

Forschungsbericht

Evaluation von Software zur Organisation sozialer Projekte

Help Here vs. Go Volunteer

Veranstaltung: Usability und Softwareergonomie

vorgelegt von:

Sarah Goerlitz

3. Semester, M. Sc. Psychologie, 1931319

Florian Oefner

3. Semester, B. Sc. MCS, 2113150

Margarete Rheindorf

3. Semester, B. Sc. MCS, 2104965

Susanne Zednik

3. Semester, B. Sc. MCS, 2115283

Würzburg, den 28. Februar 2017

Betreuer/Prüfer:

Diana Löffler, Psychologische Ergonomie, Universität Würzburg

Inhaltsverzeichnis

Einleitung	3
Analytische Evaluation	4
Durchführung	4
Ergebnisse	4
Help Here	4
GoVolunteer	5
Diskussion	5
Empirische Evaluation	6
Versuchsaufbau	6
Hypothese	6
Beschreibung der Versuchspersonen	6
Erhobene Usability-Kriterien	6
Use-Cases	7
Dokumentation der Testaufgaben	7
Eckdaten des Tests	7
Versuchsdurchführung	8
Ergebnisse	8
Effektivität	8
Effizienz	9
Zufriedenstellung	11
Diskussion	11
Interpretation der Ergebnisse	12
Effektivität	12
Effizienz	12
Zufriedenheit	12
Störvariablen	13
weiterführende Fragestellungen	14

Einleitung

Seit der großen Flüchtlingswelle 2015 rückte das Thema soziales Engagement in Deutschland immer mehr in den Vordergrund. Allein 2015 wurden rund 480.000 Asylanträge in der Bundesrepublik gestellt, 2016 waren es bereits knapp 750.000 (Bundesamt für Migration, 2016). Allein in diesem Jahr kamen ca. 890.000 Asylsuchende in die Bundesrepublik. 2016 waren es ca. 280.000. Dabei stellten sich viele die Frage, wie man diese Menschen unterstützen und bei der Integration helfen kann. In der heutigen Zeit gehört es zum Alltag sich über das Internet zu informieren. Webseiten und Apps für das Smartphone sind nicht mehr wegzudenken. Umso wichtiger ist es auch bei solchen Themen immer auf dem aktuellen Stand zu sein, um das Beste daraus zu machen. Somit war es ein nur ein nötiger Schritt Hilfe für die Flüchtlinge auch über Webseiten oder das Handy zu vermitteln. Damit befassten sich auch unter anderem zwei gemeinnützige Vereine aus Hamburg und Berlin. Der Verein aus Hamburg entwickelte eine App „HelpHere“ und der Berliner Verein eine Webseite „GoVolunteer“. Das Ziel ist es Hilfesuchenden als auch für den Hilfeanbietenden eine Möglichkeit zu geben sich schnell und einfach in Verbindung zu setzen und Anlaufstellen zu bieten. Dies versuchen diese beiden Softwares mittels zwei verschiedener Ansätze sinnvoll umzusetzen. HelpHere möchte durch die App als eine Art mobiles schwarzes Brett fungieren, um sozial Bedürftige, seien es Flüchtlinge, Obdachlose oder alte Menschen, ganz einfach über das Handy Hilfe zu vermitteln. Betreiber des sozialen Schwarzen Brettes ist der gemeinnützige Help Here e.V., der von Tim Mälzer, den Werbeagenturen Jung von Matt und Polycore sowie dem clubkinder e.V. für das Projekt neu gegründet wurde.[2] Die App kann kostenlos heruntergeladen und verwendet werden. Durch die Einstellungsmöglichkeit vieler Sprachen möchten sie möglichst viele Menschen jeder Herkunft einfach und schnell erreichen. GoVolunteer dagegen ist eine kostenfreie Online-Plattform mit bisher über 20.000 Nutzern. Sie will Ehrenamtliche mit Organisationen in der Flüchtlingshilfe zusammenbringen[4]. Es soll erreicht werden, dass immer mehr Menschen freiwillig aktiv werden, indem der Zugang zur Hilfe soweit wie möglich vereinfacht und jedem die Chance geben wird seine eigenen Ideen für Projekte einzubringen und umzusetzen.[5] Das Gründerteam des GoVolunteer e.V. besteht aus Malte Bedürftig, Henryk Seeger, Tim Dettmann und Moritz Vossenbergh. Hier setzen die beiden Produkte GoVolunteer und HelpHere an. Sie versuchen mittels neuer Technologien und Kommunikationswege wie Websites und Apps, den Zugang zu sozialen Projekten und die Vermittlung derer zu erleichtern. Beide legen den Fokus dabei auf ansprechendes Design und intuitive Benutzung, unterscheiden sich jedoch grundlegend in ihrer inhaltliche Herangehensweise. Die App „HelpHere“, aus Hamburg stammend, bezeichnet sich selbst als „digitales schwarzes Brett“ und bietet dementsprechend ein Portal zum Suchen und Anbieten. Neben klassischen Hilfsleistungen, wie z.B. den Sprachkursen oder der Kinderbetreuung, lassen sich mithilfe von HelpHere auch Gegenstände anbieten und suchen. So bietet die App zusätzlich noch eine Tauschbörse und fasst einen weiten Rahmen um das Thema „soziale Unterstützung“.

GoVolunteer hingegen ist ausschließlich als Website gedacht und wurde von den Berliner Gründern als Plattform für verschiedene soziale „Projekte“ geschaffen, beispielsweise die Betreuung von Flüchtlingskindern. Diese Projekte gliedern sich in einzelne Events, an denen ein Nutzer konkret teilnehmen kann, während das Projekt das thematische Rahmenwerk bildet. Dadurch wird sowohl die Möglichkeit eines langfristigen Engagements in einem Projekt als auch der kurzfristigen Mithilfe bei einem Event geboten.

