

---

# Interaction Design Evaluation

---

Daniel Ullrich

Mensch-Maschine-Interaktion

Institut für Informatik

Ludwig-Maximilians-Universität München

# Agenda

---

— Sind und Zweck der Evaluation

— Evaluationskategorien

- Formativ vs. Summativ
- Analytisch vs. empirisch

— Methoden

- Cognitive Walkthrough
- Heuristische Evaluation
- Fokusgruppe
- Experiment

— Methodensteckbriefe

— Statistische Verfahren

---

**Ein Beispiel...**

---

# Szenario: SD-Karte bei Amazon kaufen.



## SPORT AM LIMIT

ALLES FÜR ADRENALINJUNKIES  
IM ACTIONSPORT SHOP

### Das interessiert Kunden aktuell



Amazon nutzt Cookies. [Was sind Cookies?](#)

**amazonbuyvip** Der Amazon Shopping Club

Ab **39,90€**

**Levi's**

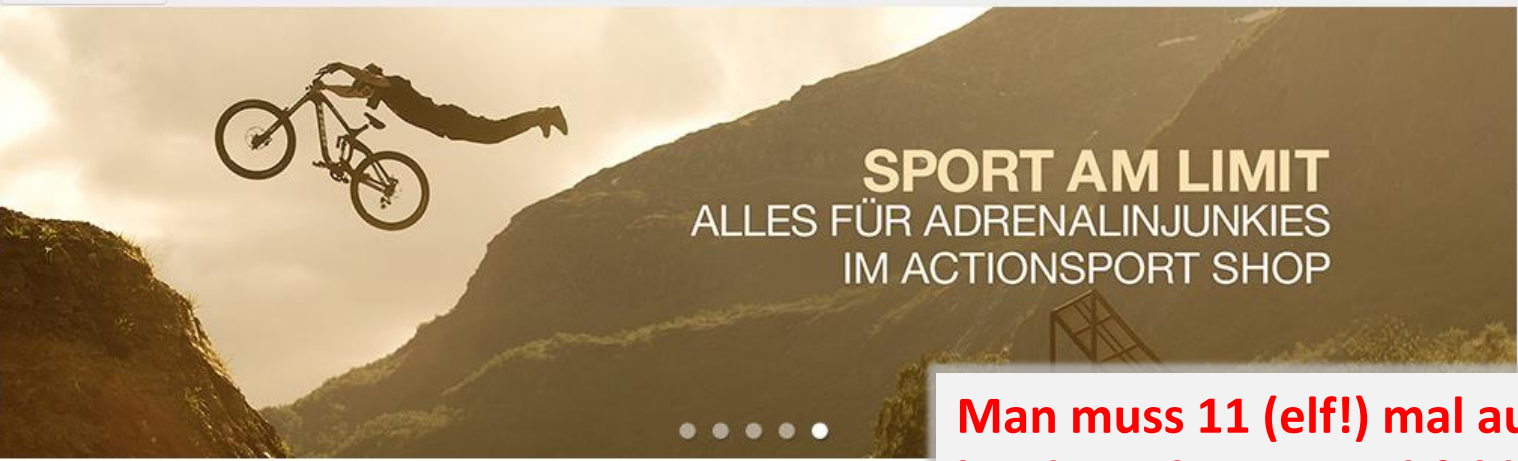
Vom 8.6. – 12.6.2015

[> Jetzt anmelden](#)

### Kaffee für Genießer: Beliebte Vollautomaten [Mehr](#)



[Anzeige-Feedback](#)



SPORT AM LIMIT ALLES FÜR ADRENALINJUNKIES IM ACTIONSPORT SHOP

Man muss 11 (elf!) mal auf TAB drücken, bis der Fokus im Suchfeld ist

Das interessiert Kunden aktuell



amazonbuyvip Der Amazon Shopping Club Ab 39,90€ Levi's Vom 8.6. - 12.6.2015 > Jetzt anmelden

Kaffee für Genießer: Beliebte Vollautomaten Mehr





# Amazon Fashion DER SOMMER IST DA!

Für Sommertage am Pool, Gartenpartys  
oder den Urlaub zu Hause: Wir haben alles,  
was Sie dafür brauchen



## Das interessiert Kunden aktuell



Amazon nutzt Cookies. Was sind Cookies?

amazonbuyvip Der Amazon Shopping Club

Ab **39<sup>90</sup>€**

**Levi's**

Vom 8.6. – 12.6.2015

> Jetzt anmelden

## Kaffee für Genießer: Beliebte Vollautomaten Mehr



Anzeige-Feedback



# Amazon Fashion DER SOMMER IST DA!

Für Sommertage am Pool, Gartenpartys oder den Urlaub zu Hause: Wir haben alles, was Sie dafür brauchen



## Das interessiert Kunden aktuell



Amazon nutzt Cookies. Was sind Cookies?

amazonbuyvip Der Amazon Shopping Club

Ab 39,90€

**Levi's**

Vom 8.6. – 12.6.2015

> Jetzt anmelden

## Kaffee für Genießer: Beliebte Vollautomaten Mehr



Anzeige-Feedback

- Alle Kategorien
- Amazon Instant Video
- Musik-Downloads
- App-Shop für Android  
Eine Gratis-App. Jeden Tag.
- Amazon Cloud Drive
- Kindle eReader & Bücher
- Fire-Tablets
- Amazon Fire TV
- Fire Phone
- Bücher
- Filme, Serien, Musik & Games
- Elektronik & Computer
- Haushalt, Garten, Baumarkt
- Beauty, Drogerie & Lebensmittel
- Spielzeug & Baby
- Kleidung, Schuhe & Uhren
- Sport & Freizeit
- Auto & Motorrad
- Alle Kategorien



Hallo! Anmelden Mein Konto Testen Sie Prime Einkaufswagen Wunschzettel



# Amazon Fashion DER SOMMER IST DA!

Für Sommertage am Pool, Gartenpartys oder den Urlaub zu Hause: Wir haben alles, was Sie dafür brauchen.



Gerät man auf eines der zahlreichen Dropdowns, klappt es auf und klaut den Fokus...

amazonbuyvip Der Amazon Shopping Club

Ab 39<sup>90</sup>€  
**Levi's**  
Vom 8.6. – 12.6.2015  
[Jetzt anmelden](#)

Kaffee für Genießer: Beliebte Vollautomaten [Mehr](#)





- Alle Kategorien
- Amazon Instant Video
  - Musik-Downloads
  - App-Shop für Android  
Eine Gratis-App. Jeden Tag.
  - Amazon Cloud Drive
  - Kindle eReader & Bücher
  - Fire-Tablets
  - Amazon Fire TV
  - Fire Phone
  - Bücher
  - Filme, Serien, Musik & Games
  - Elektronik & Computer
  - Haushalt, Garten, Baumarkt
  - Beauty, Drogerie & Lebensmittel
  - Spielzeug & Baby
  - Kleidung, Schuhe & Uhren
  - Sport & Freizeit
  - Auto & Motorrad
  - Alle Kategorien

# Amazon Fashion DER SOMMER IST DA!

Für Sommertage am Pool, Gartenpartys oder den Urlaub zu Hause: Wir haben alles, was Sie dafür brauchen.

Gerät man auf eines der zahlreichen Dropdowns, klappt es auf und klaut den Fokus...



Kaffee für Genießer: Beliebte Vollautomaten Mehr



**Gefixt!**  
(nach 5 Jahren...)

1-16 von 125.770 Ergebnissen in "sd-karte" Abteilung zum Sortieren auswählen

Ergebnisse anzeigen für

Verwandte Suchbegriffe: micro sd karte, sd karte 64 gb, sd karte 8gb.

