

Research Methods

Jörg Cassens

Contextual Design of Interactive Systems



1 Introduction

Assignment 2.1: D. Norman

- Required reading for week 1
 - Norman, Donald A. “Human-centered design considered harmful.” interactions 12, no. 4 (2005): 14-19.
- The text will be discussed in the tutorial 16.04.2019
- Course readings can be downloaded in the learnweb
- Every text has a wiki-page in the learnweb
 - Use it to describe the text
 - Use it to link the text to the course
- Results of the discussion may also be written up

Assignment 3.1: Pruitt & Grudin, Chapman & Milham

- Required reading for week 2
 - Pruitt, John, and Jonathan Grudin. “Personas: practice and theory.” In Proceedings of the 2003 conference on Designing for user experiences, ACM, 2003.
 - Chapman, Christopher N., and Russell P. Milham. “The personas’ new clothes: methodological and practical arguments against a popular method.” In Proceedings of the human factors and ergonomics society annual meeting, vol. 50, no. 5, pp. 634-636. Sage Publications: Los Angeles, CA, 2006.
- The texts will be discussed in the tutorial 30.04.2019
- Course readings can be downloaded in the learnweb
- Every text has a wiki-page in the learnweb
 - Use it to describe the text
 - Use it to link the text to the course
- Results of the discussion may also be written up

2 Usability

Gebrauchstauglichkeit

- Was ist Gebrauchstauglichkeit (Usability) und wie kann man sie messen?
- Manche Systeme sind gut zu benutzen, andere nicht.
 - Beispiele haben wir gesehen... weitere?
 - Woran liegt es?
- Geringe Gebrauchstauglichkeit:
 - Aufgaben können nicht oder nur schwer erfüllt werden
 - Die Arbeit mit dem System macht keinen Spaß
 - Benutzer weiß zumeist nicht, ob es am System liegt oder an ihm selbst
- Zielrichtung:
 - Die Benutzung der Systeme vereinfachen, nicht den Menschen ändern

Qualitätsfaktoren

- Annahme: Qualität = Funktionalität + X
- Was ist X?
- "Ease of Use" ist nur einer von mehreren Qualitätsfaktoren
 - Korrektheit (keine Fehler)
 - Verfügbarkeit (für wie lange darf das System nicht erreichbar sein?)
 - Performanz (z.B. antwortet in 3 sec.)
 - Sicherheit (z.B. Schutz vor Ausfällen der Netzinfrastruktur oder vor Angriffen aus dem Netz)
 - Wartbarkeit (wie schwer zu warten)
- Multivariate Optimierung
 - Fast, reliable, cheap; chose two

Usability Probleme

Problemtypen

- Programmfehler (Bug)
 - arbeitet nicht wie vom Entwickler vorgesehen
- Fehlende Funktionalität
 - Aufgabe unlösbar
- Bedienbarkeitsproblem
 - die eigentlichen Usability Probleme
 - * System funktioniert und Aufgabe ist lösbar,
 - * aber Benutzer weiss nicht wie oder muss Umwege gehen oder...

Definition

- DIN EN ISO 9241 Teil 110
- Anforderungen an die Gebrauchstauglichkeit – Leitsätze
 - "Das Ausmaß, in dem ein Produkt durch bestimmte Benutzer in einem bestimmten Nutzungskontext genutzt werden kann, um bestimmte Ziele
 - effektiv,
 - effizient und
 - mit Zufriedenheitzu erreichen."

Definition (contd.)

- **Nutzungskontext:** "Die Benutzer, Arbeitsaufgaben, Arbeitsmittel (Hardware, Software und Materialien) sowie die physische und soziale Umgebung, in der das Produkt genutzt wird."
- **Effektivität:** "Die Genauigkeit und Vollständigkeit, mit der Benutzer ein bestimmtes Ziel erreichen."
- **Effizienz:** "Der im Verhältnis zu Genauigkeit und Vollständigkeit eingesetzte Aufwand, mit dem Benutzer ein bestimmtes Ziel erreichen."
- **Zufriedenstellung:** "Freiheit von Beeinträchtigungen und positive Einstellungen gegenüber der Nutzung des Produkts."

Usability nach Ben Shneiderman

- Effektivität mit SW-Engineering-Methoden gut meßbar
- Effizienz und Zufriedenheit sind nicht direkt meßbar
- Kriterien zur praktischen Evaluation, z.B. Usability Measures von [Shneiderman and Plaisant \[2005\]](#)
 - **Lernzeit:** Wie lange braucht der typische Benutzer, um mit dem System relevante Aufgaben lösen zu können?
 - **Aufgabeneffizienz:** Wie lange dauern bestimmte Aufgaben?
 - **Fehlerrate** der Benutzer: Welche und wie viele Fehler machen die Benutzer?
 - **Erinnerungsvermögen:** Wie gut behalten die Benutzer, was sie gelernt haben?
 - **Zufriedenheit:** Wie gefällt den Benutzern das Interface?

3 Evaluation: Grundlagen

Evaluation: Zielrichtung und Methoden

- Meßbarkeit
 - Wir wissen jetzt, was es ist, aber wie kann ich es messen? Stichwort Evaluation.
- Unterscheidung nach Art/Zielrichtung
 - Formativ *vs.* Summativ
- Unterscheidung nach Methoden
 - Analytisch *vs.* Empirisch
 - Qualitativ *vs.* Quantitativ
 - Feldstudien *vs.* Laboruntersuchungen

Formativ und Summativ

- **Formative Evaluation**
 - Während des Prozesses, "wo stehen wir?"
 - Hinweise auf konkrete Mängel bzw. Produktmerkmale, die Mängel verursachen
 - Ausgangspunkt für konstruktive Verbesserungsvorschläge
 - während der Entwicklung von Produkten
 - als Grundlage für Neuentwürfe
- **Summative Evaluation**
 - Am Ende, "sind die Anforderungen erfüllt?"
 - Abschließende Beurteilung
 - Produktvergleich
 - Überprüfung der Einhaltung von Kriterien
 - Zumutbarkeit von Softwaresystemen
 - Zertifizierung

