

Usability Engineering

Vorlesung 5: Usability Testing (Teil 2)

VU 183.123

Christoph Wimmer

Übersicht

- Usability Testing:
 - Vorbereitung & Planung
 - Durchführung
 - Ablauf der Test-Sessions
 - Guidelines für Teammitglieder
 - Datenerfassung
 - Auswertung
 - Kommunikation der Ergebnisse

Vorbereitung und Planung

- Umfang festlegen
- Testziel festlegen
- Metriken und gesammelte Daten festlegen
- Zeitplan erstellen
- Taskliste und Szenarien ausarbeiten
- Benutzerprofil definieren und Testpersonen rekrutieren
- Testraum und Equipment vorbereiten
- Testablauf planen und Test-Team einweisen

Testmaterial (1)

- Orientierungsskript
 - Vorstellung
 - Zweck des Tests
 - Hinweis, dass das Produkt und nicht die Testperson getestet wird
 - Hinweis, dass das Produkt noch fehlerhaft sein kann
 - Auf Video- und Audioaufnahmen hinweisen
 - Testperson aufklären, dass sie jederzeit aufhören oder Fragen stellen kann
- Background Fragebogen
 - Demographische Daten
 - Computererfahrung, ...
- Pre-Test Fragebogen
 - Eindrücke vor dem eigentlichen Test
 - Erfahrung mit diesem oder einem ähnlichen Produkt

Testmaterial (2)

- Task-Liste
 - Realistische Szenarien
 - Jede Aufgabe auf einem eigenen Blatt
 - Beschreibung des Ziels, nicht der einzelnen Schritte
- Trainingsmaterial (optional)
 - Walk-through von Beispielaufgaben
 - Demonstration der Oberfläche (GUI)
 - Demonstration der Interaktionsmöglichkeiten
- Non-Disclosure Agreement
- Einverständniserklärung zur Aufzeichnung
- Logging Sheets

Testmaterial (3)

- Post-Test Fragebogen
 - Wording
 - Navigation
 - Angezeigte Informationen ausreichend
 - Allgemeiner Eindruck
 - Standardisierte Fragebögen (SUS, CSUQ, ...)
- Debriefing
 - Themen, die nach dem Test noch geklärt werden sollen
 - Strukturiertes Interview
- Checkliste
 - Chronologische Liste des Testablaufs
 - Immer selber Ablauf

Durchführung der Test Sessions

Durchführung

1. Begrüßung

Der erste Kontakt mit der Testperson sollte in einer angenehmen Umgebung stattfinden und nicht gleich im Testlabor. Die Testperson soll sich nicht unwohl fühlen oder nervös sein.

2. Ausfüllen von einleitenden Dokumenten

- Non-disclosure Agreement
- Erlaubnis zur Aufzeichnung von Daten
- Background Fragebogen

3. Orientierungsskript vorlesen

Dies dient der Erklärung des Testablaufs und der Betonung von Methoden (z.B. Thinking-Aloud), die verwendet werden. Wichtiger Hinweis an dieser Stelle: Es wird das Produkt, nicht die Testperson getestet!

Durchführung

4. Ausfüllen von Pre-Test Fragebögen

Wenn sinnvoll, ggf. mit Background Fragebogen kombinierbar

5. Vorbereitung im Usability Labor

- Vorstellung der BeobachterInnen
- Erklärung der Testeinrichtung

6. Vorbereitendes Training

Training kann notwendig sein, um die Testperson mit dem Produkt vertraut zu machen.

Durchführung

7. Ausgeben der Task Szenarien

Die Tasks sollten einzeln, schriftlich und sequentiell ausgegeben werden. Die Testperson soll die Möglichkeit haben jederzeit die Aufgabenstellung nachzulesen.

8. Startzeit aufnehmen, Testperson beobachten und kritische Daten aufnehmen

Alle Instrumente zur Datensammlung sollten nun gestartet sein!

Durchführung

9. Ausfüllen der Post-Test Fragebögen

Vor jeglicher Diskussion über den Test soll der Fragebogen ausgefüllt werden. Damit wird eine Beeinflussung der Testperson vermieden.

10. Debriefing

Fragen, die während des Tests aufgetaucht sind, können nun behandelt werden.

11. Bedanken und Verabschieden

12. Daten und Logging Sheets sammeln

Alle Daten, Formulare und Kommentare, auch von den BeobachterInnen, sollten gesammelt in einem Ordner abgelegt werden.

