

Evaluationsmethoden

Ilhan Aslan, Chi Tai Dang, Björn Bittner, Katrin Janowski,
Elisabeth André



Human Centered Multimedia

Institute of Computer Science

Augsburg University

Universitätsstr. 6a

86159 Augsburg, Germany

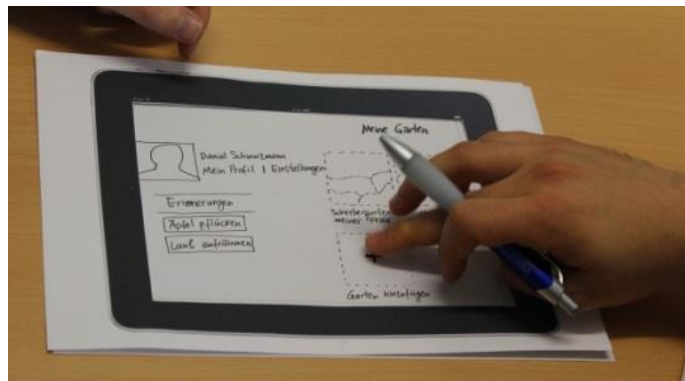
Evaluationen – Was kann man alles evaluieren?



Konzepte, Ideen



Alte Systeme



Prototypen

Fertige Produkte



- Wann wird das System evaluiert?
 - Während des Prozesses (Formativ)
 - Nach dem Prozess (Summativ)
- Wie misst man die Daten?
 - Qualitativ
 - Quantitativ
- Wie ist die Datenerhebung?
 - Subjektiv
 - Objektiv
- Wer evaluiert das System?
 - Nutzer (Empirische Evaluationen)
 - Experten (Analytische Evaluationen)

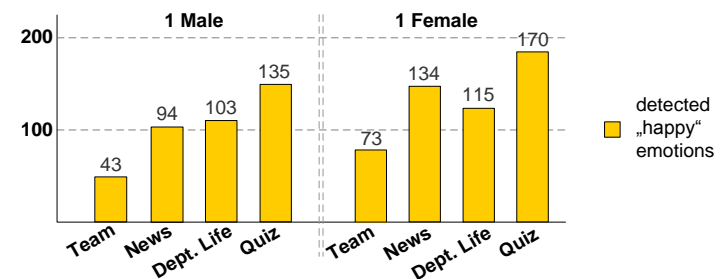
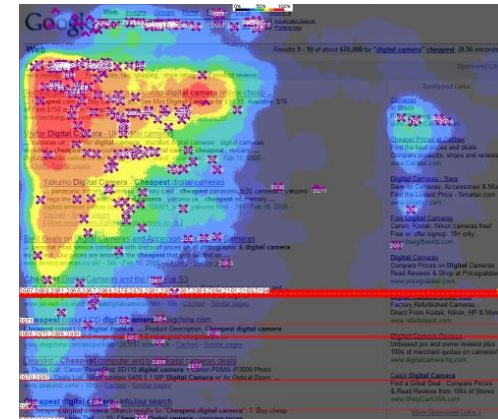
- **Formative Evaluation**

- Untersucht Zwischenergebnisse
- Bewertung und Verbesserung eines Prozesses oder Prototypen
- Vorab definierter Zeiträume (z.B. Alle X Wochen, Milestones)
- Vorab definierte Kriterien (z.B. Effizienz, Effektivität, Zufriedenheit)
- Ziel: Erhöhung der Wahrscheinlichkeit, dass festgelegte Ziele erreicht werden.
- Beispiele:
 - Vollständigkeit eines Konzeptes
 - Güte/Qualität eines Prototypen/Systems

- **Summative Evaluation**
 - Untersucht das Ergebnis eines Prozesses
 - Abschließende Bewertung eines fertigen Produkts (vgl. Qualitätskontrolle)
 - Vergleich: festgelegte vs. erreichte Ziele
 - Beispiele:
 - Erfolg eines Verfahrens
 - Güte/Qualität eines vollständig entwickelten Produkts/Systems