Analytisch und Empirisch

"If you want to evaluate a tool, say an axe, you might study the design of the bit, the weight distribution, the steel alloy used, the grade of hickory in the handle, etc., or you might study the kind and speed of the cuts it makes in the hands of a good axeman." [Scriven, 1966]

- Die Evaluation der Charakteristik der Axt ist eine **analytische** Evaluation
 - Experten bewerten das System
 - Checklisten, Cognitive Walkthroughs
- Die Evaluation der Benutzung durch den Handwerker ist **empirisch**
 - Nutzer benutzen das System
 - Benutzertests, Benutzerbefragungen
- Diese Unterscheidung ist orthogonal zur Aufteilung formativ/summativ

Qualitativ und Quantitativ

- **Qualitativ**
 - Konkretes Feedback, aber nicht in Form von Meßgrößen
 - * Kommentare, Eindrücke, subjektive Bewertungen in Benutzerbefragungen
 - * Detaillierte Ergebnisse einiger weniger Benutzertests
- **Quantitativ**
 - Erhebung von Meßgrößen in kontrollierten Umgebungen
 - * Messung von Fehlerraten, Dauer der Interaktion, Anzahl der Interaktionsschritte
 - * In erster Linie zur Messung der Effizienz des Systems

Feldstudien und Laboruntersuchungen

- **Feldstudien**
 - Am Ort des Einsatzes
 - (Teilnehmende) Beobachtung der Nutzung des Systems im tatsächlichen Kontext
- **Laboruntersuchungen**
 - In speziellen Usability-Labs
 - Unter kontrollierten Bedingungen werden einzelne Arbeitsschritte durchgeführt

Feldstudien sind in der Regel wesentlich aufwendiger in der Durchführung, liefern aber auch realistischere Resultate

4 Analytische Evaluation

Experten-Evaluation

- 1-3 Usability-Experten sehen sich das System an und suchen gezielt nach Usability-Problemen
 - benutzen Wissen und Erfahrung und/oder heuristische Richtlinien
 - Dabei immer die vorgesehene Aufgabe und die vermuteten Benutzereigenschaften in den Vordergrund stellen
- Alternative: Die Entwickler betrachten das System fortlaufend anhand von Checklisten und Styleguides
 - Aber: Regelwerke sind nie vollständig
 - Entwickler zumeist nicht geschult
 - Externe Kontrolle besser (Voreingenommenheit)

4.1 Shneiderman

8 goldene Regeln von Shneiderman

- Konsistenz
 - Verwende Styleguides und weitere schriftliche Konventionen
- Berücksichtige unterschiedliche Erfahrungen
 - Eine Benutzungsschnittstelle sollte jeder NutzerIn möglichst eine passende Interaktionsform anbieten
 - * Anfänger: über Menüs
 - * Abkürzungen für erfahrene Benutzer
- Rückmeldungen auf Aktionen des Benutzers
 - Aktion bei der Software angekommen
 - * Insbesondere, wenn die Aktion spät ein Ergebnis liefert
 - * Akustisch, visuell, taktil
- Abgeschlossene Operationen
 - Schritte einer Operation im Zusammenhang darstellen
- Fehler verhindern
 - Darstellung eindeutig
 - Auswahlalternativen anbieten
- Einfache Rücksetzmöglichkeiten (undo)
 - Selbstsicherheit des Benutzers steigt stark an
 - Exploratives Lernen
- Benutzer bestimmt den Kontrollfluss
 - Gefühl, die Anwendung steuern, kontrollieren zu können
 - * "I am in control"
- Geringe Belastung des Kurzzeitgedächtnisses
 - Aufbau von Menüs besser breit statt tief

[Shneiderman and Plaisant \[2005\]](#)

4.2 Nielsen

Usability-Heuristiken von Nielsen

- Einfache und natürliche Dialoge
 - dem Lösungsweg der Aufgabe angepaßt
 - an den Erwartungen des Benutzers orientiert
- Ausdrucksweisen des Anwenders
 - Fachsprache des Anwendungsgebiets
- Minimale mentale Belastung des Benutzers
 - "Don't make me think" (Steve Krug)
- Konsistenz
 - Darstellung
 - Dialoge folgen immer der gleichen Logik
- Rückmeldungen
 - über Annahme der Aktion

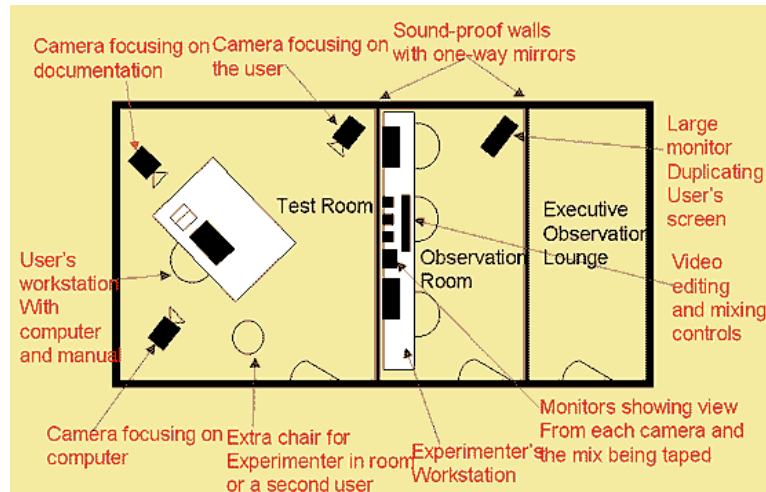
- insbesondere, wenn die Aktion länger dauert
- Klare Auswege
 - bei falscher Navigation
 - bei falscher Aktion
- Abkürzungen
 - Standardwerte, History-Funktionen
 - für geübte Benutzer: shortcuts
- Gute Fehlermeldungen
 - konstruktive Rückmeldungen
- Fehlervermeidung
 - besser als Fehlerbehandlung
- Hilfe und Dokumentation
 - bei Anwendung und Einarbeitung unterstützen.
 - vollständig und übersichtlich
 - korrekt und auf dem aktuellen Stand

