Universität Siegen

Wirtschaftswissenschaften
Wirtschaftsinformatik
Wirtschaftsrecht

Einheit 10:
Evaluation von Informationssystemen

Willkommen zur Vorlesung: **Wirtschaftsinformatik II**

uni-siegen.de

Prof. Dr. Rainer Wieching (vertr.)

5. April 2024



Wirtschaftsinformatik II
Roadmap

- Ethische und Soziale Fragen
- Gestaltungsrelevante Aspekte
- Organisationsstrukturen und –entwicklung
- Systementwicklung: Vorgehensweise
- ~~Systementwicklung: Modellierungsmethoden~~
- Kooperations- und Koordinationssysteme
- ~~Kommunikationssysteme~~
- Wissensmanagement
- Informationsmanagement
- **Evaluation von Informationssystemen**
- **Wrap-Up**

Die Bedeutung von Artefakten und IS in der Wirtschaftsinformatik

Der Gegenstand der Wirtschaftsinformatik ist wesentlich durchdie (sozio-technischen) Artefakte geprägt, um nicht zu sagen konstituiert:

- Informationsbestände in unterschiedlichen Ausprägungen,
- Betriebliche Anwendungssysteme,
- Daten(bank)-Architekturen,
- Software-Werkzeuge,
- Kommunikations-Standards,
- Informationsmodelle, wie auch Modellierungssprachen und diverse Analyse- und Systementwurfsmethoden

Evaluation: Software/Funktionen

Es ist fast unmöglich, alle Fehler in großen Programmen zu beseitigen und deren Ernsthaftigkeit ist unbekannt

- Hauptfehlerursache ist die Komplexität des Codes
- Validierung: Wurde das richtige System gebaut?
- Verifikation: Wurde das System richtig gebaut?
- Aufgabe des Testens erfordert Experten
- Andere Fehlerarten: Provozierte Fehler (buffer overflow), falsche Werkzeuge

