

# From a Distance: Usability Testing aus der Ferne

## From a Distance: Remote Usability Test

Remote Usability, qualitativ, quantitativ, online, synchron, asynchron, Usability Testing, Usability Labor, User Experience, Online Befragung, Tracking, Logfiles

Unter dem Begriff Remote Usability Testing werden seit Jahren eine Fülle von Ansätzen und Methoden entwickelt, um unterschiedlichste Konzepte und Interaktionsformen nutzerbasiert zu testen, ohne dabei die Nutzer in ein Usability Labor einladen zu müssen. Testleiter und Probanden befinden sich demnach an unterschiedlichen Orten. Es qualifiziert sich jedoch nicht jede technisch vermittelte Testmethode als Remote Usability Test. Einige Ansätze, die viele Nutzerdaten liefern aber dem Testleiter kein Verständnis des Nutzerverhaltens ermöglichen, fallen nicht darunter. Andere Ansätze können je nach Erkenntnisinteresse eine interessante Alternative oder Ergänzung zum Usability Labor darstellen.

For the last few years a lot of technical approaches and methods have been developed to evaluate concepts and to test the interaction of users with a systems without the need of inviting them to a usability lab. Users and usability experts are separated in place (and time). These techniques are summarised under the heading „remote usability testing“.

Not every method is suitable to conduct a remote usability test. If the method only gathers user data, it's impossible for usability experts to determine the intention of the user. However, there are some methods and approaches that are interesting alternatives to the common usability test in the lab.

Summary. Eine kurze Zusammenfassung in Englisch.

## 1. Remote Usability Testing wird immer attraktiver

Die Verbreitung von performanten Breitband-Internetanschlüssen und die Entwicklung von VoIP-, Audio-Video-Conferencing Tools, Live Collaboration Tools, Remote Desktop Viewing, etc. schreiten immer weiter voran. Die räumliche und zeitliche Unabhängigkeit wird dadurch stetig erweitert und ermöglicht in jüngerer Zeit eine interessante Verfeinerung der Remote Usability Testmethoden.

### 1.1 Wann ist ein Online-Test ein Remote Usability Test?

Alle Usability Tests haben in der Regel eins gemeinsam: Sie zielen darauf ab, den Nutzer bzw. die Interaktion des Nutzers mit einer Anwendung zu verstehen, um Usability relevante Daten zu erhalten, die zu einer Optimierung der Anwendung führen. Um Nutzer verstehen zu können, sind Methoden der Verhaltensbeobachtung, des Lauten Denkens und die Befragung der Probanden im Rahmen des Usability Tests besonders wichtig (vgl. Hamborg et al. 2006). Usability Tests und folglich auch den Remote Usability Tests ist damit eine qualitative Komponente inhärent, ohne die es nicht möglich wäre, neue und unerwartete Erkenntnisse zu generieren. Es geht also nicht nur darum, bekannte Probleme oder Fragestellungen, z.B. geläufige Usability Heuristiken in Bezug auf eine Anwendung, durch skalierte Antwortmöglichkeiten zu bewerten. Vielmehr richtet sich auch die Qualität der Remote Usability Ansätze danach, wie gut die Interaktion, das mentale Model der Nutzer und deren Zielsetzung durch die Testleiter verstanden werden kann.

Es gibt eine Reihe von Methoden, bei denen Erkenntnisse zur Optimierung der Usability einer Anwendung nur durch die Interpretation von gesammelten, meist quantitativen, Daten möglich ist. Dazu gehört neben der klassischen Online-Befragung auch die Analyse der Logfiles eines Webserver. Die Datenerhebung kann durch einen dazwischen geschalteten Proxyserver (sofern kein Zugriff auf den Webserver möglich ist) erfolgen. Werden spezielle clientseitige Logging Mechanismen, wie z.B. der Ergobrowser o.ä. eingesetzt, können auch Mausspuren und Screenshots aufgezeichnet werden. Bei diesen Methoden wird in Erfahrung gebracht, welche Pfade Benutzer durch eine Anwendung nehmen, wie lange sie wo verweilen, was geklickt wurde und an welcher Stelle sie die Anwendung verlassen. Das WebTango Projekt der University of Berkley führte den Logging Ansatz weiter und lässt anstelle von Benutzern eine Anwendung durch einen Web-Crawler spiders, der die Anwendung auf voreingestellte Usability Heuristiken wie z. B. Textlängen, das Vorhandensein eines Hilfe-Links, etc. überprüft. Leider wird bei diesen Methoden, kein Feedback der Benutzer erfasst. Es kann also nicht festgestellt werden, weshalb Benutzer z.B. gerade einen bestimmten Navigationspfad durch eine Anwendung wählen und weshalb sie die Anwendung an einer anderen Stelle beenden. Diese Methoden der quantitativ ausgerichteten Datenerhebung oder reine Dokumentation von Nutzerinteraktionen, können aufgrund der extrem eingeschränkten Möglichkeiten, die Nutzer wirklich verstehen zu können, nicht als echte Remote Usability Testansätze bezeichnet werden.