Analytische Evaluation

Verfasst von Susanne Zednik, 2115283

Zur analytische Evaluation der beiden Produkte HelpHere und GoVolunteer eignet sich das Testverfahren nach den zehn Heuristiken von Nielsen. Die Funktionen der App und der Website sind überwiegend gleich, das Hauptziel beider Anbieter liegt im Anlegen und Finden von Projekten, wodurch keine Vielzahl an Use Cases entsteht.

Durchführung

Verfasst von Susanne Zednik, 2115283

Bei einem Vorgehen nach der heuristischen Evaluation erfolgt zuerst eine grobe Einschätzung der Produkte durch unsere Expertengruppe, bestehend aus 51 Personen. Für beide zu evaluierenden Anbieter werden unabhängig voneinander erste negative sowie positive Eindrücke gesammelt. Nach erster Einschätzung und Besprechung unter den Experten ergibt sich daraus eine Tendenz. Anschließend erfolgt die eigentliche Evaluation nach den zehn Heuristiken, welche unabhängig voneinander stattfindet. Jeder Experte nimmt sich für HelpHere gleichermaßen wie für GoVolunteer eine unbegrenzte Zeit um alle Funktionen, die die App und die Webseite bieten, durchzugehen. Während dieses Ablaufes notiert sich jedes Gruppenmitglied aufkommende Verstöße gegen die Heuristiken. Zusätzlich werden die Fehlerquellen in Schweregrade eingeteilt, diese reichen von "kein Usability Problem"(0) bis Usability Katastrophe"(4). Nachdem sich alle Experten ausführlich mit den beiden Systemen auseinander gesetzt haben werden die Ergebnisse der heuristischen Evaluation jeweils für HelpHere und GoVolunteer in der Gruppe zusammengetragen, nach dem Schweregrad der Verletzung eines Usability Kriteriums gegliedert und untereinander diskutiert. Bei auftauchenden Doppelnennungen, wurden die Punkte zusammengefasst und gemeinsam zu einem Konsens gefunden. Der letzte Teil der heuristischen Evaluation beinhaltet das Ausarbeiten von Verbesserungsvorschlägen.

Ergebnisse

Im Folgenden werden die Ergebnisse, aufgeteilt in die jeweilig zu untersuchenden Anbieter, dargestellt.

Help Here.

Nach der Evaluation und der Zusammentragung der Ergebnisse jedes Experten ergaben sich für HelpHere 28 Verstöße gegen die Heuristiken nach Nielsen. Hierbei wurden 6 kosmetische Probleme der Stufe 1 gefunden. 17 Probleme in Stufe 2, welche keine dringliche Verbesserung bedürfen. Verbesserungsdringlichere Usabilityprobleme wurden in Kategorie 3 eingeteilt, mit einer Anzahl von 5 Verstößen. Die App HelpHere wies keine Usability Katastrophe, der Stufe 4 auf. Die Übereinstimmung unter den Experten betrug 76%. Sehr auffällig war die häufige Verletzung der zweiten und fünften Heuristik (?Übereinstimmung zwischen System und realer Welt? und ?Fehlervermeidung?).

GoVolunteer.

Für Go Volunteer ergaben sich ebenfalls insgesamt 30 Fehler nach der heuristischen Evaluation. Hier erschloss sich eine interrate Übereinstimmung von 63%. Die überwiegend verletzten Heuristiken waren ?Sichtbarkeit des Systemstatus? (1), Flexibilität und effiziente Nutzung (7) und Unterstützung beim Erkennen, Verstehen und Bearbeiten von Fehlern. Go Volunteer zeigte eine Usability Katastrophe (Stufe 4) auf. In Stufe 1 wurden acht Verstöße gefunden, als kleine Usability Probleme wurden 17 eingestuft und als größere Usabilityprobleme drei.

Diskussion

Aufgrund des modernen Designs der Webseite GoVolunteer haben Nutzer den Eindruck, dass diese Seite einfach und intuitiv zu Bedienen ist. In unserer Expertengruppe ergab sich ebenfalls nach erster Einschätzung beider Produkte die Erwartung, dass GoVolunteer bei der analytischen Evaluation besser abschneiden wird als HelpHere. Entgegen unserer Annahme tauchten bei HelpHere weniger Probleme der höheren Kategorien auf, wobei zu beachten ist, dass sich beide Anbieter in der heuristischen Evaluation nicht entscheidend voneinander unterscheiden. Bei beiden Anbietern ergab sich eine gleiche Anzahl an zu behebenden Fehlerquellen, wobei die Usability Katastrophe nur bei GoVolunteer auftrat. Bei diesem schwerwiegenden Problem, welches laut Kategorie 4 mit größter Dringlichkeit behoben werden sollte, handelte es sich um einen fehlenden Lösch-Befehls bei der Erstellung von Projekten.

Die Webseite GoVolunteer bot zwar mit eingebundenen Videos und Bildern ein dynamisches und übersichtliches Interface auf, konnte aber durch zu größten Teilen mangelndes Feedback an den Nutzer nur bedingt ?genossen? werden. Jeweilige schwerwiegende Usability Probleme sind von Entwicklungstechnischer Sicht leicht umgänglich gewesen, weshalb es in der Expertengruppe zum Teil zu großem Unverständnis kam. Dadurch wurden einzelne Kriterien möglicherweise unbewusst strenger gehandhabt als bei Problemen der zu vergleichenden App HelpHere. Diese Software wies ein eher spartanisch gehaltenes Interface auf, mit teilweise mangelnden Informationen. Nach der analytischen Methode jedoch kamen weniger Probleme bei der Interaktion mit dem System zustande. Ein Erlebnisgefühl wie bei GoVolunteer kam hierbei gar nicht erst auf, dafür war es benutzerfreundlicher, bezogen auf eine heuristische Analyse.