Nielsen [1994]

5 Empirische Evaluation

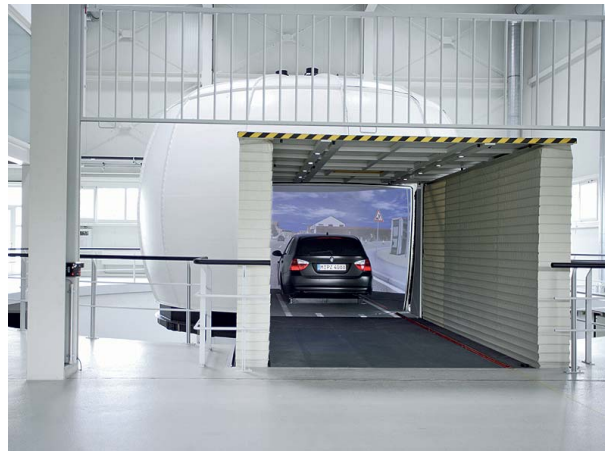
5.1 Benutzertest

Usability Labor I



Quelle: M. Herzceg

Usability Labor II



Quelle:  BMW

Usability Labor III



Quelle:  silverbackapp.com

Think-aloud



Quelle: M. Herzceg

- Variante: retrospektives Lautes Denken
 - Video-Aufzeichnung der Interaktion (VAI)
 - Video-Konfrontation mit VAI mit der Aufforderung laut zu denken:
 - * Was haben Sie in dem Moment gedacht?
 - * Warum haben Sie Interaktion x ausgeführt?

Planen

- Testpersonen auswählen
 - typische Nutzer
 - * d.h. i.d.R. keine Informatiker
 - * definitiv keine Mitentwickler!
- Testaufgaben stellen
 - “echte” Arbeitssituation
 - (in sich) abgeschlossene Arbeitsaufgabe

- Testaufgabe angepasst an
 - * Testperson,
 - * Arbeitsaufgabe,
 - * Testartefakt (Gerät, Programm, etc.)

Durchführen: Schritte

Schritte bei der Durchführung

- Wichtig: der Testperson den Sinn des Tests erklären
 - wir möchten Probleme finden
 - nicht die Person testen
 - das System ist schuld, nicht die Person
 - es gibt keine richtigen/falschen Antworten!
- Aufgabe erklären
- Aufforderung laut zu denken: erklären, was sie tut und warum
 - Aufforderung muß i.d.R. öfter wiederholt werden
 - Rationalisierungen vermeiden

Durchführen: Aufgaben

- Aufgaben bei der Durchführung
 - beobachten
 - zuhören
 - aufschreiben
 - (nachfragen)
 - * Video-Konfrontation
 - * retrospektives lautes Denken
- Selbst das System studieren
 - die Versuche des Benutzers verstehen
 - warum tut der Benutzer was er tut

Berichten/Analysieren

- Berichten
 - Probleme aufschreiben
 - * Methode: post-its, post-its, post-its...
 - beim oder kurz nach dem Test
 - * Daumenregel: innerhalb von 12h
- Auswerten/ Analysieren
 - kann u.U. fast beliebig lang dauern
 - * insbesondere bei Video- oder Blickbewegungsanalyse
 - abhängig von der Fragestellung
- Ziele
 - die Versuche des Benutzers verstehen
 - warum tut der Benutzer was er tut
 - was läuft schief
 - was irritiert
 - was könnte leichter gehen

5.2 Benutzerbefragung

Benutzerbefragung

- Voraussetzung
 - Ausreichende Anzahl von ausreichend geübten Benutzern steht zur Verfügung
- Durchführung
 - Strukturiertes Interview: Benutzer werden von Experten vor Ort befragt
 - Fragebögen: Erfahrene Benutzer beantworten standardisierte Fragebögen zur ergonomischen Qualität selbst
- Auswertung
 - Statistische Analysen
- Ziel
 - Feststellung der Zufriedenstellung der Benutzer (Grobanalyse)
 - eher für die summative Evaluation geeignet
 - ebenfalls sinnvoller Einsatz: Analysephase

Strukturiertes Interview

- direkte Befragung durch einen Experten
 - strukturiert anhand von Leitfaden
- Vorteile
 - direkter Eingriff bei Mißverständnissen und Unklarheiten möglich
 - qualitativ hochwertige Daten
- Nachteile
 - hoher sozialer Druck
 - hohe soziale Kompetenz der Befragten und Befragenden notwendig
 - Große Fehlermöglichkeit durch (unfreiwillig) suggestive Fragen
 - sehr aufwendig

Durchführung

- Fragen
 - Habe ein Ziel
 - * Was will ich wissen?
 - aber formuliere die Fragen offen
 - * Keine Suggestivfragen, kein A oder B
 - Sei bereit, von Deinem Ablaufplan abzuweichen
 - Lasse Pausen zu
 - Dokumentiere die Antworten
 - * Notizblock
 - * Kamera
 - Mache einen Test bevor es ernst wird
- Durchführung
 - Wer bin ich? Was will ich?
 - Fragen & Antworten
 - Abschluß

Fragebogen

- Standardfragebogen oder produktspezifisch
 - Aber: Fragebogenkonstruktion ist nicht trivial
- asynchrone Beantwortung, freie Zeiteinteilung
- Vorteile:
 - kostengünstig
 - schnell
- Nachteile:
 - Datenschutzbedenken können zu schwächeren Aussagen führen
 - hohe Fehleranfälligkeit bei der Konstruktion und Auswertung der Fragebögen
 - Rücklaufquote häufig enttäuschend

Standard-Fragebögen

- Es können Standard-Fragebögen oder speziell für die Untersuchung erstellte Fragebögen zum Einsatz kommen
- Beispiele für Standard-Fragebögen:
 -  ISONORM-9241/10
 -  SUMI
- Probleme:
 - Kaum Zusammenhänge zwischen Zufriedenheit und objektiver Leistung
 - Selbstauskunftsdaten sind oft problematisch:
 - * soziale Erwünschtheit
 - * Tradieren von Vorurteilen

5.3 Feldstudien

Geschichte



Bronisław Malinowski, Trobriand Islands, 1918

Feldstudien

Was sind Feldstudien?