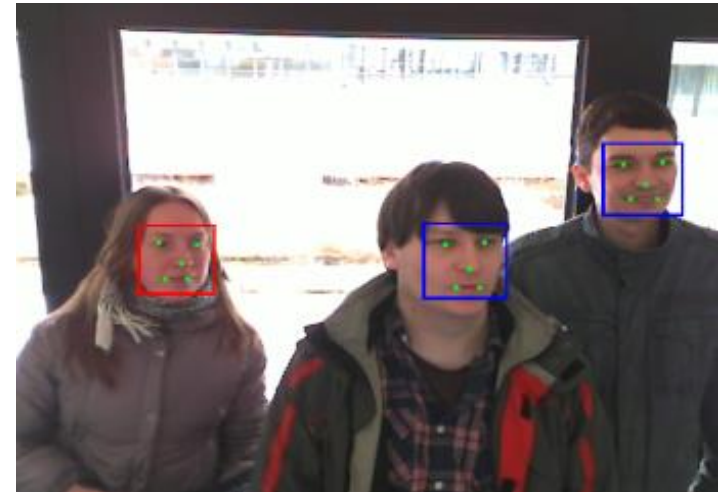
• Quantitative Messung

- Direkt vergleichbare zählbare Werte (z.B. Zeit für Aufgabe, Verweildauer auf Objekt, Fehleranzahl während Aufgabe)
- Beispiele:
 - Logfiles, Eyetracking
 - Geschlossene Fragen bei Fragebögen (z.B. „Bewerten sie das Design der Webseite auf einer Skala von 1 = „hässlich“ bis 5 = „schön“.“)
 - abhängige Variablen bzw. Messwerte bei Experimenten (Beispiele siehe oben)



- **Qualitative Messungen**

- Indirekt vergleichbare zählbare Werte
 - Messbarkeit erst durch Annotation bzw. anhand von Analyseschemas (qualitative -> quantitative Daten)
 - Beispiele:
 - Videoaufzeichnung
 - Offene Fragen bei Fragebögen (z.B. „Was hat Ihnen an diesem System besonders gefallen?)
- + mehr Informationen
(z.B. Gründe für Fehler)
- zeitaufwendige und fehleranfällige Auswertung



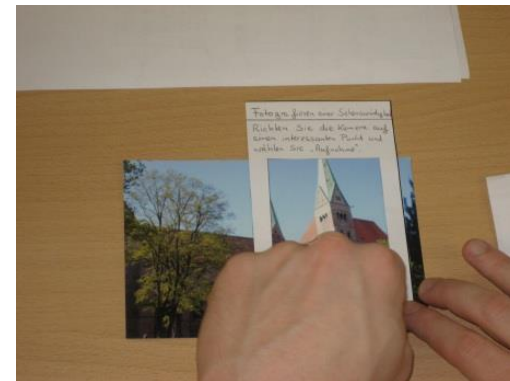
- **Subjektive Datenerhebung**

- Ergebnisse zur **Meinung** von Nutzern
- Beispiele:
 - Befragungstechniken (z.B. Fragebogen, Interview)



- + Einfach zu erhaltende Ergebnisse
- oft mangelnde Validität
(Stimmen Antworten und tatsächliches Verhalten überein?)

- **Objektive Datenerhebung**
 - Ergebnisse zum **Verhalten** von Nutzern
 - Beispiele:
 - Beobachtungstechniken (z.B. Videoaufzeichnung oder physiologische Daten wie Puls und Hautleitwert)



- Aufwendiger zu erhaltende Ergebnisse
- + oft mit höherer Validität

- Usability Engineering:
 - **Empirische Evaluation** mit Nutzern
 - **Analytische Evaluation** mit (Usability-) Experten
- Sowohl summativ, als auch formativ durchführbar
- Können aufeinanderfolgend durchgeführt werden
- Beispiel:
 1. Evaluation mit Experten: Erkennen und Beheben offensichtlicher, großer Probleme bei der Bedienung
 2. Evaluation mit Nutzern, Erkennen und Beheben weiterer nutzerspezifischer Probleme bei der Bedienung

Analytische Evaluationen

- Mit Experten (z.B. aus dem Projektteam)
- Identifiziert grundlegende (offensichtliche) Probleme
- Nutzung von Richtlinien, Checklisten oder Modellen
- Objektive Ergebnisse und in erster Linie quantitative Messwerte
- Arten von analytischen Evaluationsmethodiken
 - **Inspektionsbasierte Evaluation:** Durchlauf der Tasks und Überprüfung anhand von Richtlinien und Checklisten (z.B. Heuristische Evaluation, Cognitive Walkthrough)
 - **Modellbasierte Evaluation:** Durchlauf anhand von empirisch ermittelnden Modellen (z.B. GOMS)

Empirische Evaluationen

- Mit tatsächlichen Nutzern der Zielgruppe
- Identifiziert nutzerspezifische Probleme
- Arten von empirischen Evaluationen
 - **Befragungstechniken:**
Befragung der Nutzer nach/vor/während der Benutzung des Systems (z.B. Interviews, Fragebogen)
 - **Beobachtungstechniken:**
Aufzeichnung der Nutzung des Systems durch den Nutzer (z.B. Videoaufzeichnung, Protokolle)