Elektronik & Foto >

- Speicherkarten
- Micro SD Speicherkarten
- SecureDigital-Cards
- + Weitere

Prime Instant Video >

- Prime Instant Video Serien
- Prime Instant Video Filme

Computer & Zubehör >

- USB-Sticks
- + Weitere

+ Alle 31 Kategorien

Filtern nach

Versandoption (Was ist das?)  
 Kostenlose Lieferung ab EUR 29 Bestellwert

Speicherkartenkapazität

- Bis 1,9 GB
- 2 GB
- 4 GB
- 8 GB
- 16 GB
- 32 GB
- 64 GB
- 128 GB & mehr

Marke

- SanDisk
- Samsung
- Transcend
- Kingston
- Qumox



Transcend Ultimate-Speed SDHC Class 10 UHS-1 32GB Speicherkarte (b 90MB/s Lesen) [Amazon Frustrfreie Verpackung]

von Transcend  
EUR 17,49 ✓Prime  
Lieferung bis zum **Dienstag, 9. Juni**

★★★★★ 799

Bestseller Nr. 1 in SecureDigital-Card  
Kostenlose Lieferung möglich.  
Elektronik & Foto: Alle 74.125 Artikel ansehen



SanDisk SDSUN-016G-FFP Ultra SDHC 16GB UHS-I Class 10 Speicherkarte bis zu 40MB/Sek. lesen [Amazon Frustrfreie...]

von SanDisk  
EUR 8,99 ~~EUR 10,99~~ ✓Prime  
Lieferung bis zum **Dienstag, 9. Juni**

★★★★★ 106

Kostenlose Lieferung möglich.  
Elektronik & Foto: Alle 74.125 Artikel ansehen



Transcend TS32GUSDHC10E Class 10 Extreme-Speed microSDHC 32GB Speicherkarte mit SD-Adapter [Amazon Frustrfreie...]

von Transcend  
EUR 12,49 ✓Prime  
Lieferung bis zum **Dienstag, 9. Juni**

★★★★★ 2.768

Kostenlose Lieferung möglich.  
Elektronik & Foto: Alle 74.125 Artikel ansehen



Kingston SDC4/8GB microSDHC Karte - Klasse 4 - 8GB mit SD-Adapter

von Kingston  
EUR 4,89 **Plus Produkt**  
Fügen Sie ihn zu einer qualifizierten Bestellung hinzu, um ihn am **Mittwoch, 10. Juni** zu erhalten

★★★★★ 975

Kostenlose Lieferung möglich.  
Produktmerkmale  
... SD Card Association ... Size SDHC-t

1-16 von 125.770 Ergebnissen in "sd-karte" Abteilung zum Sortieren auswählen

- Ergebnisse anzeigen für
- Elektronik & Foto >
    - Speicherkarten
    - Micro SD Speicherkarten
    - SecureDigital-Cards
    - + Weitere
  - Prime Instant Video >
    - Prime Instant Video Serien
    - Prime Instant Video Filme
  - Computer & Zubehör >
    - USB-Sticks
    - + Weitere
  - + Alle 31 Kategorien

- Filtern nach
- Versandoption (Was ist das?)  
 Kostenlose Lieferung ab EUR 29 Bestellwert
- Speicherkartenkapazität  
 Bis 1,9 GB  
 2 GB  
 4 GB  
 8 GB  
 16 GB  
 32 GB  
 64 GB  
 128 GB & mehr
- Marke  
 SanDisk  
 Samsung  
 Transcend  
 Kingston  
 Qumox

Verwandte Suchbegriffe: micro sd karte, sd karte 64 gb, sd karte 8gb.



Transcend Ultimate-Speed SDHC Class 10 UHS-1 32GB Speicherkarte (b  
90MB/s Lesen) [Amazon Frustrfreie Verpackung]  
von Transcend

EUR 17,49 ✓Prime  
Lieferung bis zum **Dienstag, 9. Juni**

★★★★★ 799  
Bestseller Nr. 1 in SecureDigital-Card



SanDisk SDSUN-016G-FFP Ultra Speicherkarte bis zu 40MB/Sek. lese  
von SanDisk

EUR 8,99 ~~EUR 10,99~~ ✓Prime  
Lieferung bis zum **Dienstag, 9. Juni**

★★★★★ 106  
Kostenlose Lieferung möglich.  
Elektronik & Foto: Alle 74.125 Artikel ansehen



Transcend TS32GUSDHC10E Class 10 Extreme-Speed microSDHC 32GB Speicherkarte mit SD-Adapter [Amazon Frustrfreie...]  
von Transcend

EUR 12,49 ✓Prime  
Lieferung bis zum **Dienstag, 9. Juni**

★★★★★ 2.768  
Kostenlose Lieferung möglich.  
Elektronik & Foto: Alle 74.125 Artikel ansehen



Kingston SDC4/8GB microSDHC Karte - Klasse 4 - 8GB mit SD-Adapter  
von Kingston

EUR 4,89 **Plus Produkt**  
Fügen Sie ihn zu einer qualifizierten Bestellung hinzu, um ihn am **Mittwoch, 10. Juni** zu erhalten

★★★★★ 975  
Kostenlose Lieferung möglich.  
Produktmerkmale  
... SD Card Association ... Size SDHC-t

**125k Ergebnisse. Aber sortieren funktioniert nur in einer Sub-Kategorie...**

1-16 von 125.770 Ergebnissen in "sd-karte" Abteilung zum Sortieren auswählen

- Ergebnisse anzeigen für
- Elektronik & Foto >
    - Speicherkarten
    - Micro SD Speicherkarten
    - SecureDigital-Cards
    - + Weitere
  - Prime Instant Video >
    - Prime Instant Video Serien
    - Prime Instant Video Filme
  - Computer & Zubehör >
    - USB-Sticks
    - + Weitere
  - + Alle 31 Kategorien

Verwandte Suchbegriffe: micro sd karte, sd karte 64 gb, sd karte 8gb.



Transcend Ultimate-Speed SDHC Class 10 UHS-1 32GB Speicherkarte (b 90MB/s Lesen) [Amazon Frustrfreie Verpackung]  
von Transcend

EUR 17,49 ✓Prime  
Lieferung bis zum **Dienstag, 9. Juni**

★★★★★ 799  
Bestseller Nr. 1 in SecureDigital-Card



SanDisk SDSDUN-016G-FFP Ultra Speicherkarte bis zu 40MB/Sek. lesen  
von SanDisk

EUR 8,99 ~~EUR 10,99~~ ✓Prime  
Lieferung bis zum **Dienstag, 9. Juni**



Transcend TS32GUSDHC10E Class 10 Extreme-Speed microSDHC 32GB Speicherkarte mit SD-Adapter [Amazon Frustrfreie Verpackung]  
von Transcend

EUR 12,49 ✓Prime  
Lieferung bis zum **Dienstag, 9. Juni**



Kingston SDC4/8GB microSDHC Karte - Klasse 4 - 8GB mit SD-Adapter  
von Kingston

EUR 4,89 **Plus Produkt**  
Fügen Sie ihn zu einer qualifizierten Bestellung hinzu, um ihn am **Mittwoch, 10. Juni** zu erhalten

★★★★★ 975  
Kostenlose Lieferung möglich.  
Produktmerkmale  
... SD Card Association ... Size SDHC-t

