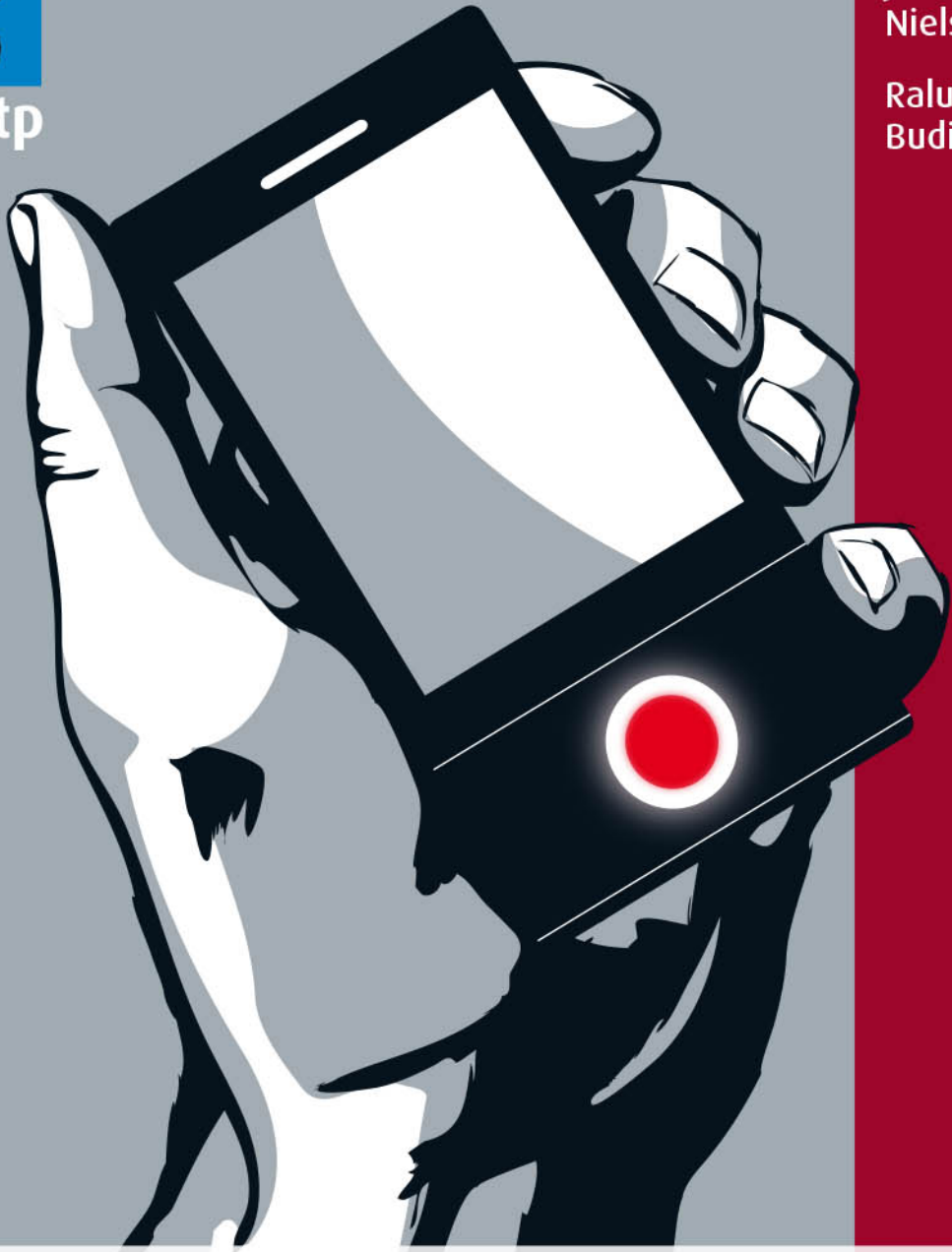




Jakob
Nielsen

Raluca
Budiu



Mobile Usability

Für iPhone, iPad, Android, Kindle

Unsere Forschungsarbeit: Aufbau unserer Usability- Studien

Dieses Buch unterscheidet sich speziell in einer Hinsicht entscheidend von fast allen anderen Büchern zu diesem Thema: Unsere Erkenntnisse, Schlussfolgerungen und Empfehlungen basieren auf empirischen Untersuchungen. Wir legen nicht unsere eigene Meinung zugrunde und wollen Sie auch nicht dazu bewegen, die Art von mobilen Websites und Apps zu gestalten, die wir persönlich bevorzugen und selbst am liebsten nutzen würden.