Weiterführende Ansätze versuchen hingegen, der qualitativen Dimension möglichst gut gerecht zu werden. In so einem Fall können Remote Usability Tests und laborbasierte Usability Tests durchaus zu einer vergleichbaren Ergebnisqualität hinsichtlich der zentralen Usability Probleme führen, obwohl zwei vollkommen unterschiedlichen Testdesigns und Settings zum Tragen kommen. Dass jedoch weniger das Setting als vielmehr die unterschiedlichen Arten von Usability Problemen zu den unterschiedlichen Ergebnissen von Usability Labor und Remote Usability Test beitragen, zeigt bereits eine Studie von 2002 (vgl. Tullis et al. 2002). Binden Sie keine Abbildungen und komplexen Tabellen in das Dokument ein, sondern speichern Sie diese jeweils in eigenen Dateien ab. Stellen Sie die Bild- oder Tabellenunterschrift an die Stelle im Dokument, an der dann später auch das Bild erscheinen soll.

## **1.2 Ansätze für Remote Usability Tests**

Wenn Testleiter und Probanden direkt miteinander kommunizieren bzw. interagieren können, spricht man von einem synchronen Remote Usability Test. Dieser zeitgleiche Remote Usability Test ist am deutlichsten mit dem Usability Labor vergleichbar, da er einen direkten Austausch zwischen Testleiter und Proband, quasi live (Live Collaboration), ermöglicht. Synchroner Remote Usability Tests werden gerne mit Online-Konferenzsystemen, wie z.B. WebEx, Netmeeting, etc. unternommen. Den Probanden können in unterschiedlichen Ausprägungen z.B. Testmaterialien online präsentiert werden und der Testleiter kann die Interaktion mit einer Anwendung per Videoübertragung live beobachten. Dabei steht ein Proband über eine Telefonverbindung oder VoIP im direkten Kontakt mit einem Usability Experten. Wie im Labor, werden Probanden angehalten während des Testens ihre Aktionen und Gedankengänge zu kommentieren. Alternativ können Usability Experten über Online-Meeting-Tools, Collaborative Whiteboards oder Remote Desktop Viewing den Probanden praktisch über die Schulter blicken und ihre Benutzerinteraktionen mitverfolgen. Wird auf Probandenseite zusätzlich eine Videokamera eingesetzt kann per Videokonferenz die Mimik der Probanden erfasst werden.

Erfolgt das Testen zeitlich versetzt, so dass Testleiter und Probanden nicht in einem direktem Kontakt stehen, spricht man von einem asynchronen Remote Usability Test. Bei dieser Variante werden möglichst viele Nutzerdaten (Screens, Mauspfade, Klickpfade, etc.) online aufgezeichnet, um eine spätere Auswertung durch den Testleiter zu ermöglichen. Um das Verständnis der Nutzungsprobleme zu erhöhen, wird für die Probanden eine Möglichkeit geschaffen, noch während der Nutzung ein (in der Regel schriftliches) Feedback abzugeben. Dabei wird das Feedback durch die Probanden entweder selbst initiiert oder systemseitig in bestimmten Situationen der Interaktion ausgelöst. Wenn Probanden z. B. einen für die Anwendung typischen Pfad verlassen, kann eine Abfrage der Beweggründe für dieses spezielle Nutzerverhalten automatisiert dazwischen- oder nachgeschaltet werden. Asynchrone Remote Usability Tests eignen sich gut, um Online-Befragungen anzuschließen.

Eine der effizientesten asynchronen Remote Usability Ansätze ist eine Kombination aus User Reported Critical Incident Methode in Verbindung mit Online Fragebögen und (proxybasiertem) Logging. Diese führt zu einer Datenqualität, deren

Analyse zu sehr validen Ergebnissen führt (z.B. LeoTrace, WebQuilt). Sie ist allerdings nur für Internetseiten einsetzbar, die über das HTTP-Protokoll erreichbar sind. Kernstück dieser Methode ist die Erkenntnis, dass Probanden (nach einer kurzen Schulung) durchaus in der Lage sind, Critical Incidents, also positive oder negative Ereignisse während ihrer Interaktion mit einem zu testenden System, selbst identifizieren und berichten zu können. Testleitern dient die Auswertung dieser Daten später zur Optimierung des Systems (vgl. Hartson et al. 1996, Castillo 2001, Nufer 2006).

