

Der VisAWI im Praxiseinsatz: Best Practices, neue Benchmarks und neue Entwicklungen

Meinald T. Thielsch
Westfälische Wilhelms-Universität Münster
Fliednerstr. 21
48149 Münster
thielsch@uni-muenster.de

Jan-Hendrik Spieth
Audials AG
Erbprinzenstraße 27
76133 Karlsruhe
jh.spieth@gmail.com

Marie Jahn
BASF Coatings GmbH
Glasuritstraße 1
48165 Münster
marie-christin.jahn@basf.com

Gerrit Hirschfeld
Deutsches Kinderschmerzzentrum
Dr.-Friedrich-Steiner-Str. 5
45711 Datteln
gerrit.hirschfeld@gmail.com

Franz Koller
User Interface Design GmbH
Wilhelm-Bleyle-Straße 10-12
71636 Ludwigsburg
franz.koller@uid.com

Abstract

Die visuelle Ästhetik beeinflusst zentrale Aspekte der Webseitenwahrnehmung wie Usability, Zufriedenheit, Wiederbesuchs- oder Weiterempfehlungsbereitschaft. Um die Ästhetik von Websites adäquat zu erfassen wurde der VisAWI (Visual Aesthetics of Websites Inventory, Moshagen & Thielsch, 2010) entwickelt. In diesem Beitrag möchten wir den VisAWI sowie zwei Best Practices zum Einsatz des Fragebogens darstellen: die Optimierung der Karrierewebsites der BASF Coatings und die Evaluation der PC-Software „Audials“. In letzterem Best Practice zeigt sich, wie der VisAWI jenseits des Webs angewendet werden kann. Zudem berichten wir aktualisierte Benchmarks auf Basis von $n = 5766$ Bewertungen mit dem VisAWI, bzw. $n = 6797$ für den VisAWI-S, sowie kritische Schwellenwerte (Cut-Points) für die beiden VisAWI-Varianten. Abschließend zeigen wir kurz aktuelle Entwicklungen des VisAWI auf.

Keywords:

VisAWI, Benchmarks, Ästhetik, Schönheit, User Experience

1. Hintergrund

Die Forschung zur User Experience im World Wide Web hat die erlebte visuelle Ästhetik als bedeutendes Konstrukt herausgestellt (Moshagen & Thielsch, 2010). Website-Ästhetik kann als unmittelbare angenehme subjektive Wahrnehmung des Webobjekts definiert werden, die wenig durch schlussfolgernde Prozesse beeinflusst wird (Moshagen & Thielsch, 2010). Ästhetik bereitet Freude – die klassische Ästhetikforschung spricht hier basierend auf den Arbeiten von Fechner (1876) von einem „subjektivem Wohlgefallen“. Die Ästhetik einer Website beeinflusst nicht nur besonders den Ersteindruck eines Nutzers (vgl. Lindgaard et al., 2006; Thielsch & Hirschfeld, 2012; Tuch et al., 2012;), sie beeinflusst auch eine Reihe weiterer Aspekte wie Usability (Lee & Koubek, 2012; Moshagen, Musch, & Göritz, 2009; Sonderegger & Sauer, 2010), Zufriedenheit (Cyr, Kindra & Dash, 2008; Lindgaard & Dudek, 2003), Kaufbereitschaft (Parboteeah, Valacich & Wells, 2009; Porat & Tractinsky, 2012) oder Wiederbesuchs- und Weiterempfehlungsbereitschaft (Mahlke, 2002; Thielsch, Blotenberg & Jaron, 2014; Yoo & Donthu, 2001).

Dementsprechend wichtig ist, dass die erlebte Ästhetik adäquat erfasst wird. Hierzu werden meistens die Nutzer nach ihren subjektiven Urteilen gefragt. Andere Methoden wie beispielsweise Paarvergleiche, Checklisten-Evaluationen, oder Cognitive Walkthroughs sind ebenso möglich, aber bisher nur an wenigen Stellen dokumentiert (siehe Thielsch, 2008, S. 53). Fragebogenmethoden sind beliebt, vergleichsweise leicht umzusetzen und entsprechend verbreitet. Bewertungen anhand von nur einer einzelnen Frage (bspw. „Wie schön ist die Website?“) können allerdings durch Messfehler und Antwortverzerrungen verfälscht sein (siehe Schmidt & Hunter, 1996). Daher ist es von Vorteil, mehrere Fragen zur Ästhetik in Form einer Skala oder eines standardisierten Fragebogens zu stellen. Grundsätzlich muss zudem bei jeder metrischen Erfassung von Website-Qualitäten die Validität (inhaltliche Gültigkeit) des Instrumentes geprüft werden. Wäre ein Ästhetik-Fragebogen beispielsweise sehr abhängig von der Stimmung eines Befragten oder stark beeinflusst von Usability und/oder Inhaltsqualität einer Website, wäre eine valide Messung von Ästhetik erschwert.

Die vergleichsweise wenigen existierenden Ästhetik-Fragebögen sind jedoch zum Teil ad hoc für eine einzelne Untersuchung gebildete Skalen, deren Validität nicht geprüft wurde und damit fraglich ist (Bargas-Avila et al., 2011). Ausnahmen stellen das von Lavie und Tractinsky (2004) entwickelte Instrument sowie der VisAWI (Visual Aesthetics of Websites Inventory; Moshagen & Thielsch, 2010 & 2013) dar. Während der Fragebogen von Lavie und Tractinsky (2004) zwei Faktoren, klassische und expressive visuelle Ästhetik annimmt, ergänzt der VisAWI dieses Modell und geht von vier Facetten aus: Einfachheit, Vielfalt, Farbigkeit und Kunstfertigkeit. Der VisAWI soll im Folgenden im Detail dargestellt werden.

