

Klassischer Labor-UX-Test vs. unmoderierter Remote-UX-Test:

Wie unterscheiden sich die Ergebnisse bei parallelisierten Aufgabestellungen?

Eine Vergleichsstudie im Bereich On Premise Business-Software

Autoren:

Dr. Siegfried Olschner, DATEV eG, Fürther Straße 212, 90429 Nürnberg,
siegfried.olschner@datev.de

Doreen Feindt, GfK, Burchardstraße 19, 20095 Hamburg, doreen.feindt@gfk.com

Ulf Schubert, DATEV eG, Fürther Straße 212, 90429 Nürnberg, ulf.schubert@datev.de

Tina Mohr, GfK, Burchardstraße 19, 20095 Hamburg, tina.mohr@gfk.com

Abstract

Um eine größere Anzahl an Anwendern in den Softwareentwicklungsprozess einzubinden, bieten sich online durchgeführte, unmoderierte User Experience Tests an. Diese lassen sich im Vergleich zu Labor-Tests schnell und kostengünstig umsetzen. Während im Web-Bereich sogenannte Crowd-Tests eine gewisse Akzeptanz gefunden haben, stellt sich die Frage, welche Daten und vor allem welche Datenqualität ein unmoderierter User Experience Test im Bereich On Premise-Business-Software liefert.

Um diese Fragestellung zu beantworten, wurde eine Vergleichsstudie durchgeführt. Ziel war es, die Methode der unmoderierten Remote-Tests allgemein zu validieren, die Vorzüge und Schwachstellen zu identifizieren und einen Vergleich zu den klassischen User Experience Tests im Labor herzustellen.

Hierzu wurden zwei Gruppen rekrutiert, die jeweils einen klar umrissenen Bereich innerhalb der DATEV-Business-Software testen und bewerten sollten. Identisch zum 'moderierten Labor-UX-Test' (N=10) bearbeitete die Gruppe 'unmoderierter UX-Test' (N=20) dieselben Fragestellungen und Aufgabenszenarien, innerhalb so weit wie möglich parallelisierter Rahmenbedingungen.

Die Ergebnisse der Studie und unsere Erfahrungen werden im Beitrag vorgestellt und diskutiert. Außerdem werden Tipps zur praktischen Umsetzung gegeben.

Keywords

User Experience Test, unmoderierter User Experience Test, Crowd-Test, Methodenvergleich, Revelation

1 Relevanz von moderierten und unmoderierten UX-Tests

DATEV eG hat das Ziel, ihre Kunden effizienter in die Entwicklungsphasen einzubeziehen. Durch eine frühzeitige realistische Bewertung mittels direkten Kundeneinbezug im Entwicklungsprozess sollen die Produkte so früh wie möglich in der praktischen Anwendung durch die Anwender getestet werden.

Um Schwachstellen bei der Bedienung von neuen oder weiterentwickelten Softwareprodukten aufzudecken und Optimierungspotential zu identifizieren, werden bisher klassische User Experience Tests (UX-Tests) im Labor durchgeführt. In Einzelinterviews werden die Probanden gebeten, entlang typischer Aufgabenszenarien das Produkt zu nutzen und währenddessen laut zu denken. Der Moderator beobachtet dabei, exploriert an geeigneter Stelle und reflektiert zusammen mit dem Probanden die Nutzungsprobleme. Die Methode hat sich bewährt, da sie eine umfassende Rückmeldung zur User Experience des Produktes auf qualitativer Ebene liefert und in verschiedenen Phasen des Entwicklungsprozesses mit unterschiedlichsten Arten von Testsystemen (Low Level Prototyp bis Software im Auslieferungszustand) durchzuführen ist.

Der UX-Test im Labor ist allerdings sehr zeit- und personalintensiv und damit auch teuer. Vor allem die Feldphase erfordert die permanente Anwesenheit mehrerer an der Testdurchführung beteiligter Personen. Neben dem Interviewmoderator ist ein Protokollant anwesend, der Probandenaussagen und Beobachtungen für die Auswertung notiert. Zusätzlich empfiehlt es sich, dass alle Interviews von den zuständigen Produktmanagern und -entwicklern über einen Einwegspiegel verfolgt werden. Dies hat den Vorteil, dass diese die Interaktion mit ihrem Produkt und eventuelle Nutzungsprobleme "live" miterleben können. Jedoch müssen sie dafür ihren Arbeitsalltag für die 2-3 tägige Feldphase komplett unterbrechen. Auch kann sich die Terminierung der Feldphase als schwierig erweisen, wenn mehrere Projekt-Stakeholder den Interviews "live" beiwohnen wollen. Die Zeit- und Personalintensität der Feldphase führt häufig zu einer Restriktion der Stichprobe auf N=8 bis N=16 Probanden. Ein weiterer Nachteil ist die lokale Gebundenheit. Es können nur Zielgruppen rekrutiert werden, die im Umkreis der Testlabore verfügbar sind.

Um zukünftig eine größere Anzahl an Anwendern in die Produktentwicklung einzubinden sowie zeit- und kosteneffiziente Rückmeldungen zu einem Entwicklungsstand zu erhalten, stellt sich die Frage, ob unmoderierte UX-Tests (Tests ohne Interaktion zwischen Testleiter und Proband) eine adäquate Alternative zu den klassischen Labor-Tests darstellen können.