Empirische Evaluationen

Befragungstechniken:

- Meinung der Nutzer wird abgeklärt
 - Subjektive Ergebnisse
 - in erster Linie qualitative Messwerte
- Quantitative Messwerte durch:
 - geschlossene Fragen
 - Annotation anhand eines Analyseschemas für qualitative Messwerte (z.B. offene Fragen)

Empirische Evaluationen

Beobachtungstechniken:

- Verhalten der Nutzer wird abgeklärt
 - Objektive Ergebnisse
 - in erster Linie qualitative Messwerte
- Quantitative Messwerte durch
 - Logging und Protokollierung
 - Annotation anhand eines Analyseschemas für qualitative Messwerte (z.B. Videos, Audio)

Empirische Evaluationsmethoden: Befragungstechniken

Stephan Hammer, Ilhan Aslan,
Andreas Seiderer, Simon Flutura
Elisabeth André



Human Centered Multimedia

Institute of Computer Science

Augsburg University

Universitätsstr. 6a

86159 Augsburg, Germany

- Interviews
- Fokusgruppen
- Fragebögen

Wichtig!

Die Meinung der Nutzer muss nicht ihr tatsächliches Verhalten widerspiegeln!

- Allgemeines Vorgehen:
 - Interviewer stellt Fragen und Nutzer antwortet
 - Ein zweiter Interviewer macht schriftliche Notizen
 - Aufzeichnung mit Audio / Video
 - Spätere Analyse der Daten
- Wann führt man Interviews durch?
 - Analysephase:
 - Grundlegende Befragung unabhängig von einem System
 - Nach/vor der Vorstellung eines Konzeptes oder einer Idee
 - Nutzerzentriertes Prototyping:
 - Nach/vor der Nutzung eines Systems

- Ergebnis:
 - Vielzahl an hauptsächlich qualitativen Daten
 - Aufwendig auszuwerten
 - Sehr wertvolle Daten
 - Nachhaken bei Unklarheiten
 - Starten von Diskussionen
 - Direktes Feedback vom Endnutzer!
 - Welche Bedürfnisse / Wünsche hat der Nutzer?
 - Was kann und was will der Nutzer?
 - Was gefällt? Was gefällt nicht?
 -

Unstrukturierte Interviews

- Kein vorgegebener Fragenkatalog
- Interviewer kann sich anpassen
- + Sehr flexible Methodik
- + Wertvolle Daten, weil man auf den Nutzer besser eingehen kann
- + Erklärungen / Interpretationen leicht möglich
- Fehleranfällig (z.B. Vergessen wichtiger Aspekte)
- Schwer reproduzierbar (z.B. bei Analyse der Antworten)
- Schlechte Vergleichbarkeit der Antworten der Nutzer

Strukturierte Interviews

- Fragebogen mit festen Fragen
 - Fragekategorien und Fragen werden vor dem Interview festgelegt
 - Keine Anpassbarkeit / Flexibilität des Interviewers erlaubt
- + Reproduzierbar
- oft nicht besonders aussagekräftig, weil der Interviewer nicht nachhaken darf

Semi-Strukturierte Interviews

- Kombination unstrukturierter und strukturierter Interviews
- Fester Fragenkatalog
- + Nachvollziehbarkeit aller Interviews
- Zusätzliche Fragen und Nachhaken erlaubt
- + Flexibilität und Anpassbarkeit an bestimmte Situationen
- + Wertvolle Informationen über Gründe für Antworten

Durchführung

1. Vorstellung

- Wer seid ihr?
- Was sind die Ziele des Interviews? (nicht zu viel Verraten!!)
- Darf das Interview aufgezeichnet werden?