2. Der VisAWI

Der VisAWI basiert auf dem Modell, dass Nutzer einen Generalfaktor Ästhetik wahrnehmen, der aus vier zugrundeliegenden Facetten besteht:

- Die Items zur **Einfachheit** fragen, wie übersichtlich und strukturiert das Layout einer Website wahrgenommen wird.
- Die Skala **Vielfalt** hingegen bittet um eine Bewertung von Originalität und Dynamik des Webdesign.

- **Farbigkeit** umfasst Aspekte von ästhetischer Farbauswahl, -einsatz und -kombination.
- Die vierte Skala **Kunstfertigkeit** bezieht sich auf die Aktualität, Durchdachtheit und Professionalität einer Website.

Dem Nutzer werden mit jeder Skala Aussagen zu Gestaltungsmerkmalen einer Website präsentiert (z. B. „Die Farben haben eine angenehme Wirkung.“), welchen der Nutzer dann zustimmen oder ablehnen kann (siehe Tabelle 1).

Die Konstruktion und Validierung des VisAWI stützt sich auf sieben Studien mit insgesamt 2027 Befragten (für eine detaillierte Darstellung siehe Moshagen & Thielsch, 2010). Aufbauend auf einer Analyse der vorhandenen Forschung zur Website-Ästhetik und einer Expertenbefragung wurde ein Entwurf des Fragebogens mit 96 Items geschaffen. Verschiedene Websitetests und Experteninspektionen führten zur finalen Version des VisAWI mit insgesamt 18 Items (siehe Tabelle 1).

Tabelle 1: Items des VisAWI und der Kurzversion VisAWI-S.

Item	Ästhetikfacette
(r) Das Layout wirkt zu gedrängt. Das Layout ist gut zu erfassen. Das Layout erscheint angenehm gegliedert.	Einfachheit ($\alpha = .89$)
(r) Die Seite erscheint zu uneinheitlich. Auf der Seite passt alles zusammen. *	
(r) Die Seitengestaltung ist uninteressant. Das Layout ist originell.	
(r) Die Gestaltung wirkt einfallslos. Das Layout wirkt dynamisch. Die Seite ist angenehm vielseitig. *	Vielseitigkeit ($\alpha = .87$)
Die farbliche Gesamtgestaltung wirkt attraktiv. *	
(r) Die Farben passen nicht zueinander. (r) Der Farbeinsatz ist nicht gelungen. Die Farben haben eine angenehme Wirkung.	Farbigkeit ($\alpha = .89$)
Das Layout ist professionell. *	
(r) Das Layout ist nicht zeitgemäß. Die Seite erscheint mit Sorgfalt gemacht.	Kunstfertigkeit ($\alpha = .85$)
(r) Das Layout wirkt konzeptlos.	

Anmerkung: Als Antwortskala dient eine 7-stufige Likert-Skala (von 1 „stimme überhaupt nicht zu“ bis 7 „stimme voll zu“). Mit * markiert sind die vier Items der Kurzversion VisAWI-S. (r) = vor der Bildung des Skalenmittels ist dieses Item umzukodieren. In Klammern jeweils Cronbach's Alpha der VisAWI-Skalen.

Für den VisAWI zeigen sich gute bis sehr gute **Reliabilitätswerte**, die innere Konsistenz (Cronbachs Alpha) ist bei allen vier Skalen $\alpha \geq .85$, für den Gesamtwert des VisAWI ist $\alpha = .94$.

Die hierarchische Faktorstruktur des VisAWI wurde mit exploratorischen und konfirmatorischen Faktorenanalysen bestimmt und erfolgreich an zwei unabhängigen Stichproben konfirmiert und damit kreuzvalidiert. Die konvergente, divergente, diskriminative und konkurrente **Validität** wurde in drei Studien nachgewiesen (siehe Moshagen & Thielsch, 2010). Dabei zeigten sich hohe Korrelationen zum

Messinstrument von Lavie und Tractinsky (2004) im Bereich von $.52 \leq r \leq .82$ für die Skala zur klassischen Ästhetik und $.36 \leq r \leq .80$ für die zur expressiven Ästhetik. Ebenso finden sich hohe Korrelationen ($.60 \leq r \leq .80$) zur Attraktivitätsskala aus dem AttrakDiff 1 (siehe z. B. Hassenzahl, Burmester & Koller, 2003). Schwächere Korrelationen zeigen sich zu divergenten Konstrukten, wie der Skala zur pragmatischen Qualität aus dem AttrakDiff ($.41 \leq r \leq .77$), einer Usability-Skala nach Flavián, Guinalú und Gurrea (2006, $.04 \leq r \leq .48$) oder dem WWI ($.48 \leq r \leq .62$), einem Instrument zur Bewertung der Inhaltsqualität (Thielsch, 2008). Weiterhin zeigte sich in zwei experimentellen Validierungen, dass die Facetten des VisAWI spezifisch auf bestimmte Veränderungen eines Webdesigns reagieren. So beeinflusste bspw. eine Änderung des Farbschemas ausschließlich die Bewertungen auf der Farbigekeitsfacette, nicht aber auf den anderen Facetten (vgl. Studie 6 in Moshagen & Thielsch, 2010).

Mittels des VisAWI können auch größere Sets von real existierenden Websites signifikant hinsichtlich der Ästhetik unterschieden werden (siehe Moshagen & Thielsch, 2010, S. 700), was für eine diskriminative Validität des Instrumentes spricht. Zudem korreliert der VisAWI-Gesamtwerts hoch ($r = .51$) mit der Intention eine Website erneut zu besuchen; auch die Subskalen korrelieren eindeutig mit der Wiederbesuchsbereitschaft ($.40 \leq r \leq .48$). Dies sind bedeutsame Zusammenhänge, wenn man bedenkt, dass weitere Faktoren, wie insbesondere der Inhalt, die Intention eine Website wieder zu besuchen stark beeinflussen (siehe Thielsch et al., 2014).

