

Reliabilität von Eye Tracking-Untersuchungen: Wie viele Probanden werden benötigt?

Elisabeth Lesemann
SirValUse Consulting GmbH
Schlossstr. 8g
22041 Hamburg
lesemann@sirvaluse.de
www.sirvaluse.de

Ulla Wilms
Media Consulta
Wassergasse 3
10179 Berlin
u.wilms@media-consulta.com

Abstract

Als externer Dienstleister haben wir bislang zahlreiche Eye Tracking Studien im Auftrag unserer Kunden durchgeführt. Unsere Erfahrung zeigt, dass Blickpfade und Aufmerksamkeitsverteilung zwischen verschiedenen Nutzern stark variieren. Daher stellten wir uns die Frage, wie viele Probanden eigentlich notwendig sind, um reliable Daten zur Aufmerksamkeitsverteilung zu erheben. Eine Literaturrecherche ergab, dass es bislang keine Studien gibt, die sich mit dieser Frage auseinandersetzen. Dies motivierte uns dazu, die vorliegende Untersuchung durchzuführen.

Um diese Frage zu beantworten, wurden in einer umfangreichen Studie Blickbewegungsdaten von 393 Probanden erhoben.

Zur Erhebung der Blickdaten wurden Screenshots der Startseiten von Webseiten aus drei verschiedenen Kategorien in rotierter Reihenfolge jeweils zehn Sekunden lang vorgelegt: Online-Shop (otto.de), News (spiegelonline.de) und Internetportal (t-online.de). Jeder Screenshot wurde in 10 Areas of Interest (AOI) eingeteilt. Für jede AOI und jeden Probanden wurde die absolute Dauer aller Fixationen während der gesamten, 10-sekündigen Betrachtungsdauer ermittelt (= Mittelwert der Gesamtstichprobe).

Um zu prüfen, ab welcher Stichprobengröße n die Fixationsdauer einer zufällig gezogenen Stichprobe mit der Fixationsdauer der Gesamtstichprobe über-

einstimmt, wurden die Stichprobe und die Gesamtstichprobe miteinander verglichen. Für jede Stichprobengröße wurden mehrere Stichproben zufällig aus der Gesamtstichprobe gezogen.

Da in diesem Fall eine Bestätigung der Nullhypothese, also der Gleichheit zweier Werte, erwünscht ist, wurden Äquivalenztests gerechnet. Ein Fehlerniveau von 1% wurde zugrunde gelegt.

Für die Interpretation der Ergebnisse werten wir die Gesamtstichprobe in dieser Untersuchung so, als spiegelte sie für uns die potenziellen Internetnutzer wider, mit denen Blickbewegungsdaten auf Websites erhoben werden können. Dies kann jedoch nur mit starken Einschränkungen gemacht werden, da es sich nicht um eine repräsentative Gesamtstichprobe handelte.

Unsere Ergebnisse zeigen, dass ab einer Stichprobengröße von $n=50$ die Ergebnisse von allen zufällig gezogenen Stichproben mit denen der Gesamtstichprobe äquivalent sind. Bei einer Stichprobengröße von $n=40$ waren 99,64% aller gerechneten Äquivalenztests äquivalent; bei $n=30$ waren es immerhin noch 97,9%; bei $n=20$ nur noch 91,2% und bei $n=10$ lediglich 71,01% (gemittelt über alle Webseiten-Kategorien).

Es ergeben sich keine Unterschiede zwischen den drei getesteten Webseiten-Kategorien. Dies spricht dafür,

dass die Ergebnisse dieser Studie übertragbar auf ähnliche Untersuchungen zur Aufmerksamkeitsverteilung in der ersten Orientierungsphase auf Webseiten sind.

Für andere Untersuchungen mit ähnlichem Studiendesign heißt das: Wenn 50 Probanden getestet werden, ist die beobachtete Aufmerksamkeitsverteilung in der Orientierungsphase reliabel. Man würde also die gleichen Ergebnisse erhalten, wenn man 50 andere Probanden testen würde.

Mit 40 Probanden ist die Wahrscheinlichkeit, reliable Ergebnisse zu erhalten, ebenfalls hoch genug.