**125k Ergebnisse. Aber sortieren funktioniert nur in einer Sub-Kategorie...**

**74k Ergebnisse in „Elektronik & Foto“  
37k Ergebnisse in „Computer“**

**→ Um das günstigste Angebot zu finden, muss mehrfach manuell gefiltert und verglichen werden.**

1-16 von 125.770 Ergebnissen in "sd-karte" Abteilung zum Sortieren auswählen

- Ergebnisse anzeigen für
- Elektronik & Foto >
    - Speicherkarten
    - Micro SD Speicherkarten
    - SecureDigital-Cards
    - + Weitere
  - Prime Instant Video >
    - Prime Instant Video Serien
    - Prime Instant Video Filme
  - Computer & Zubehör >
    - USB-Sticks
    - + Weitere
  - + Alle 31 Kategorien

- Filtern nach
- Versandoption (Was ist das?)  
 Kostenlose Lieferung ab EUR 29 Bestellwert
- Speicherkartenkapazität  
 Bis 1,9 GB  
 2 GB  
 4 GB  
 8 GB  
 16 GB  
 32 GB  
 64 GB  
 128 GB & mehr
- Marke  
 SanDisk  
 Samsung  
 Transcend  
 Kingston  
 Qumox

Verwandte Suchbegriffe: micro sd karte, sd karte 64 gb, sd karte 8gb.

 **Transcend Ultimate-Speed SDHC Class 10 UHS-1 32GB Speicherkarte (b 90MB/s Lesen) [Amazon Frustrfreie Verpackung]**  
von Transcend  
**EUR 17,49** ✓Prime  
Lieferung bis zum **Dienstag, 9. Juni**  
★★★★★ 799  
Bestseller Nr. 1 in SecureDigital-Card

 **SanDisk SDSDUN-016G-FFP Ultra Speicherkarte bis zu 40MB/Sek. lese**  
von SanDisk  
**EUR 8,99** ~~EUR 10,99~~ ✓Prime  
Lieferung bis zum **Dienstag, 9. Juni**

 **Transcend TS32GUSDHC10E Class 10 Extreme-Speed microSDHC 32GB Speicherkarte mit SD-Adapter [Amazon Frustrfreie Verpackung]**  
von Transcend  
**EUR 12,49** ✓Prime  
Lieferung bis zum **Dienstag, 9. Juni**

 **Kingston SDC4/8GB microSDHC Class 4 Speicherkarte mit SD-Adapter**  
von Kingston  
**EUR 4,89** Plus Produkt  
Fügen Sie ihn zu einer qualifizierten Bestellung hinzu und Sie erhalten ihn am **Mittwoch, 10. Juni** zu erhalten  
Andere Angebote

**125k Ergebnisse. Aber sortieren funktioniert nur in einer Sub-Kategorie...**

**74k Ergebnisse in „Elektronik & Foto“  
37k Ergebnisse in „Computer“**

**→ Um das günstigste Angebot zu finden, muss mehrfach manuell gefiltert und verglichen werden.**

**→ Frustration beim Nutzer  
→ Negative Einkaufs-Experience**

# Beispiel Amazon.de

---

Was ist hier falsch gelaufen?

- \_ Artikel suchen oder Ergebnismenge filtern erscheint kein total ungewöhnlicher Use-Case zu sein auf einem Einkaufsportaal...

# Beispiel Amazon.de

---

Was ist hier falsch gelaufen?

- \_ Artikel suchen oder Ergebnismenge filtern erscheint kein total ungewöhnlicher Use-Case zu sein auf einem Einkaufsportaal...

Erklärungen:

## Usability-Problem 1 (Fokus im Suchschlitz)

- Es ist von den Nutzern nicht gewollt, die Tastatur zu nutzen
- Es ist technisch nicht möglich, den Fokus beim Laden der Seite in den Suchschlitz zu setzen

# Beispiel Amazon.de

---

Was ist hier falsch gelaufen?

\_ Artikel suchen oder Ergebnismenge filtern erscheint kein total ungewöhnlicher Use-Case zu sein auf einem Einkaufsportale...

Erklärungen:

## Usability-Problem 1 (Fokus im Suchschlitz)

- Es ist von den Nutzern nicht gewollt, die Tastatur zu nutzen
- Es ist technisch nicht möglich, den Fokus beim Laden der Seite in den Suchschlitz zu setzen

## Usability-Problem 2 (Sortieren nur mit Teilmenge)

- Es ist von den Nutzern nicht gewollt, alle Ergebnisse gleichzeitig zu sortieren
- Es ist technisch nicht möglich, alle Ergebnisse gleichzeitig zu sortieren



# Beispiel Amazon.de

Was ist hier falsch gelaufen?

- Artikel suchen oder Ergebnismenge filtern erscheint kein total ungewöhnlicher Use-Case zu sein auf einem Einkaufsportaal...

Erklärungen:

Usability-Problem 1 (Fokus im Suchschlitz)

- ~~Es ist von den Nutzern nicht gewollt, die Tastatur zu nutzen~~
- ~~Es ist technisch nicht möglich, das Filter beim Laden der Seite in den Suchschlitz zu setzen~~

**Technisch spricht nichts gegen einen Umbau (theoretisch)**

**Fragt man Nutzer, finden sie die aktuellen Lösungen suboptimal**

Usability-Problem 2 (Sortieren nur mit Teilmenge)

- ~~Es ist von den Nutzern nicht gewollt, alle Ergebnisse gleichzeitig zu sortieren~~
- ~~Es ist technisch nicht möglich, alle Ergebnisse gleichzeitig zu sortieren~~

# Beispiel Amazon.de

Was ist hier falsch gelaufen?

- Artikel suchen oder Ergebnismenge filtern erscheint kein total ungewöhnlicher Use-Case zu sein auf einem Einkaufsportale...

Erklärungen:

Usability-Problem 1 (Fokus im Suchschlitz)

- ~~Es ist von den Nutzern nicht gewollt, die Tastatur zu nutzen~~
- ~~Es ist technisch nicht möglich, das Filter beim Laden der Seite in der Suchschlitz zu setzen~~

**Technisch spricht nichts gegen einen Umbau (theoretisch)**

**Fragt man Nutzer, finden sie die aktuellen Lösungen suboptimal**

Usability-Problem 2 (Sortieren nur mit Teilmenge)

- ~~Es ist von den Nutzern nicht gewollt, alle Ergebnisse gleichzeitig zu sortieren~~
- ~~Es ist technisch nicht möglich, alle Ergebnisse gleichzeitig zu sortieren~~

**Alternativerklärung: Es wurde nicht (oder zu spät) getestet, ob die Lösung mit den Wünschen der Nutzer im Einklang steht.**

---

**Ein weiteres Beispiel...**

---

FLATRON W3000H



alles aufgeräumt



Papierkorb



▼ BRIGHTNESS ▲



vorher

FLATRON W3000H

Standard Wissenschaftlich Konverter

0

M	CE	C	$\leftarrow$ X	$\div$
$\pm$	7	8	9	$\times$
$\sqrt{\quad}$	4	5	6	-
%	1	2	3	+
$1/x$	0	,	=	

LG

BRIGHTNESS

nachher

FLATRON W3000H

Standard Wissenschaftlich Konverter

0

30Zoll

M	CE	C	$\leftarrow$	$\div$
$\pm$	7	8	9	$\times$
$\sqrt{\quad}$	4	5	6	-
%	1	2	3	+
$1/x$	0	,	=	

LG

BRIGHTNESS

nachher

# Beispiel Win8

---

Branchenreport Usability 2015

Welche Entwicklung finden Sie aktuell am negativsten/ nervt Sie am meisten?