2.1. Die Kurzform VisAWI-S

In verschiedenen Situationen könnte ein sehr kurzer Fragebogen zur Erfassung von Webästhetik benötigt werden, zum Beispiel wenn andere Themen im Vordergrund einer Umfrage stehen oder insgesamt nur wenige Fragen gestellt werden können. Hierfür wurde aus dem VisAWI heraus die Kurzform VisAWI-S (Visual Aesthetics of Websites Inventory - Short; Moshagen & Thielsch, 2013) geschaffen. Der VisAWI-S erfasst allerdings nur den generellen Ästhetik-Faktor, will man etwas über die einzelnen Facetten erfahren, so empfiehlt sich der Einsatz des vollständigen VisAWI.

Der VisAWI-S wurde in drei Studien mit insgesamt 1673 Befragten entwickelt und validiert (eine umfangreiche Darstellung findet sich bei Moshagen & Thielsch, 2013). Die Kurzversion ist für Gruppenvergleiche ausreichend **reliabel** (je Studie $.76 \leq \alpha \leq .81$) und zeigt eine hohe Korrelation zur Langversion ($r = .91$). Eine konfirmatorische Faktorenanalyse offenbarte eine hervorragende Modellpassung – der VisAWI-S repräsentiert den General-Faktor Website-Ästhetik. Die Kurzversion zeichnet sich ebenso wie die Langversion durch konvergente, divergente und konkurrente **Validität** aus: Es finden sich zum einen hohe Korrelationen zur Attraktivitätsskala aus dem AttrakDiff 1 ($r = .72$). Zum anderen zeigen sich geringere Korrelationen zu divergenten Maßen wie Usability ($r = .54$), der Skala zur pragmatischen Qualität aus dem AttrakDiff ($r = .53$) und zu verschiedenen Inhaltsbewertungen aus dem WWI wie Gefallen ($r = .49$), Verständlichkeit ($r = .34$) oder Qualität und Nutzen ($r = .41$). Die Stimmung der Befragten (erhoben mittels des MDBF von Steyer et al., 1997) hatte keinerlei signifikanten Einfluss (vgl. Moshagen & Thielsch, 2013), was ebenso für eine divergente Validität spricht. Konkurrente Validität zeigt sich in einer hohen Korrelation des VisAWI-S zur Wiederbesuchsbereitschaft ($r = .52$).

2.2. Hinweise zur Anwendung des VisAWI

Verschiedene Tipps und Hinweise zur Anwendung des VisAWI finden sich bei Thielsch und Moshagen (2011). Der VisAWI wird zumeist online genutzt, kann aber auch in Papierform oder anders dargeboten werden. Falls andere grafische Interfaces als Websites beurteilt werden sollen, wurden hierfür in den Items Wörter wie "Seite" durch beispielsweise Begriffe wie "Software" ersetzt. Wichtig ist daran zu denken, vor der Auswertung die negativ gepolten Items (siehe Tabelle 1) umzukodieren, danach können Skalen- und Gesamtmittelwerte gebildet werden. Die Kombination mit anderen Fragebögen sowie das Anhängen spezifischer offener Fragen sind typisch, dies haben wir im Folgenden auch in den Fallbeispielen illustriert.

Zur Interpretation erzielter Werte auf den Skalen ist es wesentlich, den subjektiven Charakter der Bewertungen zu berücksichtigen. So bedeutet ein hoher Wert beispielsweise auf der Skala Farbigkeit nicht, dass das Design besonders bunt ist, sondern, dass die realisierte Farbgestaltung positiv bewertet wird.

3. VisAWI Best Practices

Der VisAWI ist in der Forschung häufig zitiert und in Verwendung (vgl. <http://scholar.google.com/scholar?oi=bibs&hl=de&cites=12037630614351493995>) und immer wieder erreichen uns Anfragen zur Anwendung und Interpretation aus der Praxis. Daher möchten wir im Folgenden zwei Best Practices kurz skizzieren. Dabei zeigen wir im zweiten Beispiel, wie die Anwendung des VisAWI über Websites hinausgehen kann.

3.1. Überarbeitung und Evaluation der Karrierewebsites der BASF Coatings GmbH

Die BASF Coatings GmbH hat in einem mehrstufigen Prozess (siehe Abbildung 1) ihre Karrierewebsites (www.basf-coatings.de/karriere) zunächst überarbeitet und anschließend evaluiert. Neben Best Practice Beispielen aus der Literatur (Frank & Giesen, 2004, Göritz & Moser, 2002), Rankings von Karriere-Websites (Jäger & Meser, 2010, Knabenreich, 2006), Unternehmensvorgaben sowie den Dimensionen Inhalt und Usability wurde zudem die Dimension Ästhetik unter Anwendung des VisAWI als Kriterium definiert.

Die Ergebnisse (siehe Thielsch et al., in Druck) zeigen, dass Ästhetik vor allem einen signifikanten Einfluss auf den Erst- sowie Gesamteindruck einer Website hat. Bei der Vorhersage von verhaltensnahen Komponenten wie der Wiederbesuchs- oder Weiterempfehlungsbereitschaft wird ersichtlich, dass die Dimensionen Inhalt und Ästhetik bedeutend sind, jedoch unterschiedliche Funktionen einnehmen: Der erste Eindruck eines Webnutzers wird durch hohe Ästhetik im Webdesign positiv geprägt – und mit guten Inhalten, die benutzerfreundlich dargeboten sind, verfestigt. So wirkt Ästhetik auch bei Handlungsintentionen zu Wiederbesuch oder Weiterempfehlung nach, auch wenn hier inhaltliche Komponenten klar tonangebend sind (vgl. Thielsch et al., 2014). Somit belegt dieses Anwendungsbeispiel nicht nur die Bedeutung des Konstrukts Ästhetik in der Website-Forschung, sondern zeigt auch die hohe Anwendbarkeit und Praxisrelevanz des VisAWI im Kontext von E-Recruiting auf.