Für die meisten Studien im praktischen Arbeitsfeld können auch mit 30 Probanden akzeptable Ergebnisse bei geringeren Kosten erreicht werden. Für Studien zu Forschungszwecken reichen 30 Probanden jedoch nicht aus.

Unseren Ergebnissen zufolge ist es nicht empfehlenswert, Studien mit weniger Probanden durchzuführen: Hier würde auf Kosten der Datenqualität gespart.

Keywords

Aufmerksamkeitsverteilung, Eye Tracking, Orientierungsphase, Stichprobengröße, Websites

1.0 Einleitung & Stand der Forschung

Seit der Entwicklung berührungsfreier Eye Trackingsysteme (auch *Remote Eye Tracker*) werden Eye Tracking-Studien immer häufiger in Marktforschungs- und User Experience-Studien eingesetzt. Laut Jacob und Karn (2002) setzen rund ein Drittel aller Usability Consulting Unternehmen seit dem Jahr 2000 Verfahren zur Blickbewegungsmessung zur Evaluation von Webseiten ein. Man hofft, dadurch Einblicke in bewusste und unbewusste Informationsverarbeitungsprozesse zu erhalten.

Ein häufiges Einsatzszenario ist die Erfassung der Aufmerksamkeitsverteilung auf der Homepage eines Internetauftritts. So kann beispielsweise überprüft werden, welchen Elementen die Nutzer Beachtung schenken und ob wichtige Informationen schon frühzeitig beachtet werden. Das Wissen um die Blickbewegungen der Nutzer kann den Betreibern einer Website wertvolle Hinweise für deren Gestaltung liefern. Ein Überblick über den Nutzen von Eye Tracking-Untersuchungen im Usability-Kontext findet sich bei Bojko (2005).

Als externer Dienstleister haben wir bislang zahlreiche Eye Tracking Studien im Auftrag unserer Kunden durchgeführt. Unsere Erfahrung zeigt, dass Blickpfade und Aufmerksamkeitsverteilung zwischen verschiedenen Nutzern stark variieren. Daher stellten wir uns die Frage, wie viele Probanden notwendig sind, um reliable Daten zur Aufmerksamkeitsverteilung zu erheben.

Laut Heinsen und Scheier (2003) liegt die typische Stichprobengröße bei Eye Tracking-Untersuchungen im Usability-Kontext bei 10 bis 15 Probanden. Inwieweit diese Anzahl ausreichend ist, wird von den Autoren jedoch weder untersucht noch diskutiert. Ein Überblick über die Anzahl der Teilnehmenden bei

Eye Tracking-Untersuchungen im Bereich Usability findet sich bei Hyönä (2003). Je nach Fragestellung und Untersuchungen wurden in den ausgewählten Studien 7 bis 30 Probanden getestet.

Vereinzelt finden sich Studien, in denen größere Stichproben getestet wurden, um die Validität und Reliabilität der Ergebnisse zu erhöhen. In der zweijährlich stattfindenden Poynter-Studie, die sich mit der Wahrnehmung von Nachrichten auf Internetseiten beschäftigt, nahmen 67 im Jahr 2002 und 45 Probanden in 2004 teil (Poynter, 2002; Outing & Ruel, 2004). In einer Untersuchung zum Orientierungsverhalten auf Internetseiten wurden Daten von 70 Probanden analysiert (Hamborg et al., 2005).

Unsere Literaturrecherche ergab, dass es bislang keine Studien gibt, die sich mit der Frage auseinandersetzen, wie viele Probanden notwendig sind, um reliable Eye Tracking-Daten zu erheben. Dies motivierte uns dazu, die vorliegende Untersuchung durchzuführen.

Für unsere Kunden stellt sich häufig die Frage, wie die eigene Webseite von Nutzern wahrgenommen wird. Bevor eine Webseite beispielsweise in einem neuen Design gelauncht wird, möchte der Kunde wissen, welche Elemente besonders frühzeitig und intensiv von den Nutzern wahrgenommen werden, und ob eine Designvariante wichtige Elemente stärker in den Aufmerksamkeitsfokus der Nutzer rückt als eine andere.

Die vorliegende Untersuchung beschäftigt sich daher mit der Fragestellung, ab welcher Stichprobengröße die Blickbewegungen in den ersten zehn Sekunden der Orientierungsphase auf Webseiten die Grundgesamtheit repräsentativ abbilden.