Thinking-Aloud

- Testpersonen sprechen bei der Erledigung von Aufgaben ihre Überlegungen und Gedanken laut aus
- Concurrent: Die Testperson denkt direkt bei der Durchführung der Tasks laut mit
- Retrospective: Nach dem Task wird mit der Testperson der Task noch einmal durchgegangen (z.B. anhand eines Screenrecordings)
- Vorteile
 - Auffinden von vielen Usability Problemen
 - Man erfährt warum diese Probleme auftreten
 - Überlegungen der Testperson werden beobachtbar
- Nachteile
 - Ungewohnt für die Testperson
 - Testpersonen denken eventuell intensiver über die Aufgabe nach (bei Concurrent)
 - Die Arbeitsgeschwindigkeit wird möglicherweise herabgesetzt (bei Concurrent)
 - Höherer Zeitaufwand (bei Retrospective)

Thinking-Aloud

- Interessant: Was im Kopf der Testperson vorgeht und wie sie Probleme lösen
 - Ich will ... machen.
 - Ich glaube ... passiert wenn ich hier draufdrücke.
 - Das ist nicht was ich erwartet habe, ich dachte ...
 - Das hat länger gedauert als erwartet.
- Weniger interessant: Subjektive Meinungen, Designvorschläge, Spekulation
 - Ich finde der Button sollte rot sein.
 - Ich glaube nicht, dass andere Benutzer das so machen würden.

Co-Discovery

- Variante der Thinking-Aloud Methode
- Zwei Testpersonen führen den Test gemeinsam durch
- Thinking-Aloud wird durch Kommunikation ersetzt

- Vorteile
 - Natürliche Gesprächssituation (Dialog)
 - Geringe Hemmschwelle (im Gegensatz zum Thinking-Aloud)
- Nachteile
 - Man benötigt doppelt so viele Testpersonen
 - Nicht so valide Ergebnisse, da die Zielgruppe das Interface (in der Regel) nicht zu zweit benutzt

Tipps für die Durchführung

- Eine angenehme Atmosphäre schaffen
- Die TeilnehmerInnen mit Respekt behandeln
- Neutral bleiben
- Den TeilnehmerInnen nicht sofort helfen
- Entscheiden, wie viel Hilfestellungen gegeben werden soll
- Gute Aufzeichnungen führen
- Aussagen der TeilnehmerInnen in deren Worten aufzeichnen
- Je besser die Aufzeichnungen, desto einfacher die Analyse

Guidelines für Teammitglieder

Rollen der Teammitglieder

- TestleiterIn
- ModeratorIn
- ProtokollantIn
- ZeitnehmerIn
- Video / Audio OperatorIn
- Produktspezialist / technischer SpezialistIn
- Zusätzliche Testrollen
- TestbeobachterInnen

Interaktion mit der Testperson

- Da jede Interaktion mit der Testperson den Test stört und ggf. das Ergebnis verfälscht, sollte man sich über folgendes klar sein:
 - Wer spricht während des Tests mit der Testperson?
 - Wann und wie sollten die Testpersonen unterbrochen werden?
 - Wie vermeidet man eine Einflussnahme auf die Testpersonen?
 - Wer entscheidet, wann eine Testperson eine Aufgabe vollendet hat?
- Daumenregel: So wenig Interaktion wie möglich, so viel wie nötig

Ansprechperson(en)

- ModeratorIn ist der/die einzige AnsprechpartnerIn für die Testperson
- Im Labor kann, wenn vorhanden, eine Gegensprechanlage zur Kommunikation genutzt werden
- Ausnahme: Wenn es weitere Testrollen gibt, so können diese auch mit der Testperson sprechen
 - Beispielsweise: Help Desk, Kunde, ...

Testpersonen unterbrechen

- Die Testperson soll die Aufgaben alleine lösen
- Die Testpersonen bestimmen wie und in welcher Zeit sie die Aufgabe lösen
- Mögliche Gründe für Unterbrechungen:
 - Erinnerung an „Thinking-Aloud“
 - Wenn die Testpersonen keine Pause zwischen den Aufgaben machen oder zu schnell arbeiten
 - Technische Probleme
- Bei Tests mit frühen Prototypen kann man nach alternativen Lösungswegen fragen