Vielmehr analysieren wir, auf welche Art und Weise Websites und Apps von einer Vielzahl durchschnittlicher Anwender weltweit genutzt werden. Was gefällt ihnen und was nicht? Welche Designs ermöglichen ihnen eine problemlose Handhabung und welche bereiten ihnen Schwierigkeiten? Wir kennen die Antworten auf diese Fragen, denn wir haben sie genauestens untersucht.

Sie, verehrte Leserin und verehrter Leser, gehören vermutlich nicht zu den durchschnittlichen Anwendern, sondern interessieren sich wahrscheinlich vor allem für die Mobilgerätenutzung als solches und das Design von Benutzeroberflächen. Warum sonst sollten Sie dieses Buch in Händen halten, wo Sie sich doch für das gleiche Geld ein paar richtig gute Krimis hätten kaufen können?

Wir (Jakob und Raluca) haben beide einen akademischen Hintergrund sowie jahrzehntelange Erfahrung im Hightech-Business und der User-Experience-Forschung. Zudem leben wir im Silicon Valley. Was uns persönlich gefällt, ist somit vollkommen irrelevant, weil wir uns von der Mehrheit der weltweiten Nutzer deutlich unterscheiden. Orientieren Sie sich bei Ihrer Designarbeit also nicht an unseren Ansprüchen – und ebenso wenig an Ihren eigenen.

Passen Sie Ihr Design vielmehr den Anforderungen Ihrer Klientel an.

Wir können Ihnen nur wärmstens empfehlen, auch eigene Usability-Testreihen mit geeigneten Nutzern durchzuführen, um Ihr Design zu testen. Jede Testreihe wird Ihnen wertvolle neue Erkenntnisse liefern. In diesem Kapitel erfahren Sie, wie Sie zur Untersuchung der Mobile Usability am besten vorgehen.

Betrachten wir für den Anfang zunächst einmal einige der von uns bereits durchgeführten Studien. Die in diesem Buch beschriebenen Untersuchungsergebnisse wurden durch die Beobachtung des Verhaltens echter Nutzer im Rahmen zahlreicher, über Hunderte von

Stunden hinweg geführte Testreihen gewonnen, deren Aufbau auf den folgenden Seiten genau beschrieben ist. Sollten Sie weniger am Aufbau als vielmehr an den Ergebnissen der Studien interessiert sein, können Sie die folgenden Seiten auch überspringen und gleich mit Kapitel 2 fortfahren.

Auftragsstudien

In der Hauptsache befassen wir uns mit zwei Arten von Nutzerforschungsprojekten: Studien im Kundenauftrag und Studien, die von unserem Unternehmen, der Nielsen Norman Group, eigenfinanziert werden. Das oberste Consulting-Gebot lautet, dass alle Kundendaten vertraulich bleiben müssen – daher können wir weder Informationen über unsere Consulting-Klientel noch über die aus den Auftragsstudien gewonnenen Untersuchungsergebnisse preisgeben. Neben den im Kundenauftrag unternommenen Studien führen wir allerdings ergänzend noch eine Vielzahl von unabhängigen Testreihen durch. Und da wir diese selbst finanzieren, steht es uns natürlich frei, deren Abläufe und Resultate vorbehaltlos zu veröffentlichen – was wir in diesem Buch tun.

Doch selbst wenn wir Stillschweigen über die Einzelheiten der Auftragsstudien bewahren müssen, machen wir im Laufe solcher Testreihen natürlich auch Beobachtungen, die nicht unmittelbar Bestandteil des Kundenauftrags sind. Und insofern fließen diverse Erkenntnisse aus den Auftragsstudien auch in die allgemeingültigen Richtlinien ein. Denn wenn sich im Rahmen einer Auftragsstudie etwas Ungewöhnliches oder völlig Unerwartetes ergibt, führen wir häufig noch eigene weiterführende Testreihen durch, um das betreffende Nutzerverhalten eingehender zu untersuchen. Werden die Ergebnisse einer unserer älteren öffentlichen Studien hingegen durch eine aktuelle Auftragsstudie bestätigt, erübrigt sich die weitere Untersuchung des jeweiligen Sachverhalts, weil die bisherigen Erkenntnisse in diesem Fall nachweislich weiterhin Bestand haben.