Um dies herauszufinden, hat DATEV eG zusammen mit der User Experience Abteilung von GfK beide Methoden in einer Vergleichsstudie gegenübergestellt. Ziel war es, die Vorzüge und Schwachstellen der unmoderierten UX-Tests im Vergleich zu klassischen UX Tests im Labor zu identifizieren. Die neue Methode sollte bezüglich Effizienz der Durchführung, Qualität der Ergebnisse und Skalierbarkeit für eine größere Menge von Produkten untersucht werden. Im Kapitel 2 wird die Durchführung der Studie näher beschrieben. Im Kapitel 3 wird darauf eingegangen, in welchen Punkten sich beide Methoden unterscheiden, dabei ist u.a. die Ergebnisqualität von Interesse. Abschließend berichten wir in Kapitel 4 über unsere Erfahrungen und geben Tipps zur praktischen Anwendung der unmoderierten UX-Tests.

2 Methodisches Vorgehen

2.1 Test beider Methoden gegeneinander bei gleichbleibenden Rahmenbedingungen

Um die unmoderierte Remote-Variante des UX-Tests mit der Labor-Variante vergleichen zu können, wurde ein und dasselbe DATEV-Produkt mit beiden Methodenvarianten getestet. Die Fragestellungen an die Teilnehmer waren dabei soweit wie möglich identisch. Auch alle weiteren Rahmenbedingungen wurden möglichst parallelisiert, wie z.B. die Stichprobenszusammensetzung, die Testumgebung und die Aufgabenszenarien. Aufgrund technischer Beschränkungen waren allerdings leichte Unterschiede in der zu testenden Software nicht vermeidbar. Diese Unterschiede waren jedoch vernachlässigbar.

Außerdem musste der Interviewleitfaden des Labortests auf die Bedürfnisse des unmoderierten Tests angepasst werden.

Beim klassischen UX-Test im Labor wurden N=10 qualitative Einzelinterviews à 60 Minuten durchgeführt. Am unmoderierten UX-Test nahmen insgesamt N=20 Probanden für jeweils eine Woche teil. Beide Methoden sind in Tabelle 1 gegenübergestellt.

Tabelle 1: Überblick über das Untersuchungsdesign

	Klassischer UX-Test im Labor (Kontrollgruppe)	Unmoderierter UX-Test (Experimentalgruppe)
		
Durchführungsort	UX-Labor	Online
Anwesenheit eines Moderators	Ja	Nein
Inhalte des UX-Tests	Teilnehmer nutzen die neue Software anhand parallelisierter klassischer Use Cases und beantworten Fragen rund um die Neuerungen	
Moderator- und Teilnehmer-Aktivitäten	<ul style="list-style-type: none"> • Teilnehmer "denken laut" während Aufgabenbearbeitung • Moderator beobachtet die Bearbeitung und exploriert tiefgehend an geeigneten Stellen • Nutzungsprobleme werden gemeinsam reflektiert • Moderator und Teilnehmer können sich direkt austauschen 	<ul style="list-style-type: none"> • Teilnehmer dokumentieren eigenständig ihr Vorgehen und ihre Beobachtungen während der Aufgaben in einem Online-Tool • Nutzungsprobleme werden von den Teilnehmer alleine reflektiert • Teilnehmer können sich zeitversetzt mit anderen Teilnehmern austauschen
Austauschmöglichkeit zwischen den Teilnehmern	Nein, da Einzelinterview	Ja, nach der Durchführung und Dokumentation jeder Aufgabe über das Online-Tool
Feldzeit	2 Tage	1 Woche (5 Arbeitstage)
Dauer pro Interview	60 min	Jeweils ca. 15-20 min pro Tag an insgesamt 5 aufeinanderfolgenden Tagen (Gesamtdauer ca. 75-100 min)
Stichprobengröße	N=10	N=20
Stichprobenzusammen	<ul style="list-style-type: none"> • ca. 75 % Steuerberater, 	

- ca. 25 % Mitarbeiter in Steuerberatungskanzleien
- Unterschiedliche Kanzleirollen berücksichtigt
 - Alle verwenden bereits eine Vorgängerversion des zu testenden Produktes, jedoch mit unterschiedlicher Nutzungshäufigkeit

2.2 Unmoderierter UX-Test: Remote-Nutzung der neuen Software und eigenständige Dokumentation

Beim unmoderierten UX-Test wurde die zu testende DATEV-Software über einen Windows Terminal Server (WTS) remote zur Verfügung gestellt. Auf diesen konnten sich die Probanden mit Hilfe des Programms Teamviewer aufschalten.

Die Teilnehmer wurden angehalten, die neue DATEV-Software über den Remote-Zugang zu testen, in dem sie vorgegebene Aufgaben bearbeiteten. Während der Aufgabenbearbeitung sollten die Teilnehmer einerseits ihr Vorgehen dokumentieren und andererseits alles, was ihnen positiv und negativ aufgefallen ist, inklusive aufgetretener Nutzungsprobleme.

Die Aufgabendokumentation fand über die Webplattform Revelation statt. Über Revelation wurden den Probanden täglich neue Fragen und Aufgaben rund um das Produkt gestellt. Auch alle Informationen zum Ablauf des unmoderierten UX-Tests und zur Nutzung des Remote-Zugangs wurden den Probanden hier zur Verfügung gestellt. Zu Beginn der Feldphase haben die Probanden eine kurze Schulung in Form von zwei Videos über die Webplattform erhalten. In einem Video wurde erklärt, wie eine Verbindung zu dem Remote-Zugang hergestellt werden kann. Das andere Video hat anhand eines Beispiels aufgezeigt, auf was bei der Aufgabenbearbeitung geachtet und welche möglichen positiven und negativen Beobachtungen von den Teilnehmern dokumentiert werden sollten.

