

Sandra Burger, Michael Burmester und Andreas Selter

Formatives Remote Usability Testing

Der vorliegende Artikel beschäftigt sich mit dem Thema Remote Usability-Testing. Im Vergleich zu einem Test im Labor können die Testpersonen bei einem Remote-Test von zu Hause oder von ihrem Arbeitsplatz aus teilnehmen, da ein solcher Test über das Internet stattfindet. Es wurde die Durchführung von Remote Usability-Tests mit der Kommunikationssoftware „virtual team room“ (vitero) untersucht. Es sollte die Frage geklärt werden, ob es möglich ist mit vitero eine weitestgehend entsprechende Testsituation wie im Labor zu erreichen. Weiterhin sollte geprüft werden, ob auch Testobjekte getestet werden können, die nicht online erreichbar sind, wie zum Beispiel Papier-Prototypen und mobile Geräte.

1. Einleitung

Formatives Usability-Testing mit repräsentativen Nutzern, die realistische Aufgaben mit einem zu testenden interaktiven System ausführen, dabei laut denken und deren Verhalten beobachtet wird, gilt als sehr effektive Methode um Usability-Probleme zu identifizieren (Nielsen 1993, S.165; Nielsen 1995, Dumas 2003). Allerdings ist es häufig schwierig, Testteilnehmer für einen Usability-Test im Labor zu gewinnen, da dies unter Umständen einen großen Zeitaufwand mit sich bringt. Remote Evaluation wird von Hartson et al. folgendermaßen definiert: „We define remote evaluation to be usability evaluation wherein the evaluator, performing observation and analysis, is separated in space and/or time from the user.“ (Hartson et al., 1996, S. 228). Mit Hilfe von Remote Usability-Testing (RUT) können Kosten für Anreise und Arbeitsausfall reduziert werden. Eine Studie von Tullis et al. (2002) zeigt, dass Testpersonen eher bereit sind an einem RUT teilzunehmen, als an einem Labor-Test. Ein

Grund dafür ist der geringere Aufwand, den ein RUT im Vergleich zu einem Labor-Test für die Testpersonen nach sich zieht (Dumas 2003). Als weiterer Vorteil ist die Testumgebung zu nennen. Da Benutzer in ihrer normalen Arbeitsumgebung an dem Test teilnehmen können, ist diese Situation realistischer als im Labor (Brush et al. 2004). Die Testsituation ist erwartungsgemäß nicht vollständig natürlich, jedoch weitaus natürlicher als im Labor. Ein besonderer Vorteil ist, dass spezielle Fachleute, die oft schwer zu rekrutieren sind, eher an einer zeitsparenden Lösung interessiert sind. Mit RUT können auch Nutzer aus anderen Kulturen einfacher erreicht werden.

RUT kann in zwei Kategorien unterteilt werden: synchron und asynchron. Bei einem synchronen RUT befinden sich der Testleiter und die Testperson zur gleichen Zeit an einem anderen Ort. Der Test wird moderiert und es findet eine Kommunikation zwischen dem Testleiter und der Testperson statt (Dray und Siegel 2004). Speziell diese Variante ist wichtig für formatives RUT.

Studien, die einen Labor-Test mit einem RUT vergleichen, belegen, dass sowohl mit synchronen als auch mit asynchronen RUT quantitative und qualitative vergleichbare Ergebnisse erreichbar sind (Dumas 2003). Auch in der Anzahl und Schwere der gefundenen Usability-Probleme konnten keine signifikanten Unterschiede festgestellt werden (Brush et al. 2004; Bolt und Peters 2005). Häufig wurden durch einen RUT sogar mehr Usability-Probleme aufgedeckt (Thompson et al. 2004). Testteilnehmer empfinden RUT zudem angenehmer als Labor-Tests (Brush et al. 2004).

In Usability-Tests werden unterschiedliche Typen von Artefakten untersucht, wie zum Beispiel mobile Systeme, Papier-Prototypen, Produktsimulationen, Websites, Softwareprodukte, Geräte etc. In dieser Studie sollte geklärt werden, wie praktikabel die Durchführung von RUT

über die Kommunikationssoftware „virtual team room“, vitero, ist und ob ein formativer Usability-Test über vitero mit einer entsprechenden Testsituation im Labor vergleichbar ist. Zudem sollte geprüft werden, ob RUT mit unterschiedlichen Testobjekten möglich ist. Drei verschiedene Testartefakte wurden ausgewählt: Papier-Prototyp, PowerPoint-Prototyp und Website.