1.1 Tagebuchstudien

Um uns zunächst einen Überblick über das Spektrum der von Handynutzern durchgeführten Aktivitäten zu verschaffen, begannen wir unser Mobile-Usability-Projekt mit zwei separaten Tagebuchstudien. An der ersten Studie nahmen 14 Nutzer unterschiedlicher Handyausführungen – Barren-Handys und Smartphones mit und ohne Touchscreen – aus sechs verschiedenen Ländern (Australien, den Niederlanden, Rumänien, Singapur, Großbritannien und den USA) teil. Und in der zweiten Studie wurden die Daten von 13 iPhone-Besitzern in den USA analysiert.

Da wir bei der ersten Tagebuchstudie weniger an spezifischen Usability-Problemen interessiert waren, wählten wir in diesem Fall Probanden mit relativ fortgeschrittenen technischen Fähigkeiten und verzichteten auf die typischen beruflichen Ausschlusskriterien (beispielsweise keine IT-bezogenen Tätigkeiten), die für die Teilnehmer traditioneller Usability-Testverfahren gelten. Für die zweite Studie rekrutierten wir hauptsächlich durchschnittliche Nutzer, die nicht im Marketing- oder IT-Bereich tätig waren.

In beiden Tagebuchstudien wurde eine Variante der sogenannten Snippet-Technik angewandt, um eine möglichst einfache und unterbrechungsfreie mobile Protokollierung zu

gewährleisten: Die Probanden zeichneten über einen Zeitraum von ein bis zwei Wochen hinweg täglich jede Aktivität mit ihrem Handy (außer Telefonaten und Textnachrichten) via Twitter auf. Anhand der Tweets konnten sie am Ende des Tages zuverlässig nachvollziehen, wann und wofür sie ihr Handy eingesetzt hatten, und den entsprechenden Kontext jeder einzelnen Mobilaktivität problemlos in einem Fragebogen beschreiben.

Die Snippet-Technik wurde erstmals von J. Brandt, N. Weiss und S. Klemmer in ihrem Artikel »*txt 4 l8r: Lowering the burdens for diary studies under mobile conditions*« (CHI, 2007) vorgestellt.

Twitter ist ein Mikroblogging-Dienst, der es seinen Usern ermöglicht, Kurznachrichten zu posten und an ihre Twitter-Freunde zu übermitteln, die zuvor ausdrücklich angegeben haben, aktuelle Nachrichten von dem betreffenden User erhalten zu wollen (sprich, die diesem User »folgen«).

Zum Abschluss der Tagebuchstudie führten wir noch eine kurze Telefonbefragung mit den Teilnehmern durch – im Falle der zweiten Studie wurden die Probanden allerdings nicht telefonisch befragt, sondern gebeten, für eine normale Usability-Testsitzung in unser Labor zu kommen.

1.2 Usability-Tests

Wir haben im Laufe der Jahre zahlreiche Mobile-Usability-Studien in den Vereinigten Staaten, aber auch in anderen Ländern (Australien, Hongkong, den Niederlanden, Rumänien und Großbritannien) durchgeführt. In allen Fällen handelte es sich dabei um traditionelle Usability-Untersuchungen unter Anwendung des Think-Aloud-Protokolls, d.h., die Tests fanden in Einzelsitzungen (nicht mit Fokusgruppen) statt, wobei die Probanden ihre Gedanken laut formulieren, sprich laut denken sollten.

Zweck dieser Studien war die Untersuchung der typischen Usability-Probleme, die beim Internetzugriff bzw. der App-Nutzung mit verschiedenen Handyvarianten auftreten – inklusive Barren-Handys und Smartphones mit und ohne Touchscreens.