Design spielt in der Usability Bewertung jedoch eine große Rolle, sodass wir trotz widerlegter Hypothese weiterhin davon ausgingen, dass GoVolunteer in der Empirie besser abschneiden wird als HelpHere.

Empirische Evaluation

Versuchsaufbau

verfasst von Margarete Rheindorf, 2104965

Hypothese.

Es werden folgende empirische Hypothesen im Hinblick auf die Usability untersucht. Wenn GoVolunteer eine höhere Effektivität als HelpHere hat, dann messen wir eine größere Prozentzahl gelöster Aufgaben und zählen weniger abgebrochene Aufgaben. Wenn GoVolunteer eine höhere Effizienz als HelpHere hat, dann messen wir eine kürzere Zeit pro Aufgabe sowie eine geringere mentale Belastung. Wenn GoVolunteer eine höhere Zufriedenheit als HelpHere hat, dann messen wir einen höheren QUESI-Score und eine Mehrheit der Präferenzurteile für GoVolunteer.

Beschreibung der Versuchspersonen.

Für die Studie wurden 26 Personen eingeladen. Das Alter der Versuchspersonen betrug zwischen 18 und 47 Jahren, wobei 13 Personen weiblich und 13 Personen männlich waren. 17 der 26 Versuchsteilnehmer haben sich in den letzten zwei Jahren sozial engagiert. Doch nur vier Probanden konnten Vorerfahrungen mit Webseiten, die soziale Projekte vermitteln aufweisen. Die Probanden waren sowohl Studenten, als auch berufstätige Personen aus dem Freundes- und Bekanntenkreis. Um einen Reihenfolgeeffekt zu vermeiden wurde die Reihenfolge der Systeme randomisiert durchgeführt.

Erhobene Usability-Kriterien.

In der vorliegenden empirischen Evaluation wurden die Usability-Kriterien Effektivität, Effizienz und Zufriedenstellung gemäß DIN EN ISO 9241-11 erhoben. Effektivität beschreibt dabei die Genauigkeit und Vollständigkeit, mit der Benutzer ein bestimmtes Ziel erreichen. (ISO 9241-11). Diese wurde bei der hier vorliegenden empirischen Evaluation durch die Anzahl gelöster Aufgaben [%] und die Anzahl der Abbrüche operationalisiert. Effizienz beschreibt den im Verhältnis zur Genauigkeit und Vollständigkeit eingesetzten Aufwand, mit dem Benutzer ein bestimmtes Ziel erreichen. (ISO 9241-11). Diese wurde durch die Bearbeitungsdauer [s] und die mentale Beanspruchung bei der Bearbeitung der einzelnen Aufgaben durch die Skala zur Erfassung subjektiv erlebter Anstrengung (SEA) operationalisiert. Die SEA-Skala umfasst einen Wertebereich von 0 bis 110, die Skala ist in Zehnerschritten unterteilt und mit verbalen Ankerpunkten versehen. Die Probanden bewerteten ihre subjektiv erlebte mentale Beanspruchung, indem sie einen Wert auf der Skala nannten. Das Kriterium Zufriedenstellung beschreibt die Freiheit von Beeinträchtigung und positive Einstellung gegenüber der Nutzung des Produkts. (ISO 9241-11). Es wurde mittels des QUESI Fragebogens (Naumann & Hurtienne 2010), sowie durch die Anzahl der Präferenzurteile im Anschluss an die Bearbeitung aller Aufgaben gemessen. Der QUESI-Fragebogen (Questionnaire for Intuitive Use) (Naumann, Hurtienne, 2010) diente im Anschluss an das jeweilige Produkt zur Erfassung der Zufriedenheit der Probanden. Der Fragebogen legt den Schwerpunkt auf persönlich wahrgenommene Intuitivität. Durch verschiedene Fragen werden 5 Subskalenwerte erfasst. Diese messen die wahrgenommene kognitive Beanspruchung, Zielerreichung, Fehlerrate, Lernaufwand

und Vertrautheit mit dem System. Durch Mittelung der 5 Werte erhält man den abschließenden QUESI-Score. Dieser kann einen Wert von 1 (kaum intuitives System) bis 5 (sehr intuitives System) annehmen.

Use-Cases.

Den Versuchspersonen wurden für die beiden Systeme jeweils vier Aufgaben gestellt. Da HelpHere und GoVolunteer leicht unterschiedliche Voraussetzungen aufweisen, alleine schon durch den Standort der Vereine und damit im Zusammenhang die Anzahl der Projektangebote in dem jeweiligen Gebieten, mussten die Use-Cases teilweise entsprechend angepasst werden. Die Use-Cases wurden im Team besprochen und ausgewählt mit dem Ziel die wichtigsten und häufigsten Anwendungen der beiden Systeme zu untersuchen. Voraussetzung war, dass sie von beiden Systeme angeboten wurden, um einen entsprechenden Vergleich über die Softwares ziehen zu können.

1. Projektsuche in einem bestimmten Gebiet zu einer vorbestimmten Thematik
2. Projekterstellung nach einer zu Verfügung gestellten Vorlage
3. Informationssammlung über Organisatoren und Gründer der Softwares
4. Sprachwechsel der Systeme

Dokumentation der Testaufgaben.

Für den Versuchsablauf ebenso wichtig ist die schnelle, übersichtliche und fehlerfreie Aufnahme der Daten. Um dies zu gewährleisten wurde ein Protokollsheet für die jeweiligen Versuche verfasst.