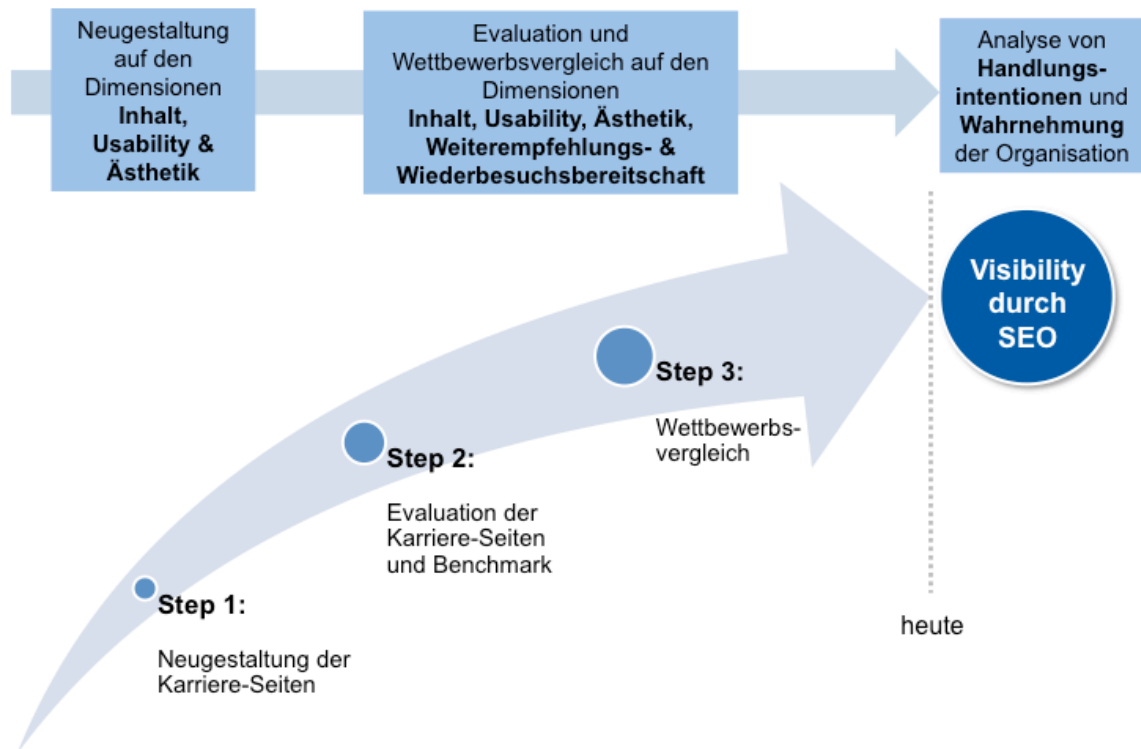


Abbildung 1: Überarbeitungs- und Evaluationsprozess der Karriereseiten der BASF Coatings

3.2. Einsatz von VisAWI und VisAWI-S im UX-Methodenmix der Audials AG

Die Software „Audials PC“ (www.audials.com) der Audials AG wird von Endanwendern eingesetzt, um Musik, Videos und Filme aus verschiedensten Quellen aufzunehmen, und die eigene Musiksammlung zu verwalten. Ein besonderes Merkmal der Software sind grafische Themes, die vom Standarddesign des Betriebssystems abweichen.

Seit 2011 setzt die Audials AG den Fragebogen VisAWI in Kurz- und Langform im UX-Methodenmix ein, um die visuelle Qualität der Oberflächendesigns zu evaluieren. Die Kurzform VisAWI-S wurde unter anderem eingesetzt, um grafische Entwürfe für das Design einer neuen Produktgeneration vergleichend zu bewerten. Dazu wurde ein elektronischer Fragebogen erstellt, der den VisAWI-S mit Kommentarfeldern und einem Ranking kombinierte (siehe Abbildung 2). Die Items des VisAWI wurden entsprechend angepasst. Der Begriff „Benutzeroberfläche“ wurde als Ersatz für „Seite“ verwendet.


Die Ergebnisse der Befragung gaben wertvolle Impulse für die endgültige Ausarbeitung des neuen Designs. Es zeigte sich auch, dass die Items des VisAWI Auswirkung auf die Beschreibung in den offenen Fragen haben können. Die Kommentare waren sehr sachbezogen und klar verständlich.

audials English

User Interface Design 3

Here's a preview of the third user interface design proposal. Take a look. Click the image to enlarge it in a new browser window.

Scroll down to answer questions about this design.



***Please select how much you agree or disagree with the following statements about the design above:**

	strongly disagree		neutral		strongly agree	
The color composition is attractive.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The layout is pleasantly varied.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Everything goes together in this user interface.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The layout appears professionally designed.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[Next >>](#)

Abbildung 2: Bewertung eines Designentwurfs der Audials AG mit VisAWI-S, umgesetzt mit LimeSurvey (www.limesurvey.org).

Kurz- und Langform des VisAWI setzt die Audials AG seit 2011 dauerhaft in Umfragen zur Produktqualität ein. Hierbei werden in Befragungen mehrere standardisierte Fragebögen, zum Beispiel die System Usability Scale (SUS), mit dem VisAWI kombiniert. Diese Kombination ermöglicht ein Bereinigen der Daten, und macht einen Vergleich der Ästhetikbewertungen über Produktgenerationen hinweg möglich.