In der Regel brachten die Nutzer ihre eigenen Handys oder Tablets zu den 60- bis 90-minütigen Einzeltestsitzungen in unserem Labor mit. (Ausnahme: Ein paar Teilnehmer der Tablet-Studien verwendeten von uns bereitgestellte Tablets.) An jeder Sitzung nahmen jeweils der Proband, ein Testleiter und in manchen Fällen auch ein oder zwei Beobachter teil. Zunächst erläuterten die Nutzer, wie sie ihre jeweiligen Geräte normalerweise bedienen, und demonstrierten uns einige ihrer darauf installierten Apps. Anschließend stellten wir jedem Teilnehmer ein paar Aufgaben, die entweder einen Websitezugriff oder die Nutzung einer mobilen App beinhalteten. Dabei sollten sie Folgendes kommentieren:

- was sie gerade suchten oder lasen
- was ihnen an der Site/App gefiel und was nicht
- was für die Aufgabenbewältigung hilfreich oder hinderlich war

Unsere erste iPad-Studie fand einen Monat nach der Markteinführung der ersten iPad-Generation statt, unsere erste Kindle-Fire-Studie zwei Wochen nach Produkteinführung. Da die Geräte zu diesem Zeitpunkt naturgemäß noch nicht sehr verbreitet waren, mussten wir auf Probanden zurückgreifen, denen zwar die entsprechenden Plattformen (iOS und Android) geläufig waren, die aber noch kein eigenes Tablet besaßen.

Einige der gestellten Aufgaben bezogen sich auf bestimmte Websites oder Apps (beispielsweise: »Schauen Sie auf der Seite *espn.go.com* nach, ob der Fernsehsender ESPN heute Abend irgendwelche NBA-Basketballspiele überträgt.«). Andere enthielten dagegen deutlich weniger Vorgaben (z.B.: »Ermitteln Sie die Herkunft des Wortes »Dollar.«.«), so dass die Probanden zunächst einmal eine oder mehrere passende Websites oder Apps ausfindig machen mussten, bevor sie das eigentliche Problem angehen konnten. Sie mussten also entweder eine geeignete App finden und installieren oder, im anderen Fall, über ihr Handy auf die benötigte Webseite zugreifen.

Tabelle 1.1 zeigt einige Beispielaufgaben. Bei den bedingungslosen Aufgaben konnten die Probanden selbst entscheiden, welche App oder Website sie ans Ziel bringen sollte – bei den bedingungsgebundenen Aufgaben war die zu verwendende App bzw. Website dagegen vorgegeben.

An unserer Handy-Usability-Studie nahmen 124 Nutzer teil und weitere 35 an der Tablet-Testreihe, insgesamt waren es somit 159 Probanden. Jeder von ihnen benutzte sein Handy mehrmals pro Woche für andere Aktivitäten als das Versenden von Textnachrichten oder Telefongespräche. Nutzer mit technischen Vorkenntnissen und Mitarbeiter der Usability- oder Marketingbranche waren von der Teilnahme ausgeschlossen, weil sie nicht zur Zielgruppe der von uns getesteten Websites gehörten und aufgrund ihrer Vorkenntnisse häufig ein atypisches Verhalten zeigten.

Bei den meisten Tests wurde die Interaktion der Teilnehmer mit ihren Handys mittels einer Dokumentationskamera aufgezeichnet. In einigen wenigen Testsitzungen verwendeten wir allerdings eine mobile Gerätekamera, die direkt auf dem Handy angebracht war. Beide Kameraausführungen gestatteten den Probanden, das Gerät in der Hand zu halten.

Bedingungsfreie Aufgaben	Bedingungsgebundene Aufgaben
Informieren Sie sich über die Symptome der Schweinegrippe und finden Sie heraus, wie man sich gegen das Virus schützen kann.	Stöbern Sie mithilfe von <i>ye1p.com</i> Kundenbewertungen für das Restaurant »Absinthe« in San Francisco auf.
Schauen Sie sich den lokalen Wetterbericht für heute Abend an.	Sie haben 50 Dollar zur Verfügung, um sich etwas zum Anziehen zu kaufen. Suchen Sie sich mithilfe der <i>JC Penney</i> -App ein Kleidungsstück aus, das Ihnen gefällt.
Sie möchten nach einem Kinobesuch noch eine Kleinigkeit essen und trinken gehen. Finden Sie ein Lokal, in dem auch nach 22 Uhr noch kleine Mahlzeiten serviert werden.	Sie möchten etwas Pasta, Tomatenwürfel und Eis kaufen. Erstellen Sie mithilfe der <i>Coles</i> -App eine entsprechende Einkaufsliste.