- Das Feld ist der reale Einsatzort eines Systems.
- Feldstudien sind ein Instrumentarium aus qualitativen und quantitativen Methoden, um Informationen zu sammeln über:
 - Nutzungskontext eines Systems

- Arbeitsprozesse oder andere Aktivitäten mit einem System
- Feldstudien sind Verhaltensbeobachtungen.
- Sie sollen den Gestaltungs- und Optimierungsprozess für ein interaktives System auf allen Gestaltungsebenen anleiten:
 - konzeptuelle (strukturelle) Ebene
 - logische (semantische) Ebene
 - Interaktionsebene

Ziele

- Erfassung von kontextualisierter Information
 - Tätigkeiten und zeitlichen Abläufen (Prozeß),
 - Beziehungen von Akteuren und ihrer Umgebung,
 - Verhaltensweisen und Kommunikationsweisen,
 - Umgebungsbedingungen und Arbeitsmittel,
 - Fähigkeiten und Fertigkeiten
- durch Beobachtung zur Konzeption von
 - effektiven,
 - effizienten und
 - zufriedenstellenden
 interaktiven Systemen.

Das Feld

Vorgehensweise

- kläre die Anforderungen an die Auswahl geeigneter Endbenutzer für die Interviews;
 - akzeptiere keine Benutzer, die diese Anforderungen nicht erfüllen
- führe Interviews am Arbeitsplatz der Benutzer durch;
 - dieses Setting für das Interview mag für die Benutzer ungeeignet wirken; führe es trotzdem dort durch
- betreibe den Aufwand, zu den Benutzern zu reisen;
 - es gibt keine Alternative zu Feldstudien

Feldstudien

- grundsätzliche Methoden
 1. naturalistische Beobachtungen
 2. teilnehmende Beobachtungen
 - Befragung im Kontext (Contextual Inquiry)
 3. Artefact Walkthroughs
- Die Methoden unterscheiden sich in
 - Echtzeitinformation oder Retrospektion
 - Umfang der Interaktion mit den Benutzern
 - Fokus auf den Kontext und die Tätigkeit *vs.* Fokus auf die Interaktion mit Technologie

Methoden für Feldstudien I

Naturalistische Beobachtungen

- Beobachtung von Aktivitäten ohne Einflußnahme und ohne mit den Benutzern zu sprechen.
- Vorgehensweise:
 1. Abstimmung über den Fokus der Beobachtungen
 2. Orientierung über vorhandene Rollen und aktuell laufende Aktivitäten
 3. Auswahl der zu beobachtenden Rollen und der Beobachtungsorte
 4. Einstieg in die Beobachtungen
 5. Beobachtung der Aktivitäten (Notizen, Video)
 6. Nachbesprechung und Verabschiedung (Debriefing)
 7. Auswertung des Beobachtungsmaterials, Validierung der Hypothesen und Planung der weiteren Vorgehensweise

Methoden für Feldstudien II

Artefact Walkthroughs

- Analyse von Artefakten am Arbeitsplatz und in der Arbeitsumgebung (Arbeitsmittel, Arbeitsgegenstände)
- Analyse von Zielen, Prozesse, Tätigkeiten und Verhalten.
- Typische, zu untersuchende Artefakte:
 - Berichte und andere Dokumente
 - Zeichnungen und Diagramme
 - persönliche Notizen auf Dokumenten, Post-Its
 - Organisations- und Flußdiagramme
 - E-Mail- und Telefonnotizen
 - Kalender und Zeitpläne
 - Ordner und andere Verzeichnisse
 - Bücher und anderes Bezugsmaterial
 - persönliche Konfigurationen von Systemen
- Artefakte sind beobachtbare Evidenz von Handlungen.
 - Hinterfragen der Entstehungsprozesse
 - Demonstration der Produktion und des Umgangs mit den Artefakten

Methoden für Feldstudien III

Teilnehmende Beobachtung (Contextual Inquiry)

- Gespräche mit Benutzern während diese ihre Tätigkeiten ausführen.
- Vorgehensweise:
 1. Besprechung des Beobachtungsfokus
 2. Aufbau einer vertrauensvollen Gesprächssituation (Rapport)
 3. Dialog mit dem Benutzer während der Ausübung der Tätigkeit
 4. Wrap-Up: Zusammenfassung der Erkenntnisse und Beantworten der Fragen der Benutzer
 5. Auswerten und Besprechung der Erkenntnisse
 6. Planung der weiteren Vorgehensweise

Artikulation

- Für die Ergebnisse aus Feldstudien ist die Art der Fragestellung und die Artikulationsfähigkeit von entscheidender Bedeutung.
- Erwarte nicht, daß die Benutzer ohne Hilfe artikulieren können was sie tun
 - je mehr Routine sie haben, desto weniger können sie artikulieren was sie tun und warum sie es tun
 - Tazites Wissens
 - * sie tun ihre Arbeit; sie denken nicht über ihre Arbeit nach
- Vorgehensweise:
 - Benutzen der “magic words”, d.h. handlungsbezogener Fragen:
 - * “Wann haben Sie zum letzten Mal ...”
 - * “Zeigen Sie mir doch, wie Sie das genau gemacht haben.”
 - * “Jetzt haben Sie es anders gemacht als vorhin, was war der Grund”
 - * ...