Die Erfahrungen der Audials AG mit dem VisAWI sind durchweg positiv: Er ist praktikabel einsetzbar und leicht auswertbar. Quantitative und qualitative Ergebnisse aus Umfragen unterstützen maßgeblich bei der Arbeit an den produkteigenen Designs und Themes.

4. Benchmarks und Schwellenwerte

Auf der German UPA Tagung 2011 wurden erste Benchmarks für den VisAWI präsentiert (Thielsch & Moshagen, 2011). Diese sind nun deutlich erweitert worden (derzeit insgesamt n = 7668 Bewertungen) und liegen auch erstmals für die Kurzversion VisAWI-S (derzeit n = 6797 Bewertungen) vor. Darüber hinaus berichten wir im Folgenden kurz die Ergebnisse von Schwellenwertanalysen. Diese ermöglichen es, anhand von gefundenen Cut-Points (basierend auf ROC-Analysen) für den VisAWI und VisAWI-S Websites einzuschätzen, für die (noch) keine Benchmarks vorliegen. Dies könnte insbesondere im Prototyping Anwendung finden. Eine ausführliche Darstellung hierzu findet sich bei Hirschfeld und Thielsch (under revision).

4.1. VisAWI

Im Folgenden sollen zunächst die Benchmarks für deutschsprachige Websites und mobile Apps, die mit der VisAWI-Vollversion getestet wurden, dargestellt werden; ausgeschlossen wurden Prototypen und nicht-deutschsprachige Websites. Nach diesen Ausschlüssen verbleiben im Benchmark $n = 5766$ Bewertungen zu $m = 162$ Websites. Von diesen Bewertungen wurden 57,1 Prozent von Frauen abgegeben, das Alter der Befragten reichte von 14 bis 83 Jahre ($M = 32,21$, $SD = 12,69$).

Wie bereits bei den früheren Benchmarks (Thielsch & Moshagen, 2011) zeigt sich der VisAWI weitgehend robust gegenüber Geschlechtseffekten: So beträgt beispielsweise der Unterschied im VisAWI-Gesamtmittelwert zwischen Männern und Frauen nur 0,08 Skalenpunkte ($M_{\text{Männer}} = 4,46$; $M_{\text{Frauen}} = 4,54$). Dieser Unterschied wird zwar aufgrund der Stichprobengröße signifikant ($F_{1, 2842} = 5,49$, $p = .02$, $\eta^2 < .01$), eine Effektgröße¹ von $d = 0,06$ zeigt aber an, dass dieser Geschlechtsunterschied praktisch unbedeutend ist. Bei der VisAWI-Skala Einfachheit wird der Geschlechtseffekt gar nicht signifikant, bei den anderen drei Skalen zeigen sich nur sehr kleine Effekte ($d_{\text{Vielseitigkeit}} = 0,06$; $d_{\text{Farbigkeit}} = 0,07$; $d_{\text{Kunsthierarchie}} = 0,06$). Geschlechtseffekte können damit auch auf Skalenebene vernachlässigt werden.

Die Korrelationen zwischen dem VisAWI und dem Alter sind ebenfalls signifikant aber gering (Gesamtwert: $r = .12$; Einfachheit: $r = .09$; Vielseitigkeit: $r = .15$; Farbigkeit: $r = .08$; Kunstfertigkeit: $r = .09$, alle $ps < .01$). Diese Alterseffekte sind zwar in Ihrer Effektgröße als klein anzusehen, es ist aber zu beachten, dass ältere Befragte zu leicht positiveren Urteilen kommen.

Deutlichere Unterschiede finden sich, wenn man die bewerteten Websites nach Kategorien sortiert und diese vergleicht (Informationen zum Kategorisierungsschema finden sich bei Thielsch, 2008, S. 86f.): Sowohl für den Gesamtwert ($F_{9, 5753} = 60,51$, $p < .01$, $\eta^2 = .09$), als auch für die Subskalen ($F_{36, 23012} = 41,04$, $p < .01$, $\eta^2 = .06$) ergeben sich bedeutende Unterschiede (Effekte von Alter und Geschlecht waren hierbei kontrolliert). Eine Darstellung der Benchmarkwerte findet sich im Appendix in Tabelle A. Hier sind auch erstmals die Bewertungen von mobilen Apps aufgeführt, dabei gehen VisAWI-Bewertungen zu verschiedenen Finanzdienstleistern ein.

Falls eine Website getestet werden soll, die nicht in dieses Benchmarkschema passt: Eine Schwellenwertanalyse für den VisAWI (Hirschfeld & Thielsch, under revision) zeigt, dass ab einer Gesamtbeurteilung von 4,5 im VisAWI Personen eine getestete Website als insgesamt eher positiv erleben. Zum Beispiel bei einer Neuentwicklung wäre es somit das Ziel einen Befragungswert von 4,5 zumindest leicht zu überschreiten. Diese Information kann bspw. im Prototyping hilfreich sein, wenn keine Ressourcen vorhanden sind die Ästhetik zu maximieren, aber immerhin ein als akzeptabel erlebter Wert erreicht werden soll.

4.2. VisAWI-S

Für die Kurzversion des VisAWI sollen Benchmarks ebenfalls nur für deutschsprachige Websites dargestellt werden. Hier liegen insgesamt $n = 6797$ Bewertungen zu $m = 290$ Websites vor. Von diesen Bewertungen wurden 72,7 Prozent von Frauen abgegeben, das Alter der Befragten reichte von 14 bis 88 Jahre ($M = 29,04$, $SD = 10,45$).

¹ Nach Cohen (1998) zeigt eine Effektgröße von $d = 0,2$ einen kleinen Effekt an, $d = 0,5$ einen mittleren und ab $d = 0,8$ wird von einem großen Effekt gesprochen.