Die Einladung zur Nutzung der Webplattform Revelation haben die Teilnehmer per E-Mail erhalten. Dort war ein entsprechender Link angegeben. Die Teilnehmer wurden dazu angehalten, sich im Verlauf der Arbeitswoche jeden Tag in die Webplattform Revelation einzuloggen und die dort gestellten Fragen zu beantworten sowie Aufgaben mit der neuen Software auszuführen. Über die Webplattform konnten die Teilnehmer außerdem zeitversetzt mit anderen Teilnehmern diskutieren, um sich über das DATEV-Produkt und ihre Beobachtungen während der Aufgaben auszutauschen.

2.3 Rekrutierung der Testpersonen

Für beide Tests wurden die Testpersonen auf Basis von entsprechenden Rekrutierungskriterien telefonisch rekrutiert. Somit war sichergestellt, dass die Testpersonen über die notwendigen Kenntnisse und Erfahrungen verfügen, um den Testgegenstand gut bewerten zu können.

3 Unterschiede zwischen den beiden Methoden

3.1 Detaillierterer Befragungsleitfaden beim unmoderierten UX-Test

Der Interviewleitfaden des UX-Tests im Labor musste an die Anforderungen des unmoderierten UX-Tests angepasst werden. Da beim unmoderierten UX-Test kein Moderator vorhanden ist, müssen alle Fragen und Aufgaben sehr detailliert und möglichst unmissverständlich formuliert sein. Auch müssen einige der im Labor-Interviewleitfaden aufgeführten Moderationshinweise als Fragen gestellt werden. Bei den Hinweisen handelt es sich meist um Beobachtungsschwerpunkte, auf die der Moderator während der Aufgabenbearbeitung achten soll. Dies kann z.B. eine gezielte Beobachtung sein, welcher Einstieg von den Teilnehmern in den Prozess verwendet wird, ob ein bestimmter Bereich der Software wahrgenommen wird oder ob der Umgang mit einem bestimmten Softwareelement Probleme bereitet. Nach diesen Punkten muss beim unmoderierten UX-Test explizit gefragt werden, da nicht

davon ausgegangen werden kann, dass die Teilnehmer solche Beobachtungen von selbst nach der Aufgabenbearbeitung reflektieren und zurückmelden (Tabelle 2). Die Vorgabe der einzelnen Aufgaben in Fragen hat gewisse Ähnlichkeit mit einem Online-Fragebogen.

Tabelle 2: Beispiel der Leitfadenanpassung zwischen den beiden Testmethoden

Fragen und Beobachtungshinweise im Interviewleitfaden des Labor-Tests	Ausformulierte Fragestellungen im Leitfaden des unmoderierten Online-Tests
<p>Aufgaben 1: Bitte rufen Sie bitte das Programm XY über die Programmsuche auf. <i>Beobachtungshinweise: Wird der Einstieg problemlos gefunden? Welche Einstiegsmöglichkeit wird verwendet? Wie gut kommt der Proband mit dem neuen Dialog der Programmsuche zurecht?</i></p> <p>// Nach Aufruf des Dialogfensters 'Programmsuche' Was erwarten Sie, wenn Sie an dieser Stelle einen Doppelklick auf das Icon XY machen?</p>	<p>Aufgabe 1: Bei der ersten Aufgabe, die ich Ihnen stelle, brauchen Sie sich die neue Software noch nicht über das Testsystem anzuschauen. Sondern beschreiben Sie einfach, was Sie glauben, was passiert. Stellen Sie sich einmal vor, Sie haben das Programm XY über die Programmsuche gesucht und Ihnen wird folgendes Fenster angezeigt:</p> <div data-bbox="821 840 1125 907" style="border: 1px solid gray; background-color: #cccccc; text-align: center; padding: 5px; margin: 10px auto; width: fit-content;">Screenshot</div> <p>Was erwarten Sie, wenn Sie an dieser Stelle einen Doppelklick auf das Icon XY machen?</p> <p>Aufgabe 2: 2.1 Bitte rufen Sie nun die neue Software im Testsystem auf und Öffnen Sie das Programm XY über die Programmsuche. Beschreiben Sie mir bitte danach genau die einzelnen Schritte Ihres Vorgehens. Uns interessiert, wie gut Sie mit dem neuen Einstieg zurechtkommen. 2.2 Welche Probleme sind aufgetreten? 2.3 Was hat Ihnen gut gefallen, was weniger gut? 2.4 Was könnte man aus Ihrer Sicht verbessern?</p>

Um die Fragen beim unmoderierten UX-Test so verständlich wie möglich zu formulieren, wurden Screenshots zur Veranschaulichung verwendet. War beispielsweise von Interesse, wie die Teilnehmer einen bestimmten Dialog bewerten, den sie während der Aufgabenbearbeitung verwendet hatten, so wurde der gesamte Prozess im Revelation-Tool noch einmal mit Screenshots nachgestellt (siehe Beispiel in Abbildung 1).