2. vitero als RUT-Plattform

vitero ist eine Kommunikationssoftware, die es ermöglicht Besprechungen und Schulungen über das Internet abzuhalten. Die Teilnehmer einer vitero-Sitzung werden durch einen Avatar repräsentiert (vitero 2006). vitero sieht drei Rollen vor: Den Moderator, den Co-Moderator und den Teilnehmer. Es werden Funktionen zur Kommunikation und gemeinsamen Bearbeitung von Dokumenten bereitgestellt. Die für den RUT benötigten Funktionen werden im Folgenden erläutert.

In der Mitte des Bildschirms befindet sich der virtuelle Besprechungstisch (vgl. Bild 1). Darüber steht den Moderatoren eine Leiste mit Werkzeugen zur Verfügung. Die Avatare der Teilnehmer sind um den Tisch angeordnet. Die Moderatorenplätze sind an den beiden Mikrofonensymbolen zu erkennen, die den Moderatoren während der gesamten Sitzung zur Verfügung stehen. Weiterhin gibt es ein Mikrofon, welches herangereicht werden kann und so den einzelnen Teilnehmern das Sprechen ermöglicht. Mit der Funktion Application Sharing können Dokumente gemeinsam angesehen und bearbeitet werden. Auf dem virtuellen Besprechungstisch wird der Desktop der Person angezeigt, die Application Sharing gestartet hat. Diese Person kann nun jede beliebige Anwendung starten. Die verwendeten Programme und Anwen-



Bild 1: vitero Besprechungstisch (vitero 2007)

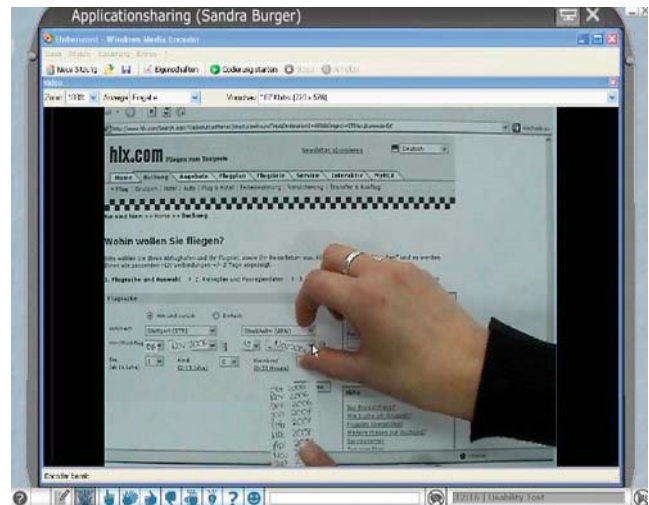


Bild 2: Ausschnitt aus vitero bei Änderung des Papier-Prototyps aus Sicht des Testteilnehmers (vitero 2006)

dungen müssen dazu nicht auf dem Rechner der Teilnehmer installiert sein. Sobald Application Sharing gestartet wird, wird die virtuelle Maus eingeblendet. Durch diese wird einem Teilnehmer das Fernsteuerrecht übergeben, das heißt, er kann die Tastatur und den Mauszeiger der Person steuern, die Application Sharing gestartet hat.

3. Ablauf der Studie

Die sechs Testteilnehmer hatten noch nie an einem Usability-Test teilgenommen und auch noch keine Erfahrungen mit vitero. Alle Teilnehmer wurden durch eine E-Mail über den Test informiert und erhielten einen Link zum Download der ausführbaren vitero-Applikation. Zur Kommunikation über vitero wird ein Headset benötigt, welches den Testpersonen zugeschickt wurde. Um sicherzustellen, dass die Computer aller Testpersonen über die geforderten technischen Voraussetzungen verfügen, wurde bei allen ein Technik-Check durchgeführt.

Im Test wurde die Website www.hlx.com untersucht. Für den Papier-Prototyp wurden Screenshots der Website ausgedruckt. Dropdown-Menüs und Auswahlmöglichkeiten wurden aus Papier erstellt und bei Bedarf auf die Screenshots gelegt, um Interaktionen zu simulieren. Der PowerPoint-Prototyp enthielt Screenshots der Website, aber keine aktiven Links. In einem Pilottest wurde der Ablauf einer Sitzung geprüft und anschließend optimiert.