Wie der VisAWI zeigt sich auch der VisAWI-S robust gegenüber Geschlechtseffekten: Der Unterschied im Gesamtwert zwischen Männern und Frauen beträgt nur 0,09 Skalenpunkte ($M_{\text{Männer}} = 4,47$; $M_{\text{Frauen}} = 4,38$). Wie zuvor wird der Effekt aufgrund der Stichprobengröße signifikant ($F_{1, 2842} = 5,49$, $p = .02$, $\eta^2 < .01$) – die Effektgröße von $d = 0,06$ zeigt aber an, dass dieser auch hier praktisch unbedeutend ist. Die Korrelationen zwischen dem VisAWI-S und dem Alter sind zwar signifikant aber sehr gering ($r = .05$; $p < .01$).

Für verschiedene Website-kategorien finden sich Unterschiede in den VisAWI-S Daten ($F_{36, 6779} = 39,36$, $p < .01$, $\eta^2 = .05$; Effekte von Alter und Geschlecht kontrolliert). Es findet sich eine Darstellung der Benchmarkwerte für den VisAWI-S im Appendix in Tabelle B. Die Schwellenwertanalyse für den VisAWI-S (Hirschfeld & Thielsch, under revision) zeigt, dass ebenso wie in der Vollversion ab einer Gesamtbeurteilung von 4,5 im VisAWI-S Personen eine getestete Website als insgesamt eher positiv erleben.

5. Aktuelle Entwicklungen zum VisAWI

Abschließend wollen wir kurz einen Ausblick auf aktuelle Entwicklungen des VisAWI geben: Dies umfasst insbesondere eine von der User Interface Design GmbH geschaffene Webplattform zur vereinfachten Nutzung und Anwendung des VisAWI. Ab voraussichtlich Herbst 2014 werden unter www.VisAWI.de der VisAWI und weiterführende Informationen in einer neugestalteten Webpräsenz offen verfügbar sein. Der VisAWI kann direkt über die Plattform zur Evaluation eingesetzt werden. Zudem wird der VisAWI derzeit in weitere Sprachen übersetzt und validiert. Dies sind jeweils relativ langandauernde Datenerhebungen, aber sobald neue Fragebogenversionen in weiteren Sprachen verfügbar sind, werden sich diese online auf www.VisAWI.de finden. Die verschiedenen VisAWI-Versionen sollen ergänzt werden um ein kurzes Manual mit weiteren Informationen und einer gebündelten Darstellung des Instrumentes.

Wir hoffen, dass die verschiedenen Fortentwicklungen, die wir in unserem Beitrag dargestellt haben, den erfolgreichen Einsatz des VisAWI in Forschung und Praxis fördern. Der VisAWI bereichert so hoffentlich verschiedenste Evaluationen durch eine standardisierte Erfassung der subjektiven Ästhetik für die verschiedene Interpretations- und Anwendungshilfen vorliegen.

6. Literaturverzeichnis

- Bargas-Avila, J., & Hornbæk, K. (2011). Old Wine in New Bottles or Novel Challenges? A Critical Analysis of Empirical Studies of User Experience. *CHI '11 Proceedings of the 2011 annual conference on Human factors in computing systems*, 2689-2698.
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. Hillsdale: Erlbaum.
- Cyr, D., Kindra, G. S. & Dash, S. (2008). Web site design, trust, satisfaction and e-loyalty: the Indian experience. *Online Information Review* 32, 773–790.
- Fechner, G. T. (1876). *Vorschule der Ästhetik*. Leipzig: Breitkopf & Härtel.
- Flavián, C., Guinalú, M. & Gurrea, R. (2006). The role played by perceived usability, satisfaction and consumer trust on website loyalty. *Information & Management*, 43(1), 1–14. Elsevier.
- Frank, G. P. & Giesen, B. (2004). Personalmarketing im Internet. In G. Hertel & U. Konradt (Hrsg.), *Human Resource Management im Inter- und Intranet* (S.33–54). Göttingen: Hogrefe.
- Göritz, A. S. & Moser, K. (2002). Personalmarketing im Internet – Eine Untersuchung des Auftritts der 100 größten deutschen Unternehmen. *Zeitschrift für Personalpsychologie*, 3, 141–148.