Fehlende Details

- Befragungen und die Antworten hinterlassen Lücken in der Feldstudie.
 - Benutzer sprechen oft in Generalisierungen, die als solche erkannt werden müssen
 - Abstraktionen sind hilfreich, wichtige Details gehen jedoch dabei verloren
- Vorgehensweise:
 - nicht nach Generalisierungen fragen
 - (“normalerweise”, “im Allgemeinen”, “typischerweise”)
 - am konkreten Beispiel arbeiten und dieses im Interview “ausleben”
 - Schlüsselfragen stellen (siehe Artikulation)
 - Wiederholen der Interviews in Teilen oder im Ganzen, falls Lücken oder Fehler vermutet werden.

Ehrlichkeit

- Benutzer scheuen sich aus verschiedenen Gründen oft davor Klartext zu sprechen.
- Sie verstecken ihre eigentliche Antwort in mehrdeutigen Phrasen.
 - wenn Benutzer “nein” sagen, meinen sie nicht immer nein
- Vorgehensweise:
 - Erkenne “nein”:
 - * “Was?”
 - * “Ah, vielleicht.”
 - * “ja” mit weiteren Ausführungen
 - Einhaken:
 - * “Sie würden es gut finden?”
 - * auf nicht-verbale Zeichen achten
 - * Kopfschütteln,
 - * abschweifender Blick,
 - * abwertende Handbewegungen,
 - * ...

Ineffektive Interviewstile

- Der Interviewstil beeinflusst wesentlich die Qualität der erhobenen Daten. Ungeeignet sind:
 1. Der traditionelle Interviewer
 - “ich frage – Du antwortest”
 2. Der Gerichtsreporter
 - alles notieren, ohne den Versuch Verständnis herzustellen
 3. Der Polizeibeamte
 - Grillen des Benutzers
 4. Der Agenda-Arbeiter
 - Fokus auf das Verfolgen des eigenen Plans, anstatt Fokus auf den Benutzer
 5. Der höfliche Interviewer
 - läßt den Benutzer ausschweifen oder traut sich nicht neugierig zu sein
 6. Der verrückte Erfinder
 - entwickelt und äußert ständig neue Ideen und Lösungen
- Vorgehensweise:
 - erkennen ungeeigneter Stile und Ändern der Vorgehensweise
 - Zurückgehen auf die “magic words”

Effekte bei Feldstudien I

- **Hawthorne-Effekt:**
 - Seit den 1920er-Jahren bekanntes Phänomen
 - Die Teilnehmer an einer Studie ändern ihr Verhalten wenn Sie unter Beobachtung stehen
 - * Alle Arbeiter auf der Baustelle tragen ihren Schutzhelm solange der Beobachter dabei ist
- **Novelty-Effekt:**
 - Die stärkste psychologische Antwort auf eine (potentiell gefährliche) Situation gibt es bei den ersten Begegnungen mit der Situation
 - * Wird eine neue Technologie eingeführt so kann sich die Leistung allein aufgrund der Tatsache steigern daß es eine neue Technologie ist

Effekte bei Feldstudien II

- **Observer-Expectancy-Effekt:**
 - Die unbewußte Beeinflussung des Tests durch die Erwartungen des Beobachters
 - * Der “Kluge Hans”, ein Pferd welches anscheinend zählen konnte, aber allein auf die Reaktionen des menschlichen Publikums reagierte
- **Subject-Expectancy-Effekt:**
 - Der umgekehrte Fall
 - Das Testsubjekt hat eine Erwartung, die sein Verhalten beeinflusst
 - * Placebo-Effekt

Widerstände

- Die Einführung neuer Technologien in Unternehmen kann von starken Widerständen der Betroffenen begleitet werden
- Dies gilt selbst dann, wenn die Einführenden glauben, daß die neue Technologie die Situation der Betroffenen objektiv verbessern würde.
- Nicht ohne Grund
 - Die Einführung von CAD-Systemen vereinfacht die Arbeit der Designer und Ingenieure, hat aber zu einer meßbaren Devaluierung der Arbeit geführt
- Diese Widerstände können zumindest teilweise durch *Kognitive Dissonanz* erklärt werden

Kognitive Dissonanz

- **Loss Aversion Bias:**
 - Die Tendenz, eher einen Verlust zu vermeiden, als Gewinne zu erzielen (Teil des Status Quo Bias)
- **Negativity Bias:**
 - Negativen Informationen und Erfahrungen wird stärkeres Gewicht beigemessen als positiven
- **Selective Perception Bias:**
 - Erwartungen beeinflussen die Wahrnehmung
- **Confirmation Bias:**
 - Die Tendenz, Informationen so auszuwählen, daß die eigenen Erwartungen bestätigt werden
- **Reactance Bias:**
 - Das Gegenteil von dem tun zu wollen was jemand anders von einem erwartet um das Gefühl zu haben, frei entschieden zu haben

Praxis



youtu.be/9tHgNXzS2EY 4:12

6 Quantitative Messungen

Quantitative Messungen

- Bisher vorgestellte Herangehensweisen (teilweise) sowohl für qualitative wie quantitative Messungen geeignet
 - Usability-Lab
 - Fragebögen
- Bisher haben wir die besonderen Herausforderungen quantitativer Messungen nicht weiter beachtet
 - Warum quantitative Daten erheben?
 - Was wird gemessen?
 - Wie werden die Messungen durchgeführt?
 - Wie werden die Daten aufbereitet?
- Im folgenden grob anreißen, ohne das dies eine Einführung in Quantitative Methoden ersetzen kann

6.1 Empirische Messungen

Ziele

- Wozu Usability quantitativ messen?
 - Neues System: im Vorfeld alle Usability Probleme zu lösen ist oft unmöglich
 - * messen, ob das System gut genug ist
 - Kauf: Vergleich verschiedener Produkte
 - Verbesserung eines existierenden Systems
- Vorher festlegen:
 - Was sind die Anforderungen oder Ziele?
 - Was wird gemessen?
 - Wie wird die Messung ausgeführt?
 - Wie sollen die Daten ausgewertet werden?