2. Warm-Up

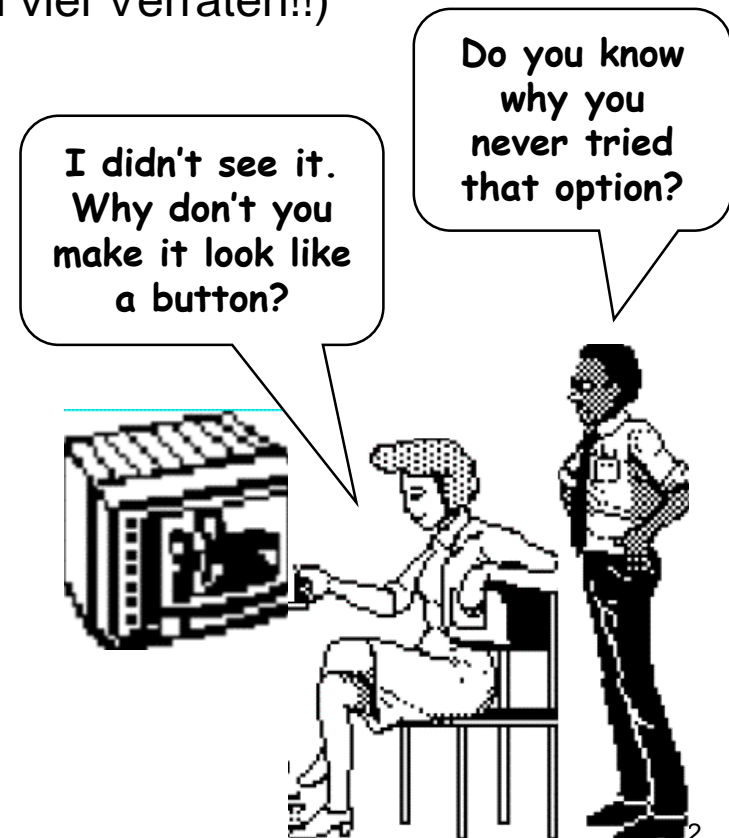
- Demographische Daten
(Geschlecht, Alter, Beruf, Interessen...)

3. Hauptteil

- (Zeigen des Konzepts oder der Aufzeichnungen, die während des Tests gemacht wurden)
- (Abarbeitung von Aufgaben)
- Einfache und allgemeine Fragen
- Detaillierte, gezielte Fragen

4. Ende

- Bedanken



Allgemeiner Leitfaden für Fragen in Interviews

- Festlegen der zentralen Fragen und eines „roten Fadens“
- Geeignete Eröffnungsfragen
- Gruppierung der Themen (Keine Themensprünge!)
- Fragen in logischer Reihenfolge anordnen
- Zusätzliche Fragen ergeben sich aus beobachteten kritischen Situationen und aufgetretenen Problemen (Critical Incident Interviews) oder aus dem Gespräch heraus



Allgemeines Vorgehen:

- “Interview in einer Gruppe”
- Gruppengröße: 4-12 Personen
- Gruppenzusammensetzung:
 - Heterogene Gruppe(n)
 - Homogene Gruppen, die sich untereinander unterscheiden
- Einer (Moderator) fragt, viele (Nutzer) antworten
- Ziel: Gesteuerte Diskussion zwischen Teilnehmern
- Ein zweiter Interviewer macht schriftliche Notizen
- Aufzeichnung mit Audio / Video



Wann führt man Fokusgruppen durch?

- Typischerweise in der Analysephase:
 - Sammeln von primär qualitativen Daten
(Bedürfnisse, Meinungen, Standpunkte, Gefühle, Ideen...)
 - Verständnis von Nutzer, Aufgaben und Kontexten
 - Ziele und Anforderungen

Was muss beachtet werden?

- Fokus auf bestimmte Themen/Fragen
- Genaue Planung für Ablauf der Diskussion
- Gute Moderation nötig:
 - Nachbohren bei interessanten Aussagen
 - Verhindern von Gruppendynamik – Meinungsmacher!
 - Jeder Teilnehmer sollte zu Wort kommen!

Beschreibung

- Ein Evaluator und eine Testperson
- Fester Fragekatalog
 - offene (qualitativ) und geschlossene (quantitativ) Fragen
- Testperson füllt Fragebogen selbstständig aus
- Auch Remote möglich (z.B. Online-Fragebogen)
 - + Schnell viele Teilnehmer bzw. Daten
 - Kein Nachfragen bei Missverständnissen möglich!
 - Wie aufmerksam und ehrlich wurde der Fragebogen ausgefüllt?

Offene Fragen

- Typischerweise “W”-Fragen:
- Beispiele:
 - Was hat Ihnen an diesem System gefallen?
 - Was hat Ihnen an diesem System nicht gefallen?
 - Welche zusätzlichen Funktionen würden Sie sich wünschen?
 - Würden Sie das System Ihren Freunden empfehlen?
- Qualitative Daten: Antworten, Begründungen und Erläuterungen

Offene Fragen

Vorteile:

- Testperson formuliert Meinung in eigenen Worten
- Begründungen für Bewertungen
- Hinweise auf zusätzliche Aspekte (z.B. Probleme)
- Kreatives Einbringen durch Ideen und konstruktive Verbesserungsvorschläge