- Hassenzahl, M., Burmester, M. & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In J. Ziegler & G. Szwillus (Eds.), *Mensch & Computer 2003. Interaktion in Bewegung* (S. 187–196). Stuttgart: B.G. Teubner.
- Hirschfeld, G. & Thielsch, M. T. (under revision). Establishing meaningful cut points for online user ratings.
- Jäger, W. & Meser, C. (Hrsg.) (2010). *Human Resources im Internet. Bewertung der HR-Websites bedeutender deutscher Arbeitgeber*. Hochschule RheinMain, Wiesbaden.
- Knabenreich, H. (2006). *Kandidatensuche im Netz*. Saarbrücken: VDM Verlag Dr. Müller.
- Lavie, T. & Tractinsky, N. (2004). Assessing Dimensions of Perceived Visual Aesthetics of Web Sites. *International Journal of Human - Computer Studies*, 60, 269-298.
- Lee, S., & Koubek, R. J. (2012). Users' perceptions of usability and aesthetics as criteria of pre- and post-use preferences. *European J. of Industrial Engineering*, 6(1), 87. doi:10.1504/EJIE.2012.044812
- Lindgaard, G. & Dudek, C. (2003). What is this evasive beast we call user Satisfaction? *Interacting with Computers*, 15, 429-452.
- Lindgaard, G., Fernandes, G., Dudek, C., & Brown, J. (2006). Attention web designers: You have 50 milliseconds to make a good first impression! *Behaviour & Information Technology*, 25,2, 115-126.
- Mahlke, S. (2002). Factors influencing the experience of website usage. In: *Proceedings of the CHI'02 Extended Abstracts on Human Factors in Computing Systems*. ACM, Minneapolis, pp. 846-847.
- Moshagen, M., Musch, J. & Göritz, A. S. (2009). A blessing, not a curse: Experimental evidence for beneficial effects of visual aesthetics on performance. *Ergonomics*, 52, 1311-1320.
- Moshagen, M. & Thielsch, M. T. (2010). Facets of visual aesthetics. *International Journal of Human-Computer Studies*, 68 (10), 689-709. doi:10.1016/j.ijhcs.2010.05.006
- Moshagen, M. & Thielsch, M. T. (2013). A short version of the visual aesthetics of websites inventory. *Behaviour & Information Technology*, 32 (12), 1305-1311. doi: 10.1080/0144929X.2012.694910
- Parboteeah, D. V., Valacich, J. S. & Wells, J. D. (2009). The influence of website characteristics on a consumer's urge to buy impulsively. *Information Systems Research* 20, 60–78.
- Porat, T., & Tractinsky, N. (2012). It's a Pleasure Buying Here: The Effects of Web-Store Design on Consumers' Emotions and Attitudes. *Human Computer Interaction*, 27, 235–276. doi:10.1080/07370024.2011.646927
- Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods*, 1(2), 199-223.
- Sonderegger, A. & Sauer, J. (2010). The influence of design aesthetics in usability testing: Effects on user performance and perceived usability. *Applied Ergonomics* 41, 403-410.
- Steyer, R., Schwenkmezger, O., Notz, P. & Eid, M. (1997). *Der Mehrdimensionale Befindlichkeitsfragebogen (MDBF)*. Göttingen: Hogrefe.
- Thielsch, M. T. (2008). *Ästhetik von Websites: Wahrnehmung von Ästhetik und deren Beziehung zu Inhalt, Usability und Persönlichkeitsmerkmalen*. Münster: MV Wissenschaft.
- Thielsch, M. T., Blotenberg, I. & Jaron, R. (2014). User evaluation of websites: From first impression to recommendation. *Interacting with Computers*, 26 (1), 89-102. doi:10.1093/iwc/iwt033
- Thielsch, M. T., Grobien, F. R., Jahn, M. & Schmidt-Bussmann, S. (in Druck). Gestaltung und Evaluation von Karriere-Webseiten – Inhalt, Usability und Ästhetik. In P. Mehlich, T. Brandenburg & M. T. Thielsch (Hrsg.), *Praxis der Wirtschaftspsychologie III: Themen und Fallbeispiele für Studium und Praxis*. Münster: MV Wissenschaft.
- Thielsch, M. T. & Hirschfeld, G. (2012). Spatial frequencies in aesthetic website evaluations – explaining how ultra-rapid evaluations are formed. *Ergonomics*, 55 (7), 731-742. doi: 10.1080/00140139.2012.665496
- Thielsch, M. T. & Moshagen, M. (2011). Erfassung visueller Ästhetik mit dem VisAWI. In H. Brau, A. Lehmann, K. Petrovic & M. C. Schroeder (Hrsg.), *Usability Professionals 2011* (S. 260-265). Stuttgart: German UPA e.V..
- Tuch, A. N., Presslauer, E. E., Stöcklin, M., Opwis, K., & Bargas-Avila, J. a. (2012). The role of visual complexity and prototypicality regarding first impression of websites: Working towards understanding aesthetic judgments. *International Journal of Human-Computer Studies*, 70(11), 794–811. doi:10.1016/j.ijhcs.2012.06.003
- Yoo, B. & Donthu, N. (2001). Developing a scale to measure the perceived quality of an Internet shopping site (SITEQUAL). *Quarterly Journal of Electronic Commerce* 2, 31–46.

7. Appendix

Tabelle A: Benchmarking des VisAWI: Gesamt- und Skalenmittelwerte in Abhängigkeit von der Website-kategorie.

Kategorie	VisAWI-Gesamtwert		Einfachheit		Vielseitigkeit		Farbigkeit		Kunstfertigkeit	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Download & Software (m = 8; n = 96)	3,67	0,97	3,43	1,34	3,04	1,23	4,15	1,24	4,05	1,07
E-Commerce (m = 15; n = 194)	4,05	1,17	3,76	1,47	3,64	1,26	4,33	1,50	4,45	1,25
Entertainment (m = 7; n = 201)	3,90	1,15	3,86	1,42	3,64	1,23	3,70	1,47	4,40	1,40
E-Learning (m = 5; n = 70)	4,43	1,26	4,59	1,35	3,88	1,47	4,51	1,60	4,74	1,42
E-Recruiting & E-Assessment (m = 11; n = 241)	4,21	1,24	4,25	1,47	3,64	1,28	4,36	1,51	4,60	1,39
Information (m = 24; n = 628)	4,08	1,26	4,33	1,41	3,35	1,35	4,28	1,61	4,34	1,47
Portale (m = 10; n = 1505)	4,72	1,20	4,80	1,37	4,36	1,35	4,76	1,29	4,95	1,28
Präsentation & Selbstdarstellung (m = 39; n = 1407)	4,47	1,15	4,42	1,40	3,81	1,35	4,72	1,41	4,95	1,26
Weblogs und Social Sharing (m = 12; n = 178)	3,81	1,23	3,64	1,44	3,43	1,31	4,14	1,50	4,01	1,42
Suchmaschinen (m = 12; n = 291)	4,02	1,13	4,54	1,29	3,13	1,28	4,22	1,49	4,19	1,34
Mobile Apps (m = 19; n = 955)	5,18	0,98	5,35	1,01	4,84	1,19	5,18	1,08	5,35	1,05
Gesamtstichprobe (m = 162; N = 5766)	4,51	1,22	4,58	1,42	4,00	1,41	4,64	1,41	4,81	1,33

Anmerkung: *M* = Mittelwert, *SD* = Standardabweichung, m = Anzahl der beurteilten Websites in der Kategorie, n = Anzahl der Beurteiler. In die Kategorie Mobile Apps gehen VisAWI-Bewertungen zu verschiedenen Finanzdienstleistern ein.