Kostenaufwand Systemfehler

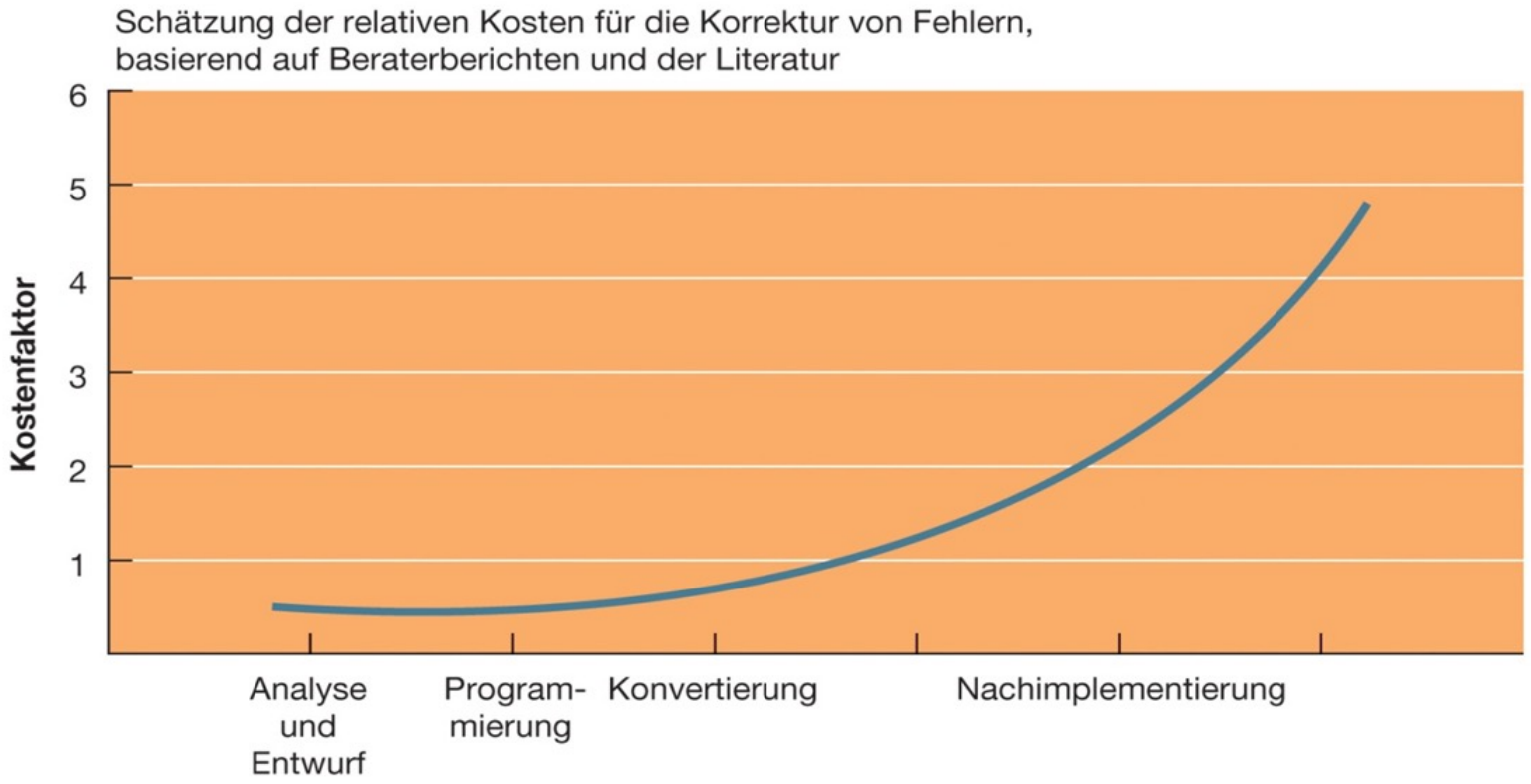


Abbildung 15.8: Kostenaufwand für Fehler im Systementwicklungszyklus
Die häufigsten, schwersten und teuersten Systemfehler entstehen in der frühen Entwurfsphase. Sie basieren oft auf einer fehlerhaften Anforderungsanalyse. Fehler in der Programmlogik oder Syntax sind weniger häufig, weniger schwerwiegend und weniger kostenaufwendig bei der Reparatur als Entwurfsfehler.

Evaluation - ein vielschichtiges Phänomen

- Evaluation wird gemeinhin verstanden als die **gezielte Bewertung** von materiellen oder immateriellen Gegenständen unter Rückgriff auf **Kriterien und Verfahren**, deren **Angemessenheit** erläutert oder begründet werden sollte
- Evaluationen dienen
  - der Vorbereitung und nicht zuletzt der Legitimation von Entscheidungen
  - zur Instrumentalisierung von Interessen
  - Zum Bemühen um *Objektivität*
    - Evaluation setzt Wahrnehmung und Urteilsvermögen voraus

Evaluation: Usability, User Experience, Akzeptanz

- Ein System, das nicht nutzbar ist, ist nicht nützlich
- Arbeitsplatz-Ergonomie
- Usability-Kriterien und –Kataloge
  - Für Bildschirmarbeitsplätze
  - Einzelplatzsysteme
  - Groupware-Systeme
  - Web-Anwendungen
- Sociability-Kriterien
  - Zu Community-Building
  - Zu nutzergenerierten Inhalten

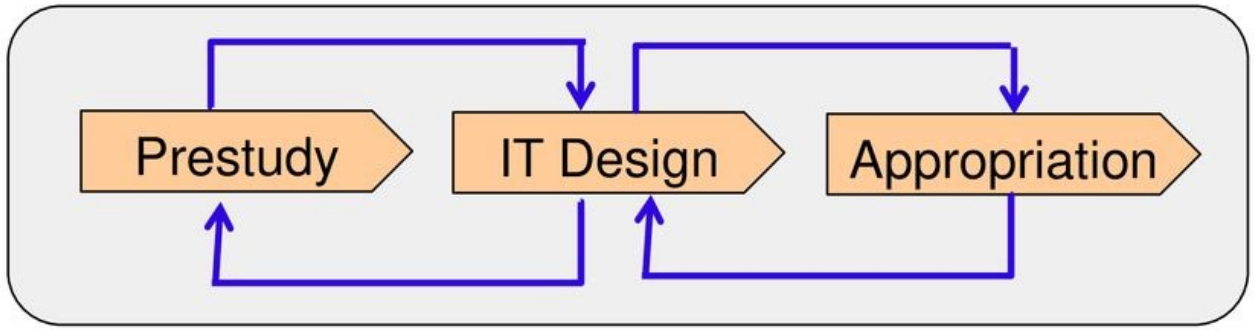
Evaluation: Nutzungsbedingungen

- Ein System, das nicht genutzt wird, ist nicht nützlich
- Motivation -> Bedürfnisse
- Kompetenz -> Qualifikation/Schulung
- Vertrauen -> Aufklärung/Transparenz
- Ängste: De-Qualifizierung, Rationalisierung, Macht-/ Kontroll-/Autonomieverlust
- Privacy -> Datenschutz, Usable Privacy, Privacy by Design
- Werte, Normen: Egoismus, Altruismus, Kollektivismus etc.
- Zufriedenheit -> Spaß, Wohlbefinden, Experience.