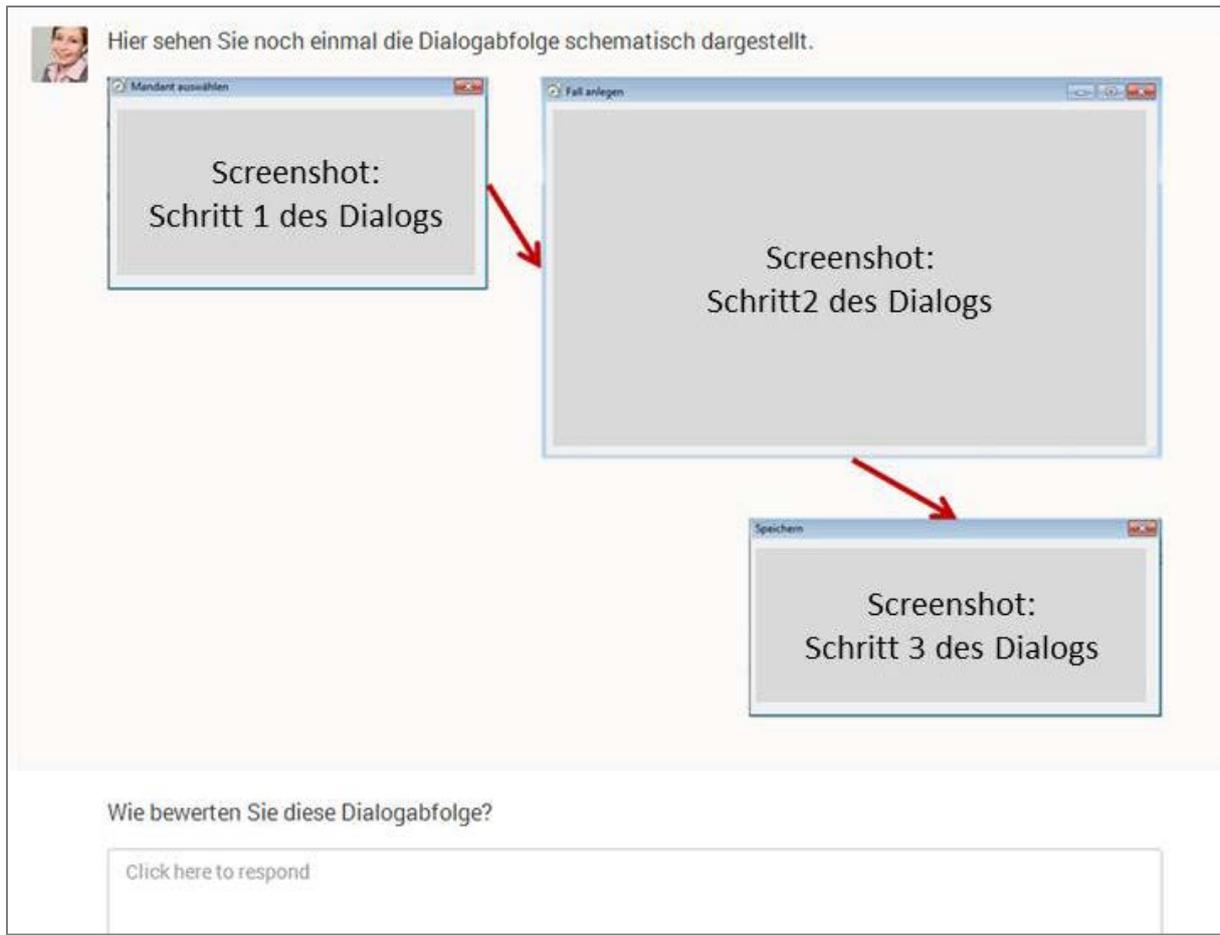


Abbildung 1: Beispielfrage mit Screenshots aus dem unmoderierten UX-Test

3.2 Über richtigen Lösungsweg erst am Folgetag beim unmoderierten UX-Test informiert

Aufgrund der fehlenden Präsenz eines Moderators beim unmoderierten Test erhalten die Probanden während der Aufgabenbearbeitung keine Hilfestellung. Allerdings sollte es auch im unmoderierten Test gewährleistet sein, dass die Teilnehmer über die richtige Lösung der Aufgabe aufgeklärt werden. Die Kenntnis des Lösungswegs ist teilweise Voraussetzung für spätere Aufgabenbearbeitungen. Außerdem kann das Scheitern an einer Aufgabe die Teilnehmer frustrieren und im schlimmsten Fall zum Abbruch der Studie durch einzelne Teilnehmer führen.

Den Teilnehmern des unmoderierten Tests wurden die Aufgabelösungen anhand von Screenshots über das Revelation-Tool angezeigt, allerdings erst am Folgetag. So sollte vermieden werden, dass sich die Lösung bereits vor Durchführung der Aufgabe angesehen wird. Aus diesem Grund durften die Aufgaben innerhalb eines Tages nicht direkt aufeinander aufbauen. Auch bei Nichtlösung einer Aufgabe sollte es möglich sein, alle weiteren Aufgaben des Tages zu bearbeiten.

Ob sich die Teilnehmer tatsächlich die Aufgabenlösungswege ansehen, kann beim unmoderierten Test nicht kontrolliert werden. Dadurch kann, anders als beim Test im Labor, nicht sichergestellt werden, dass alle Teilnehmer die richtige Lösung kennen und diese bei Folgeaufgaben berücksichtigen.

3.3 Ergebnisse aus beiden Methoden weitestgehend identisch, Datenqualität beim Labor-Test etwas besser

Die Studienergebnisse aus dem Labor-Test und dem unmoderierten Test sind zu großen Teilen identisch. Die Hauptnutzungsprobleme der neuen DATEV-Software konnten mit beiden Testmethoden identifiziert werden. Zusätzlich gibt es bei beiden Testmethoden kleinere Erkenntnisse, die bei der jeweils anderen Methode nicht zur Sprache kamen. Aufgrund der geringen Fallzahl kann es sich hierbei um zufällige Schwankungen handeln, die unabhängig von der Testmethode sind.

Allerdings ist festzuhalten, dass sich die Qualität und Tiefe der Ergebnisse zwischen den beiden Testmethoden deutlich unterscheidet.