Ein typischer Sitzungsverlauf sah folgendermaßen aus: Nach einer kurzen Einführung in die benötigten vitero-Funktionen wurde der jeweilige Testablauf (s. Kapitel 3.1, 3.2, 3.3) erklärt und geübt. Pro Testobjekt wurden zwei Tests mit jeweils einer Testperson durchgeführt. Die Testteilnehmer sollten während der Aufgabenbearbeitung laut denken, was durch die bestehende Audioverbindung möglich war. Die Testaufgaben wurden auf dem virtuellen Besprechungstisch präsentiert und durch den Moderator gestellt. Um zum Beispiel die Website gemeinsam zu nutzen wurde Application Sharing verwendet. Das Testobjekt konnte durch den Vollbild-Modus in vitero nahezu in Bildschirmgröße angezeigt werden. Durch Starten von Application Sharing erscheint die virtuelle Maus, die es dem Teilnehmer ermöglicht, die aufgerufene Anwendung zu bearbeiten. Um die nächste Aufgabe zu stellen, wurde Application Sharing und somit die Anwendung geschlossen. Da die Testperson den Mauszeiger des Moderators steuert, und dieser immer nur von einer Person verwendet werden kann, war es notwendig die „Übergabe der Maus“ abzusprechen.

Nach dem Usability-Test wurden die Teilnehmer über ihren Eindruck zum Testablauf mit vitero befragt. Die Fragen wurden über Application Sharing in einer PowerPoint-Präsentation geöffnet und gleichzeitig durch den Interviewer akustisch gestellt. Die einzelnen Sitzungen dauerten rund eine Stunde und wurden zur Dokumentation und Auswertung aufgezeichnet.

3.1 Remote Usability-Test mit einem Papier-Prototyp

Der Papier-Prototyp wurde mit einer Kamera gefilmt und über Application Sharing in vitero übertragen. Um eine bessere Lesbarkeit zu erreichen wurden die verwendeten Screenshots auf DIN A3 Papier ausgedruckt. Aufgrund der hohen Bildinformationen, kommt es je nach verfügbarer Bandbreite zu Verzögerungen bei der Übertragung des Films. Dies wurde den Testpersonen in der Einführung genau erklärt und demonstriert. Mittlerweile konnte das Problem durch eine spezielle Webcam mit hoher Auflösung gelöst werden. Zur Bearbeitung der Aufgaben erhielt der Teilnehmer die virtuelle Maus. Mit dem Mauszeiger konnte er im Kamerabild zeigen und gleichzeitig erläutern, auf welches Element er klicken würde. Bild 2 zeigt das Kamerabild aus Sicht des Testteilnehmers.

3.2 Remote Usability-Test mit einem PowerPoint-Prototyp

Bei dem PowerPoint-Prototyp konnten die Testteilnehmer, ähnlich wie beim Papier-Prototyp, nirgends klicken. Der Teilnehmer sollte auf das entsprechende Element zeigen und erläutern, was er tun würde, um die gewünschte Information zu finden. Hierbei fand sehr häufig eine „Mausübergabe“ statt, da der Teilnehmer dem Moderator die Maus übergeben bzw. diese loslassen musste, damit der Moderator die entsprechende Folie anzeigen konnte.

3.3 Remote Usability-Test mit einer Website

Bei diesem Test konnte die Testperson selbst mit der Website interagieren. Der Unterschied zur Bearbeitung am eigenen PC bestand darin, dass der Mauszeiger ferngesteuert wurde und somit je nach Anbindung leichte Verzögerungen auftraten.

4. Ergebnisse

Alle Tests konnten erfolgreich durchgeführt, und der Testablauf konnte ohne Schwierigkeiten eingehalten werden.

Die Sprachqualität der Audioverbindung war sehr gut und wurde von allen Teilnehmern als förderlich empfunden, da sie immer die Möglichkeit hatten Fragen zu stellen, und sich dadurch sicherer fühlten. Da die Aufgaben nicht während der Bearbeitung sichtbar waren, wurden Details oft vergessen und der Moderator wiederholte die Aufgabe verbal. Wenn die Aufgabenstellung parallel zum Testobjekt gesehen werden kann, fällt es leichter sich auf die Bearbeitung zu konzentrieren. Eine mögliche Lösung wäre den Teilnehmern die Aufgaben im Voraus per E-Mail oder per Post zuzuschicken.

Die Testteilnehmer sollten nach der Sitzung angeben, wie anstrengend sie diese erlebt haben. Zur Bewertung wurde die SMEQ-Skala zur Beurteilung der subjektiv erlebten Anstrengung verwendet (Eilers, Nachreiner und Hänecke, 1986). Diese stellt eine Skala von 0 bis 220 zur Verfügung, wobei 20 „kaum anstrengend“ und 220 „außerordentlich anstrengend“ bedeutet. Drei von sechs Testteilnehmern fanden die Sitzung „kaum anstrengend“ (20). Der höchste Wert (120) wurde beim Test mit dem Papier-Prototyp vergeben. Insgesamt lässt sich sagen, dass die Sitzungen in geringem Maße anstrengend waren.