Evaluation: Aneignung und Anpassung

- Systeme werden von Benutzern **anders angeeignet als geplant (unintended use)**, erwartet, vorhergesehen
- Systeme werden in der Aneignung verändert und umgenutzt
- Per Anpassungsmaßnahmen und/oder Nachbesserung muss auf Dysfunktionalitäten reagiert werden
- Durch Aneignung und Anpassung verändert sich die Praxis und die Ausgangslage
- Prozessevaluation -> Design Case Studies (Wulf et al. 2011)

Prozess-Evaluation: Design Case Studies (Wulf 2011)



Practice > Design Space

Pre-study

- Understanding the existing practice
- Technological, organizational and social perspective
- (usually) Qualitative empirical research
- Identify problem/need/opportunity

Technology > Design

Design

- Applying technological options to a specific problem/need/opportunity
- Describe process and product
- Participatory approach
- Possible design iterations

Usage > Changing Practice

Appropriation

- Reflection with practitioners
- Repeated study to capture long-term effects
- Amalgated with pre-study efforts

Evaluation: Nutzung und Arbeitspraxis

- „*Double Hermeneutics*“ (Giddens 1984): Das System verändert die Praxis, die Praxis verändert das System
- Neue Probleme und Herausforderungen
- Prozessänderungen/Änderungscontrolling
- Re-Justierung von Strategien, Modellen und Geschäftsprozessen
- Kulturelle Faktoren: Änderung Betriebsklima, Kommunikationsstile, Konfliktmanagement
- Finanzielle Faktoren: Kosten-/Nutzenanalyse, Controlling

Kriterien für die Wahl einer spezifischen Evaluationsmethode

Orientierung der Evaluationsmethode anhand folgende Fragen (Karat 1997):

- **Was ist das Ziel der Evaluation?**

Geht es um einen abschließenden Test des Gesamtsystems, oder sollen bestimmte Aspekte überprüft werden?

- **Wer soll die Evaluation durchführen?**

Ist es besser, dass ein Experte das System inspiziert, oder sollten auch Benutzer die Evaluation durchführen?

- **Wo soll die Evaluation durchgeführt werden?**

Soll sie in einem Labor oder einem Arbeitsplatz durchgeführt werden?

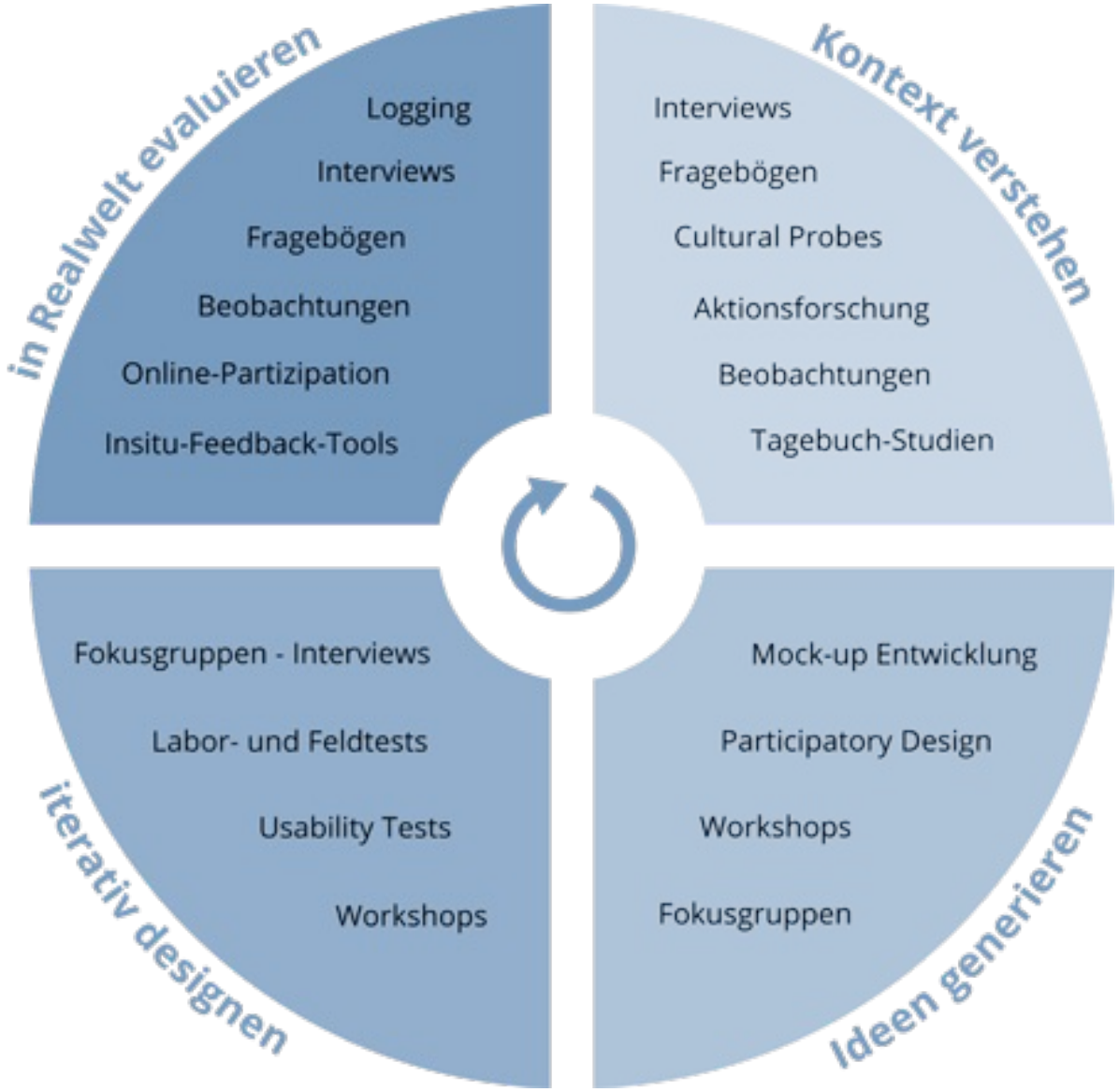
- **Welche Informationen werden gesammelt?**

