distraktoren dienten. Die fünfte Alternative lautete immer „Ich habe mich bemüht, die Antwort jedoch nicht gefunden“. Als objektive Maße wurden die bei der Aufgabebearbeitung benötigte Zeit und Korrektheit der Lösung erfasst.

### 3 Ergebnisse

Die subjektiven Usability-Bewertungen auf Basis von Screenshots die 50 ms präsentiert wurden korrelieren sehr hoch mit den Bewertungen nach 500 ms ( $r = .89, p < .05$ ) und nach 10 s Präsentation ( $r = .87, p < .05$ ). Keine dieser Bewertungen korrelierte signifikant mit den Bewertungen nach der Interaktionsbedingung und den objektiven Maßen (siehe Tabelle 1).

Es zeigt sich allerdings, dass die subjektiven Usability-Bewertungen in den verschiedenen Screenshotbedingungen sehr hoch mit Ästhetikeinschätzungen korrelieren (siehe Tabelle 1).

Usability-Bewertung in der Bedingung...	Bearbeitungs-dauer	Aufgaben-lösung	Ästhetik-bewertung
Screenshot für 50 ms	.08	.19	.75**
Screenshot für 500 ms	-.11	.07	.75**
Screenshot für 10 s	-.19	-.07	.84**
Nach Interaktion mit Website	-.86**	.62*	.05

Tabelle 1: Korrelation der subjektiven Usability-Bewertungen mit den objektiven Usability-Maßen Bearbeitungsdauer und Aufgabenlösung sowie der subjektiven Ästhetikbewertung; \* =  $p < .05$ ; \*\* =  $p < .01$

### 4 Diskussion

Die Daten legen nahe, dass schnelle Usability-Bewertungen auf Basis von Screenshots vor allem die subjektive Ästhetik widerspiegeln. Um Website-Usability valide zu erfassen, das heißt mit Ersteindrucksmessungen die auch prädiktiv für tatsächliche Performanz sind, müssen die Befragten zuvor mit der Website interagieren können.

Diese Ergebnisse stehen im Einklang mit Studien, die eine sehr schnelle Verarbeitung ästhetischer Reize annehmen (vgl. Thielsch & Hirschfeld, 2012) und die von einem Halo-Effekt von Ästhetik- auf Usability-Urteile ausgehen (vgl. Hassenzahl & Monk, 2010). Bei der Interpretation von Studien, die die Usability ausschließlich anhand von Screenshots bewerten, ist nach unseren Ergebnissen Vorsicht geboten. Die vorliegenden Befunde untermauern nicht nur die Forderung von Hornbaek (2006) gleichzeitig objektive und subjektive Usability-Indikatoren zu erfassen. Noch mehr legen sie nahe, Webusability-Tests auf echten Interaktionen zu basieren. Die Idee in einzelnen Studien, dass Probanden allein auf Basis von Screenshots anhand von Cues wie Suchfeldern oder Navigationsbenennungen Usability valide einschätzen können, scheint aus unserer Sicht nicht haltbar.

#### Literaturverzeichnis

Gediga, G., & Hamborg, K. (2002). Evaluation in der Software-Ergonomie: Methoden und Modelle im Software-Entwicklungsprozess. *Zeitschrift für Psychologie, 210*(1), 40–57.

- Hassenzahl, M., & Monk, A. (2010). The Inference of Perceived Usability From Beauty. *Human-Computer Interaction*, 25(3), 235-260.
- Hornbaek, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies*, 64(2), 79–102.
- ISO (2010). *Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems (ISO 9241-210:2010)*. Geneva: International Organization for Standardisation.
- Lindgaard, G., Dudek, C., Sen, D., Sumegi, L. & Noonan, P. (2011). An exploration of relations between visual appeal, trustworthiness and perceived usability of homepages. *ACM Transactions On Computer-Human Interaction*, 18(1), 1-30.
- Moshagen, M. & Thielsch, M. T. (2010). Facets of visual aesthetics. *International Journal of Human-Computer Studies*, 68, 689-709.
- Salaschek M., Holling H., Freund P.A. & Kuhn J-T. (2007). Benutzbarkeit von Software. Vor- und Nachteile verschiedener Methoden und Verfahren. *Zeitschrift für Evaluation*, 2, 1-30.
- Thielsch, M. T. & Hirschfeld, G. (2012). Spatial frequencies in aesthetic website evaluations – explaining how ultra-rapid evaluations are formed. *Ergonomics*, 55 (7), 731-742.
- Tractinsky, N., Katz, A. S. & Ikar, D. D. (2000). What is beautiful is usable. *Interacting With Computers*, 13(2), 127-145.