Tabelle 1.1: Beispielaufgaben unseres Usability-Tests

Bedingungsfreie Aufgaben	Bedingungsgebundene Aufgaben
Ihre Freundin möchte heute Abend nach 20 Uhr fernsehen. Informieren Sie sich über das abendliche Fernsehprogramm und suchen Sie einen Film aus, der sie interessieren könnte.	Ermitteln Sie mithilfe der <i>AA Stocks</i> -App den aktuellen Börsenwert des Unternehmens »China Mobile«. Wie hat sich der Aktienwert im vergangenen Monat entwickelt?
Schauen Sie sich ein <i>Tom und Jerry</i> -Zeichentrickvideo an.	Informieren Sie sich mithilfe der <i>Flipboard</i> -App für das iPad über die neuesten Nachrichten. Richten Sie die App so ein, dass die Schlagzeilen/Themen angezeigt werden, die Sie interessieren.
Es ist 18 Uhr, und Sie müssen von West Kensington zum Tufnell Park. Sie beschließen, die U-Bahn zu nehmen. Ermitteln Sie die günstigste Verbindung mit den wenigsten Umsteigestationen.	Sie möchten am Aussichtspunkt »Vista Point« ein Foto von der Golden Gate Bridge machen. Ermitteln Sie mithilfe der App <i>Light-Trac</i> den Sonnenstand für morgen um 12 Uhr mittags.

Tabelle 1.1: Beispielaufgaben unseres Usability-Tests (Forts.)

1.3 Qualitative Nutzerforschung

Sollten Sie aus dem »Oldschool-Marketing« kommen, werden Sie die Anzahl von 159 Teilnehmern an unseren Mobile-Usability-Studien vermutlich nicht unbedingt beeindruckend finden. (Natürlich wurden in unseren früheren Testreihen zur Nutzung von WAP-[Wireless Application Protocol-] und einfachen »Prä-iPhone«-Handys darüber hinaus noch deutlich mehr Probanden befragt – diese Studien sollen hier jedoch keine Rolle spielen.)

Ziel unserer Forschungsarbeit ist nicht etwa, eine Vorhersage treffen zu können, ob das iPhone oder die Android-Systeme im nächsten Quartal bei der Generation Y die Nase vorn haben werden. Ebenso wenig sind wir daran interessiert, die Bekanntheit oder Beliebtheit einer bestimmten Marke zu bemessen.

Wir betreiben *Verhaltensforschung*, keine Meinungsforschung. Wir untersuchen spezifische User-Interface-Designs und analysieren, ob die individuellen Designoptionen die Bewältigung der nutzerseitigen Aufgabenstellungen einfacher oder eher schwieriger gestalten. Und diese Art von Forschung lässt sich am besten durch die Verfolgung eines qualitativen Ansatzes bewerkstelligen statt durch die quantitätsorientierte statistische Erhebung von Nutzerdaten.

Nehmen wir zum Vergleich doch einmal die Fragestellung nach der Wahl des richtigen Bodenbelags für Krankenhausflure: Die Beobachtung, dass ältere Patienten mit Gehhilfen Probleme haben, sich über einen uneben beschaffenen Bodenbelag fortzubewegen und mitunter sogar ins Stolpern geraten, lässt zweifellos den Rückschluss zu, dass *unebene Bodenbeläge ein Usability-Problem für ältere Patienten darstellen* – und rechtfertigt somit auch die Empfehlung, Krankenhäuser besser mit glatten Bodenbelägen auszustatten.