Was soll gemessen werden, welche Daten sind von Bedeutung?

- **Wie groß sind die Ressourcen?**

Wie viel Geld und Zeit stehen für die Evaluation zur Verfügung?

Evaluation: Methoden und Assessments



Evaluation: Methoden und Assessments

- Quantitative Evaluationsmethoden
  - Mess- und Testmethoden
  - Operationalisierung: Kenn- und Vergleichswerte
  - Objektivität, Generalisierbarkeit, hohe interne Validität (niedrige externe Validität), Homogenität
- Qualitative Evaluationsmethoden
  - Explorations-/“Erforschungs“methoden
  - Verstehen und Lernen
  - Subjektivität, Situiertheit, hohe externe Validität (auf Kosten interner Validität), Inhomogenität

Ausgewählte Methoden und Assements in der Evaluation

- Seit Anfang der 1980er beschäftigt sich das Fachgebiet der Mensch-Maschine-Interaktion mit der Entwicklung von Konventionen und Regeln zur Gestaltung interaktiver Benutzerschnittstellen.

- Auch die Begriffe **Usability** und **User Experience** sind genormt und finden sich in der DIN EN ISO 9241 wieder.

- Die ISO 9241 beschreibt **Usability** als *das Ausmaß, in dem ein Produkt durch bestimmte Benutzer in einem bestimmten Anwendungskontext genutzt werden kann, um bestimmte Ziele effektiv, effizient und zufriedenstellend zu erreichen*.

Ausgewählte Methoden und Assements in der Evaluation

- Auch die Begriffe **Usability** und **User Experience** sind genormt und finden sich in der DIN EN ISO 9241 wieder.

- Der Begriff **User Experience** wird in der ISO Norm 9241 - 210 definiert. Demnach beschreibt die User Experience „A person's perceptions and responses that result from the use and/or anticipated use of a product, system or service.”

- User Experience beinhaltet also alle **Effekte**, die die Nutzung einer Bedienoberfläche vor, während und nach der Nutzung auf einen Nutzer hat.

UX Testing & Nutzerforschung

- **Usability-Test**
- Bei einem Usability-Test (auch Nutzertest) führen repräsentative Nutzer typische Aufgaben mit einem System aus, um Probleme, aber auch positive Aspekte des Systems aufzudecken.
- Die Interaktion des Nutzers wird dabei meist aufgezeichnet. Seine Wahrnehmungen und Meinungen werden im Interview mit Fragebögen oder mit der Methode des Lauten Denkens erfasst.