Anforderungen an die Messung

- Zentrale Usability Faktoren:
 - Wozu soll das System benutzt werden?
 - * Aufgabe
 - Wer soll das System benutzen?
 - * Zielgruppe
 - Wo soll das System benutzt werden?
 - * Kontext
 - Welche Strukturen beeinflussen den Einsatz?
 - * Organisation
 - Welche Gegenstände werden noch benutzt?
 - * Artefakte
- Usability ist multifaktoriell:
 - sind alle Faktoren (gleich) relevant?
 - * "Leichte Erlernbarkeit" kann z.B. durch die Anzahl der Usability-Probleme beim Erstkontakt gemessen werden

Gemessene Faktoren

- Häufig betrachtete Faktoren:
 - Bearbeitungszeit messen
 - Eingaben zählen
 - Probleme zählen
 - Verständnistest

Verwendete Methoden

- Verschiedene Methoden für
 - verschiedene Evaluationsziele (summativ *vs.* formativ)
 - verschiedene Phasen im Lebenszyklus
- Beispiel Usability Labor
 - Bearbeitungszeit
 - Anzahl Klicks, Mausbewegungen
 - Eyetracking
- Beispiel Fragebogen
 - Verständnistest

Aufbereitung der Daten

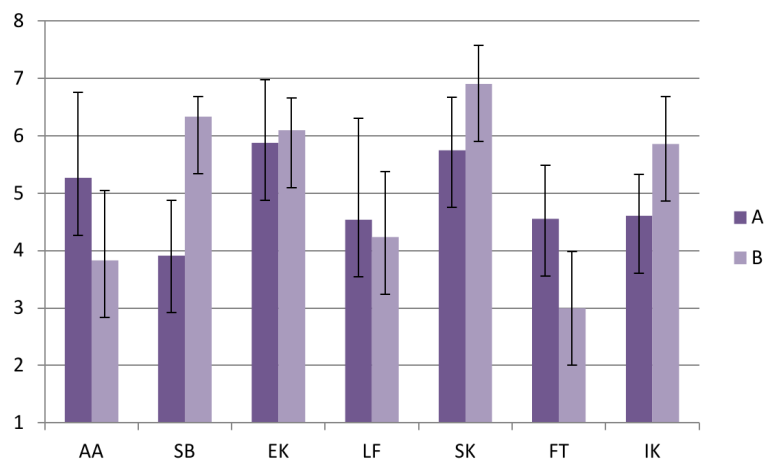
- Kategorische Variablen
 - Beispiele
 - * Geschlecht
 - * Rollen
 - Aufbereitung
 - * Häufigkeitscharts, Prozentzahlen, Proportionen
 - * Bar charts
- Ordinale Variablen
 - Beispiele
 - * Zustimmung zu Aussagen in Fragebögen
 - * Bearbeitungszeiten
 - Aufbereitung
 - * Histogramme, Häufigkeitscharts
 - * Intervalle mit gleichmäßiger Aufteilung
 - * Statistische Beschreibung (Mittelwerte, Median)
 - * Standardabweichung angeben

Beispiel: Tabelle

Faktoren	Gruppe A		Gruppe B	
	M	SD	M	SD
Aufgabenangemessenheit	5,33	0,69	5,00	0,48
Selbstbeschreibungsfähigkeit	3,92	0,96	6,33	0,35
Erwartungskonformität	5,88	1,10	6,10	0,56
Lernförderlichkeit	4,54	1,76	4,24	1,13
Steuerbarkeit	5,75	0,92	6,90	0,67
Fehlertoleranz	4,56	0,93	3,00	0,98
Individualisierbarkeit	4,61	0,72	5,86	0,82

Mittelwerte (M) und Standardabweichungen (SD) der Faktoren ermittelt mit Hilfe eines Fragebogens ISONORM 9241/110-S für Gruppen A und B (N=20).