2.0 Methode

2.1 Stichprobe

Um diese Frage zu beantworten, wurden in einer umfangreichen Studie Blickbewegungsdaten von 393 Probanden erhoben.

Die Probanden wurden über öffentliche Aushänge angeworben sowie durch eine professionelle Agentur rekrutiert und nahmen freiwillig und anonym teil.

Die Datenerhebung fand an der Hochschule für Angewandte Wissenschaften in Hamburg (HAW) sowie in Hamburg und Kiel in den Räumen des Marktforschungsinstitutes MW Research statt.

Mit einem Sehtest wurde überprüft, dass alle Probanden über eine normale Sehstärke verfügten. Aufgrund technischer Probleme beim Eye Tracking mussten Personen mit bifokalen oder besonders starken Brillen ausgeschlossen werden.

Alle Probanden nutzen regelmäßig das Internet. Die Gruppe der 20- bis 29-jährigen war im Vergleich zur ARD/ZDF Online-Studie des Jahres 2005 (Eimeren & Frees, 2005) stark überrepräsentiert aufgrund der Datenerhebung an einer Hochschule.

2.2 Materialien & Ablauf

Als Stimulusmaterial wurden Screenshots der Startseiten von Otto (www.otto.de), Spiegel Online (www.spiegelonline.de) und T-Online (www.t-online.de) verwendet. Die Screenshots wurden im Juni 2006 wenige Tage vor der Erhebung der Blickbewegungsdaten erstellt. Das Material wurde in rotierter Reihenfolge jeweils zehn Sekunden lang dargeboten.

Zur Blickbewegungsaufzeichnung wurde ein Tobii 1750 Eye Tracker verwendet. Dabei handelt es sich um einen 17-Zoll-Monitor mit integrierter Blickbewegungskamera, die mit einer Frequenz von 50 Hz Blickdaten aufzeichnet. Der Eye

Tracker wurde an einen PC mit 3,2 GHz CPU, 2x 512 MB RAM und einer Radeon X1300 Grafikkarte mit 256 MB angeschlossen, auf dem die benötigte Software, der TET-Server 2.9.9 sowie ClearView 2.6.3, installiert wurde.

Darbietung der Stimuli und Blickbewegungsaufzeichnung wurden mit Hilfe der Software ClearView 2.6.3 gesteuert. Die Blickbewegungskamera wurde für jeden Probanden zunächst kurz kalibriert. Direkt danach wurden die drei Screenshots für jeweils 10 Sekunden rotiert vorgelegt.

Im Anschluss daran wurden demographische Daten in einem kurzen Fragebogen erfragt.

2.3 Blickbewegungsdaten

Grundlage der Datenauswertung bildeten die erfassten Fixationen. Dazu wurde als Fixationskriterium eine Minimaldauer von 100 ms festgelegt.

Auf jedem der drei Screenshots wurden nach inhaltlichen Gesichtspunkten Areas of Interest (AOI) festgelegt. Es wurde davon abgesehen, ein gleichmäßiges Gitter für die Festlegung der AOIs zu verwenden, um beispielsweise dasselbe Bild nicht in zwei AOIs zu unterteilen. Für jeden Screenshot wurden 10 AOIs festgelegt, so dass sich für die Datenanalyse insgesamt 30 AOIs ergaben (Abb. 1).

Für jede AOI und jeden Probanden wurde die absolute (aufsummierte) Dauer aller Fixationen während der gesamten, 10-sekündigen Betrachtungsdauer ermittelt. Für die statistische Datenanalyse wurden diese Werte in die prozentuale Blickdauer der 10-sekündigen Gesamtbetrachtungsdauer umgerechnet.



Abb. 1: Areas of Interest auf den Screenshots (von oben nach unten): Otto, Spiegel Online, T-Online.

3.0 Datenanalyse

3.1 Vorbemerkungen

In dieser Untersuchung werten wir die Gesamtstichprobe so, als spiegele sie für uns die potenziellen Internetnutzer wider, mit denen Blickbewegungsdaten auf Websites erhoben werden können. Dies kann jedoch nur mit der Einschränkung gemacht werden, dass es sich nicht um eine repräsentative Gesamtstichprobe handelt.