Offene Fragen

Nachteile

- Aufwändige und schlecht automatisierbare Auswertung
- Teilweise große Unterschiede in Formulierung:
 - Umfang (z.B. redselige vs. introvertierte Probanden)
 - Stil (z.B. motivierte vs. faule Probanden)
 - Inhalt (z.B. viel bzw. wenig Wissen in der Domäne)
 - Schlechte Vergleichbarkeit erschwert Ableitung genereller Trends
- Aufwand und Anstrengung für Probanden
 - Vergleichsweise wenige Fragen möglich
 - Probanden können Interesse/Motivation verlieren

Geschlossene Fragen

Arten geschlossener Fragen:

- Ja/Nein-Fragen
 - Direkte Antworten (z.B. Wie alt sind sie?)
 - Likert Skalen (z.B. 1 = „sehr gut“ bis 5 = „sehr schlecht“)
 - Multiple Choice (Eine oder mehrere Antwortoptionen auswählbar)
 - Ranking
- Quantitative Daten

Geschlossene Fragen

Likert Skalen (nach R. Likert 1932)

- Skala mit 3 bis 9 Optionen
- Typischerweise 5er und 7er Skalen
- Gerader Wert als Mittelwert (neutraler Wert)
- Auch gerade Anzahl an Optionen möglich
 - Teilnehmer müssen sich entscheiden
- Abfrage der Zufriedenheit oder der Übereinstimmung mit einer Aussage
- Anmerkung: Slider nicht empfehlenswert, da kontinuierliche Werte schwer zu interpretieren sind.

Zeichen auf dem Bildschirm waren:

Schwer zu lesen						Leicht zu lesen
1	2	3	4	5	6	7



Schwer zu lesen

Leicht zu lesen

Geschlossene Fragen

Likert Skalen – Anforderungen:

1. Klar voneinander unterscheidbare Optionen

- Bildung semantischer Differenziale (gegensätzliche Wortpaare) als Enden der Likert-Skala
 - Beispiele:
 - schlecht \leftrightarrow gut
 - unverständlich \leftrightarrow verständlich
 - schwer \leftrightarrow einfach
- Auswahl der Wortpaare hängt vom Bewertungsgegenstand und den Bewertungsdimensionen ab:
 - z.B. funktionale vs. ästhetische Aspekte eines Produkts

Geschlossene Fragen

Likert Skalen – Anforderungen:

2. Konsistente Polarität der Skalen

- Standard bei horizontalen Skalen:
 - linkes Skalenende = negativ / wenig / schlecht / ...
 - rechtes Skalenende = positiv / viel / gut / ...
- Standard bei vertikalen Skalen:
 - oberes Skalenende = positiv / viel / gut / ...
 - unteres Skalenende = negativ / wenig / schlecht / ...
- Gefahr des „blinden“ Ankreuzens:
 - Unterschiedlich formulierte Fragen bei gleicher Polarität
- Beispiel:
 1. Ich denke ich würde das System regelmäßig nutzen.
 2. Ich halte das System für unnötig komplex.

Geschlossene Fragen

Likert Skalen – Anforderungen:

3. Doppelte Verneinungen vermeiden!



Mir hat das Design des Systems nicht gefallen.
trifft nicht zu O - O - O - O - O - O trifft zu

4. Nur ein sachlicher Inhalt / Gedanke pro Frage!



Ich arbeite sehr gerne und sehr lange am Computer
trifft nicht zu O - O - O - O - O - O trifft zu


Geschlossene Fragen

Multiple Choice

1. Hinweis, falls mehrere Optionen gewählt werden können
2. Anzahl der Optionen vernünftig wählen. Lieber weniger, aber dafür eindeutig und einfach zu beantwortende Optionen.

Welches Textverarbeitungsprogramm nutzen sie? (Mehrfachauswahl möglich)

- ☐ LaTeX
- ☐ Word
- ☐ FrameMaker
- ☐ OpenOffice



Wie häufig nutzen Sie einen Computer auf der Arbeit?