- \_ Kacheln & Kachel-Apps
- \_ Flat-Design

# Beispiel Win8

---

Branchenreport Usability 2015

Welche Entwicklung finden Sie aktuell am negativsten/ nervt Sie am meisten?

- \_ Kacheln & Kachel-Apps
- \_ Flat-Design

**Erklärung: Design >> Usability/UX**

- Nutzer wollen keine Kacheln auf dem Desktop
- MS verbesserte mit Win8.1 nach (Startmenü; Boot to Desktop)



---

**Eine Untersuchung...**

---

# Evaluationsbeispiel - UX messen

---

Vergleich von zwei Kameras:

Wie kommen sie bei den Nutzern an?

Welche ist praktischer? Welche ist schöner?



**Samsung L100**



**Samsung i8**

# AttrakDiff

## AttrakDiff 2

- Semantisches Differential zur Messung hedonischer und pragmatischer Qualitäten

## Zentrale Frage

„Welchen Charakter vermittelt das Produkt?“

## Was wird gemessen:

- Hedonische Qualität
  - Stimulation
  - Identität
- Pragmatische Qualität
- Globale Attraktivität

### Ihr Urteil!

Bitte geben Sie mit Hilfe der folgenden Wortpaare Ihren Eindruck des <Produkt> wieder.  
Bitte kreuzen Sie nur jeweils ein Kästchen an!

Bitte ausfüllen...

	1	2	3	4	5	6	7		
menschlich	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	technisch	p FQ_1
isolierend	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	verbindend	HQ1_1
angenehm	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	unangenehm	p ATT_1
originell	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	konventionell	p HQS_1
einfach	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	kompliziert	p FQ_2
fachmännisch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	laienhaft	p HQ1_2
hässlich	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	schön	ATT_2
praktisch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	unpraktisch	p FQ_3
sympathisch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	unsympathisch	p ATT_3
umständlich	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	direkt	FQ_4
stilvoll	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	stillos	p HQ1_3
voraussagbar	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	unberechenbar	p FQ_5
minderwertig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	wertvoll	HQ1_4
ausgrenzend	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	einbeziehend	HQ1_5

# AttrakDiff

Beide Kameras wirken gleichermaßen attraktiv (ATT)

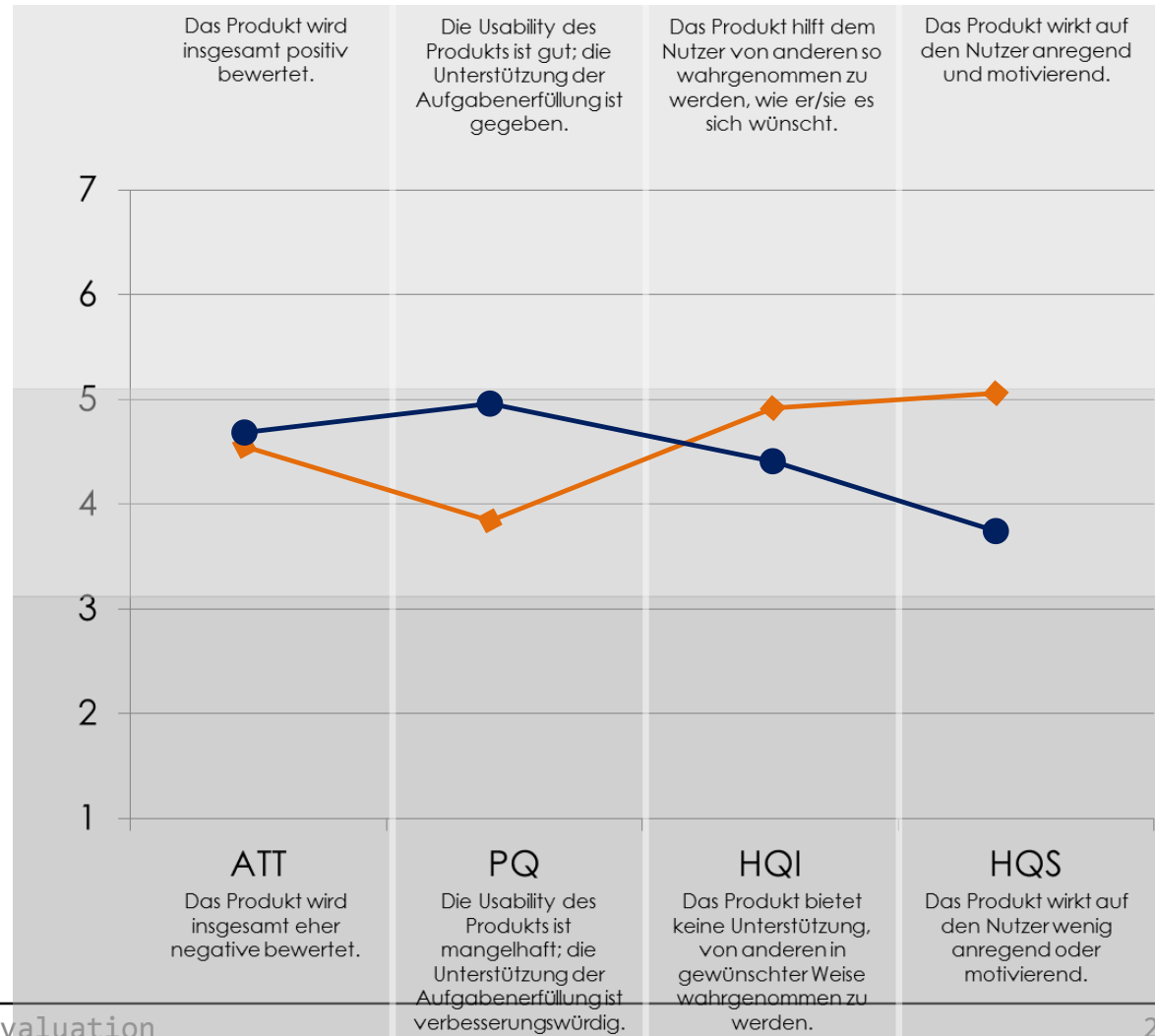
... aber unterscheiden sich im Produktcharakter (PQ vs. HQ)