Entscheidung über das Ende der Aufgabe

- Testpersonen entscheiden, wann eine Aufgabe erfüllt ist
- Es gibt zwei mögliche Situationen, die bedacht werden müssen:
 - Die Testperson hat alle Schritte durchgeführt, ist sich darüber aber nicht im Klaren
 - Beispielsweise durch fehlende Rückmeldung
 - Die Testperson sagt, dass sie die Aufgabe gelöst hat, obwohl dies nicht der Fall ist
 - In diesem Fall sind drei Dinge zu beachten:
 - Man möchte vergleichbare Tests haben ⇒ weiter machen
 - Man möchte so viel wie möglich lernen ⇒ weiter machen
 - Man hat schon genug gesehen ⇒ stoppen

Einflussnahme vermeiden

- Mögliche Quellen der Beeinflussung der Testpersonen:
 - Fragen, die man den Testpersonen stellt
 - Antworten auf Fragen und Anmerkungen der Testpersonen
- Fragen: sollten neutral formuliert sein
 - Nicht „War die Navigation verständlich?“
 - Sondern „Wie fanden Sie die Navigation?“ oder „Fanden Sie die Navigation verständlich oder verwirrend?“
- Antworten: sollten keine Hinweise geben
 - Fragen indirekt oder durch Gegenfragen beantworten
 - Frage: „Habe ich das richtig gemacht?“
 - Antwort: „Glauben Sie, dass Sie es richtig gemacht haben?“

Anforderungen an Protokollanten

- Ein/e ProtokollantIn muss
 - das System gut kennen
 - die Ziele der Testaufgaben kennen
 - zu 100% seine Aufmerksamkeit auf die Testperson richten
 - aktiv beobachten, d.h. an alle möglichen Aktionen der Testpersonen denken und die passenden Rückschlüsse daraus ziehen
 - objektive Beobachtungen von subjektiven Meinungen trennen
- Ein/e ProtokollantIn sollte nicht:
 - das Produkt in Gedanken verteidigen oder Probleme weglassen, die aus seiner/ihrer Sicht nicht lösbar sind

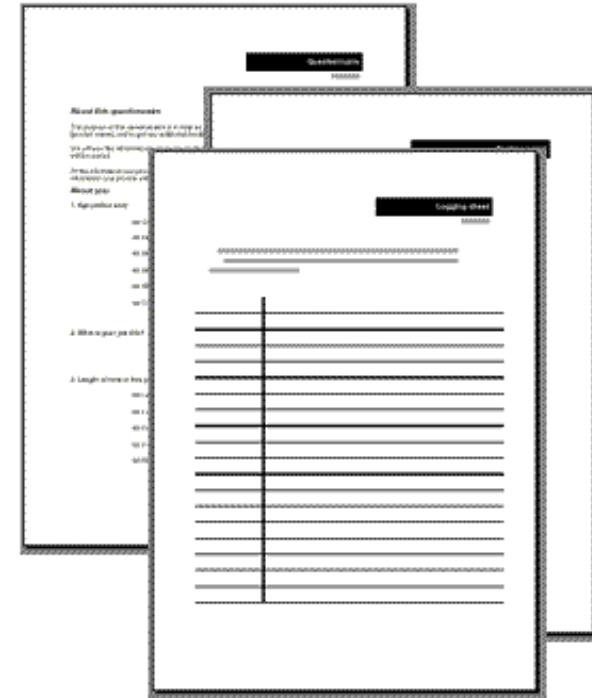
Datenerfassung

Vor Ort vs. Aufzeichnungen

- Beobachtung vor Ort:
 - Unmittelbar: Egal wie gut und umfassend die Aufzeichnung, es geht immer etwas verloren
 - Man kann Unklarheiten direkt im Anschluss an den Test mit den Testpersonen diskutieren
 - Man hat nur eine Chance (um nichts zu übersehen)
- Auswertung von Aufzeichnungen
 - Nützlich für Dokumentation und Reporting
 - Um offene Fragen zu klären (wenn man etwas übersieht)
 - Zusätzliche EvaluatorsInnen lassen sich später hinzuziehen
 - Man spart keine Zeit wenn man die Beobachtung auf später verschiebt

Logging Sheets & Notizen

- Anfangs- / Stopzeit
- Aussagen der TeilnehmerInnen
- Fehler und Probleme
 - Zeitpunkt
 - Häufigkeit
- Annahmen und Interpretationen zu den Beobachtungen
 - Klar als solche gekennzeichnet!
- Eindeutige Zuordnung zu einem bestimmten Task und Testperson muss möglich sein!
- Empfehlung: Ein eigenes Logging Sheet pro Protokollant, Teilnehmer und Task



