

MMI2 Übung 8: Mobile Evaluation

Prof. Dr. Michael Rohs, Dipl.-Inform. Sven Kratz

michael.rohs@ifi.lmu.de

MHCI Lab, LMU München

Aufgabe 1: Mobile Evaluation

- a) Beschreiben Sie die wesentlichen Unterschiede zwischen heuristischer Evaluation und der Think-Aloud-Methode.

Aufgabe 1: Mobile Evaluation

- b) Welche Aspekte machen die Evaluation mobiler Benutzungsschnittstellen zu einer besonderen Herausforderung? Geben Sie ein Beispiel.

Aufgabe 1: Mobile Evaluation

- c) In der Vorlesung wurden verschiedene Kategorien von Regeln beschrieben, die einerseits den Design-Prozess unterstützen sollen und andererseits bei der Evaluation herangezogen werden können.

Welche Kategorien lassen sich unterscheiden?

Für wie hilfreich halten Sie solche Regeln (a) als Unterstützung beim Entwurf und (b) bei der Evaluation einer mobilen Benutzungsschnittstelle?

hier ist die eigene Einschätzung gefragt

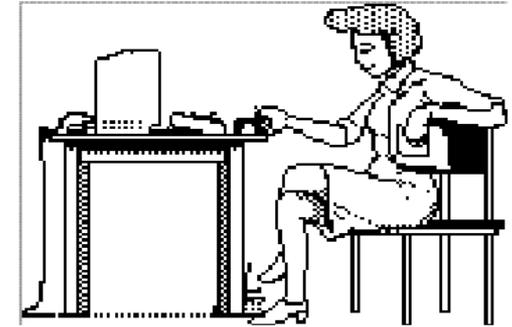
Aufgabe 2: Evaluation von Foursquare

- Lernziel: Benutzerstudie außerhalb des Labors in mobiler Nutzungssituation durchführen
- „Foursquare“ (<https://foursquare.com>)
 - standortbezogener sozialer mobiler Dienst
 - für Android, iOS, Blackberry
 - Funktionen: an Orten „einchecken“ und Punkte erhalten
 - Abzeichen („Badges“)
 - Status (z.B. „Bürgermeister“ = Benutzer mit den meisten Check-ins an einem Standort)
 - Freunden Standort und Kommentare zu Orten mitteilen
 - Anbindung an Twitter und Facebook
 - Anlegen von To-Do-Listen
 - Austausch von Tipps zu bestimmten Standorten

Mobile Think-Aloud-Studie

- reale Umgebung: auf der Straße, GPS eingeschaltet
- Testbenutzer ohne Foursquare-Erfahrung
 - Testbenutzer darf nicht Teilnehmer der Vorlesung sein!
- Aufgaben für den Testbenutzer
 1. Registrieren eines Benutzerkontos (evtl. Pseudonym für Name)
 2. Ändern der Einstellungen, so dass Name nicht in öffentlicher Liste der an einem Ort eingechekten Personen erscheint
 3. Einchecken in einem Restaurant oder Café in der Nähe, dabei einen kurzen Tipp schreiben und ein Foto machen
 4. Finden empfohlener Kultur-Events in der Nähe (weniger als 3km Entfernung)
 5. Erstellen einer To-Do-Liste mit 2 Orten in der Nähe, die der Benutzer in Zukunft noch besuchen möchte

Silent Observation



Source: Saul Greenberg

- Designer watches user in lab or in natural environment while working on one of the tasks
- No communication during observation
- + Helps discover big problems
- No understanding of decision process (that may be wrong) or user's mental model, opinions, or feelings

Think Aloud



Source: Saul Greenberg

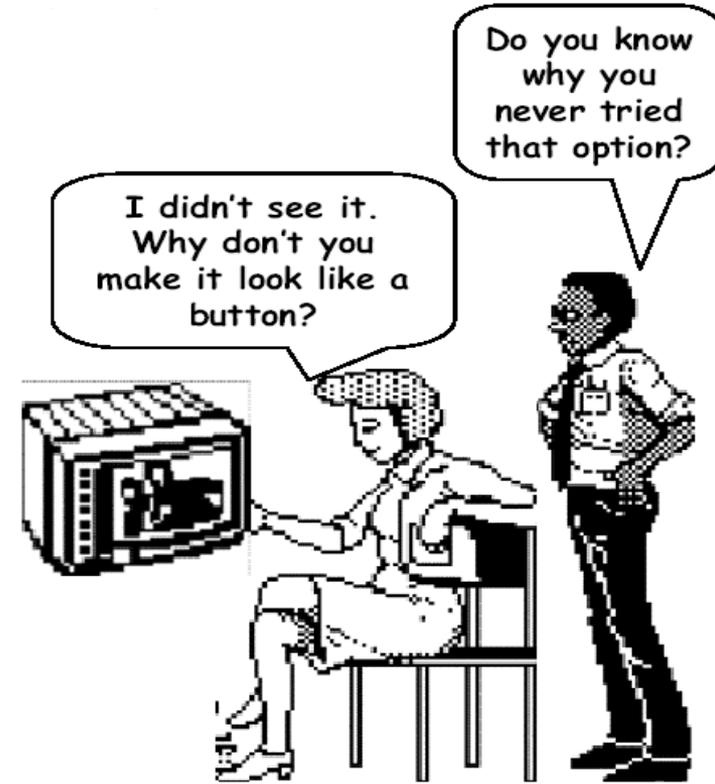
- As Silent Observation, but user is asked to say aloud
 - What he thinks is happening (state)
 - What he is trying to achieve (goals)
 - Why he is doing something specific (actions)
- Most common method in industry
- + Good to get some insight into user's thinking, but:
 - Talking is hard while focusing on a task
 - Feels weird for most users to talk aloud
 - Conscious talking can change behavior