Tabelle B: Benchmarking des VisAWI-S: Werte der Kurzversion in Abhängigkeit von der Website-kategorie.

Kategorie	<i>M</i>	<i>SD</i>	<i>m</i>	<i>n</i>
Download & Software	3,63	1,30	29	441
E-Commerce	4,57	1,39	31	431
Entertainment	3,95	1,33	30	427
E-Learning	4,39	1,27	24	318
E-Recruiting & E-Assessment	4,46	1,28	30	416
Information	4,60	1,36	35	2005
Portale	4,17	1,36	35	515
Präsentation & Selbstdarstellung	4,68	1,46	42	1370
Weblogs und Social Sharing	4,54	1,31	30	381
Suchmaschinen	3,92	1,41	33	493
Gesamtstichprobe	4,41	1,40	290	6797

Anmerkung: *M* = Mittelwert, *SD* = Standardabweichung, *m* = Anzahl der beurteilten Websites in der Kategorie, *n* = Anzahl der Beurteiler.

8. Viten der Autoren



PD Dr. Meinald T. Thielsch hat Psychologie an der Westfälischen Wilhelms-Universität Münster studiert; 2004 bis 2008 folgte ein Promotionsstudium Psychologie und Wirtschaftsinformatik, im Jahr 2013 die Habilitation im Fach Psychologie zum Thema User Experience. Seit 2004 wissenschaftlicher Mitarbeiter am Institut für Psychologie der Westfälischen Wilhelms-Universität Münster; seit 2014 als Akademischer Rat. Nebenberuflich als Referent und wissenschaftlicher Berater tätig, u.a. seit 2005 in Projekten mit der Meuter & Team GmbH. Lehraufträge an den Universitäten Bonn und Fribourg (Schweiz) sowie der Fachhochschule Münster. Seine Arbeits- und Forschungsschwerpunkte liegen in

den Bereichen Wirtschaftspsychologie, User Experience, E-Recruiting, Forschungs-Praxis-Transfer, Evaluation und Online-Forschung.

E-Mail: thielsch@uni-muenster.de

Web: www.meinald.de

Marie Jahn hat Wirtschaftspsychologie an der Universität Mannheim studiert.

Masterarbeit zum Thema Web-Based Applicant Attraction;

Seit 2014 bei der BASF Coatings; HR Strategy, Organizational Development & Change Management. Ihre Arbeitsschwerpunkte liegen in der Betreuung von Veränderungsprozessen; verantwortlich für das Karriereseiten-Projekt.

E-Mail: marie-christin.jahn@basf.com



Jan-Hendrik Spieth studierte Informatik an der Universität Karlsruhe (TH), unter anderem mit dem Schwerpunkt Mensch-Maschine-Schnittstellen. Er arbeitet seit 2009 als zertifizierter Usability Professional (CPUX-F) bei der Audials AG in Karlsruhe. Er erfasst seit mehreren Jahren die Zufriedenheit der Kunden und Nutzer der firmeneigenen Multimediasoftware mit Hilfe von standardisierten Fragebögen. Er ist bei der German UPA in den Arbeitskreisen "In-house Usability" und "User Research", und als Mitorganisator der Regionalgruppe Karlsruhe, engagiert.



Dr. Gerrit Hirschfeld hat Psychologie an der Westfälischen Wilhelms-Universität Münster studiert; 2007 bis 2010 folgte ein Promotionsstudium der Biologie. Von 2007 bis 2010 wissenschaftlicher Mitarbeiter am Institut für Psychologie der Westfälischen Wilhelms-Universität Münster; seit 2010 wissenschaftlicher Mitarbeiter am Deutschen Kinderschmerzszentrum, Datteln. Nebenberuflich als Referent und

wissenschaftlicher Berater tätig; Lehraufträge an den Fachhochschule Köln und FOM Essen. Seine Arbeits- und Forschungsschwerpunkte liegen in den Bereichen Interpretation von Patient Reported Outcomes, Schmerzdiagnostik bei Kindern und Stabilität multivariater Analysemethoden.

E-Mail: g.hirschfeld@deutsches-kinderschmerzzentrum.de

Web: www.gerrithirschfeld.de



Franz Koller ist seit 1998 Managing Director der User Interface Design GmbH (UID) und war bis 2009 Manager Consumer Solutions. 1998 gründete er UID gemeinsam mit Dr. Claus Görner. Der Diplom-Informatiker war von 1997 bis 1998 stellvertretender Leiter der deutschen Niederlassung der schwedischen User Interface Design AB. Zuvor leitete er das Marktstrategieteam „Interaktive Produkte“ am Fraunhofer-

Institut für Arbeitswissenschaft und Organisation IAO und war wissenschaftlicher Mitarbeiter und Projektleiter an der Universität Stuttgart.

Franz Koller ist Mitglied im Industriebeirat des Studiengangs Informationsdesign an der Hochschule der Medien (HdM) in Stuttgart und war von 2006 bis 2007 Lehrbeauftragter an der Hochschule Heilbronn. Er war Mitinitiator und Editor der DIN EN ISO 14915 zur Software-Ergonomie für Multimedia-Benutzerschnittstellen.

E-Mail: franz.koller@uid.com

Web: www.uid.com