**Samsung L100**



**Samsung i8**



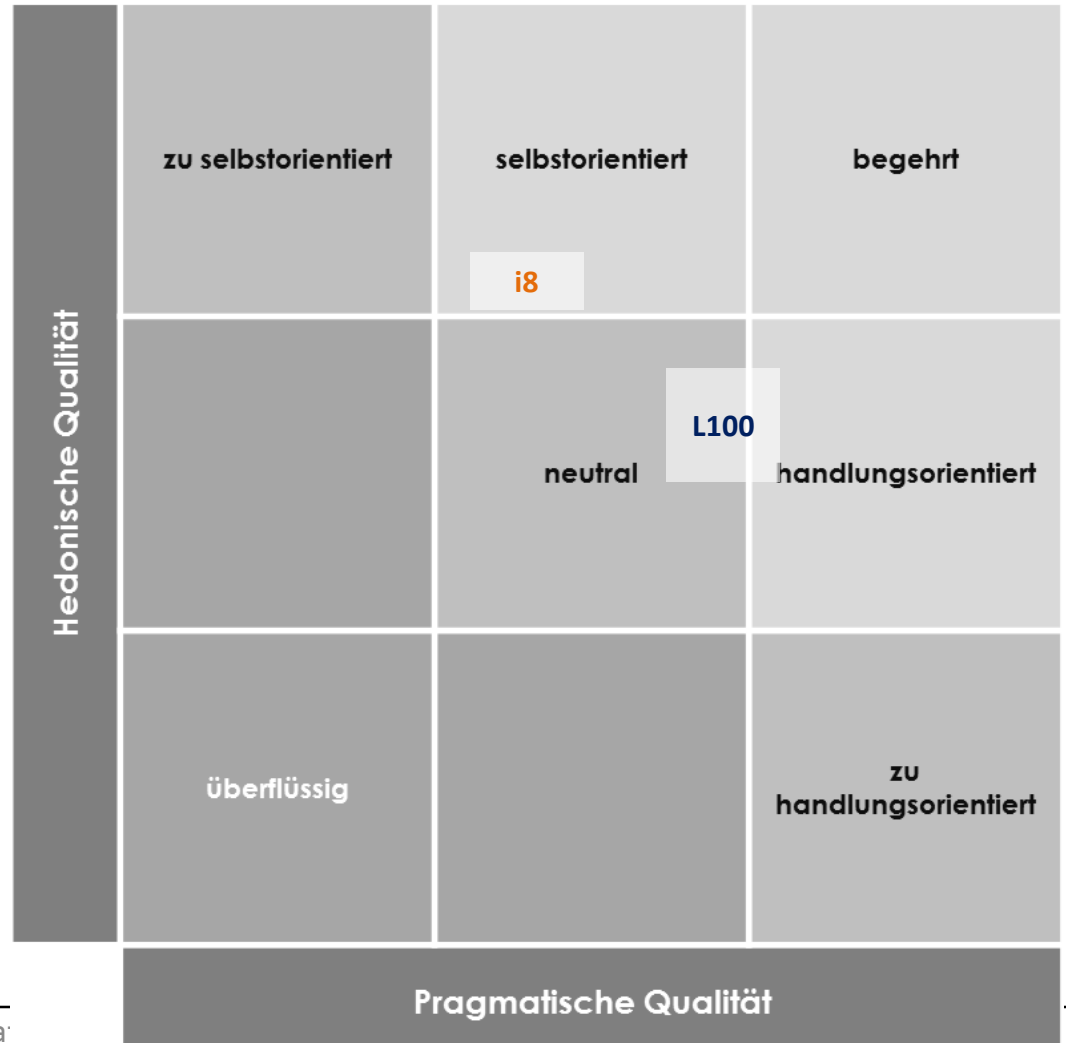
# AttrakDiff



Samsung L100

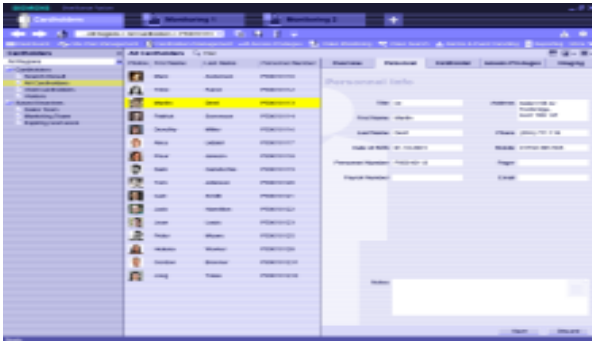


Samsung i8

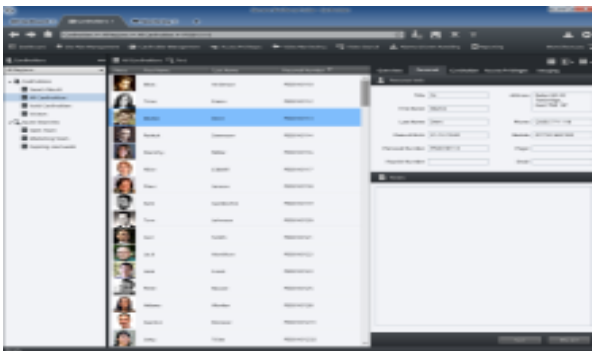


# AttrakDiff

## Vergleich von zwei Überwachungs-Tools



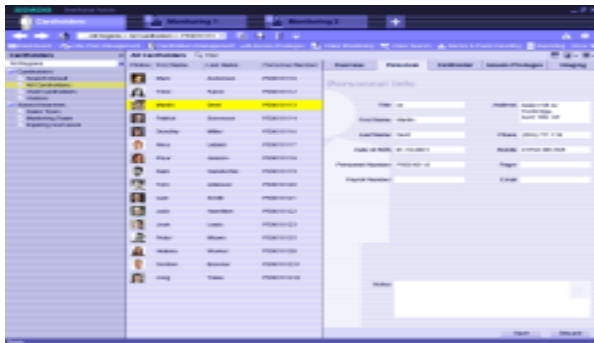
**Variante A**



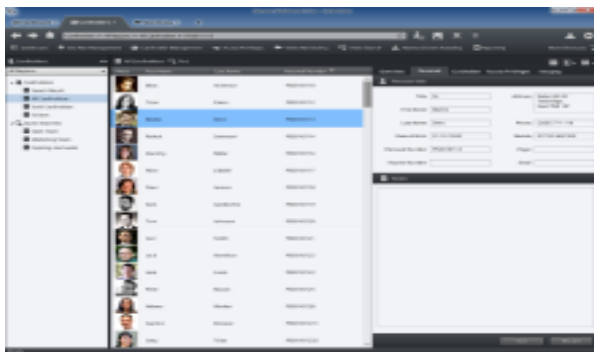
**Variante B**

# AttrakDiff

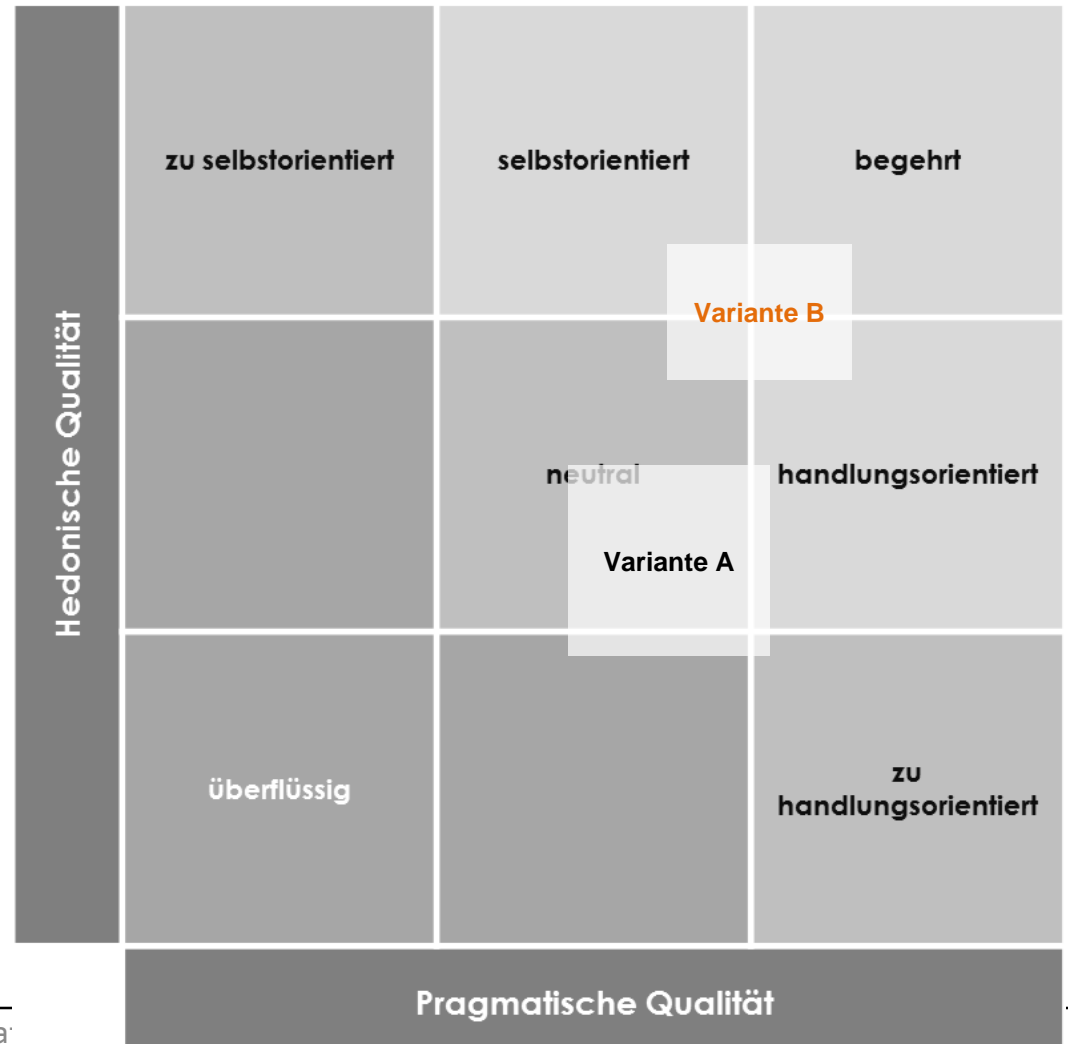
Variante B wirkt hedonischer, bzgl. Variante A größere Uneinigkeit  
Keine signifikanten Unterschiede bzgl. pragmatischer Qualität