https://www.testingtime.com/blog/usability-test/

UX Testing & Nutzerforschung

- ## Usability-Test

WANN: Frühe Prototypen, Zwischenversionen, fertige Produkte

WER: Repräsentative Nutzer + Testleiter, optional Protokollant

DAUER: Ca. 1 Std. pro Nutzer plus Auswertung

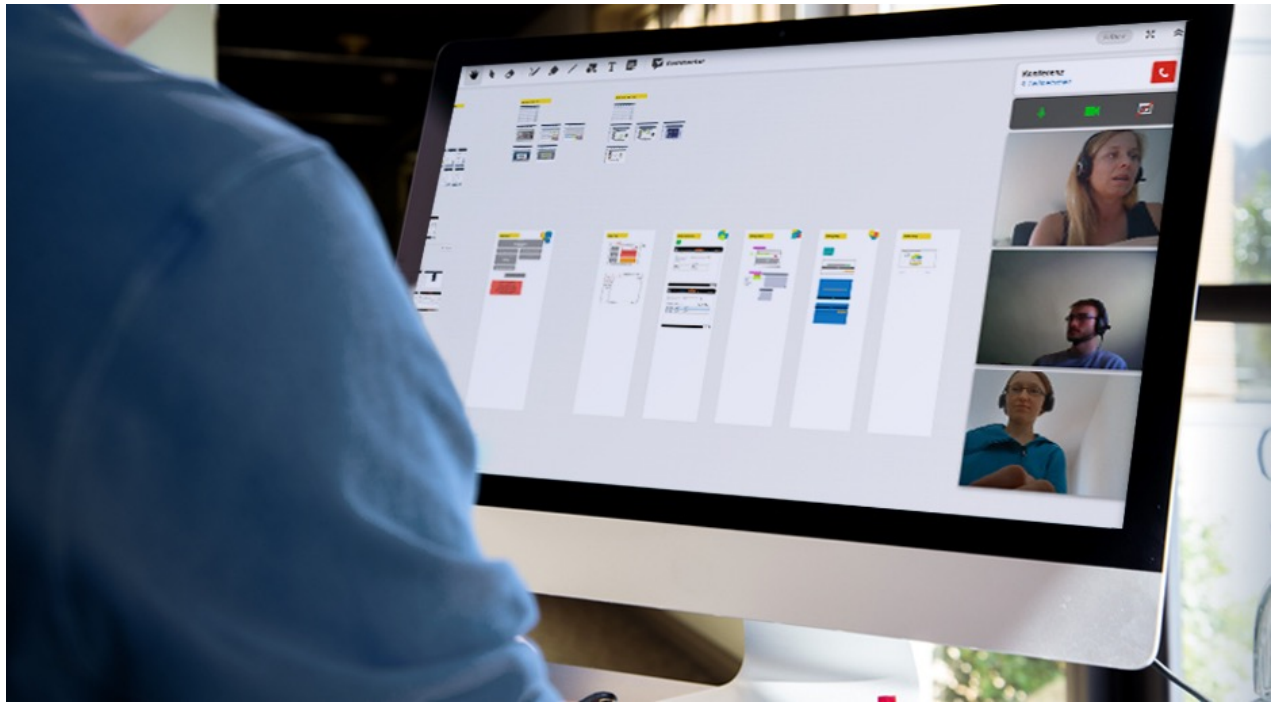
WOMIT: Optional Aufzeichnungsgeräte (Screencapture-Software, Audiorekorder, Kamera)



(Foto: YP_Studio/Shutterstock)

UX Testing & Nutzerforschung

- **Remote UX Research**
- Der Remote Usability Test (auch: Online Usability Test) ähnelt dem „klassischen“ Usability-Test stark, allerdings sind hier **Testleiter und Nutzer räumlich voneinander** getrennt.
- Häufig kann mit geringem Aufwand eine Testung von Nutzern zu Hause oder am Arbeitsplatz erfolgen.
- Unabhängig ob Smartphone, Tablet oder Desktop.



https://www.uid.com/de/schwerpunkte/remote-ux

UX Testing & Nutzerforschung

- **Cognitive Walkthrough**
- Beim Cognitive Walkthrough evaluiert ein Usability-Experte das interaktive Systems aus Nutzersicht, indem er sich in die Rolle eines Nutzers versetzt. Dabei führt er konkrete Aufgaben mit dem interaktiven System durch und prüft dabei die Erfüllung der Usability-Kriterien aus der Sicht einer Persona:
  - kann die Aufgabe ohne Umwege zufriedenstellend erledigt werden?
  - Gibt die interaktive Anwendung ausreichend Hinweise und Feedback um Fortschritt und nächste Schritte zu erkennen?



https://www.user-experience-methods.com/evaluation/cognitive-walkthrough.html

User Experience messen mit UX KPIs

- Nutzerzentrierte UX KPIs beziehen sich darauf, wie Nutzer Produkte nutzen sich verhalten und was sie sagen. Solche Metriken können zum Beispiel in einem Usability-Test gemessen oder in einer Online-Umfrage erhoben werden.



Aufgabenerfolgsrate (Task succes rate)
Anteil der Nutzer, die eine Aufgabe erfolgreich beenden.



Zeit pro Aufgabe (Time on task)
Zeit, die ein Nutzer durchschnittlich benötigt, um eine Aufgabe erfolgreich zu beenden.



Ease of Use
Standardisierte Abfrage der empfundenen Schwierigkeit bei der Lösung einer Aufgabe.



System Usability Scale (SUS)
Standardisierte Abfrage der empfundenen Nutzerfreundlichkeit mittels eines Fragebogens mit 10 Fragen.