- ☐ mehr als 5 Std. täglich
- ☐ zwischen 3 und 5 Std täglich
- ☐ zwischen 2 und 3 Std täglich
- ☐ zwischen 1 und 2 Std täglich
- ☐ weniger als 1 Std täglich

Geschlossene Fragen

Rankings

- Ermittlung von Prioritäten und Präferenzen
- Problem:
Teilweise schwer zu beantworten

Gewichten Sie

1 – Sehr hilfreich

2 – Unentschieden

3 – Nicht hilfreich

0 – Keine Meinung

___ Tutorial

___ On-line help

___ Documentation

Geschlossene Fragen

- Vorteile:
 - Verhindert unpräzise Antworten
 - Klare Interpretation der Antworten
 - Leicht quantifizierbar
 - Erleichterte statistische Auswertung
- Nachteile
 - Optionen müssen alle sinnvollen Möglichkeiten abdecken
 - Anzahl der sinnvollen Antworten variiert oft stark, man möchte jedoch möglichst einheitliche Skalen
 - Keine Möglichkeit Gewissenskonflikte zu artikulieren (z.B. für- und widersprechende Argumente)
 - Abhilfe: zusätzliches Kommentarfeld

Allgemeiner Leitfaden zur Erstellung von Fragebögen

- Aufbau:
 1. Allgemeiner Teil mit demographischen Daten (Geschlecht, Alter, Beruf, Interessen...)
 2. Hauptteil
 - a) Einfache und allgemeine Fragen
 - b) Detaillierte, gezielte Fragen
- Ausgewogenheit zwischen offenen und geschlossenen Fragen
 - Geschlossene Fragen: Bewertungen, Auswahl oder Rankings
 - Offene Fragen: Meinungen / Begründungen

Allgemeiner Leitfaden zur Erstellung von Fragebögen

- Vermeide Themensprünge
 - Klare Strukturierung
 - Von allgemeinen zu speziellen Fragen
- Vermeide lange und verschachtelte Fragen und Sätze
 - Klar, einfach und kurz
- Vermeide Umgangssprache und Stereotypen
- Vermeide Fragen, die die Antwort vorgeben
 - Beispiel: Warum MAGST du....?
 - Neutralität
 - Falls keine neutrale Fragestellung möglich, Ausgewogenheit zwischen positiven und negativen Fragen und Antworten.

Allgemeiner Leitfaden zur Erstellung von Fragebögen

- Vermeide unnötigen Fragen! Jede Frage muss sinnvoll sein und hilfreiche Informationen liefern!
 - Analysephase:
Wer sind die Nutzer? Welche Aufgaben haben sie? Wann? Wie häufig? Welche Schritte nötig? Welche Kontexte liegen vor? Welche konkreten Anforderungen haben die Nutzer an das System?
 - Prototypingphase:
Anforderungen/Ziele erfüllt? Was funktioniert oder funktioniert (noch) nicht? Alle Aktionen/Features erkennbar und vorhanden? Gibt es unnötige Aktionen/Features? Fortschritt erkennbar? Gibt es falsche/zusätzliche Anforderungen/Ziele?
 - Vergleiche 7 Handlungsschritte von Norman!!!

Problem:

- Erstellung eigener Fragebögen ist fehleranfällig
 - Verständlichkeit
 - Wiederholungen
 - Vollständigkeit
 - Eignung der Wortpaare und Polaritäten

➤ Besser:

- Verwendung erprobter und validierter Fragebögen
- Verwendung erprobter Frageblöcke für die jeweilige Problemstellung (z.B. Zufriedenheit, Vertrauen, Effizienz)

SUS (System Usability Scale)

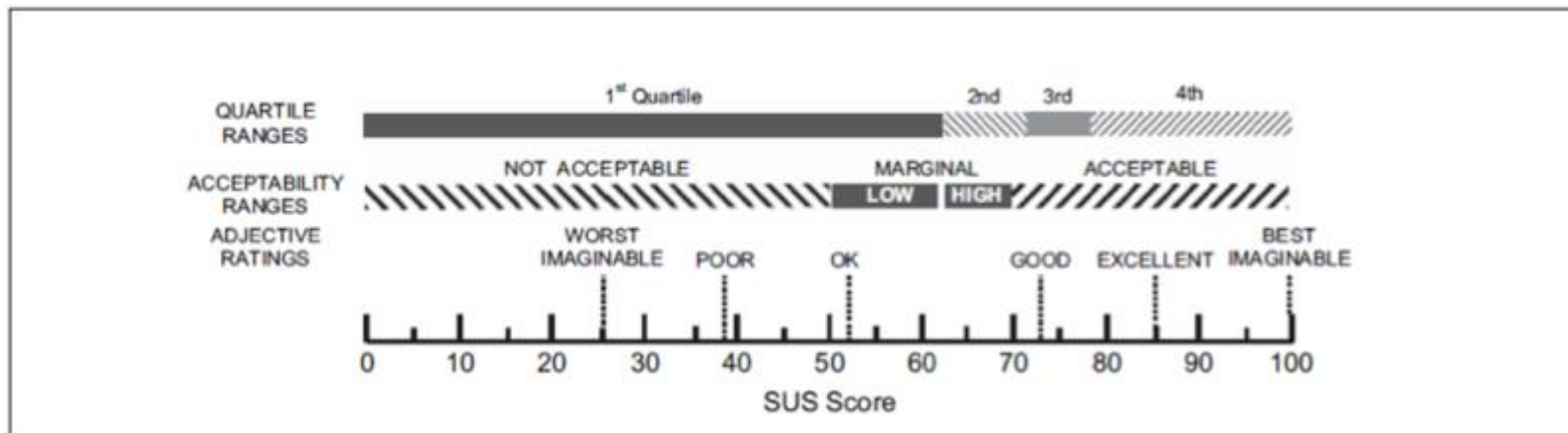
- 10 Fragen bzw. Statements
- 5er Skalen (“Strongly Disagree” bis “Strongly Agree”)
- Quellen:
 - <http://www.measuringusability.com/blog/sus-websites.php>
 - http://www.measuringusability.com/papers/Lewis_Sauro_HCII2009.pdf
- Deutsche Versionen:
 - <http://ux.stackexchange.com/questions/10181/what-is-the-standardized-german-version-of-the-system-usability-scale-sus>
 - <http://www.sapdesignguild.org/resources/sus.asp>