- Der unmoderierte Test liefert nur eingeschränkt Hintergründe zu den erhobenen Daten. So kann z.B. nicht bei allen Nutzungsproblemen nachvollzogen werden, aus welchen Gründen diese aufgetreten sind oder warum sich Teilnehmer für ein bestimmtes Vorgehen innerhalb der Aufgabe entschieden haben. Die Daten des unmoderierten Tests beruhen allein auf den Antworten der Teilnehmer, die trotz aller Bemühungen bei der Erstellung der Instruktionen nicht an die präzisen Beobachtungsprotokolle des Labor-Tests heranreichen. Im Vergleich zum Lab-Test fehlen die zusätzlichen Beobachtungsdaten und die Informationen, die der Moderator an bestimmten Stellen tiefergehend exploriert hat. Dieser Mix aus verschiedenen Daten macht es möglich, die Ergebnisse umfassender zu interpretieren als bei einem unmoderierten Test. Beim unmoderierten Test variiert die Ausführlichkeit der Teilnehmerantworten sehr stark. Während einige Teilnehmer sehr detailliert erläutert haben, wie sie Schritt für Schritt innerhalb der Aufgaben vorgegangen sind und welche positiven und negativen Beobachtungen sie dabei gemacht haben, haben sich viele Teilnehmer kurz gefasst und damit einige der für die Untersuchung relevanten Details offen gelassen. Vor allem die Problembeschreibungen ließen häufig Interpretationsspielraum. So war zum einen nicht immer klar, wo genau ein Nutzungsproblem aufgetreten ist (Ort) und zum anderen warum es zu dem Problem kam (Ursache).
- Ferner konnte auch teilweise nicht festgestellt werden, ob eine Aufgabe richtig gelöst wurde, da manchmal die entscheidenden Hinweise in der Vorgehensbeschreibung fehlten. In einigen Fällen waren die Teilnehmer selbst unsicher, ob sie erfolgreich waren oder nicht.

Zusammenfassend lässt sich festhalten, dass viele Teilnehmerantworten zwar ausführlich waren (zumindest umfangreicher als ursprünglich erwartet), jedoch nicht ausführlich genug, um sie 100 % nachvollziehen zu können. Die Datenqualität des unmoderierten UX-Tests kommt nicht an die eines Labor-Tests heran. Trotz einiger Lücken innerhalb der Daten liefern sie dennoch umfangreiche – und ausreichende - Erkenntnisse zur User Experience eines Produktes.

4 Unsere Erfahrungen und Tipps für die praktische Umsetzung

4.1 Evaluation des unmoderierten UX-Tests

Um das methodische Vorgehen des unmoderierten UX-Tests an sich zu evaluieren, wurden die Teilnehmer im Anschluss an die Studie zu Ihren Erfahrungen befragt. Dazu wurde ein Online-Fragebogen geschaltet. An der Evaluation der Methodik nahmen 15 der 20 rekrutierten Testpersonen teil.

4.2 Fast alle Teilnehmer würden erneut mitmachen

Das Feedback der Teilnehmer zur neuen Methode "unmoderierter UX-Test" ist insgesamt sehr positiv. Der Testaufbau wird als angenehm empfunden. Die Aufteilung der Aufgaben und Fragen auf mehrere Tage lässt den Test weniger umfangreich wirken ("Jeden Tag kleine Häppchen habe ich für sinnvoll

erachtet."). Als positiv an der Methode wird vor allem die freie Zeiteinteilung angesehen. So kann die tägliche Teilnahme gut in den Arbeitsablauf integriert werden.

Fast alle Teilnehmer finden, dass mit dem unmoderierten UX-Test wertvolle Erkenntnisse für die Weiterentwicklung eines Programms gewonnen werden können und würden Kolleginnen und Kollegen dazu ermuntern, ebenfalls an einem teilzunehmen. Auch würden fast alle Teilnehmer selbst noch einmal bei einer solchen Studie mitmachen. Es soll an dieser Stelle jedoch nicht verschwiegen werden, dass im Vergleich zu einem Labor-Test die Rekrutierung trotz der zeitlichen und räumlichen Vorteile für die Teilnehmer leicht aufwendiger ist, als bei einem klassischen Labor-Test. Sicherlich liegt auch eine leichte Selbstselektion vor hinsichtlich der Technik-Affinität. Wir gehen davon aus, dass die Teilnehmer des unmoderierten Tests etwas technik-affiner sind als die Teilnehmer des Labor-Tests.

4.3 Studienaufwand ähnlich zum klassischen UX-Test im Labor

Die erhoffte Ersparnis beim Erstellungsaufwand konnte nicht umgesetzt werden. Für die Durchführung eines unmoderierten UX-Tests ist ähnlich hoher Zeit- und Personalaufwand notwendig wie für einen Test im Labor. Die Zeitersparnis während der Feldphase wird durch den Mehraufwand in der Vorbereitung und Auswertung wieder kompensiert.

Viele der Aufgaben, die der Moderator im Lab während der Interviews übernimmt, müssen beim unmoderierten UX-Test explizit vorhergesehen und vorbereitet werden.

- Die Fragen und Aufgaben im Instruktions-Leitfaden müssen sehr präzise ausformuliert werden, um Missverständnisse zu vermeiden.
- Außerdem muss sich im Vorfeld Gedanken über mögliche Nutzungsprobleme gemacht werden, um diese ggf. als Fragen mit in den Leitfaden aufzunehmen.
- Zusätzlich muss die richtige Aufgabenlösung differenziert beschrieben werden (z.B. mit Screenshots), so dass sie von allen Teilnehmern nachvollzogen werden kann.
- Auch alle Informationen und Hinweise zum Testablauf, die der Moderator normalerweise während des Interviews gibt, müssen vollständig ausgearbeitet werden. Zum einen erhalten die Teilnehmer vor Beginn des Tests eine schriftliche Information zum allgemeinen Ablauf, zum anderen wird das detaillierte Vorgehen anhand von Schulungsvideos erläutert.
- Während der Vorbereitungsphase müssen außerdem die technischen Systeme aufgesetzt werden. Dies umfasst die Einrichtung des Systems, auf welchem sich die zu testende Software befindet, inklusive der einzelnen Teilnehmerzugänge.