Zu Beginn konnten die grundlegenden Daten wie Datum, Versuchsleiter, Versuchspersonencode notiert werden und welches Betriebssystem verwendet wurde, sowie die Start- und am Schluss die Endzeit. Anhand des Sheets konnte der Leiter die Aufgabenstellung und die Aufgabenlösung entnehmen. Anschließend an jede Aufgabe waren durch Ankreuzfragen und leere Felder gekennzeichnet, wo die Korrektheit der Aufgabenlösung, Use Case-Abbrüche, Werte der SEA-Skala und Zeit eingetragen werden mussten. Zudem wurden Leerzeilen zu Verfügung gestellt um wichtige Aussagen der Versuchspersonen, wie der erste Eindruck, Reflektion der Aufgabenbewältigung, Fazit und Produktpräferenz dokumentieren zu können.

Um eine höhere Kontrolle sowohl quantitativ als auch qualitativ gewährleisten zu können, wurden die Versuche zusätzlich mit der Software „Lookback“, zur Aufzeichnung von Displays, Personen und Ton überwacht. Somit konnten auch noch im Nachhinein an die Tests Daten ermittelt bzw. nochmals kontrolliert werden.

Eckdaten des Tests.

Der Versuch wurde mit insgesamt 26 Personen durchgeführt. 16 Personen ortsgebunden mit zwei Versuchsleitern an der Universität in Würzburg am Hubland Nord, Gebäude 52. Der Arbeitsplatz für die Versuchspersonen war mit einem Laptop mit kabelloser Maus und Internetverbindung ausgestattet. Zudem wurde ein Smartphone für den App-Test bereitgestellt. Die Versuchsreihe wurde

mit den Betriebssystemen Mac OS X und iOS durchgeführt. Im Anschluss an den Versuch /durfte sich der Proband zur Belohnung konnte sich der Proband an den angebotenen Süßigkeiten bedienen. Die zehn weiteren Versuche wurden von einer Person ortsunabhängig durchgeführt. Doch mit denselben Voraussetzungen eines ruhigen Arbeitsplatzes mit Laptop und Smartphone.

Versuchsdurchführung.

Eine Versuchsdurchführung wurde jedes Mal mit der entsprechenden Vorbereitung der benötigten Versuchsmaterialien, d.h. Geräten, Fragebögen und Protokolle gestartet. Ist der Proband eingetroffen, wurde dieser kurz in den groben Versuchsplan eingewiesen, wobei er im Anschluss eine Datenschutz-, Einverständniserklärung und einen kurzen Präfragebogen zum Ausfüllen vorgelegt bekam.

Der Hauptteil des Versuches startete mit Frage nach dem ersten Eindruck und den Erwartungen die der Proband an die Software hatte. Darauf folgte eine genauere Erläuterung des Ablaufes der Use Cases. Die Aufgaben sollten ohne Zeitdruck, doch auch ohne Beantwortung von eventuell auftretenden Fragen vom Probanden durchgeführt werden. Sollte dieser der Meinung sein er habe sein aufgetragenes Ziel erreicht, sollte er dies dem Versuchsleiter deutlich mitteilen. Im Anschluss an jeden Use Case erfolgte ein kurzer Rückblick des Probanden sowie eine Einschätzung der kognitiven Belastung anhand der SEA-Skala. Nachdem alle Use Cases zu einem System abgeschlossen wurden, folgte ein kurzes mündliches Resümee zur ersten Einschätzung des Produkts und der Versuchsteilnehmer füllte den QUESTI-Fragebogen aus. Dieser Vorgang wurde mit dem zweiten System wiederholt.

Nach der separaten Beurteilung folgte eine Befragung zur Präferenz zu einer Software in Verbindungen mit möglichen besonderen Anmerkungen und Verbesserungsvorschlägen. Abschließend wurde sich bei den Probanden für die Teilnahme bedankt, Süßigkeiten angeboten und letztendlich verabschiedet. Es folgte die Zusammentragung und Auswertung der ermittelten Daten.

Ergebnisse

verfasst von Sarah-Maria Goerlitz 1931319

Im Folgenden werden die Ergebnisse der empirischen Evaluation getrennt nach den drei Usability-Kriterien Effektivität, Effizienz und Zufriedenstellung berichtet. Zur Auswertung wurden jeweils zweiseitige t-Tests für abhängige Stichproben verwendet und als Effektstärke-Maß Cohens d berechnet. Das Signifikanzniveau wurde auf 5 % festgelegt.

Effektivität.

Zur Operationalisierung des Usability-Kriteriums Effektivität wurde der Anteil gelöster Aufgaben [%], sowie die Anzahl der Abbrüche herangezogen. Eine Aufgabe wurde dann als gelöst angesehen, wenn das vorher beschriebene Ziel z.B. „Finden Sie ein Projekt bei dem Sie mit alten Menschen arbeiten können“ erreicht wurde. Als nicht richtig gelöst galten Aufgaben, bei denen der Zielzustand nicht der Aufgabe entsprach (z.B. ein anderes Projekt ausgewählt wurde) oder die Aufgabe von der Versuchsperson abgebrochen wurde, weil sie diese nicht lösen konnte. Insgesamt konnten pro System vier Aufgaben richtig gelöst werden.

Bei der Benutzung von Go Volunteer wurden im Mittel 90% der Aufgaben richtig gelöst, bei der

Benutzung von Help Here 94% (GV: $m = 90.38$, $sd = 15.93$; HH: $m = 94.23$, $sd = 10.74$). Dieser Unterschied ist nicht signifikant, $t(25) = -0.891$, $p = .381$, $d = .237$.