SUS (System Usability Scale)

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.

SUS (System Usability Scale) – Auswertung:

- Positive Formulierungen (Fragen 1,3,5,7,9):
Codierung der einzelnen Bewertungen von 0 bis 4
- Für negative Formulierungen (Fragen 2,4,6,8,10):
Codierung der einzelnen Bewertungen von 4 bis 0
- Summe der Werte * 2,5



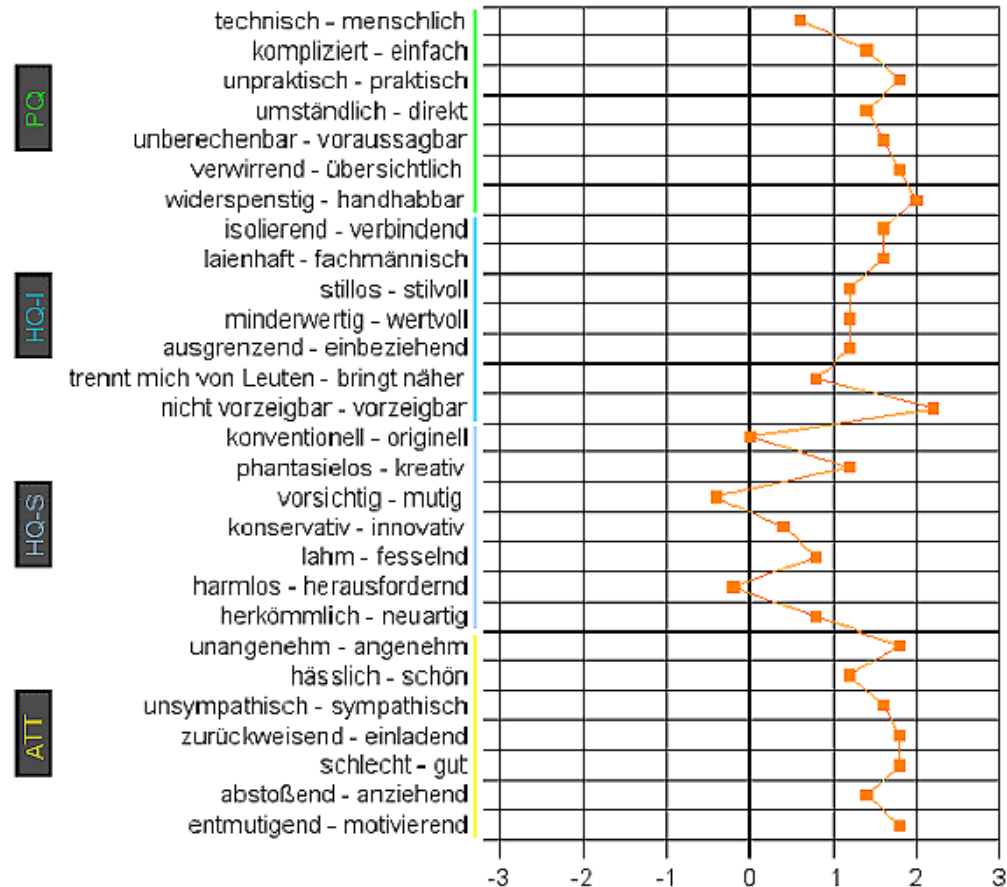
SUS (System Usability Scale)