Variante A



Variante B



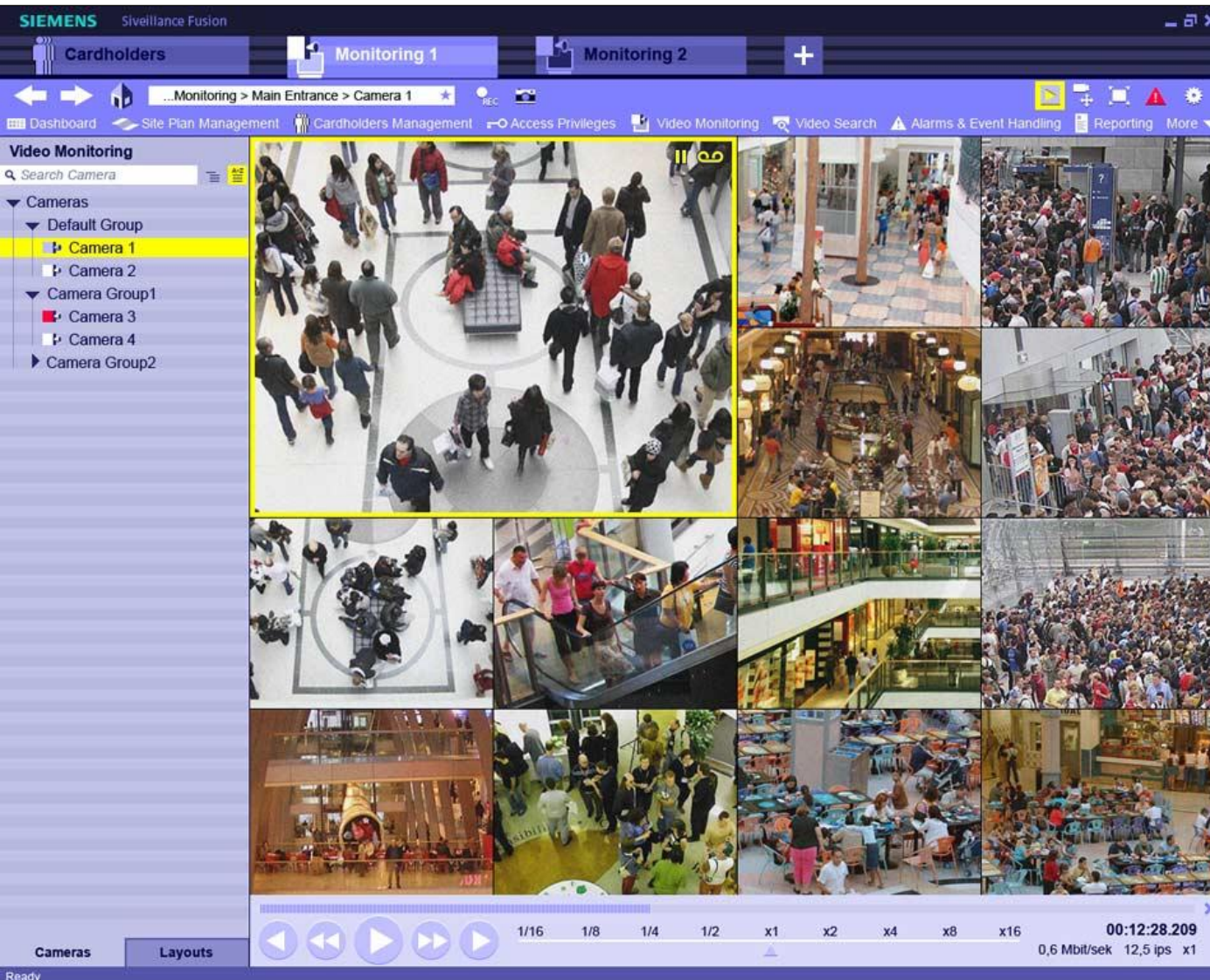
---

**Woher kommen die Unterschiede?**

---



# AttrakDiff



*"too colorful"*

*"color not right"*

*"too colorful"*

*"difficult to look at"*

*"uncomplementary colors"*

*"color choice wrong"*

*"teen style (pink would still be better than lilac)"*

*"assault on my eyes"*

*"looks like spam software, not consistent with our quality"*

*"please don't use this"*

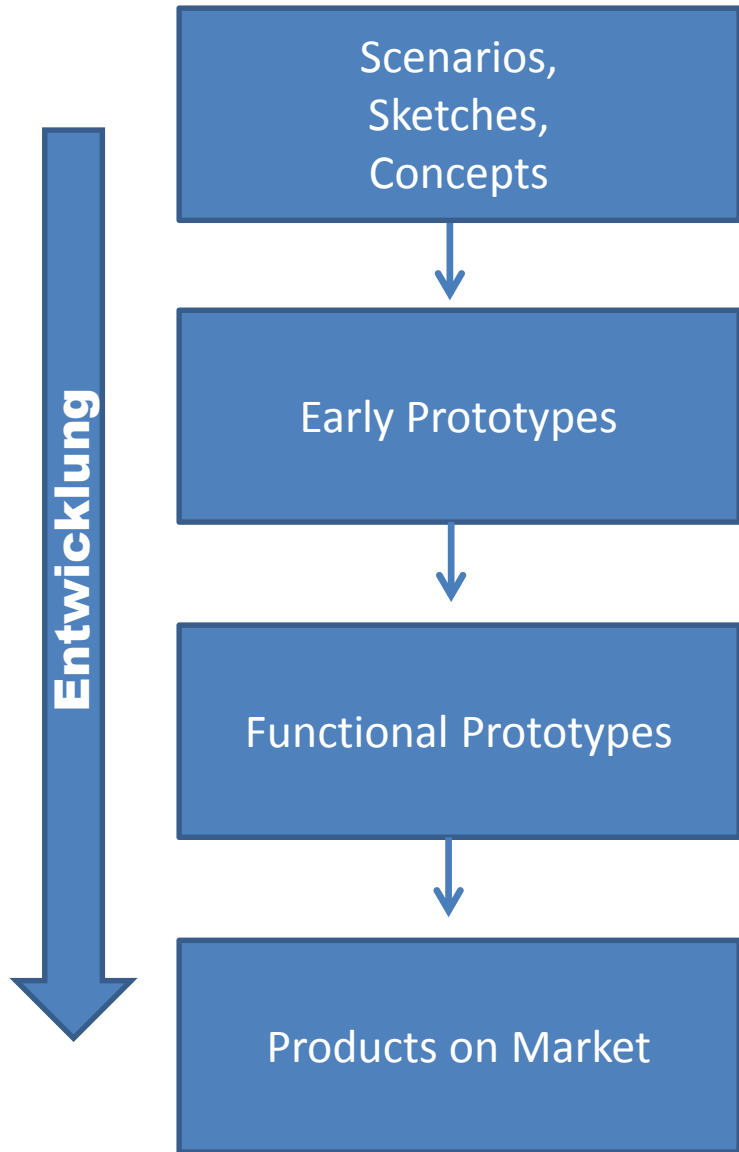
## Variante A

---

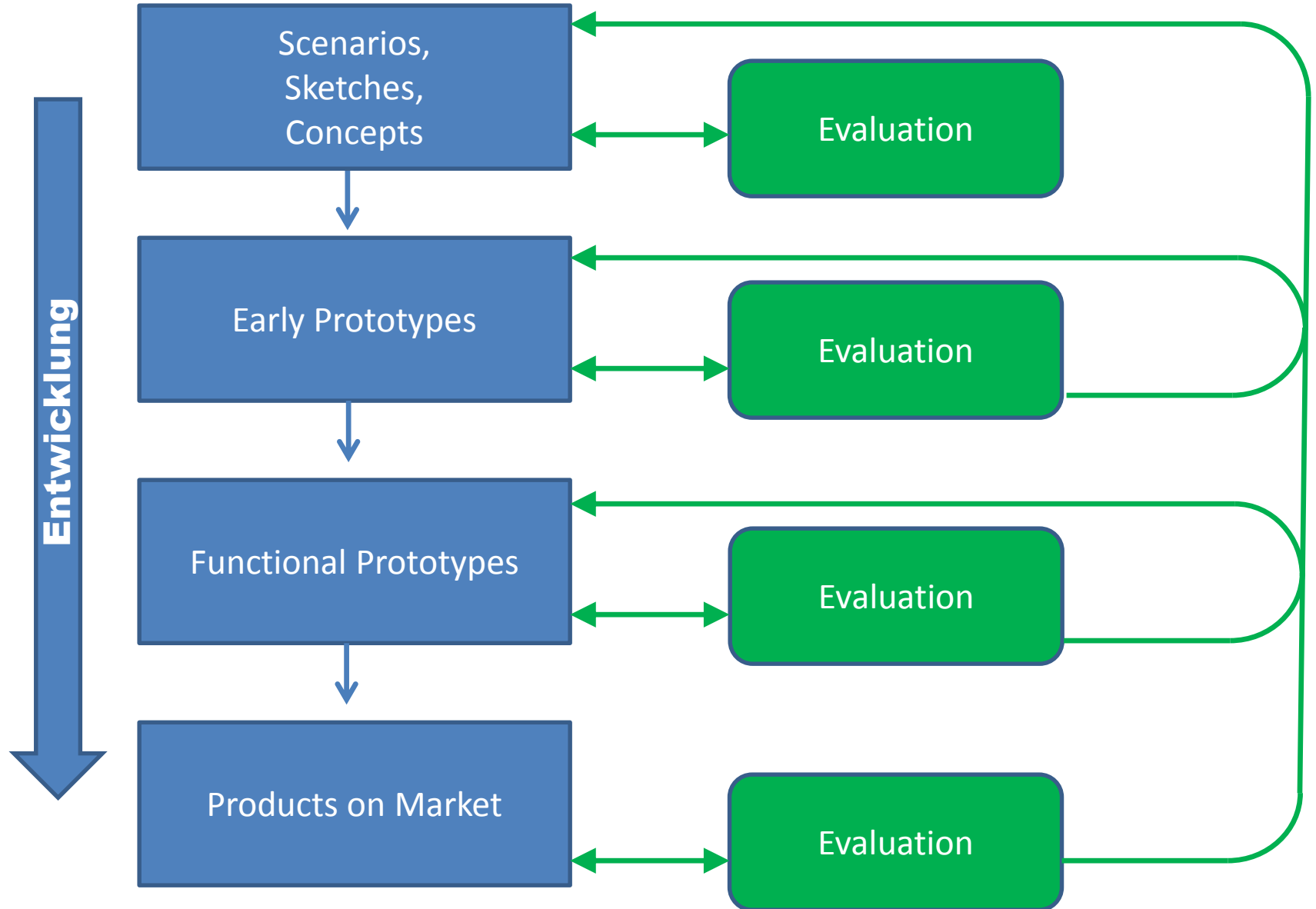
# **Evaluationskategorien**

---

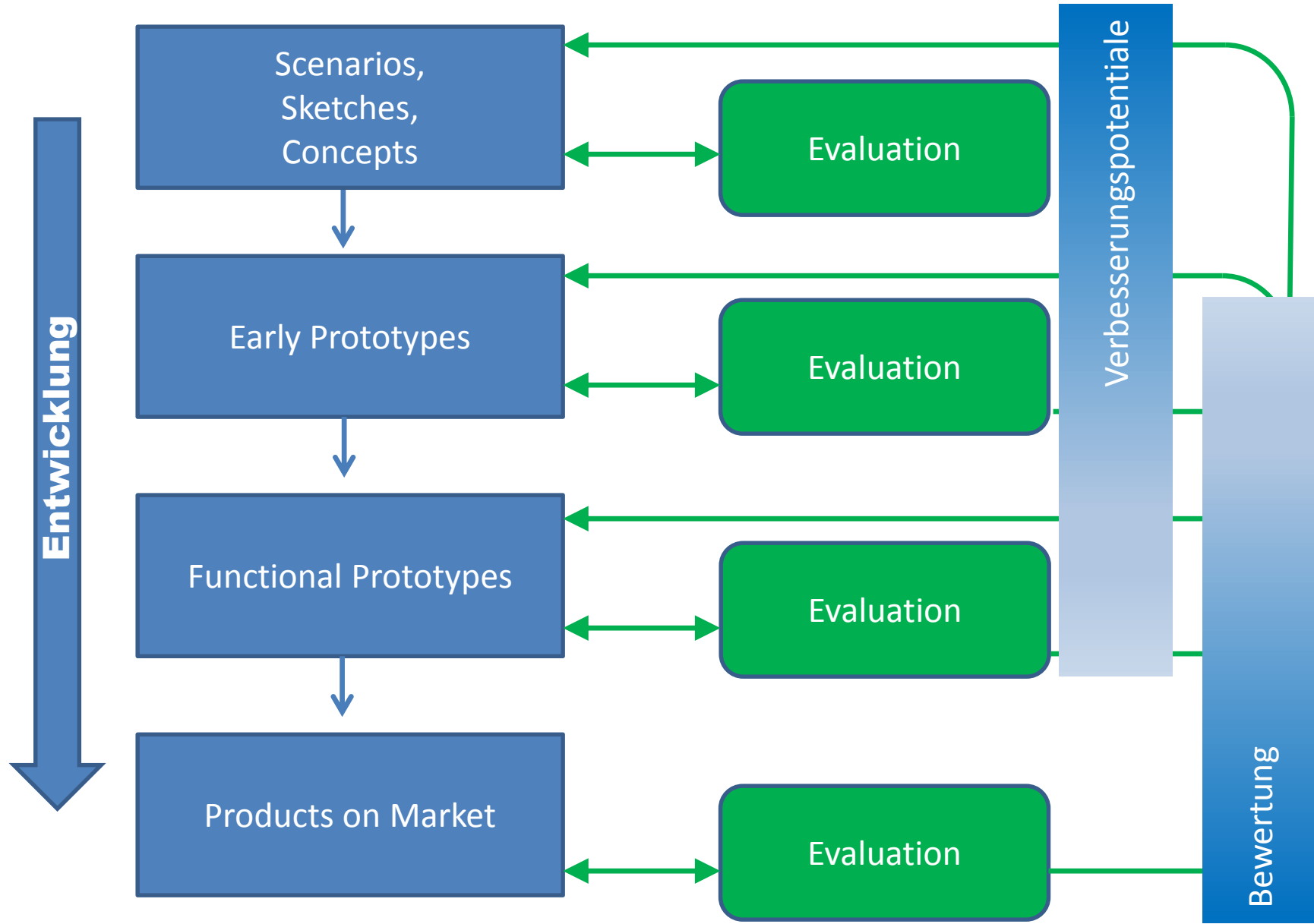
# Produkt-Entwicklungszyklus



# Produkt-Entwicklungszyklus



# Produkt-Entwicklungszyklus



# Summativ vs. Formativ

---

## Summative Evaluation

"Wie gut ist es geworden?" – Bewerten

Quantitativ

Abschließend, zusammenfassend, kriteriumsorientiert

Z.B. "Zertifizierung"

z.B.

Fragebögen

Effizienzmaße

## Formative Evaluation

"Was muss wie umgestaltet werden?" - Verstehen

Qualitativ

Prozessbegleitend, verbesserungsorientiert

Z.B. "Usability Test eines Papierprototypen"

z.B.

Personas

Card Sorting

Walkthrough

# Analytisch vs. Empirisch

---

## Analytische Evaluation

Expertenurteil, "Begutachtung"

Oft einzeln

z.b.

Walkthrough

Heuristische Evaluation

Kriterienkataloge

Zertifizierung

## Empirische Evaluation

Laienurteile, Laienperformanz

Gruppen, statistische Analyse möglich

z.b.

Usability-Tests

Fragebögen

Fokusgruppen

# Evaluationskategorien

---

	<b>Formativ</b>	<b>Summativ</b>
<b>Analytisch</b>		
<b>Empirisch</b>		



# Evaluationskategorien

	<b>Formativ</b>	<b>Summativ</b>	
<b>Analytisch</b>			Experten bewerten
<b>Empirisch</b>			Nutzer bewerten

# Evaluationskategorien

	<b>Formativ</b>	<b>Summativ</b>	