3.2 Statistische Testmethode

Um zu prüfen, ab welcher Stichprobengröße n die Fixationsdauer einer zufällig gezogenen Stichprobe mit der Fixationsdauer der Gesamtstich-

probe übereinstimmt, wurden die Stichprobe und die Gesamtstichprobe miteinander verglichen.

Da in diesem Fall eine Bestätigung der Nullhypothese, also der Gleichheit zweier Werte, erwünscht ist, wurden gemäß den Empfehlungen von Schlittgen (1996) Äquivalenztests gerechnet.

Äquivalenztests überprüfen die Übereinstimmung zweier Mittelwerte. Dazu wird ein Konfidenzintervall berechnet, das um die Differenz der Mittelwerte gelegt wird. Wenn dieses Konfidenzintervall innerhalb vorher festgelegter Grenzen liegt, stimmen die beiden Mittelwerte überein. In unserem Fall würde dies bedeuten, dass die zufällig gezogene Stichprobe mit der Gesamtstichprobe übereinstimmt.

Die statistische Datenauswertung wurde mit Hilfe des Statistikpakets SPSS 12.0 vorgenommen, indem eine eigene Syntax erstellt wurde.

3.3 Festlegung der Parameter

Für jede AOI wurde der Mittelwert M_G der absoluten Fixationsdauer über alle Probanden ermittelt (Mittelwert der Gesamtstichprobe). Die Äquivalenztests sollen für jede AOI überprüfen, ob der Mittelwert einer zufällig gezogenen Stichprobe bei einer Stichprobengröße n mit dem Mittelwert der Gesamtstichprobe übereinstimmt.

Für die Festlegung der Grenzen, innerhalb derer das Konfidenzintervall liegen soll, gab es zum Zeitpunkt der Durchführung dieser Studie keine Erfahrungswerte aus anderen Studien. Unter der Annahme, dass die Gesamtstichprobe ausreichend groß gewählt wurde, um repräsentative Mittelwerte zu liefern, liefern die Daten zudem zu jeder AOI eine Standardabweichung. Da es keine Erfahrungswerte gibt, liefert die Standardabweichung der Gesamtstichprobe eine

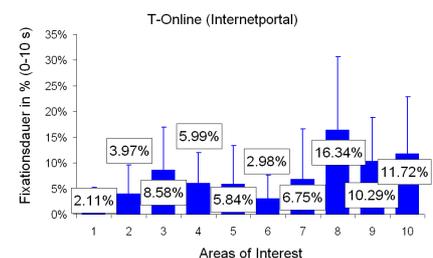
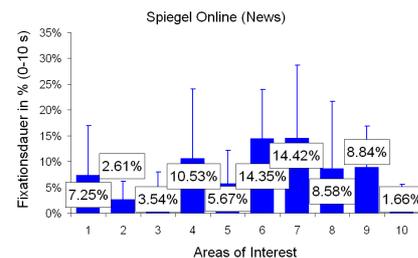
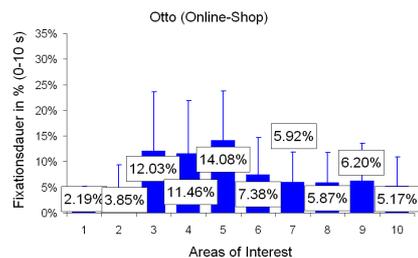


Abb. 2: Absolute Fixationsdauer in den ersten 10 Sekunden der Betrachtung: Heatmaps und Balkendiagramme für die Screenshots (von links nach rechts): Otto, Spiegel Online, T-Online. Balkendiagramme zeigen die prozentuale Fixationsdauer auf vorher festgelegten Areas of Interest, dargestellt sind Mittelwerte über alle Probanden; Fehlerbalken zeigen die Standardabweichung.

bestmögliche Approximation für die Festlegung der Grenzen.

Zur Absicherung der Ergebnisse dieser Untersuchung werden pro Stichprobengröße n nicht nur eine, sondern mehrere zufällig gezogene Stichproben mit der Gesamtstichprobe verglichen. Es werden also immer mehrere, unabhängige Äquivalenztests durchgeführt.

Um möglichst konservativ zu testen, wird für alle Äquivalenztests ein Alpha-niveau von 1% zugrunde gelegt.