Bei der Verwendung von Go Volunteer wurden insgesamt drei Aufgaben abgebrochen, da die Versuchspersonen zu keiner Lösung kamen. Dabei handelte es sich jeweils um die erste Aufgabe. Bei der Verwendung von Help Here wurden insgesamt fünf Aufgaben abgebrochen. Dabei handelte es sich um die erste (drei Abbrüche), dritte und vierte Aufgabe (jeweils ein Abbruch).

Effizienz.

Das Kriterium Effizienz wurde durch die Bearbeitungsdauer [s] und die mentale Beanspruchung bei der Bearbeitung der einzelnen Aufgaben [SEA-Skala] operationalisiert. Beide Variablen wurde für jede Aufgabe einzeln ausgewertet, da sich die Aufgaben im Umfang und Komplexität stark unterschieden.

Bei **Aufgabe 1** ging es darum ein Projekt einer bestimmten Kategorie über die Suchmaske des Systems zu finden. Bei der Benutzung von Go Volunteer benötigten die Versuchspersonen dazu im Mittel ca. eine Minute und 20 Sekunden, bei der Benutzung von Help Here annähernd zwei Minuten (GV: $m = 81.73$, $sd = 44.88$; HH: $m = 115.62$, $sd = 63.36$), was einen signifikanten Gruppenunterschied darstellt, $t(25) = -2.545$, $p = .017$, $d = .706$. Die Mentale Belastung war dabei bei der Verwendung von Help Here höher, der Effekt wurde gerade nicht signifikant (GV: $m = 39.62$, $sd = 28.12$; HH: $m = 51.62$, $sd = 32.25$), $t(25) = -1.850$, $p = .076$, $d = .513$.

Bei **Aufgabe 2** wurden die Versuchspersonen dazu aufgefordert ein neues Projekt anhand einer Projektvorlage anzulegen. Die Bearbeitungszeit lag bei beiden System bei ca. dreieinhalb Minuten und unterschied sich nicht signifikant zwischen den Gruppen (GV: $m = 202.69$, $sd = 61.61$; HH: $m = 221.96$, $sd = 145.42$), $t(25) = -0.774$, $p = .446$, $d = .215$. Der mentale Aufwand unterschied sich dagegen zwischen den Gruppen, $t(25) = -2.165$, $p = .040$, $d = .600$. Bei der Verwendung von Help Here war die mentale Beanspruchung höher, als bei Go Volunteer (GV: $m = 43.65$, $sd = 28.59$; HH: $m = 55.04$, $sd = 32.81$).

Bei **Aufgabe 3** sollten sich die Versuchspersonen über die Inhaber des jeweiligen Systems informieren. Dafür wurden jeweils ca. 40 Sekunden benötigt, die Bearbeitungszeit unterschied sich nicht signifikant zwischen den Gruppen (GV: $m = 41.77$, $sd = 23.15$; HH: $m = 47.46$, $sd = 30.65$), $t(25) = -0.965$, $p = .344$, $d = .268$. Die mentale Beanspruchung lag bei Go Volunteer im Mittel bei ca. 25 und somit leicht oberhalb des verbalen Ankers „kaum anstrengend“. Bei Help Here lag sie bei 44 und somit leicht oberhalb des verbalen Ankers „etwas anstrengend“ und unterschied sich damit signifikant zwischen den Gruppen (GV: $m = 25.65$, $sd = 18.71$; HH: $m = 44.08$, $sd = 37.29$). Dieser Unterschied ist signifikant, $t(25) = -2.865$, $p = .008$, $d = .795$.

Bei **Aufgabe 4** sollten die Probanden die Sprache des Systems umstellen. Bei der Verwendung von Go Volunteer benötigten die Probanden dafür nur wenige Sekunden, bei Help Here dagegen annähernd 30 Sekunden (GV: $m = 5.27$, $sd = 3.41$; HH: $m = 26.15$, $sd = 29.58$), $t(25) = -3.527$, $p = .002$, $d = .978$. Auch die mentale Belastung unterschied sich signifikant zwischen den Gruppen. Sie lag bei Help Here deutlich höher, als bei Go Volunteer (GV: $m = 10.50$, $sd = 17.73$; HH: $m = 27.96$, $sd = 27.74$), $t(25) = -2.699$, $p = .012$, $d = .749$.

Abbildung 1 zeigt die Bearbeitungszeit, sowie die mentale Belastung für Go Volunteer und Help

Here. Es wird deutlich, dass bei der Verwendung von Help Here in allen Aufgaben tendenziell mehr Zeit benötigt wurde und die mentale Belastung höher war, als bei der Verwendung von Go Volunteer.