Die Durchführung dieses Testansatzes beginnt für Probanden, indem ihnen auf der zu testenden Online-Anwendung (z.B. einer Internetseite) ein zufallsgesteuertes Popup Fenster oder eine DHTML Layer eingeblendet wird, welche zur Teilnahme am Remote Usability Test einlädt. Entscheiden sich Probanden teilzunehmen, werden sie auf einen Online Fragebogen geleitet, der alle im Rekrutierungsleitfaden definierten Kriterien abfragt und im Hintergrund mit einer Quote der zu rekrutierenden Probanden vergleicht. Benutzer, die nicht als Proband in Frage kommen, können alternativ für andere Fragestellungen herangezogen werden, wie z. B. die Gründe ihres Besuchs oder für Kundenzufriedenheitsbefragungen. Wird das Profil nach Abgleich mit der Quote für den Test benötigt, werden die Probanden für das selbständige Identifizieren und Berichten von Critical Incidents anhand anschaulicher Beispiele instruiert. Danach erscheint die zu testende Internetseite, die über einen Proxyserver umgeleitet wird. Dieser fügt jeder vom Original Webserver ausgelieferten Seite zusätzlich JavaScript Code zum Erfassen von Mausspuren und zum Erstellen von Screenshots hinzu. Für Probanden erscheint die Internetseite wie gewohnt und kann ihrem Vorhaben entsprechend benutzt werden. Während ihrer Interaktion mit der Internetseite sollen über ein Online Formular positive und negative Beobachtungen und Erlebnisse berichtet werden. Erfasst werden dabei drei Dimensionen: 1) Die Probanden werden gebeten ihr aktuelles Vorhaben zu schildern. 2) Es folgt das Berichten ihrer Beobachtung. 3) Die Probanden werden gebeten, die Schwere der Beobachtung selbst anhand einer Skala einzuschätzen. Dabei sollen sie beurteilen, wie sehr ihre Beobachtung oder ihr Erlebnis sie von ihrem zuerst genannten Vorhaben abgebracht bzw. gehindert haben. Usability Experten werten die Critical Incidents Mitteilungen der Probanden später zusammen mit den geloggtten Daten, Mausspuren, Screenshots, etc. aus. Diese qualitativ hochwertigeren Daten helfen Critical Incidents besser verstehen und Beobachtung eines Probanden notfalls nachstellen zu können. Da lediglich die Schlüsselerlebnisse und nicht die ganze Session von Probanden erfasst und ausgewertet wird, kann die Auswertung durch Usability Experten schneller als im Labor erfolgen. Neben Problemen, in der zu testenden Anwendung, werden auch eher beiläufige Probleme, die die gesamte Customer Experience einer Kundenbeziehung betreffen (z. B. Kundenservice, Bestellungen, Versand etc.) von Probanden mitgeteilt. Wobei in diesem Fall nicht unbedingt ein Zusammenhang zum eigentlichen Testobjekt besteht.

### **1.3 Stärken und Schwächen von Remote Usability Tests**

Die räumliche - und bei asynchron durchgeführten Remote Usability Tests auch zeitliche - Trennung zwischen Probanden und Testleitern ermöglichen einige nicht von der Hand zu weisende Vorteile. Probanden im Labor sehen sich einem Testleiter in einer für sie ungewohnten Umgebung gegenüber, während Probanden beim Remote Usability Test z. B. über ihren eigenen Computer und möglichst in einem ihnen vertrauten Setting maximal eine technisch vermittelte Kommunikation mit einem Testleiter unterhalten oder einfach über vorgegebene Oberflächen ihr Feedback zu einer Anwendung geben, ohne dass es eines konkreten Testleiters als Gegenüber bedarf.

Die Rekrutierung der Probanden wird nicht, wie bei laborbasierten Tests, auf nur eine Stadt oder Region begrenzt, sondern kann international erfolgen (Dray & Siegel 2004). Einziges Rekrutierungskriterium ist z.B. ein Breitband-Internetanschluss – somit kann bei der Rekrutierung von Probanden aus einer wesentlich größeren Stichprobe ausgewählt werden. Für Probanden entstehen keine An- und Abreisekosten. Bei der Durchführung von Remote Usability Tests ist eine größere zeitliche Flexibilität möglich, was dazu führt, dass mehr Probanden und vor allem schwer zu rekrutierende Zielgruppen leichter rekrutiert werden können (vgl. Perkins 2001). Da Probanden nicht in einer künstlichen Laboratmosphäre, sondern von ihrer gewohnten Umgebung aus am Test teilnehmen, sind diese dazu geneigt, sich normal zu verhalten und zeigen dabei für sie typische Verhaltensmuster (vgl. Perkins 2001).