Die Steuerung der Maus war bei dem Test mit der Website etwas ungewohnt. Da der Mauszeiger ferngesteuert wird, wird, aus Sicht der Testperson zusätzlich zum Mauszeiger (Mauszeiger der Maus, die ferngesteuert wird) noch ein Kreuz (Mauszeiger der eigenen Maus) angezeigt. Bei geringen Bandbreiten befindet sich der Mauszeiger etwas hinter dem Kreuz. Weiterhin konnte das Scrollrad an der Maus von den Teilnehmern nicht benutzt werden, was nicht der normalen

Nutzung entspricht. Um zu scrollen musste der Scrollbalken der Website verwendet werden, was auch durch die Anzeige von Pfeil und Kreuz nicht ganz einfach war, und weshalb manchmal unabsichtlich zu weit gescrollt wurde. Nach einer Eingewöhnungszeit konnten die Teilnehmer aber gut damit umgehen. Durch Realisierung der Nutzung des Scrollrads und Eliminierung des Kreuzes könnten diese Probleme vermieden werden.

Beim Test mit dem PowerPoint-Prototyp fand sehr oft ein „Mauswechsel“ statt, da der Moderator die virtuelle Maus übernehmen musste, um die nächste Folie anzuzeigen. Um dies zu vermeiden, könnten im Application Sharing Modus so genannte Teilnehmerpfeile eingeführt werden. Teilnehmerpfeile können durch Klicken auf eine Stelle auf dem virtuellen Besprechungstisch gesetzt werden. Somit wäre es möglich, Tests mit nicht interaktiven Testobjekten ohne die Nutzung der virtuellen Maus durchzuführen.

Die Reaktionszeit bei einem Test mit einem Papier-Prototyp wurde allgemein als langsam eingestuft, da die Teilnehmer genau beschreiben müssen, was sie tun würden, und dann abwarten müssen, bis die Änderung ausgeführt wird. Aus Sicht der Testpersonen spielt die zusätzliche Ladezeit des Bildes keine Rolle.

Die Ergebnisse zeigen, dass sich vitero sehr gut für RUT eignet. Weiterhin wird bestätigt, dass mit vitero eine Testsituation, die der im Labor weitestgehend entspricht, erreicht werden kann. Durch die bestehende Audioverbindung ist es möglich die Methode des lauten Denkens anzuwenden und auch Rückfragen zu stellen. Die Mausbewegungen und das Klickverhalten des Testteilnehmers kann durch die gemeinsame Nutzung der Anwendungen gut nachvollzogen werden. Im Anschluss an den Test kann ein Interview stattfinden. Weiterhin kann die gesamte Sitzung aufgezeichnet werden, wodurch Screenshots und die Kommentare der Testpersonen für die Auswertung und Highlight-Videos verfügbar sind.

Mit vitero ist es möglich, RUT auch mit Testobjekten durchzuführen, die nicht online erreichbar sind, wie zum Beispiel Papier-Prototypen. Gezeigt wurde, dass sowohl Testobjekte wie Websites, Online-Prototypen und Software, als auch Offline-Prototypen und mobile Geräte, durch Abfilmen mit einer Kamera, getestet werden können. Wichtig ist in diesem

Zusammenhang die Möglichkeit, über Application Sharing Anwendungen zur Verfügung zu stellen, die sich nur beim Moderator befinden.

Durch die von vitero bereitgestellten Funktionen konnten die Nutzer sich nach einer kurzen Eingewöhnungszeit in die jeweilige Testsituation hineinversetzen. Besonders wichtig sind in diesem Zusammenhang die Aspekte, dass vitero sehr anwenderfreundlich und leicht erlernbar ist und die Teilnehmer nicht über technische Kenntnisse verfügen müssen, um vitero zu nutzen. Wie die Interviews ergaben, war für die Teilnehmer ein wichtiger Punkt, dass die Tests von zu Hause aus durchgeführt werden konnten. Alle Teilnehmer würden wieder an einem solchen Test teilnehmen, da es für sie keinen großen Aufwand bedeutet hat. Auch die niedrigen Anstrengungseinschätzungen der Sitzungen durch die Testteilnehmer sprechen dafür, dass die RUT-Sitzung für die Testpersonen insgesamt angenehm war.