User Experience Questionnaire (UEQ)
Standardisierte Abfrage des Nutzererlebnisses durch 26 gegensätzliche Adjektiv-Paare.

https://www.usability.de/leistungen/methoden/ux-kpis.html

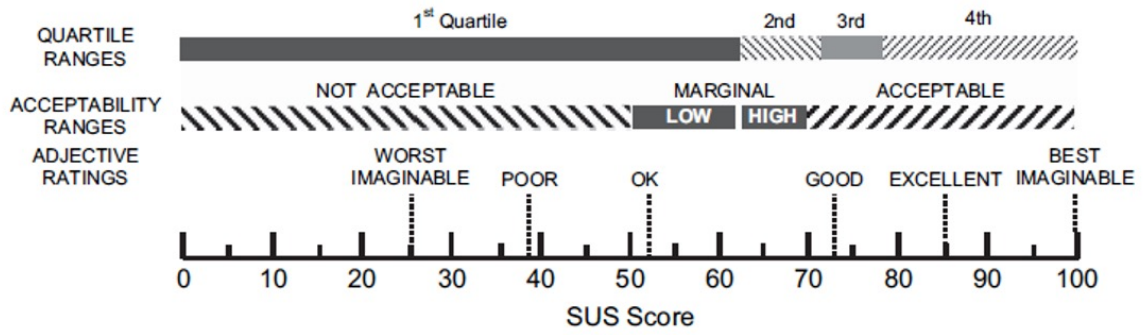
Fragebögen

- Mit Fragebögen (schriftliche Befragung, Umfrage) können leicht viele Daten erhoben und statistisch ausgewertet werden. Im Fokus stehen dabei die Meinungen und Bewertungen des Nutzers zu einem System. Im Gegensatz zum Interview können aufwandsarm viele Nutzer befragt werden. Die Antworten können leicht zwischen Nutzergruppen sowie zwischen Versionen eines Systems verglichen werden.
- **UEQ - User Experience Questionnaire, SUS - System Usability Scale, DART**
- AttrakDiff, ISONORM 9241/10, QUIS - Questionnaire for User Interaction Satisfaction
- SUMI - Software Usability Measurement Inventory
- Akzeptanzmodell von Degenhardt, Technologieakzeptanzmodell von Davis (TAM),
- Task Technology Fit Model. Dynamisches Akzeptanzmodell

Fragebögen

SUS - System Usability Scale

- Mithilfe eines einfachen Fragebogens, der aus insgesamt zehn Fragen auf Basis von Likert-Skalen mit je fünf Optionen besteht, kann ermittelt werden, wie nutzerfreundlich eine Software wahrgenommen wird.
- Der SUS-Fragebogen enthält fünf positiv und fünf negativ formulierte Aussagen zur Usability des zu bewertenden Systems



Bewertung der Bedien- und Nutzbarkeit des MobiAssist Systems durch System Usability Scale (SUS)

*Der nachfolgende Abschnitt bezieht sich auf **Gedanken und Gefühle** die Sie eventuell während der Benutzung des MobiAssist Systems empfunden haben. Bitte Kreisen Sie bei jeder der nachfolgenden Aussagen die Zahl ein, die auf Sie am ehesten zutrifft.*

Es ist jeweils nur eine Antwort möglich.

Aussage	trifft überhaupt nicht zu ⟷ trifft voll zu				
1. Ich denke, dass ich das System gerne häufig benutzen würde.	1	2	3	4	5
2. Ich fand das System unnötig komplex.	1	2	3	4	5
3. Ich fand das System einfach zu benutzen.	1	2	3	4	5
4. Ich glaube, ich würde die Hilfe einer technisch versierten Person benötigen, um das System benutzen zu können.	1	2	3	4	5
5. Ich fand, die verschiedenen Funktionen in diesem System waren gut integriert.	1	2	3	4	5
6. Ich denke, das System ist in sich unstimmig.	1	2	3	4	5
7. Ich kann mir vorstellen, dass die meisten Menschen den Umgang mit diesem System sehr schnell lernen.	1	2	3	4	5
8. Ich fand das System sehr umständlich zu nutzen.	1	2	3	4	5

Fragebögen

UEQ - User Experience Questionnaire

Attraktivität

unerfreulich / erfreulich
gut / schlecht
abstoßend / anziehend
unangenehm / angenehm
schön / hässlich
sympathisch / unsympathisch

Designqualität

Stimulation

erfrischend / einschläfernd
langweilig / spannend
uninteressant / interessant
abwechslungsreich / eintönig

Novelty

kreativ / phantasielos
neu / alt
veraltet / modern
unauffällig / auffällig

Benutzungsqualität

Effizienz

schnell / langsam
ineffizient / effizient
stockend / flüssig
aufgeräumt / überladen