Beispiel: Graphik



Gütekriterien

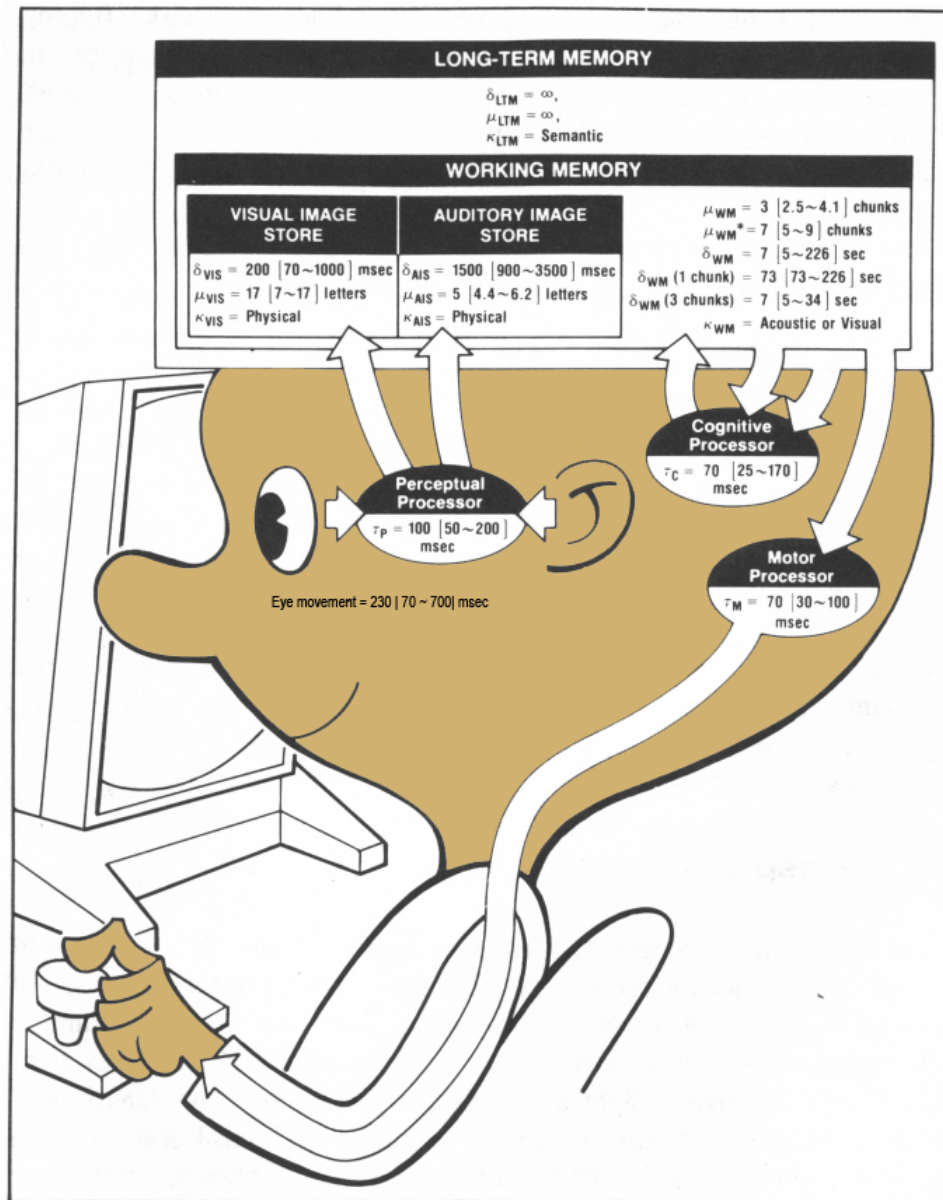
- Was ist bei der Einschätzung der verschiedenen Evaluationsansätze zu beachten?
- Reliabilität (Zuverlässigkeit):
 - Das Ausmaß, in dem wiederholte Messungen mit dem gleichen Meßinstrument bei Konstanz der zu messenden Eigenschaft die gleichen Werte liefern, bezeichnet man als Reliabilität oder Zuverlässigkeit (engl.: reliability).
- Validität (Gültigkeit):
 - Das Ausmaß, mit dem ein 'Meßinstrument' das mißt, was es messen soll, bezeichnet man als Validität oder Gültigkeit (engl.: validity).
- Folgerungen
 - Unabhängigkeit vom Prüfer
 - Vermeidung suggestiver Fragen
 - Sorgfältige Konstruktion von Benutzerbefragungen (Fragebogen)

6.2 Analytische Methoden

Analytisch und Qualitativ?

- Gerade betrachtet: empirische, quantitative Methoden
- Kann es analytische, quantitative Methoden geben?
- Rückgriff auf kognitive Modelle, wie wir sie in der Medieninformatik bereits kennengelernt haben

Model Human Processor (MHP)



Card et al. [1983]

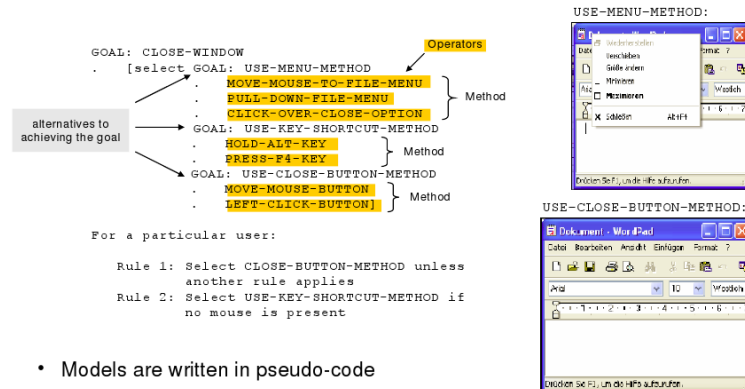
GOMS

GOMS (Goals, Operators, Methods, Selection Rules)

- Reduzieren der Interaktion auf elementare Aktionen (Operators)
- Elemente:
 - **Goals:** Was will der Benutzer erreichen
 - **Operators:** Aktionen die ausgeführt werden um das Ziel zu erreichen
 - **Methods:** Sequenz aus Operationen zur Erreichung des Ziels
 - **Selection Rules:** Auswahl aus Optionen
- Dekomposition in Teilziele
- Motivation:
 - Frühe Entscheidungen
 - Keine teuren Prototypen
 - Klare Metriken

GOMS Beispiel

- Goal: Close the window that has the focus (Windows XP)



- Models are written in pseudo-code

Beispiel: Raphael Wimmer

Varianten

- Das Original CMN-GOMS Card, Newell, and Moran (1983)
- Keystroke-Level Model (KLM-GOMS)
- Natural GOMS Language (NGOMSL)
- Cognitive, Perceptual, and Motor (CPM-GOMS)

KLM-GOMS

- Vereinfachte Variante
 - Nur Operatoren auf dem Keystroke Niveau
 - Keine Ziele
 - Keine Methoden
 - Keine Auswahlregeln
- Aussagen über die Dauer einer Aufgabe
- Dekomposition in primitive Operatoren:
 - Motorische Operationen
 - Mentale Vorbereitung
 - Systemantworten

Komponenten

- **K**, key press and release (keyboard) — 0.08-1.2 seconds
Ein typischer Wert für durchschnittlich geübte Benutzer wie Sie wäre 0.28 Sekunden.
- **P**, Point the mouse to an object on screen — 1.10 seconds
- **B**, Button press or release (mouse) — 0.10 seconds
- **H**, Hand from keyboard to mouse or vice versa — 0.40 seconds
- **M**, Mental Preparation — 1.20 seconds
- **T(n)**, Type string of characters ($n * K$ sec.)
- **W(t)**, User waiting for the system to respond — varies