Recording Observations

- Paper and pencil
 - Evaluator notes events, interpretations, other observations
 - Cheap but hard with many details (writing is slow)
 - Forms can help
- Audio recording
 - Good for speech with Think Aloud and Constructive Interaction
 - But hard to connect to interface state
- Video
 - Ideal: 2 cameras (user and screen) in 1 picture
 - Best capture, but may be too intrusive initially
- Logging
 - Log input events of the user, synchronize with audio & video

Retrospective Testing

- Additional activity after an observation
- Subject and evaluator look at video recordings together, user comments his actions retrospectively
- Good starting point for subsequent interview, looking at video avoids wrong memories
- Often results in concrete suggestions for improvement



Durchführung der Think-Aloud-Studie

- Aufruf des Grundbildschirms der App
- Prinzip Think-Aloud erklären
- Aufgabe erklären (jeweils die nächste auszuführende)
- Beobachten des Benutzers: Video+Audio oder Notizen
- Beobachtungen / Besonderheiten notieren
 - Pausen zwischen den Aufgaben für Notizen nutzen
 - Besonderheiten: Fehler, Zögern, positives / negatives Verhalten
 - Zuordnung von Beobachtungen zu Stelle in UI
- Einfluss der Umgebung
 - Ablenkungen, wann Aufmerksamkeit nicht auf UI
- Zeit und Erfolg/Misserfolg für jede Aufgabe notieren

SUS-Umfrage im Anschluss an Think-Aloud

- **Achtung:** Fehler in Vorlesungsfolien beim Berechnen des SUS-Scores
 - [Korrigierte Version online](#)
- System Usability Scale (SUS) liefert einen Wert (0..100) für subjektive Zufriedenheit des Benutzers mit System

SUS: System Usability Scale

- Developed by DEC Corporation
- 10 5-point Likert scales
- Single score (0-100)
 - Odd items: position – 1
 - Even items: 5 – position
 - Add item scores
 - Multiply by 2.5

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

Brooke. SUS: A "quick and dirty" usability scale. Usability Evaluation in Industry. London: Taylor and Francis, 1996

Example: SUS-Ratings

	Strongly disagree				Strongly agree	
1. I think that I would like to use this system frequently		X				1
	1	2	3	4	5	
2. I found the system unnecessarily complex	X					4
	1	2	3	4	5	
3. I thought the system was easy to use		X				1
	1	2	3	4	5	
4. I think that I would need the support of a technical person to be able to use this system			X			2
	1	2	3	4	5	
5. I found the various functions in this system were well integrated	X					0
	1	2	3	4	5	

pos=2: score = pos-1=1

pos=1: score = 5-pos=4

pos=2: score = pos-1=1

pos=3: score = 5-pos=2

pos=1: score = pos-1=0

Brooke. SUS: A "quick and dirty" usability scale. Usability Evaluation in Industry. London: Taylor and Francis, 1996

Example: SUS-Ratings

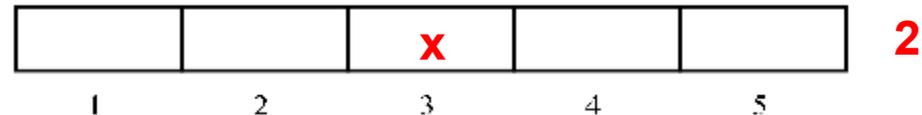
6. I thought there was too much inconsistency in this system



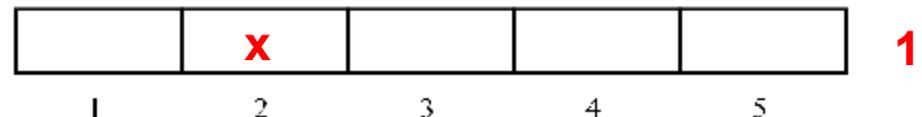
7. I would imagine that most people would learn to use this system very quickly



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



Sum = 16

SUS-Score = Sum * 2.5 = 40

Brooke. SUS: A "quick and dirty" usability scale. Usability Evaluation in Industry. London: Taylor and Francis, 1996

Bestandteile der Abgabe (Aufgabe 2)

- Demographische Angaben zum Benutzer
- Ort der Versuchsdurchführung, Kontext-Variablen
- Gefundenen Usability-Probleme
- Beobachtungen des Benutzerverhaltens
- Beschreibung des Einflusses der Umgebung
- Dauer und Erfolg der Aufgabenbearbeitung
- Bewertung der Think-Aloud-Methode
- Resultat der SUS-Umfrage

Ausleihe von Android-Handys

- Falls kein Android-Mobiltelefon oder iPhone vorhanden
- Durchführung des Versuchs für zwei Stunden ausleihen
- Foursquare-Client bereits installiert
- Zur Ausleihe bitte in einen der Zeitslots eintragen
 - maximal 2 Personen pro Zeitslot
 - <http://www.doodle.com/stbpus467wsvg2bt>

Abgabe

- Einzelbearbeitung
- PDF-Datei
- Montag, den 16.1.2012 um 12:00 Uhr
- UniWorX Portal (<https://uniworx.ifi.lmu.de/>)

- in der Übung Erfahrungen mit Think-Aloud vorstellen können!