Es ist also nicht zwangsläufig nötig, 1.000 ältere Herrschaften beim Spaziergang durch die Flure eines Krankenhauses zu beobachten und mitzuzählen, wie viele von ihnen tatsäch-

lich stolpern – und abzuwarten, bis dem Krankenhaus eine Klage ins Haus flattert, weil eine arme alte Dame gestürzt ist und sich die Hüfte gebrochen hat, empfiehlt sich hier natürlich auch nicht wirklich.

Gesicherte Statistiken, die belegen, dass beispielsweise 13,2% der Patienten aufgrund von Unebenheiten im Bodenbelag stolpern, mögen in einer Präsentation Eindruck machen und ein überzeugendes Argument für das Management sein, in bessere Bodenbeläge zu investieren. Ebenso würde auch eine Schadensersatzklage zweifellos Einfluss auf die Haltung der Verantwortlichen haben. Aber die Empfehlung, einen sichereren Bodenbelag verlegen zu lassen, wäre schon nach einer Begutachtung der Sachlage an einem einzigen Nachmittag gerechtfertigt.

Denn in dem Moment, in dem im realen Leben ein Problem festgestellt wird, herrscht Gewissheit, dass es existiert. Es muss aber nicht zwangsläufig genau bemessen werden.

Zudem lässt sich das Ausmaß eines Problems auch mithilfe einer deutlich geringeren Anzahl von Probanden ermitteln, als für eine präzise Datenerhebung erforderlich wäre. Spielt es wirklich eine Rolle, ob nun 63 oder 65% der Nutzer Schwierigkeiten im Umgang mit einem bestimmten Designelement haben? Wenn ein einfacherer Testaufbau das Ergebnis zutage fördert, dass die »meisten« Nutzer Probleme damit haben, reicht das vollkommen aus, um eben dieses Design zu überdenken bzw. zu vermeiden.

Historischer Exkurs: »The Blue-green Book«

Sein erstes Buch zum Thema Website-Usability veröffentlichte Jakob Nielsen 1999 unter dem Titel »*Designing Web Usability*« (Markt+Technik, 2001). Wegen seines auffälligen blau-grünen Einbands ist dieses Werk vielen Lesern im englischsprachigen Raum auch als »The Blue-green Book« geläufig.

»*Designing Web Usability*« basierte in etwa auf dem gleichen Usability-Testvolumen wie die in diesem Buch beschriebenen Untersuchungen: Es wurden weniger als 200 Nutzer befragt. Diese Anzahl reichte jedoch völlig aus, um alle wesentlichen Probleme im Bereich der Web-Usability zu identifizieren – die Verärgerung über lange Wartezeiten bei trägen Downloads, die Abneigung gegenüber Begrüßungsbildschirmen, das größtenteils schnelle und relativ oberflächliche Überfliegen der Websiteinhalte und so weiter.

Einige Befürworter des »cooleren« Designs standen den Aussagen in »*Designing Web Usability*« zunächst einmal skeptisch gegenüber. Heute ist es allerdings eine Tatsache, dass alle Unternehmen, deren Geschäftstätigkeit sich auf die Vermarktung im Internet konzentriert, den meisten Empfehlungen in diesem frühen Werk Folge leisten.

Selbst wenn das allein natürlich noch keine endgültige Beweiskraft hat, ist es dennoch nicht allzu abwegig, anzunehmen, dass »*Mobile Usability*« (also dieses Buch) ähnliche Wirkung zeigen wird wie seinerzeit »*Designing Web Usability*« (das Buch von 1999). Es wird zweifellos wieder einige Webdesigner geben, die den hier ausgeführten – ihren persönlichen Präferenzen widersprechenden – Erkenntnissen anfangs ablehnend gegenüberstehen werden, aber schon in ein paar Jahren werden die Ergebnisse dieser Usability-Studien als Allgemeinwissen gelten. Somit bietet Ihnen die Lektüre dieses Buches die Chance, Ihrer Zeit voraus zu sein und sich schon jetzt ganz selbstverständlich an unseren Empfehlungen zu orientieren, bevor es auch alle anderen tun.