<b>Analytisch</b>			Experten bewerten
<b>Empirisch</b>			Nutzer bewerten
	Verbesserungs- potential identifizieren	Holistische Bewertung	

# Evaluationsziele

---

Welche Konzeptkategorien können gemessen werden

- Produkturteile
- Erleben
- Affekt / Emotionen
- Leistungsdaten

# Quantitativ vs. Qualitativ

---

## Quantitativ

Vorgegebene Antwortkategorien

Schnelle Durchführung

Einfache Auswertung

„wie fühlen Sie sich auf einer Skala von 1-9?“

numerisch

## Qualitativ

Freie Antwortmöglichkeiten

Aufwändigere Durchführung

Macht Vergleiche schwierig

„wie fühlen Sie sich?“

Kann Aspekte erfassen, die bei quantitativer Messung verloren gehen könnten.

verbalisiert

# Beobachten vs. Befragen

---

Befragen – mündlich oder schriftlich

Fragebögen – "schriftliche Befragung"

Befragungsgegenstand ist das Produkt, System

Itemarten

Fakten – "Die Software bietet mir eine Wiederhol-Funktion für wiederkehrende Arbeitsschritte"

Beurteilungen – "Zur Erkundung des Systems durch Versuch und Irrtum wird ermutigt"

Gefühle – "Das System ist sehr unangenehm"

Jede Art von Urteil (auch das faktenorientierte) reflektiert persönliche Erlebnisse mit dem beurteilten System

Befragte benötigen eine Erlebnisstichprobe ("Erfahrung") als Grundlage der Beurteilung und "Erfahrung" muss kontrolliert werden

# Cognitive Walkthrough

	<b>Formativ</b>	<b>Summativ</b>	
<b>Analytisch</b>	<b>Cognitive Walkthrough</b>		Experten bewerten
<b>Empirisch</b>			Nutzer bewerten
	Verbesserungspotential identifizieren	Holistische Bewertung	

# Cognitive Walkthrough

---

## Übersicht

- \_ Aufgabenorientierte Inspektionsmethode
- \_ Usability-Experte untersucht Funktionalität eines Systems entlang definierter Aufgaben

# Cognitive Walkthrough

---

## Ablauf

### Vorbereitung

- Nutzercharakteristiken definieren