3.4 Ziehung der Zufallsstichproben

Beginnend mit einer Stichprobengröße von $n=100$, die sukzessive verringert wurde ($n=90$, $n=80$, etc.), wurden zufällige Stichproben aus der Gesamtstichprobe gezogen.

Die Anzahl der gezogenen Stichproben wurde für jedes n so gewählt, dass jeder Proband mit einer nahezu 100-prozentigen Wahrscheinlichkeit mindestens einmal in einer der zufällig gezogenen

Stichproben landet. Beispielsweise geht die Wahrscheinlichkeit bei einer Stichprobengröße von $n=10$ (aus $N=393$) gegen 1, dass jede Person in einer der zufällig gezogenen Stichproben landet, wenn 357 Stichproben gezogen werden.

4.0 Ergebnisse & Diskussion

4.1 Darstellung der Ergebnisse

Abbildung 2 zeigt für jeden Screenshot die prozentuale Aufmerksamkeitsverteilung während der 10-sekündigen Betrachtungsphase auf den 10 AOIs, jeweils gemittelt über alle Probanden.

Auf dem Screenshot von Otto.de werden insbesondere die AOI mit dem Bild der Frau und dem Titel "Special" (AOI 5: $m=14,08\%$, $sd=9,68\%$) sowie die obere Navigationsleiste (AOI 3: $m=12,03\%$, $sd=11,45\%$) betrachtet.

Die Rubrik "Exklusiv" am rechten Seitenrand (AOI 7: $m=14,42\%$,

$sd=14,29\%$) sowie die größte Schlagzeile "USA bietet Boeing-Teile an" (AOI 6: $m=14,35\%$, $sd=9,64\%$) erhalten auf dem Screenshot von Spiegel Online viel Aufmerksamkeit.

Auf dem Screenshot von T-Online erhalten vor allem die bebilderten Schlagzeilen viel Aufmerksamkeit: "Beckenbauer warnt vor..." (AOI 8: $m=16,34\%$, $sd=14,30\%$), "Süßes Tigerbaby..." (AOI 10: $m=11,72\%$, $sd=11,17\%$), "Sichern Sie sich Ihr Heimspiel..." (AOI 9: $m=10,29\%$, $sd=8,56\%$).

Gemäß der oben beschriebenen Methode wurden für unterschiedlich große, zufällig gezogene Stichproben Äquivalenztests gerechnet, um die Übereinstimmung mit der Gesamtstichprobe zu überprüfen.

Abbildung 3 fasst die Ergebnisse aller Äquivalenztests zusammen: Ab einer Stichprobengröße von $n=50$ sind die Ergebnisse von *allen* zufällig gezogenen äquivalent mit den Ergebnissen der Gesamtstichprobe.

Bei allen drei Screenshots sind bei einer Stichprobengröße von $n=40$ einzelne Stichproben nicht äquivalent zur Gesamtstichprobe: Bei Otto sind 2 von 86 gezogenen Stichproben bei jeweils einer AOI (AOI 2 bzw. AOI 9) nicht äquivalent. Bei Spiegel Online ist 1 Stichprobe bei einer AOI (AOI 8) nicht äquivalent; bei T-Online sind es 3 Stichproben (alle AOI 8). Über alle AOIs gemittelt sind bei Otto 99,76%, bei Spiegel Online 99,88% und bei T-Online 99,29% aller 86 Stichproben äquivalent mit der Gesamtstichprobe; über alle Screenshots gemittelt ergibt sich ein Wert von 99,64% (Abb. 3).

Bei $n=30$ sind bei Otto immerhin noch 98,01%, bei Spiegel Online 98,1% und bei T-Online 97,6% der Ergebnisse aller Stichproben äquivalent mit denen der Gesamtstichprobe; über alle Screenshots gemittelt ergibt sich ein Wert von 97,9% (Abb. 3).

Ab $n=20$ werden insgesamt nur noch 91,2% äquivalente Stichproben gefunden (Otto: 90,56%, Spiegel Online: 92,24%, T-Online: 90,79%); ab $n=10$ sogar nur noch 71,01% (Otto: 71,72%, Spiegel Online: 69,3%, T-Online: 72,01%).

4.2 Diskussion der Ergebnisse

Unsere Ergebnisse zeigen, dass eine zufällig gezogene Stichprobe von $n=50$ Probanden in Bezug auf die Aufmerksamkeitsverteilung in der Orientierungsphase die gleichen Ergebnisse liefern kann wie eine Datenerhebung mit 393 Probanden.