Je mehr Technik beim asynchronen Remote Usability Test zum Einsatz kommt, desto höher wird auf Probanden-Seite die technische Hürde, z.B. wenn eine Installationen am PC des Probanden oder eine Konfiguration der Firewall erforderlich ist. Dies führt dazu, dass sogar weniger Benutzer am Test teilnehmen können bzw. die Repräsentativität eingeschränkt wird, da weniger technikaffine Nutzer kaum noch in Frage kommen. Bei Testobjekten die für eine breite Zielgruppe entwickelt

wurden, muss es das Ziel sein, den technischen Aspekt so gering wie möglich zu halten. Bei asynchronen Remote Usability Tests ist die Testsituation wenig bis gar nicht kontrollierbar. Da keine Beobachtung, z.B. der Mimik von Probanden möglich ist, kann hinsichtlich des Verständnisses der Nutzerinteraktion nur auf schriftliche und damit bewusst berichtete Usability-Probleme zurückgegriffen werden. Unbewusste Anteile der Interaktion werden nicht erfasst. Der Ansatz erlaubt es dem Testleiter nicht zu intervenieren, um Themen zwecks besseren Verständnisses zu vertiefen. Aufgrund der Vielzahl von Probanden und der Datenmenge wird zudem meist auf eine Nachexploration verzichtet.

#### **1.4 Remote Usability Tests als Alternative zum Usability Labor**

Der asynchrone Testansatz kann, selbst mit einer sehr ausgereiften Technik, den Usability Test im Labor nicht vollständig ersetzen, obwohl hochwertige Ergebnisse und aufgrund der hohen Probandenzahl und vielfältigerer Nutzungsszenarien sogar mehr Usability Probleme gefunden werden können (vgl. Kehr et al 2006). Allerdings sind die Ergebnisse nicht deckungsgleich. Beobachtungen im Lab und User Reported Critical Incidents im Remote Usability Test decken teilweise unterschiedliche Usability Probleme auf (vgl. Molich 2004).

Je nach Fragestellung kann dieser Unterschied jedoch vernachlässigt werden. Benötigt der Testleiter einen möglichst breiten und vielfältigen Input an Usability Problemen, ggf. technischen Problemen, allgemeine Probleme die Customer Experience betreffend und quantitativen Daten, und dies über einen längeren Zeitraum hinweg, ist ein asynchroner Testansatz vollkommen richtig und auch effizienter. Kommt der menschlichen Komponente jedoch eine höhere Bedeutung zu, also das gesamte Erlebnis bei der Nutzung einer Anwendung, ist man mit einem Usability Labor oder einem technisch ausgereiften synchronen Remote Usability Test besser beraten. Auf jeden Fall stellen die Remote Usability Ansätze schon heute eine viel versprechende Ergänzung und Alternative zu Usability Tests im Labor da.

#### **Literatur**

- Castillo, J.: The User-Reported Critical Incident Method for Remote Usability Evaluation. Blacksburg, VA: Virginia Polytechnic Institute and State University 1997. Verfügbar unter: <http://research.cs.vt.edu/usability/publications/castillo-remote-usability.pdf> (22.02.2006)
- Dray, S.; Siegel, D. (), Remote possibilities? International usability testing at a distance, ACM interactions, 11, 2004, S. 10-17. Verfügbar unter: <http://portal.acm.org> (25.03.2006)
- Flanagan, C.: The Critical Incident Technique. Psychological Bulletin, 51/4, 1954, S. 327-358
- Hamborg, K.-C.; Hoemske, T.; Ollermann, F.: Qualitätssicherung im Usability-Testing – zur Reliabilität eines Klassifikationssystems für Nutzungsprobleme. In: Heinecke, A.M.; Paul, H.: Mensch & Computer 2006: Menschen und Computer im StrukturWandel. München 2006, S. 115-124
- Hartson, R.; Castillo, J.; Kelso, J.; Kamler, J.; Neale, C.: Remote Evaluation: The Network as an Extension of the Usability Laboratory. In: Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '96), 1996, S. 228–235
- Kehr, S.; Bosenick, T.; Kuehn, M.; Nufer, S.: Remote Usability Tests – eine Erweiterung der Usability-Toolbox für Online-Shops. In: Bosenick, T.; Hassenzahl, M.; Müller-Prove, M.; Peissner, M. (Hrsg.): Usability Professionals 2006, S. 104-108
- Molich, R.; Ede, M.; Kaasgaards, K.; Karyukin, B.: Comparative usability evaluation. In: Behaviour & Information Technology, 2004 - 23/1, 65-74
- Nufer, S.: Weiterentwicklung einer asynchronen Remote Usability Testmethode mit Bewertung und Evaluation der Ergebnisse. Diplomarbeit in Kooperation mit otto.de und der Fachhochschule Kaiserslautern, 2006
- Perkins, R.: Remote Usability Evaluation over the internet. In: Branaghan, R. (ed.): Essays on Usability. Usability Professionals Association 2001
- Tullis, T.; Fleischmann, S.; McNulty, M.; Cianchette, C.; Bergel, M.: An Empirical Comparison of Lab and Remote Usability Testing. In: Proceedings of Usability Professionals Association Conference, 2002