Ferner ist das Einpflegen aller Fragen, Aufgaben, Videos, etc. in das Revelation-System und das Testen des Ablaufs ebenfalls zeitintensiv. Betrachtet man die Auswertungsseite, müssen beim unmoderierten UX-Test zunächst alle Antworten der Teilnehmer durchgelesen werden, um einen Gesamtüberblick zu erhalten. Um die Teilnehmeraussagen und beschriebenen Nutzungsprobleme nachvollziehen zu können, muss z.T. die Testsoftware aufgerufen und das Szenario nachgestellt werden. Beim Labor-Test hingegen sind dem Projektteam nach Beendigung der Feldphase bereits alle Ergebnisse durch die Beobachtung bekannt und zum Großteil schon analysiert, so dass diese nur noch inhaltlich in Form eines Berichts strukturiert werden müssen.

Tabelle 3: Überblick über den Arbeitsaufwand eines unmoderierten UX-Tests im Vergleich zum Labor-Test

Zeitaufwand im Vergleich zum Labor-Test		Gründe
Vorbereitung		<ul style="list-style-type: none"> • Höherer Rekrutierungsaufwand, da mehr Teilnehmer

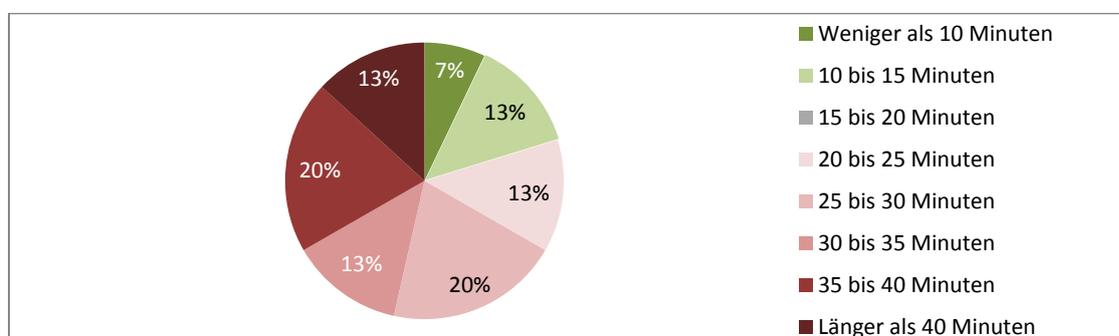
		<ul style="list-style-type: none"> • Einrichtung der WTS-Zugänge zum Testsystem • Aufwendigere Leitfadiskonzeption, u.a. aufgrund detaillierterer Fragestellungen • Beschreibung der Aufgabenlösung anhand von Screenshots • Einspielen der Fragen, Aufgaben und Lösungen aus dem Leitfaden in das verwendete Online-Tool • Erstellung eines Schulungsvideos für die Teilnehmer • Erneute Kontaktierung der Teilnehmer und Information über Studienablauf, Zugangsdaten, etc.
Durchführung	↓	<ul style="list-style-type: none"> • Permanente Anwesenheit während der Feldphase entfällt • Nur Support bei technischen Problemen und Fragen
Auswertung	↑	<ul style="list-style-type: none"> • Zunächst Durchlesen aller Teilnehmerantworten notwendig für Gesamtüberblick (danach erst ähnlicher Wissensstand wie nach Durchführung oder Protokollierung aller Interviews im Labor-Test) • Aufwendigere Detailauswertung, da Probleme bzw. Aussagen im Testsystem nachvollzogen werden müssen

4.4 Zeitlicher Aufwand für die Teilnehmer höher als ursprünglich geplant

80 % der Teilnehmer haben mehr Zeit für die Studie aufgewendet als ursprünglich angesetzt war. Geplant waren 15-20 min täglich. Tatsächlich hat fast jeder zweite Teilnehmer länger als 30 min benötigt, um die täglichen Fragen zu beantworten und die Aufgaben durchzuführen.

Soll der Umfang eines 60-minütigen Labor UX-Tests auf einen unmoderierten Test übertragen werden, so ist mit einem Gesamtzeitaufwand für die Teilnehmer von ca. 150 min zu rechnen (5 Tage für jeweils 30 min). Diese Bearbeitungszeit sollte den Teilnehmern vor Beginn der Studie mitgeteilt werden, um ihnen eine realistische Planung zu ermöglichen. Idealerweise sollte der tägliche Zeitaufwand auf 20 min reduziert werden, da die jetzige Bearbeitungsdauer von einigen Teilnehmern als grenzwertig eingeschätzt wurde. Vor allem, wenn technische Probleme zu lösen sind, die einen längeren Kontakt erfordern, wird dies von Teilnehmern aufgrund der "verschwendeten Zeit" als kritisch betrachtet.

Tabelle 4: Täglicher Zeitaufwand der Teilnehmer (n=15)



4.5 Persönlicher Kontakt von den Teilnehmern vermisst

Während der Feldphase haben einige Teilnehmer den Kontakt zu einem Moderator vermisst und dies auch zurückgemeldet ("Im Ganzen war das Online-Benutzerlabor sehr gut. Allerdings fehlt die persönliche Komponente.").