Auch mit $n=40$ Probanden können mit einer hohen Wahrscheinlichkeit die gleichen Ergebnisse erhoben werden. Schaut man sich die Daten genauer an, so stellt man jedoch fest, dass 3 der 86 Stichproben bei T-Online nicht äquivalent sind, weil die Aufmerksamkeitsverteilung auf einer AOI nicht übereinstimmen, die eine besondere Relevanz für die Seite hat: Sie wurde im Mittel am

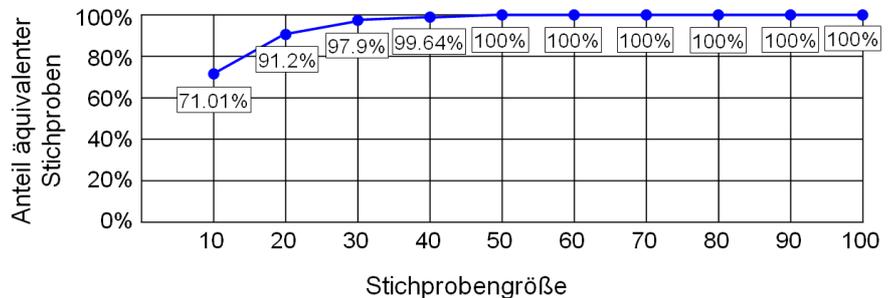


Abb. 3: Zusammengefasste Ergebnisse aller Äquivalenztests: Ab einer Stichprobengröße von $n=50$ waren die Ergebnisse von allen (100%) der zufällig gezogenen Stichproben äquivalent mit denen der Gesamtstichprobe. Die Anzahl der gezogenen Stichproben wurde für jedes n so gewählt, dass jeder Proband mit einer nahezu 100-prozentigen Wahrscheinlichkeit mindestens einmal in einer der zufällig gezogenen Stichproben landet. Als Irrtumswahrscheinlichkeit wurde ein Alphaniveau von 1% festgelegt.

längsten betrachtet. Es handelt sich um die erste Schlagzeile mit Bild und Text im redaktionellen Teil der Webseite. Dies zeigt auf, dass auch bei einer Stichprobengröße von 40 Probanden vereinzelte Fehler in den Ergebnissen nicht ganz auszuschließen sind.

Für die Interpretation der Ergebnisse werten wir die Gesamtstichprobe in dieser Untersuchung so, als spiegelte sie für uns die potenziellen Internetnutzer wider, mit denen Blickbewegungsdaten auf Websites erhoben werden können. Dies kann jedoch nur mit starken Einschränkungen gemacht werden, da es sich nicht um eine repräsentative Gesamtstichprobe handelte.

Unter Beachtung dieser Einschränkungen deuten unsere Ergebnisse darauf hin, dass bei einem ähnlichen Studiendesign 40 Probanden getestet werden müssen, um reliable Ergebnisse zu erhalten. Man würde also mit 40 anderen Probanden die gleichen Ergebnisse erhalten. So kann man davon ausgehen, dass die erhobenen Daten zur Aufmerksamkeitsverteilung alle Nutzer widerspiegeln.

Schaut man sich die Ergebnisse bei $n=30$ Probanden an, so lassen sich mit geringen Abstrichen ebenfalls reliable Ergebnisse erheben. Für die meisten Studien im praktischen Arbeitsfeld der User Experience können mit 30 Probanden akzeptable Ergebnisse bei geringeren Kosten erreicht werden.

Unseren Ergebnissen zufolge können wir unseren Kunden nicht empfehlen, Studien mit weniger als 30 Probanden durchzuführen. Hier würde auf Kosten der Datenqualität gespart.

Für Studien, in denen ein Irrtum als schwerwiegend bewertet würde, beispielsweise Forschungsstudien, sollten Daten von mindestens 40, besser 50 Probanden erhoben werden.

5.0 Diskussion / Ausblick

Diese Studie kann Hinweise dafür liefern, wie viele Probanden im Allgemeinen für Eye Tracking-Studien zur Aufmerksamkeitsverteilung benötigt werden. Je nach Fragestellung und Studiendesign kann die notwendige Probandenanzahl jedoch noch von weiteren Faktoren abhängen. Möchte man beispielsweise Ergebnisse von verschiedenen Nutzergruppen kontrastieren oder

Geschlechterunterschiede beim Betrachten von Internetseiten feststellen, müssten deutlich mehr Probanden einbezogen werden.