(»*Designing Web Usability*« war zu der damaligen Zeit ein brillantes Buch – wenn wir das einmal selbst behaupten dürfen. Aber es stammt eben aus dem Jahr 1999, als das Internet noch in den Kinderschuhen steckte. Mehr zu unseren neueren Forschungsergebnissen erfahren Sie in unseren aktuelleren Büchern »*Web Usability*« (Co-Autorin: Hoa Loranger; Addison-Wesley, 2006) und »*Eyetracking Web Usability*« (Co-Autorin: Kara Pernice; New Riders Publishing, 2010), die auf Forschungsreihen mit mehreren Tausend Testnutzern basieren.)

Betrachten wir doch noch ein weiteres Beispiel. Angenommen, auf einer Website soll ein Artikel über das H5N1-Grippevirus unter der Überschrift »Vogelgrippe in Honkong ausgebrochen« veröffentlicht werden. Normalerweise würde der Redakteur den Tippfehler in dem Wort »Vogelgrippe« bemerken und daher eine Korrektur in »Vogelgrippe« empfehlen. Das bedeutet: Selbst wenn nur eine einzige Person den Text noch mal Korrektur liest, reicht dies aus, um den Fehler festzustellen und seine Behebung zu veranlassen – es ist also nicht nötig, erst 1.000 Redakteure oder Korrekturleser zu befragen, um festzustellen, wie viele von ihnen der Meinung sind, dass »Vogelgrippe« mit einem »o« geschrieben wird und nicht mit einem »u«.

Zugegeben, es gibt auch Usability-Probleme, die derartig subtil sind, dass sie bei einmaligem Hinsehen nicht gleich auffallen. Wenn man allerdings 159 Testpersonen in vier Kontinenten beobachtet, genügt dies durchaus, um größere Probleme aufzuspüren. Im Rahmen dieses Buches werden wir lediglich die schwerwiegendsten Problematiken erörtern, für die unser qualitativer Forschungsansatz mehr als ausreicht, um die Kriterien für gutes oder schlechtes Design zu ergründen.

Wir forschen also nicht nach geringfügigen Designproblemen, die lediglich 1% des Geschäftswerts einer mobilen Website oder App ausmachen – wir befinden uns auf Großwildjagd – und haben daher die großen Kaliber eingepackt!

Natürlich kann ein Geschäftswert von 1% für große Unternehmen schnell gleich Hunderttausende von Dollar bedeuten, so dass diese durchaus in weiterführende Untersuchungen investieren sollten, um auch die weniger gravierenden Usability-Probleme in ihren Designs zu lösen. Trotzdem ist auch das größte Unternehmen gut beraten, mit kleineren qualitativen Studien selbst in die Mobile-Usability-Forschung einzusteigen – in deren Verlauf zweifellos sehr schnell einige größere Usability-Problematiken zutage befördert werden, derer man sich annehmen sollte, bevor es an das Finetuning der Benutzeroberfläche geht.

Und hier noch zwei letzte Hinweise für alle, die der qualitativen Nutzerforschung immer noch skeptisch gegenüberstehen:

- Die empirische Beweisführung zahlloser Projekte der letzten 23 Jahre zeigt, dass Studien mit einer Handvoll Nutzern absolut ausreichend sind, um den Großteil der wesentlichen Usability-Probleme in einem Design aufzuspüren. Dabei wird sicher nicht jede Schwachstelle erfasst, die maßgeblichen Mängel werden aber deutlich. Für dieses Buch war es nicht wichtig, jedes noch so kleine Usability-Problem auf den getesteten Websites und Apps zu finden. Ausschlaggebender war hier der Abgleich der wichtigsten Testergebnisse mit verschiedenen Designs, um eine genauere Übersicht zu erhalten und allgemeingültige Usability-Richtlinien aufstellen zu können, die jeder beherzigen sollte.

- Fragen Sie sich doch mal, in welcher Relation unsere 159 Studienteilnehmer zu der Anzahl der Endverbraucher stehen, die von Mitarbeitern Ihres Unternehmens persönlich in Einzelsitzungen bei der Bewältigung von Testaufgaben beobachtet wurden. Für die meisten Unternehmen gilt, dass die Anzahl unserer Probanden im direkten Vergleich höher liegt. Und insofern kommt den aus unserer Forschungsarbeit gewonnenen Erkenntnissen im Verhältnis zu den empirischen Erhebungen der meisten Designteams eine umso größere Bedeutung zu.