Während im Labor-Test die Fragestellungen individuell auf die Teilnehmer zugeschnitten werden, ist dies beim unmoderierten UX-Test nicht möglich. Es können keine Bezüge zu vorherigen Antworten hergestellt oder Fragen, die irrelevant für einen Teilnehmer sind, weggelassen werden (wenn er bspw. eine komplett andere Arbeitsweise hat oder sich schon dazu geäußert hat). Dadurch fühlten sich einige Teilnehmer partiell von "unnötigen" Fragen genervt ("Weniger gut fand ich auf den ersten Blick, dass sich einige Fragestellungen scheinbar wiederholten."). Sie vermissten die individuelle Ansprache.

4.6 Höhere Abbruchquoten machen Überrekrutierung notwendig

Es ist zu vermuten, dass aufgrund der fehlenden Teilnehmermotivierung durch einen Moderator beim unmoderierten UX-Test eine höhere Abbruchquote zu verzeichnen ist. Die Teilnahmequote beim klassischen UX-Test im Labor liegt erfahrungsgemäß bei 90-100%. Beim unmoderierten UX-Test haben 19 der 20 rekrutierten Teilnehmer die Studie begonnen, aber nur 14 bis zum letzten Tag durchgehalten (70 % erfolgreiche Teilnahmequote). Pro Tag sank die Teilnehmeranzahl im Schnitt um ca. einen Teilnehmer. Auch innerhalb eines Tages war eine steigende Abbruchquote zu beobachten. Die ersten Fragen des Tages wurden weitaus häufiger und ausführlicher beantwortet als die letzten.

Aus diesem Grund ist eine Überrekrutierung beim unmoderierten UX-Test notwendig. Auch um die geringere Datenqualität zu kompensieren, sollten mindestens doppelt so viele Teilnehmer zum unmoderierten UX-Test eingeladen werden im Vergleich zum Labor-Test.

Um die Abbruchquote zu reduzieren, erscheint es außerdem sinnvoll, die Feldphase auf 4 Arbeitstage zu beschränken statt bisher 5 Tage. Die Teilnehmer des unmoderierten UX-Tests nahmen mehrheitlich von ihrem Arbeitsplatz aus an der DATEV-Studie teil. Dies führt dazu, dass die Bearbeitung der Fragen sich in den Arbeitsalltag integrieren muss. Sofern an einem Tag keine Zeit zur Beantwortung der Fragen bleibt, wird dies oftmals am nächsten Tag nachgeholt. Da das Online-Tool Revelation nur für eine bestimmte Anzahl von Arbeitstagen (in diesem Fall 5 Tage) zur Verfügung steht, müssen alle Fragen bis zum letzten Verfügbarkeitstag des Online-Tools beantwortet sein. Ein Verschieben auf den Folgetag (Tag 6) ist nicht möglich.

4.7 Kein Test von komplexen Fragestellungen und Aufgaben möglich

Aus der Vergleichsstudie wurde deutlich, dass der unmoderierte UX-Test einige Restriktionen gegenüber dem Labor-Test hat.

- Es ist nicht möglich, Software oder Prototypen zu testen, die sehr erklärungsbedürftig sind, weil bspw. einige Funktionen noch nicht richtig funktionieren.
- Auch sind komplexe Fragestellungen und Aufgaben nur schwer im unmoderierten UX-Test zu verproben. Dadurch, dass Fragen leicht missverstanden werden können und es schwer ist, diese im Nachhinein zu präzisieren, sollten möglichst einfache Fragestellungen gewählt werden. Eine hohe Komplexität von Aufgaben ist nicht ratsam, da den Teilnehmern während der Bearbeitung nicht gezielt geholfen werden kann.

4.8 Bisherige technische Umsetzung funktioniert, ist aber noch optimierungsfähig

DATEV eG stellt hohe Datenschutzerfordernungen an die verwendeten Test-Systeme. Schließlich nutzen die meisten Teilnehmer ihren Arbeitsplatz zur Teilnahme. Während des unmoderierten UX-Tests gab nur wenige technische (Konfigurations-) Probleme. Fast alle Teilnehmer sind gut mit dem Teamviewer und dem WTS zurechtgekommen. Es gefällt, dass durch den Teamviewer kein extra Programm installiert werden muss. Das Herunterladen und Ausführen der Software wurde, vor allem aufgrund der Erläuterungen im Einführungsvideo, als einfach empfunden. Hier sollte zukünftig sichergestellt werden, dass tatsächlich alle Teilnehmer Zugang zum Einführungsvideo haben. Einige Teilnehmer konnten keinen Ton am Arbeitsplatz abspielen und hätten sich per E-Mail einen Nutzungsleitfaden für das Video gewünscht, um entsprechende Vorbereitungen treffen zu können.

Bei der Nutzung des Online-Tools Revelation kam es im Test teilweise zu Problemen beim Einloggen und Anzeigen von Videos, wenn eine ältere Browserversion installiert war. Daher sollten die Nutzer in Folgestudien präzise auf die technischen Voraussetzungen hingewiesen werden, bevor sie den Test beginnen.

Da es immer wieder zu kleineren technischen Problemen kommen kann, muss eine Betreuung der Teilnehmer in Form von E-Mail- und / oder Telefonkontakt sowohl auf Seiten der DATEV als auch GfK gewährleistet sein. Der Support des Online-Tools Revelation reicht hier nicht aus, da dieser zum einen keine Einflussmöglichkeiten auf den WTS-Server der DATEV oder den TeamViewer hat. Und zum anderen sitzt der Support des Online-Tools Revelation in den USA, so dass die sprachliche Barriere und die Zeitverschiebung die Kommunikation erschweren.