Auch wenn wir unsere Ergebnisse für die Interpretation so werten, als spiegelte unsere Gesamtstichprobe die potenziellen Internetnutzer wider, mit denen Blickbewegungsdaten auf Websites erhoben werden können, kann dies nur mit starken Einschränkungen einhergehen, da die Gesamtstichprobe die tatsächliche Grundgesamtheit der deutschen Internetnutzer nicht repräsentativ abbildet.

So kann diese Untersuchung nicht klären, ob sich die errechnete notwendige Stichprobengröße bei einer anderen Zusammensetzung der Grundgesamtheit ändern würde. Will eine Untersuchung also Aussagen über alle deutschen Internetnutzer treffen, so kann unsere Untersuchung nur deutliche Hinweise dafür liefern, dass diese mit 40 Probanden getroffen werden können.

Ebenso bleibt die Frage offen, ob weniger Probanden benötigt werden, wenn die Zielgruppe, mit der Daten erhoben werden sollen, stark fokussiert ist. Bei einer Internetseite, die beispielsweise nur von Frauen mittleren Alters frequentiert wird, die sich bezüglich ihrer Internetvorerfahrung und ihres inhaltlichen

Wissens über das Themengebiet der Internetseite ähnlich sind, können eventuell auch mit weniger Probanden reliable Ergebnisse erhoben werden.

Darin bietet sich ein interessanter Ausblick für weitere Forschungsarbeiten: Welche Faktoren beeinflussen die Art, wie Nutzer eine Internetseite betrachten? Gibt es Faktoren, wie bspw. Vorerfahrungen, die das Betrachten beeinflussen? Und wie wirken sich diese Faktoren auf die Stichprobengröße aus?

Danksagung

Wir danken der Otto GmbH & Co KG, der Spiegel Online GmbH und der Deutschen Telekom AG dafür, dass sie uns Screenshots ihrer Webseiten für diese Untersuchung zur Verfügung gestellt haben.

Der HAW Hamburg und dem Marktforschungsinstitut MW Research danken wir, dass wir in ihren Räumen testen konnten.

Ganz besonders bedanken wir uns bei Claudia Weitz und Benjamin Jöckel, die uns mit viel Tat und Rat bei der Datenauswertung unterstützt haben.

6.0 Literaturverzeichnis

Bojko, Agnieszka (2005): Eye tracking in user experience testing: How to make the most of it. *Proceedings of the UPA 2005 Conference*.

Eimeren, B. van; Frees, B. (2005): ARD/ZDF Online-Studie 2005: Nach dem Boom. *Media Perspektiven* 8.

Hamborg, Kai-Christoph; Ollermann, Frank, Reinecke, Stefan (2005): Visuelles Orientierungsverhalten. *I-Com*, 1, 20-25.

Heinsen, Sven; Scheier, Christian (2003): Aufmerksamkeitsanalyse. In: S. Heinsen, P. Vogt (Hrsg.): *Usability praktisch umsetzen: Handbuch für Software, Web, Mobile Devices und andere interaktive Produkte*. München: Hanser.

Hyönä, Jukka (Hrsg., 2003): *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*. Amsterdam: Elsevier Science.

Jacob, Robert; Karn, Keith (2002): Commentary on section 4. Eye Tracking in human-computer interaction and usability research. In J. Hyönä (Hrsg.): *The mind's eye: Cognitive and applied aspects of eye movement research*. Amsterdam: Elsevier Science.

Outing, Steve; Ruel, Laura (2004): The best of eyetrack III: What we saw when we looked through their eyes. (Verfügbar unter: <http://www.poynterextra.org/eyetrack2004/main.htm>, letzter Zugriff am 05.05.2006).

Poynter (2002): Stanford Poynter Project: Introductory Highlights (Verfügbar unter: <http://www.poynterextra.org/et/i.htm>, letzter Zugriff am 03.05.2006).

Schlittgen, Dr. R. (1996): *Statistische Inferenz* (S. 306ff). München: R. Oldenbourg Verlag.