Standardisierte Fragebögen

- Usability:
 - SUS - System Usability Scale
<https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>
 - CSUQ – Computer System Usability Questionnaire
<http://hcibib.org/perlman/question.cgi>
 - SUMI - Software Usability Measurement Inventory
<http://sumi.uxp.ie/>
 - QUIS - Questionnaire for User Interaction Satisfaction
<http://www.cs.umd.edu/hcil/quis/>
- Affekt & Emotion (Nützlich für UX)
 - AttrakDiff - <http://attrakdiff.de>
 - I-PANAS-SF - International Positive and Negative Affect Schedule Short Form
<http://jcc.sagepub.com/content/38/2/227.abstract>
- Workload:
 - NASA Task Load Index (TLX)
<https://humansystems.arc.nasa.gov/groups/TLX/>
- ...

Fragebogen: Beispiel SUS

1. I think that I would like to use this system frequently
2. I found the system unnecessarily complex
3. I thought the system was easy to use
4. I think that I would need the support of a technical person to be able to use this system
5. I found the various functions in this system were well integrated
6. I thought there was too much inconsistency in this system
7. I would imagine that most people would learn to use this system very quickly
8. I found the system very cumbersome to use
9. I felt very confident using the system
10. I needed to learn a lot of things before I could get going with this system

Strongly disagree

Strongly agree

1	2	3	4	5

1	2	3	4	5

1	2	3	4	5

1	2	3	4	5

1	2	3	4	5

1	2	3	4	5

1	2	3	4	5

1	2	3	4	5

1	2	3	4	5

1	2	3	4	5

Audioaufnahmen

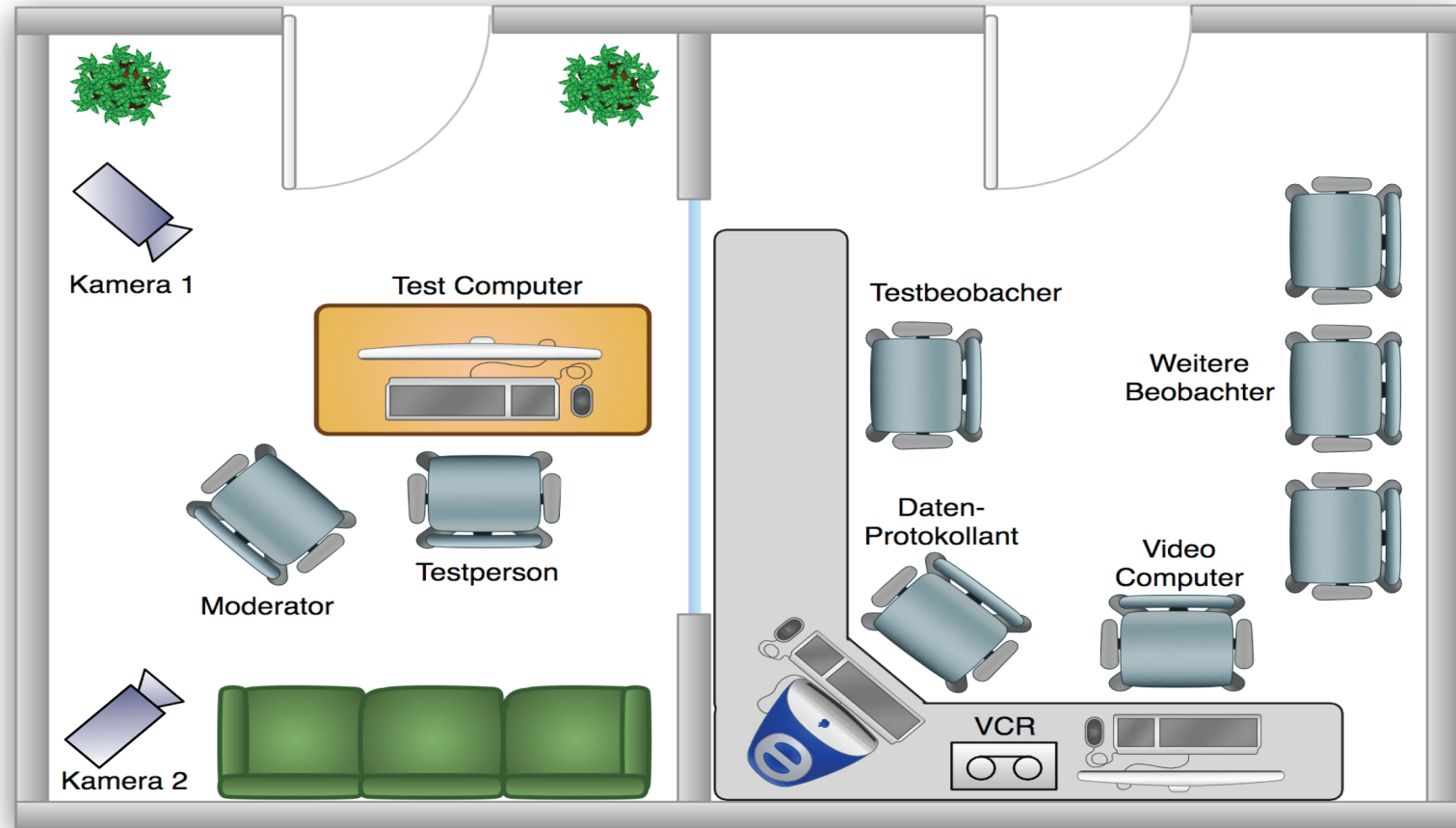
- Sinnvoll bei „Thinking aloud“
- Kann auch unabhängig von Videoaufnahmen zum Einsatz kommen
- Aufzeichnung von:
 - Anmerkungen und Fragen der Testpersonen während der Tests
 - Anmerkungen und Antworten des Moderators
 - Interviews
- Dient zur Arbeitserleichterung der ProtokollantInnen