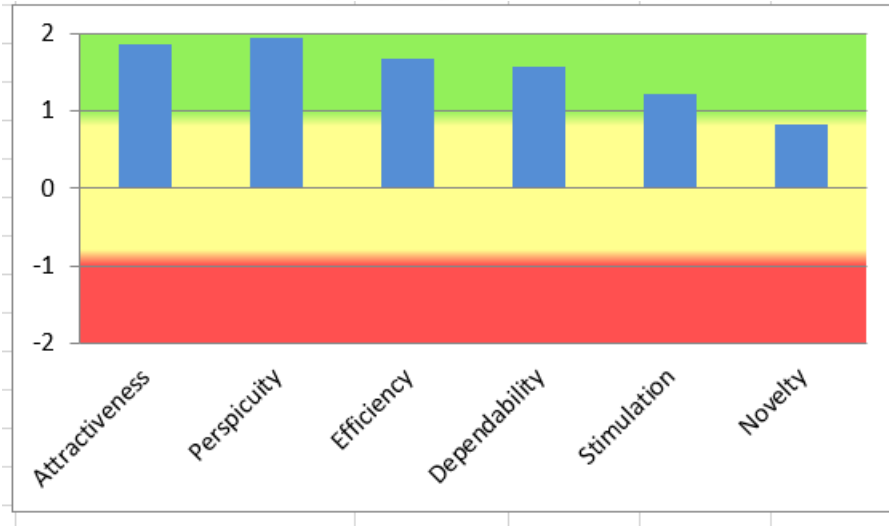
Durchschaubarkeit

unverständlich / verständlich
leicht zu lernen / schwer zu lernen
kompliziert / einfach
übersichtlich / verwirrend

Verlässlichkeit

unberechenbar / voraussagbar
bedienbar / unbedienbar
vorhersagbar / unvorhersagbar
zuverlässing / unzuverlässig

leicht zu lernen		schwer zu lernen 4
wertvoll		minderwertig 5
langweilig		spannend 6
uninteressant		interessant 7
unberechenbar		voraussagbar 8
schnell		langsam 9
originell		konventionell 10
behindernd		unterstützend 11
gut		schlecht 12
kompliziert		einfach 13
abstoßend		anziehend 14
herkömmlich		neuartig 15
unangenehm		angenehm 16
sicher		unsicher 17
aktivierend		einschläfernd 18
erwartungskonform		nicht erwartungskonform 19
ineffizient		effizient 20
übersichtlich		verwirrend 21
unpragmatisch		pragmatisch 22
aufgeräumt		überladen 23
attraktiv		unattraktiv 24
sympathisch		unsympathisch 25
konservativ		innovativ 26



Fragebögen

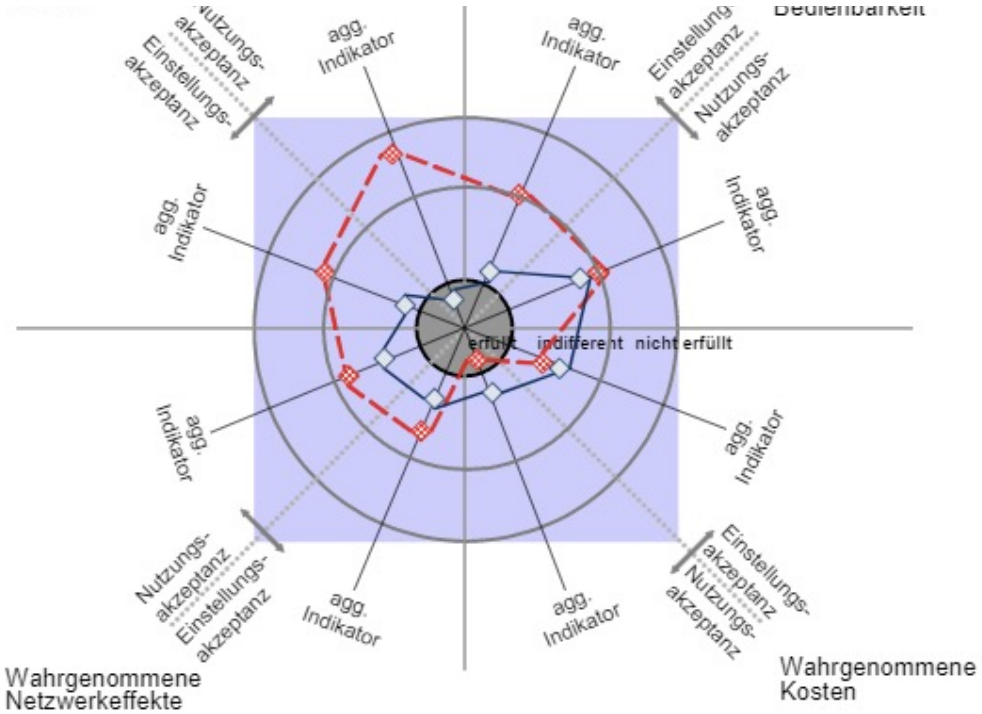
DART - Dynamisches Akzeptanzmodell zur Re-Evaluierung von Technologiebasierten Anwendungen (Amberg 2003)

- Zur Bewertung der Benutzerakzeptanz wird im DART ein **ausgewogenes Verhältnis** von individuell messbaren Akzeptanzkriterien verwendet.
- Ausgerichtet auf die jeweilige Anwendung werden die relevanten Einflussfaktoren systematisch identifiziert, unabhängig davon, ob sie direkt oder indirekt, rational oder emotional auf die Benutzerakzeptanz wirken.
- Die ausgewogene Betrachtung von **mehreren akzeptanzrelevanten Kriterien** fahrt zu einer differenzierten Bewertung der Benutzerakzeptanz. Der DART-Ansatz kann als **stetiges Steuerungsinstrument** für die wiederholte Anpassung einer Anwendung an die Bedürfnisse der Anwender aufgefasst werden.