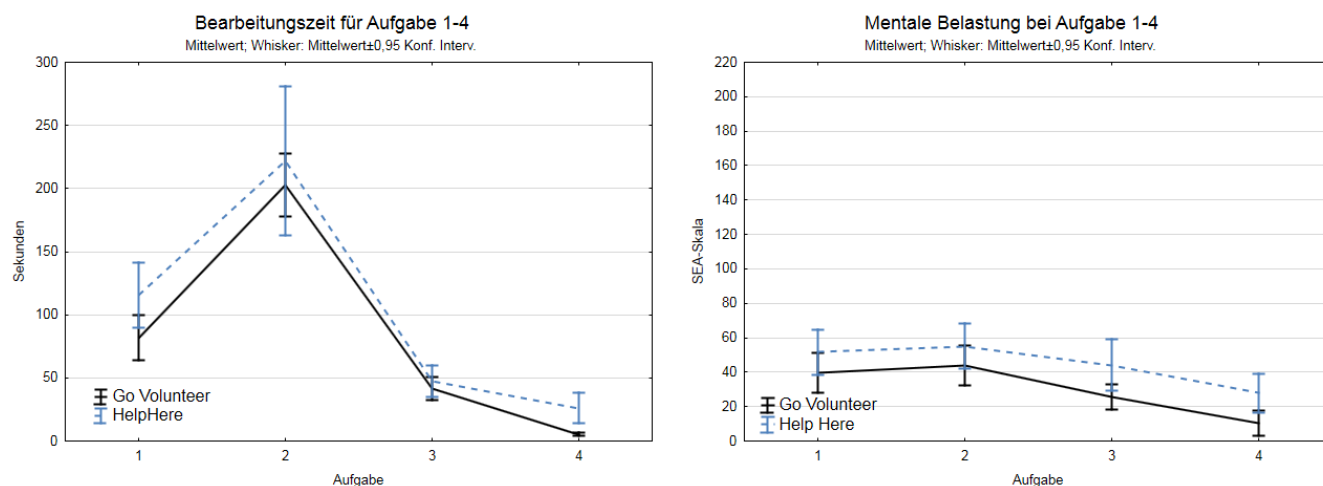


Abbildung 1. Mittlere Bearbeitungszeit [s] und mentale Belastung [SEA-Skala] in den Aufgaben 1- 4 bei der Verwendung von Go Volunteer (durchgezogene schwarze Linie) und Help Here (gestrichelte blaue Linie).

Zufriedenstellung.

Das Kriterium Zufriedenstellung wurde zum Einen durch den QUESI-Fragebogen zur intuitiven Benutzung, zum Anderen durch die Erhebung der Präferenzurteile im Anschluss an die Bearbeitung der vier Aufgaben je System operationalisiert.

QUESI-Fragebogen besteht aus sechs Subskalen je drei Items. Es wurde ein Mittelwert für jede Subskala berechnet, sowie ein Gesamtscore aus den Mittelwerten der Subskalen. Der QUESI-Gesamtscore beträgt bei Go Volunteer 3,99 Punkte ($sd = 0.55$), bei Help Here 3,11 Punkte ($sd = 0.73$). Daraus ergibt sich ein hochsignifikanter Gruppenunterschied, $t(25) = 5.708$, $p = .000$, $d = 1.583$. Abbildung 2 zeigt die Mittelwerte der Subskalen - es fällt auf, dass alle Kategorien gleichermaßen zu dem Effekt beitragen. In allen Subskalen zeigen sich höhere Mittelwerte für Go Volunteer.

Die Auswertung der **Präferenzurteile** zeigt, dass sich die klare Mehrheit der Probanden nach der Verwendung beider Systeme für Go Volunteer entschied. Insgesamt präferierten 22 Versuchspersonen Go Volunteer, vier Versuchspersonen entschieden sich für Help Here.

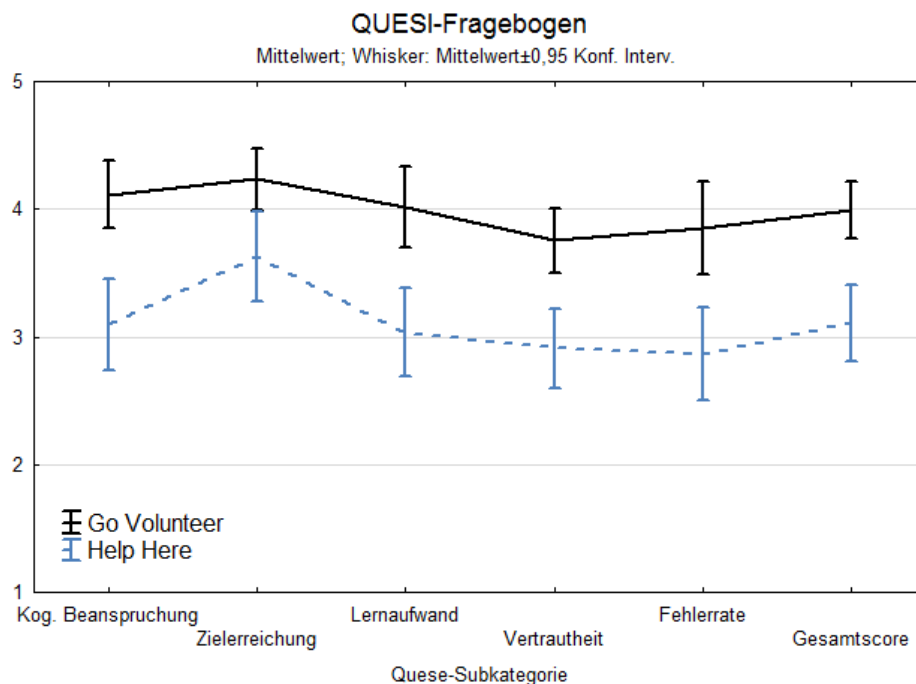


Abbildung 2. Mittlere Punktzahlen der Subskalen des QUESI-Fragebogens. Der Fragebogen wurde jeweils nach der Verwendung von Go Volunteer und Help Here ausgefüllt.

Diskussion

verfasst von Florian Oefner 2113150 Zum Abschluss versuchen wir die Ergebnisse anhand der qualitativen Erkenntnisse zu erklären bzw. interpretieren. Darauf folgt eine kurze Beschreibung der potentiellen Störvariablen zusammen mit unseren Maßnahmen diese unter Kontrolle zu halten. Anschließend wird noch einmal auf noch offene oder neu entstandene Fragen eingegangen.

Interpretation der Ergebnisse

Insgesamt interpretieren wir unsere Ergebnisse als Bestätigung unserer Hypothese und damit als Vorteil zugunsten GoVolunteer. Vor allem in den Maßen der Effizienz und der Zufriedenheit finden sich deutlicher und sogar signifikante Resultate, während bei der Effektivität kein Unterschied zwischen beiden Services zu erkennen war.