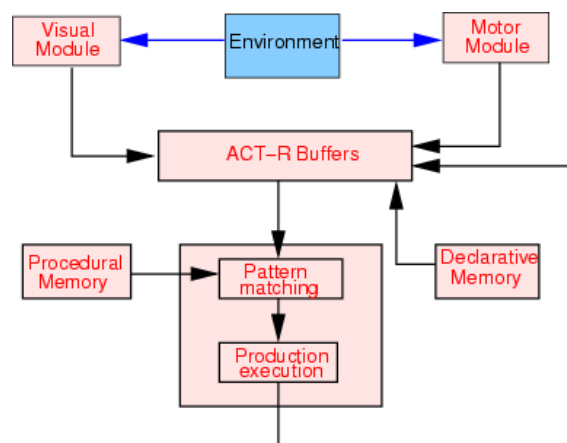
GOMS: Anwendung

- Aufgrund einer Aufgabenanalyse kann der Aufwand abgeschätzt werden, den die Bedienung eines Softwaresystems beansprucht
- Mit dem KLM-GOMS-Modell ist eine modellbasierte Analyse von Benutzungsschnittstellen möglich
- Dies beinhaltet auch Voraussagen über die Bedienbarkeit neuer Software bereits im Designstadium

ACT-R: Grundlagen

- [Anderson \[2009\]](#) Kognitive Architektur
- Theorie zur Simulation und zum Verständnis der menschlichen Kognition
- Zwei Sichtweisen
 - Ein Laufzeitsystem mit eigener Programmiersprache
 - Explikation von Annahmen über die menschliche Kognition
- Lauffähiges Modell des menschlichen Benutzers
- Hybrides System:
 - Produktionsregeln (Symbolisch)
 - Mathematische Gleichungen (Subsymbolisch)

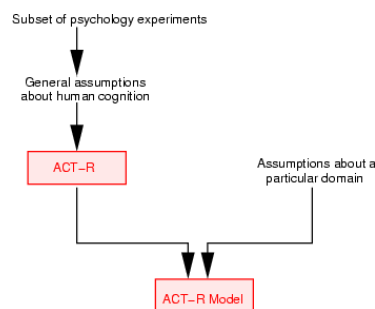
ACT-R: Buffer

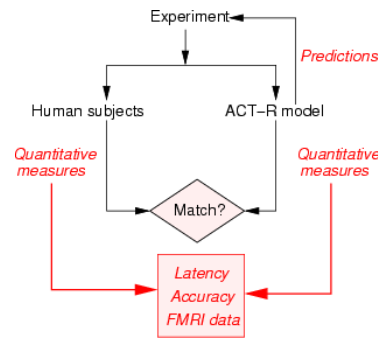


ACT-R buffers (Anderson)

ACT-R: Modell und Methode

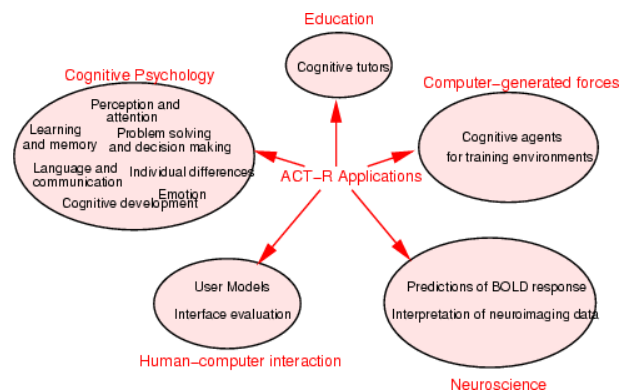
- Erstellung des Modells
- Test des Modells gegen menschliche User





ACT-R model, method (Anderson)

ACT-R: Anwendungen



ACT-R applications (Anderson)

Assignment 3.2: Evaluation

- Form groups of 3-6
- Sie möchten verschiedene Email-Clients auf Ihre Tauglichkeit für den Einsatz in einem Startup evaluieren
 - Desktop-Client (Thunderbird, Outlook)
 - Web-Client (GMail)
 - Mobile-Client (GMail for ...)
- Welche Optionen zur Evaluation haben Sie?
- Skizzieren Sie eine analytische und eine empirische Evaluationsmethode

Assignment 3.3: GOMS

- Form groups of 3-6
- Sie möchten verschiedene Email-Clients auf Ihre Tauglichkeit für den Einsatz in einem Startup evaluieren
 - Desktop-Client (Thunderbird, Outlook)
 - Web-Client (GMail)
 - Mobile-Client (GMail for ...)
- In der Vorlesung haben Sie das GOMS-Modell kennengelernt
- Eignet sich das KLM-GOMS-Modell für die Evaluierung mobiler Clients?
 - Wenn ja: wie können Sie KLM-GOMS auf das mobile System anwenden?
 - * Beschreiben Sie das Vorgehen und skizzieren die Resultate beispielhaft
 - Wenn nein: Wie müßten Sie das Modell evtl. erweitern?
 - * Designen Sie eine Methode, KLM-GOMS auf mobile Geräte zu erweitern

References

Literatur

- Anderson, J. R. (2009). *How can the human mind occur in the physical universe?* Oxford University Press.
- Card, S. K., Newell, A., and Moran, T. P. (1983). *The Psychology of Human-Computer Interaction*. L. Erlbaum Associates Inc., Hillsdale, NJ, USA.
- Chapman, C. N. and Milham, R. P. (2006). The personas' new clothes: methodological and practical arguments against a popular method. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 50, pages 634–636. SAGE Publications Sage CA: Los Angeles, CA.
- Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 152–158. ACM.
- Norman, D. A. (2005). Human-centered design considered harmful. *interactions*, 12(4):14–19.
- Pruitt, J. and Grudin, J. (2003). Personas: practice and theory. In *Proceedings of the 2003 conference on Designing for user experiences*, pages 1–15. ACM.
- Scriven, M. (1966). The methodology of evaluation. Technical Report SEC-FUB-110, Social Science Education Consortium.
- Shneiderman, S. B. and Plaisant, C. (2005). *Designing the user interface (fourth edition)*. Pearson Addison Wesley.