Die Ergebnisse dieser Arbeit zeigen, dass Remote Usability-Tests mit ganz unterschiedlichen Testobjekten, auf eine für den Benutzer unkomplizierte Art und Weise mit einem System wie vitero durchgeführt werden können. Durch die laufende Realisierung der Verbesserungsvorschläge zur Optimierung von vitero wird die Durchführung weiter optimiert.

Literatur

- Bold I Peters (2005): Remote Testing versus Lab Testing. (Verfasser J. Houck-Whitaker) <http://boltpeters.com/articles/versus.html>. (Datum des Zugriffs: 18.11.2006).
- Brush, B.A.J.; Ames, M.; Davis, J.: A Comparison of Remote and Local Usability Studies for an Expert Interface. Vienna, Austria: CHI (2004) 1179–1182.
- Dumas, J.C.: Usability Evaluation from your Desktop. In: Association for Information Systems (AIS) SIGCHI Newsletter (November 2003) 6–8.
- Dray, S.; Siegel, D.: Remote Possibilities? International Usability-Testing at a Distance. Interactions (März + April 2004) 10–17.
- Eilers, K.; Nachreiner, F.; Hänecke, K.: Entwicklung und Überprüfung einer Skala zur Erfassung subjektiv erlebter Anstrengung. Zeitschrift für Arbeitswissenschaft 40[1] (1986) 215–224.
- Hartson, H.R.; Castillo, J.C.; Kelso, J.; Neale, W.C.: Remote Evaluation: The Network as an Extension of the Usability Laboratory. CHI (1996) 228–235.
- Nielsen, J.: Usability Engineering. Boston: Academic Press, 1993.

Nielsen, J.: Getting usability used. In K. Nordby, P. Helmersen, D. Gilmore & S.A. Arnesen (Eds.), Human Computer Interaction - Interact '95 (pp. 3-13). London: Chapman & Hall, 1995.

Snyder, C.: Paper Prototyping. Amsterdam: Morgan Kaufmann, 2003.

Tullis, T.; Fleischman, S.; McNulty, M.; Cianchette, C.; Bergel, M.: An Empirical Comparison of Lab and Remote Usability-Testing of Web Sites. Usability Professionals Association Conference, 2002.

Thompson, K.E.; Rozanski, E.P.; Haake, A.R.: Here, There, Anywhere: Remote Usability-Testing That Works. SIGITE'04, S. 132-137. ACM, USA, 2004.

vitero: vitero Gruppenraum Plus. Stand Juni 2006.

vitero: vitero Gruppenraum Plus. (2007)



1



2



3

1 Sandra Burger ist im Bereich Usability Engineering bei der User Interface Design GmbH (UID) in München tätig. Sie hat Informationswirtschaft an der Hochschule der Medien in Stuttgart studiert. Bereits während des Studiums sammelte sie umfangreiche Erfahrungen im Bereich Usability von Webanwendungen. Mit dem Thema Remote Usability Testing unter Verwendung der Kommunikationssoftware „virtual team room“ (vitero) hat Sandra Burger sich im Rahmen ihrer Diplomarbeit eingehend beschäftigt.

E-Mail: sandra.burger@uidesign.de

2 Prof. Dr. Michael Burmester, Diplom-Psychologe, bis Ende 1996 als wissenschaftlicher Mitarbeiter am Fraunhofer-Institut für Arbeitswirtschaft und Organisation (IAO) in Stuttgart, danach im Fachzentrum User Interface Design der Siemens AG. Von 2000 bis 2002 leitete er den Bereich Usability-Engineering und die Geschäftsstelle München der User Interface Design GmbH und übernahm anschließend die Professur für Ergonomie und Usability im Studiengang Informationsdesign an der Hochschule der Medien in Stuttgart. Hauptarbeitsgebiete: Benutzerzentrierte Gestaltung, Attraktivität interaktiver Produkte, Eye-Tracking.

E-Mail: burmester@hdm-stuttgart.de

3 Andreas Selter arbeitet seit 2001 bei UID als Usability Engineer und ist dort seit 2004 Key Account Manager für den Bereich Web Solutions. Der Diplom-Informationswirt (FH) hat u.a. beim Usability Center der Siemens AG in den USA praktische Erfahrungen gesammelt. Andreas Selter hat reichhaltige Erfahrungen in der Leitung und Durchführung von großen Nutzerstudien auch auf internationaler Ebene. Andreas Selter hat zudem einen Lehrauftrag für Medienergonomie an der FH Ulm sowie für Informationsdesign an der FH Joanneum in Graz (Österreich).

E-Mail: andreas.selter@uidesign.de