Effektivität.

Hier muss noch zitiert werden

Beim Anteil gelöster Aufgaben unterschieden sich beide deskriptiv minimal, jedoch nicht signifikant innerhalb eines t-Tests. Wir führen dies darauf zurück, dass beide ihren Service auf etablierten Designpattern aufgebaut haben, wie z.B. die Navigationsleiste am unteren Ende einer App. Diese Pattern werden teils explizit beschrieben, wie unter anderem in den iOS Human Interface Guidelines (<https://developer.apple.com/ios/human-interface-guidelines/overview/design-principles/>), teils haben sie sich über die Zeit in einer Vielzahl von erfolgreichen Apps bewährt, wie z.B. das Hamburger-Icon (<https://de.wikipedia.org/wiki/Hamburger-Menü>). Für den geringen Unterschied in der Anzahl an Abbrüchen zugunsten von GoVolunteer konnten wir keinen schlüssigen Grund finden, vermuten ihn jedoch darin, dass die Testpersonen längere Zeit an Websites als an Apps gewöhnt sind.

Effizienz.

Betrachtet man das erste Maß der Effizienz, die Zeit pro Aufgabe, liegen die Werte von GoVolunteer stets unter denen von HelpHere und sind bei der ersten und letzten Aufgabe signifikant, ein klares Indiz, dass die gesamte Effizienz zugunsten von GoVolunteer unterscheidet. Wir sehen dies begründet darin, dass der fehlende Platz zur Interaktion bei HelpHere, bedingt durch das Medium Smartphone, nicht angemessen kompensiert wurde. Durch den vielen Platz den die Website GoVolunteer zur Verfügung hat, kann sie nicht nur mehr Funktionen auf einmal darstellen, sondern bleibt übersichtlich und aufgeräumt. Dadurch kann sich der Benutzer leichter zurecht finden und schneller sein Ziel erreichen, was unsere Probanden durch Beschreibungen wie "klar", "aufgeräumt", "einfach" bestätigten. HelpHere hingegen hat es verpasst, den fehlenden Platz durch gute Informationsarchitektur aufzufangen, sodass weiterhin die Übersichtlichkeit gewährleistet wird. Diese Vermutung wird bekräftigt durch die Resultate des zweiten Maßes, der mentale Belastung, welches eher repräsentiert, wie viel Aufwand der Benutzer hatte, die Struktur zu verstehen und zu navigieren. Hier sind alle Vergleiche signifikant für GoVolunteer, ausgenommen der ersten Aufgabe.

Zufriedenheit.

Bei den Ergebnisse der Zufriedenheit, konkret dem QUESI-Score sowie einem Präferenzurteil, finden sich die deutlichsten Ergebnisse. Beim QUESI-Score lässt der signifikante Unterschied darauf schließen, dass neben den Vorteilen in Bereichen der Effizienz GoVolunteer auch intuitiver zu bedienen ist. Ein klares Präferenzurteil von 22 zu vier zugunsten von GoVolun-

teer deutet außerdem daraufhin, dass die Website im Gesamtpaket zufriedenstellender war. Unsere qualitativen Untersuchungen ergaben, dass dies beides vor allem im Unterschied des Mediums, also Website im Gegensatz zu App, verwurzelt ist. So gaben Probanden an “für so etwas keine App nutzen” zu würden, da es “nicht alltägliches” sei. Des Weiteren vermuten wir, dass die Farbgebung eine Rolle spielt, da sich HelpHere größtenteils auf sogenannte “Nichtfarben” (<http://www.figunetik.com/2013/12/24/farben-und-nichtfarben-der-farbkreis-nach-johannes-itten/>) , also Schwarz, Weiß und Grautöne beschränkt. GoVolunteer verwendet hingegen vermehrt ein sanftes Grün (<http://www.webdesign-journal.de/psychologische-wirkung-von-farben-farbkombinationen/Gruen>) , dem eine beruhigende Wirkung nachgesagt wird, und setzt zudem viele Bilder mit Menschen ein.

Störvariablen

Der Versuchsablauf wurde mithilfe eines Wordingsheets und eines Ablaufplans standardisiert, um die Wahrscheinlichkeit von Versuchsleitereffekten auf ein Minimum zu reduzieren. Weiterhin wurden die erste Hälfte der Probanden von zwei Leitern betreut, um zu verhindern, dass aufgrund der vielen ungewohnten Aspekte die Dokumente nicht eingehalten werden. Auch mögliche Störvariablen bedingt durch die Umgebung wie z.B. Lärm traten bei keiner Durchführung in Erscheinung. Das Kriterium, alle Versuche am selben Ort durchzuführen, wurde bewusst aufgeweicht, um der Stichprobe, die weit mehr als nur Studenten aus dem Raum Würzburg umfasst, gerecht zu werden. So konnten wir Menschen höheren Alters aus verschiedenen Bundesländern für die Versuche gewinnen, ohne den logistischen Aufwand erheblich zu erhöhen, indem wir einen Teil der Probanden ortsunabhängig getestet haben. Es wurde stets darauf geachtet, trotz der Ortsunabhängigkeit einen ruhigen Raum, mit Stuhl, Schreibtisch und natürlichem Licht zu wählen. Neben dem Ort der Testung könnte sich das Betriebssystem auf dem getestet wurde als Störvariable ausgewirkt haben. Dies könnte vor allem bei der Testung auf dem Smartphone von Bedeutung gewesen sein. Möglicherweise konnten die Versuchspersonen mit ihrem gewohnten Betriebssystem schneller durch die App navigieren, oder bewerteten diese unbewusst positiver, als Personen, die nicht an das Betriebssystem gewöhnt waren. Um den Einfluss post hoc prüfen zu können wurde abgefragt, welches Betriebssystem die Versuchspersonen normalerweise benutzen. Auch die Reihenfolge der Systeme könnte sich auf Bearbeitungszeit oder die Beantwortung des Quesi-Fragebogens ausgewirkt haben. Die Struktur der Angebote bei einem System könnte eine Erwartungserhaltung für das kommende System geschaffen haben. Um den Einfluss der Störvariablen zu prüfen, haben wir eine multiple Regression mit den drei identifizierten Störvariablen “Ort” (Labor $N = 15$, Feld $N = 11$), “Betriebssystem” (IOS gewöhnt $N = 13$, IOS nicht gewöhnt $N = 13$) und “Reihenfolge” (StartHH $N = 13$; StartGV $N = 13$) als Prädiktoren durch geführt. Als abhängige Variable haben wir zunächst die Differenzen der Gesamt-Bearbeitungszeiten von Go Volunteer und Help Here betrachtet. Es zeigt sich, dass die Störvariablen 11,9% der Varianz aufklären, und damit keinen signifikanten Einfluss haben, $F(2,22) = 0.988$; $p = .417$. Zieht man die Differenz der Quesi-Scores von GoVolunteer und Help Here zur Auswertung heran zeigt sich ein ähnliches Bild. Die Störvariablen klären hier 20,5% der Varianz auf, $F(3,22) = 1.88$, $p = .161$.