Screen Recording & Videoaufzeichnung

- Spezialisierte Tools
 - Noldus Observer / Viso, Techsmith Morae (retired), ...
- Vielzahl an Alternativen für Screen Recording:
 - Im Betriebssystem integriert: Quicktime (Mac OS), Game Bar (Windows 10)
- **Tip: Open Broadcaster Software (OBS)**
 - Synchronisation unterschiedlicher Video- und Audioquellen (Webcam, Screen Recording, zusätzliche Kameras, Mikrofone, ...)
 - <https://obsproject.com/>



Testumgebung: Aufbau eines Usability Labors





Mobile Usability Testing

- Herausforderung durch
 - Technische Einschränkungen der Plattform
 - Tools in Funktionalität eingeschränkt oder erfordern direkten Zugriff auf Quellcode (Einbindung von Logging-Frameworks)
 - Plattformen haben inzwischen viel nützliche Funktionalität zur Aufzeichnung integriert (Screenrecording, Screencasting)
 - Ergonomie (kleiner Bildschirm)
 - Screencasting für besseren Überblick
 - Verlust der Kontrolle über die Test-Umgebung im „Feld“
- Trade-off:
Kontrollierte Umgebung vs. realistischer Anwendungskontext

Mobile Usability Testing

- Protokollierung am Device:
 - Teilweise bereits in Betriebssystemen integriert
 - Screen Recording, Screen Sharing
 - Teilweise über Apps möglich
 - z.B. UXCam
- Zusätzliche Hardware:
 - Nicht nur Screen, auch Bedienung
 - Abstriche bei Ergonomie
 - Meist nicht „feld-tauglich“



Quelle: usability.gov

Remote Testing

- Ablauf von moderierten Remote Usability Tests weitgehend analog zu persönlichen Tests, mit dem Unterschied, dass die Testperson nicht direkt vor Ort anwesend sein muss
- Es wird ein Beobachtungskanal benötigt, wie z.B. Tools für Screen Sharing, Teleconferencing, Sprachkommunikation, ...
 - VNC-Lösungen, Zoom, GoToMeeting, Skype, MS Teams, Google Meet, ...
- Herausforderung: Remote Tests benötigen intensivere Vorbereitung und Planung
 - Inbetriebnahme und Überprüfung der Kommunikationstools
 - Ablauf während der Test-Sessions: Weniger Raum für spontane Improvisation
 - z.B.: Wie werden die Tasks an die Testperson ausgegeben? Verbal? Im Chat?

Remote Testing

- Vorteile
 - Kein Test Labor notwendig
 - Flexiblere Organisation (Zeit, Ort)
 - Test in natürlicher Umgebung

- Nachteile
 - Distanz zur Testperson
 - Weniger Überblick, Unmittelbarkeit der Beobachtung geht verloren
 - Weniger Kontrolle über die Testumgebung
 - Höheres Risiko technischer Probleme
 - Schwierigkeit der Behebung technischer Probleme



deco.inso.tuwien.ac.at