- Vorteile:
 - Sehr kurze Bearbeitungszeit
 - Alle Software- und Hardwaresysteme sind vergleichbar
 - Sehr etablierter Fragebogen
- Nachteile:
 - Keine qualitativen Aussagen über System
 - Ergänzen durch offene Fragen

AttrakDiff <http://attrakdiff.de/>

Online-Werkzeug zur empirischen Attraktivitätsbestimmung von Produkten

1. Online-Befragungen (AttrakDiff Server)
2. Auswertung der Antworten
3. Reporter und übersichtliche Präsentation der Ergebnisse



Weitere Beispiele:

- QUIS (Questionnaire for User Interface Satisfaction)
<http://www.lap.umd.edu/QUIS/index.html>
- CSUQ (Computer System Usability Questionnaire)
- Words
- SUMI (Software Usability Measurement Inventory)
<http://www.ucc.ie/hfrg/questionnaires/sumi/index.html>

Objektivität

- Unabhängigkeit der Ergebnis vom Interviewer
- Konsistente Bedingungen bei der Durchführung (z.B. gleiche Zeitvorgaben)
- Konsistente Auswertung der Fragebögen (z.B. von der selben Person)
- Geringer Spielraum für Interpretationen der Ergebnisse

Verlässlichkeit (Reliability)

- Grad der Genauigkeit, mit dem eine Befragung ein bestimmtes Merkmal (z.B. Zufriedenheit) misst
- Vor der tatsächlichen Befragung Pilottest zur Überprüfen der Verlässlichkeit eines Fragebogens (z.B. Verständlichkeit und Eindeutigkeit der Fragen)
 - Teile des Fragebogens
 - Tests unter verschiedenen Bedingungen (Stabilitätstest)

Validität

Zweckdienlichkeit der Fragen für die untersuchte Problemstellung

- keine unnötigen Fragen nach nicht benötigten Daten
- hinreichend viele Aspekte (Indikatoren) für untersuchte Merkmale (z.B. „Bedienfreude“)

Problem: Ehrlichkeit der Antworten unklar

- fake good: Vorgabe von normalerweise nicht gezeigtem Verhalten
- fake bad: Verschleiern von normalerweise gezeigtem Verhalten
- Strategien zur Abhilfe:
 - Aufklärung über Gründe der Befragung
 - Aufklärung über Nutzung der Daten
 - Zusicherung von Anonymität
 - Art der Befragung: z.B. Multiple Choice statt Ja/Nein
 - Ausreichende Anzahl an Teilnehmern

Strategien zur Abhilfe:

- Validitätsskalen zur Erfassung sozial erwünschter Antworten
 - Vorgabe sozial erwünschter, aber selten zutreffender Antworten
 - Beispiel: „Ich bin immer objektiv.“ (Fake Good)
 - Vorgabe sozial unerwünschter, aber häufig zutreffender Antworten
 - Beispiel: „Manchmal bin ich ungeduldig.“ (Fake Bad)
 - Bei entsprechender Beantwortung („Ja“ und „Nein“) Annahme, dass der Proband im gesamten Test sozial erwünscht geantwortet hat.

Interviews

- Schwere Auswertung (z.B. Reproduzierbarkeit)
- Aufwendig an Material zu kommen
- Wenig Daten, aber qualitativ am hochwertigsten
- Vor allem geeignet, um konkrete Konzepte zu überprüfen, da man leicht nachhaken kann

Fokusgruppen

- Leichter und schneller an viel Material zu kommen
- Wegen Gruppeneffekten ist die Qualität der Daten etwas schlechter als bei Interviews

Fragebögen

- Sehr leicht an viele Daten zu kommen (z.B. Onlinefragebogen)
- Sehr schwer gutes Material zu bekommen, wenn Interesse der Umfrage spezifisch ist (z.B. bestehendes Konzept überprüfen)
- Keine Möglichkeit nachzufragen, wenn Antwort unklar ist.
- Geeignet bei eher allgemeinen Befragungsinteresse (grundsätzlich Interessen etc.) und keinem konkreten Konzept

- Befragungstechniken liefern subjektive Daten!
 - Meinung des Nutzers wird abgefragt!
 - Beispiel: Würde der Nutzer das System wirklich nutzen?
- Aber:
 - Sagt der Befragte wirklich die Wahrheit?
 - Wie ist sein tatsächliches Verhalten?
 - Subjektive Daten sind oft weniger valide als objektive Daten.
- Beobachtungstechniken liefern objektive Daten
 - Verhalten des Nutzers wird ermittelt!