weiterführende Fragestellungen

Am Anfang des Experimentes stellten wir die Hypothese auf, GoVolunteer sei besser, rein aufgrund des optischen Eindrucks. Diese hat sich in der analytischen Methode kaum bestätigen lassen, in der empirische Methode überwiegend schon. Ist GoVolunteer nun der bessere Anbieter und wird sich über lange Zeit durchsetzen? Um diese Frage, die versucht den Kern dieses Experimentes zusammenzufassen, zu beantworten, müssen wir tiefer eintauchen. Rein aus dem Blickwinkel der Usability bzw. User Experience, welche als Differenzierungsmerkmal zwischen Produkten bzw. Firmen immer wichtiger wird, würde man argumentieren GoVolunteer setze sich durch. Auch unsere Beobachtungen hinsichtlich der Aktivität auf beiden Plattformen, insbesondere das Erstellen neuer Angebote, untermauern dies. So konnten wir bei GoVolunteer über die Zeit immer mehr und neuere Angebote entdecken, während HelpHere scheinbar stillstand. Außerdem gaben viele unserer Probanden an, dass sie das Medium App für eine solche Thematik als ungeeignet empfinden, HelpHere also selbst mit verbesserter Usability wenig Chancen hätte. Man sollte jedoch nicht den Rahmen vergessen, in dem beide Produkte geschaffen wurden. Während GoVolunteer in Berlin tätig ist, hat HelpHere seinen Hauptsitz in Hamburg, weshalb eine räumliche Aufteilung des Marktes ebenfalls möglich wäre. Das ist vor allem deshalb plausibel, weil die Inhalte eine lokal gebundene Verbindung zur realen Welt haben. So mag GoVolunteer zwar benutzbarer sein, für einen Menschen in Hamburg jedoch nutzlos, also ohne Effektivität, da es keine Angebote bei ihm gibt. Des Weiteren steckt mit der Agentur Jung von Matt Deutschlands erfolgreichste Werbeagentur[<http://www.handelsblatt.com/unternehmen/dienstleister/studie-jung-von-matt-deutschlands-beste-werbeagentur/7685014.html>] hinter HelpHere, die viel Expertise, ebenso viel Investitionskraft und daher Raum für Entwicklung mitbringt. Es kann also gut sein, dass HelpHere bezüglich der Usability schon in näherer Zukunft große Fortschritte macht. Außerdem spielt die Zeit für HelpHere, da die heutigen Generationen mit Smartphones aufwachsen und sehr vertraut im Umgang mit diesen sind. So könnte sich das Blatt wenden und in zehn Jahren das Medium "Website" ungeeignet sein für eine solche Thematik. Gleichzeitig, und das ist wohl die interessanteste Entwicklung, kann mit der Zeit der Bedarf für bzw. den Nutzen von Services wie GoVolunteer oder HelpHere im Allgemeinen sinken oder gar vollständig verschwinden, wenn die Anzahl an Flüchtlingen und der damit verbundene logistische Aufwand zurückgeht. Denn beide Services wurden geboren, um der großen Anzahl an ehrenamtlicher Arbeit, die notwendig wurde, Herr zu werden. Ist diese Arbeit nicht mehr notwendig, stellt sich die Frage, ob sich solche Services aufrecht erhalten können. Denn klassische ehrenamtliche Arbeit in Vereinen organisiert sich schon immer ohne digitale Hilfe.

Abschließend ist festzuhalten, dass der weitere Erfolg beider Services maßgeblich von diesen zwei Fragen abhängt: Wie groß wird der Bedarf für Unterstützung bei sozialen Projekten in Zukunft sein? Wie entwickelt sich das Verhältnis der Nutzer zu den beiden eingesetzten Medien weiter.

Anhang A

*

Abbildungsverzeichnis

- 1 Mittlere Bearbeitungszeit [s] und meltale Belastung [SEA-Skala] in den Aufgaben 1-4 bei der Verwendung von Go Volunteer (durchgezogene schwarze Linie) und Help Here (gestrichelte blaue Linie). 10
- 2 Mittlere Punktzahlen der Subskalen des QUESI-Fragebogens. Der Fragebogen wurde jeweils nach der Verwendung von Go Volunteer und Help Here ausgefüllt. 11

Anhang B

*

Tabellenverzeichnis