4.9 Komplette unmoderierter UX-Test nicht sinnvoll

Unsere Erfahrungen zeigen, dass ein komplett unmoderierter UX-Test, d. h. der vollständige Verzicht auf die Moderation, nicht sinnvoll ist. Um einen reibungslosen Ablauf der Feldphase zu gewährleisten, sollte das Projektteam mindestens zwei Stunden pro Tag zum Durchlesen der Teilnehmerantworten einplanen. Nur so kann schnell reagiert werden, wenn

- ein Fehler im Testsystem vorliegt,
- mehrere Teilnehmer eine Frage falsch verstanden haben,
- Teilnehmer sonstige Fragen haben,
- Teilnehmerantworten unverständlich sind

Die Interaktion des Moderators mit den Teilnehmern steigert die Qualität der Ergebnisse, da auf Basis der Beobachtungen gezielt nachgefragt werden kann. Zudem erhöht die persönliche Komponente die Wahrscheinlichkeit, dass die Teilnehmer den Test nicht vorzeitig abbrechen.

5 Fazit

Der unmoderierte UX-Test ist mit den oben genannten Einschränkungen eine geeignete Methode, um die User Experience eines Produktes zu untersuchen und die größten Probleme während der Produktnutzung zu identifizieren. Aufgrund der fehlenden Beobachtungs- und Nachfragemöglichkeit gibt er allerdings nur eingeschränkt Aufschluss über die Hintergründe und Ursachen von Nutzungsproblemen. Weswegen ein klassischer Labor-UX-Test nur in Teilen durch einen unmoderierten Remote-Test ersetzt werden kann.

Eine 100prozentige unmoderierte Durchführung des UX-Tests erscheint wenig sinnvoll. Der Moderator der Studie sollte über das Online-Tool regelmäßig mit den Teilnehmern in Kontakt treten, um auf

Fragen zu antworten oder selbst nachzufragen. Dadurch könnte zum einen die Ergebnisqualität gesteigert werden. Zum anderen würde die Interaktion mit einem Moderator die Teilnehmermotivation erhöhen und somit die Abbruchquote verringern.

Die Online-Variante des UX-Tests ist dann der Labor-Variante vorzuziehen, wenn ein großer potentieller Anwenderkreis einbezogen werden soll und Testgegenstand sowie Fragestellung nicht zu komplex sind. Große Stichproben (ab $N=30$) sind online besser zu managen und weniger kostenintensiv als im Labor. Für die Labor-Variante sollte sich in jedem Fall entschieden werden, wenn die Fragestellungen und Aufgabenszenarien komplex sind und wenn das Testsystem sehr erklärungsbedürftig ist.

Viten der Autoren



Dr. Siegfried Olschner ist Mitglied des User Experience Design -Teams der DATEV eG Nürnberg. Neben allgemeinen Beratungsthemen und Workflow- bzw. Dialoggestaltung innerhalb des User Centered Design führt er vor allem User Research-Projekte und Methoden-Beratungen durch. Er beschäftigt er sich mit Fragen zur Verbesserung und Effizienzsteigerung von Datenerhebungen (z. B. beim innerhalb der DATEV eingesetzten User Experience Questionnaire) und mit Prozessfragen, die die Integration der UX-Ergebnisse in den Entwicklungsprozess betreffen.



Doreen Feindt ist Diplom Wirtschaftspsychologin (FH) mit dem Schwerpunkt Arbeits- und Ingenieurpsychologie. Seit 2004 arbeitet sie im User Experience Team von GfK in Hamburg (früher SirValUse Consulting GmbH). Als Senior Consultant leitet sie dort zahlreiche qualitative User Experience Projekte, die inhaltlich von Konzepttests und Rapid Prototyping über Akzeptanztests bis hin zu multi-nationalen Studien reichen. Zur Unterstützung eines ganzheitlichen Nutzererlebnisses führt sie auch Studien durch, die einen qualitativen und quantitativen Ansatz sinnvoll kombinieren.

Ihr besonderes Interesse gilt den digitalen Produkten der Finanzbranche, für die sie als Account Manager im Bereich User Experience bei GfK mitverantwortlich ist.



Ulf Schubert ist Leiter User Experience bei DATEV eG in Nürnberg. Zu seinen Aufgabenbereichen gehören u. a. die anwenderorientierte Oberflächengestaltung und die User Experience der DATEV-Produkte. Zuvor arbeitete er mehrere Jahre als Usability Consultant und User Experience Designer u. a. bei SirValUse Consulting in Hamburg. Er berichtet in seinem User Experience Blog über Neuigkeiten der Branche und den Erfahrungen aus seiner täglichen Arbeit (www.ux-blog.de).



Tina Mohr studierte Wirtschaftspsychologie an der Leuphana Universität Lüneburg und arbeitet seit Anfang 2009 für das User Experience Team von GfK in Hamburg (früher SirValUse Consulting GmbH). Sie ist Consultant im Bereich User Experience und leitet in dieser Position jährlich die Durchführung zahlreicher, vornehmlich qualitativer Studien. Sie verfügt über einen breiten Erfahrungshorizont an durchgeführten Methoden: von User Feedbackdays in der frühen Entwicklung,

über Expertengutachten, Gruppendiskussionen und klassischen Usability-Studien, sowohl während der Entwicklung mit Prototypen als auch nach dem Launch mit dem bereits fertigen Produkt. Ihre Expertise liegt dabei in der Planung, Durchführung und Auswertung von Eye-Tracking-Studien.