1.2 Wie wichtig sind Ihnen folgende Punkte in Bezug auf das Gesamtsystem?

Das System sollte…

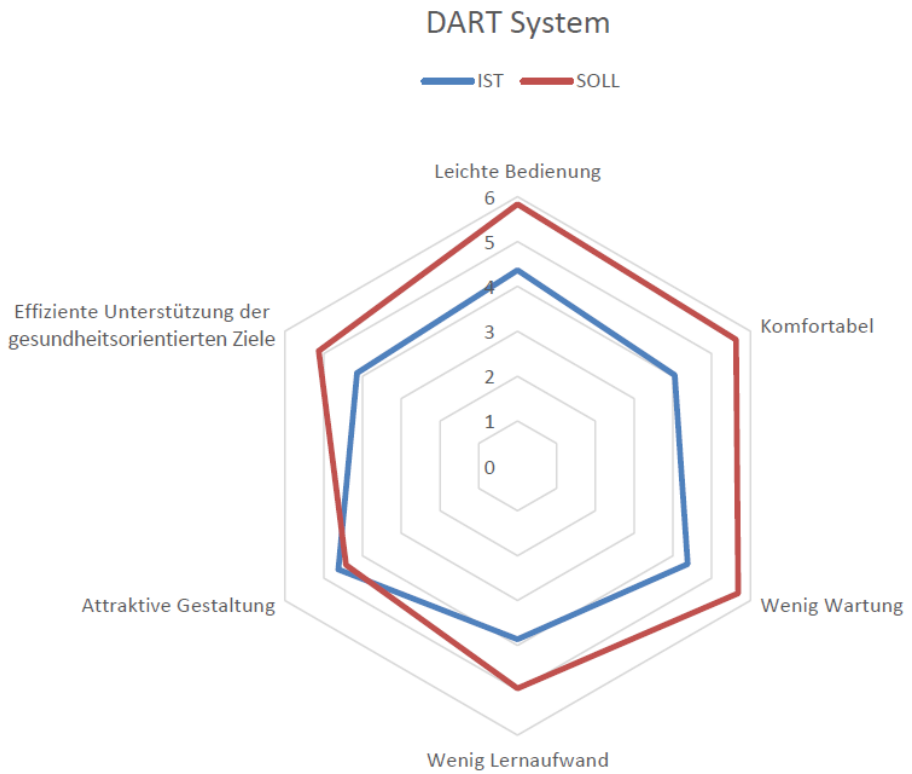
	Völlig unwichtig					Sehr wichtig
-leicht zu bedienen sein	☐	☐	☐	☐	☐	☐
-komfortabel / bequem sein	☐	☐	☐	☐	☐	☐
-wenig Wartung benötigen (z.B. Ein- und Ausschalten, Aufladen des Gerätes)	☐	☐	☐	☐	☐	☐
-wenig lernaufwendig sein	☐	☐	☐	☐	☐	☐
-attraktiv gestaltet sein	☐	☐	☐	☐	☐	☐
-meine Fähigkeiten verbessern	☐	☐	☐	☐	☐	☐
-meine gesundheitsorientierten Ziele effizient unterstützen	☐	☐	☐	☐	☐	☐



1.3 Inwieweit hat das Gesamtsystem die oben bewerteten Punkte erfüllt?

Das System…

	Überhaupt nicht					Voll und ganz
-ist leicht zu bedienen	☐	☐	☐	☐	☐	☐
-ist komfortabel / bequem	☐	☐	☐	☐	☐	☐
-benötigt wenig rüstungssaufwand (z.B. Ein- und Ausschalten, Aufladen des Gerätes)	☐	☐	☐	☐	☐	☐
-ist wenig lernaufwendig	☐	☐	☐	☐	☐	☐
-ist attraktiv gestaltet	☐	☐	☐	☐	☐	☐
-ist nicht unangenehm oder peinlich beim Tragen	☐	☐	☐	☐	☐	☐
-unterstützt meine gesundheitsorientierten Ziele effizient	☐	☐	☐	☐	☐	☐



Vielen Dank!

Kontakt

Fakultät III
Kohlbettstraße 15
57223 Siegen

rainer.wieching@uni-siegen.de