- Aufgabenstellung und Einzelschritte spezifizieren
- Idealen Aktionspfad zur Erledigung der Aufgaben definieren

### Analyse

- Jede Aufgabe wird einzeln analysiert
- Bewertung der Aktionen und Systemreaktionen
  - Ist die korrekte Aktion zur Ausführung einer Handlung klar? Weiß der Nutzer, was er tun soll?
  - Sind die korrekten Aktionen verfügbar? Ist den Nutzern bewusst, dass sie verfügbar sind?
  - Erhalten die Nutzer System-Feedback? Verstehen die Nutzer das System-Feedback?

## Protokollierung und Ableitung von Verbesserungsvorschlägen



# Cognitive Walkthrough

---

## Vorteile

- \_ Schnell und einfach durchführbar
- \_ Einsetzbar in frühen Entwicklungsstadien
- \_ Fehler im Ansatz erkennbar

## Nachteile

- \_ Künstliches Setting
- \_ Evaluator kann den Benutzer nicht realistisch einschätzen

# Heuristic Evaluation

	<b>Formativ</b>	<b>Summativ</b>	
<b>Analytisch</b>	Cognitive Walkthrough	<b>Heuristic Evaluation</b>	Experten bewerten
<b>Empirisch</b>			Nutzer bewerten
	Verbesserungspotential identifizieren	Holistische Bewertung	

# Heuristic Evaluation

---

## Übersicht

- \_ Mehrere Experten prüfen, ob ein User Interface mit bestimmten anerkannten Prinzipien des Usability Engineerings übereinstimmt
- \_ Heuristiken = Richtlinien / Faustregeln  
(auf Basis empirischer Ergebnisse entwickelt)

# Heuristic Evaluation

---

## Wer evaluiert?

- \_ Mehrere Experten arbeiten unabhängig voneinander
- \_ Faustregel: 5 Experten finden 75% der Probleme
- \_ Usability-Experten
  - besser: Usability + Domänenexperte
  
- \_ Falls kein Domänenexperte vorhanden ist:  
Szenario bereitstellen mit Schritten, die ein Benutzer gehen würde um eine reale Aufgabe zu erfüllen (vorab Aufgabenanalyse notwendig)

# Heuristic Evaluation

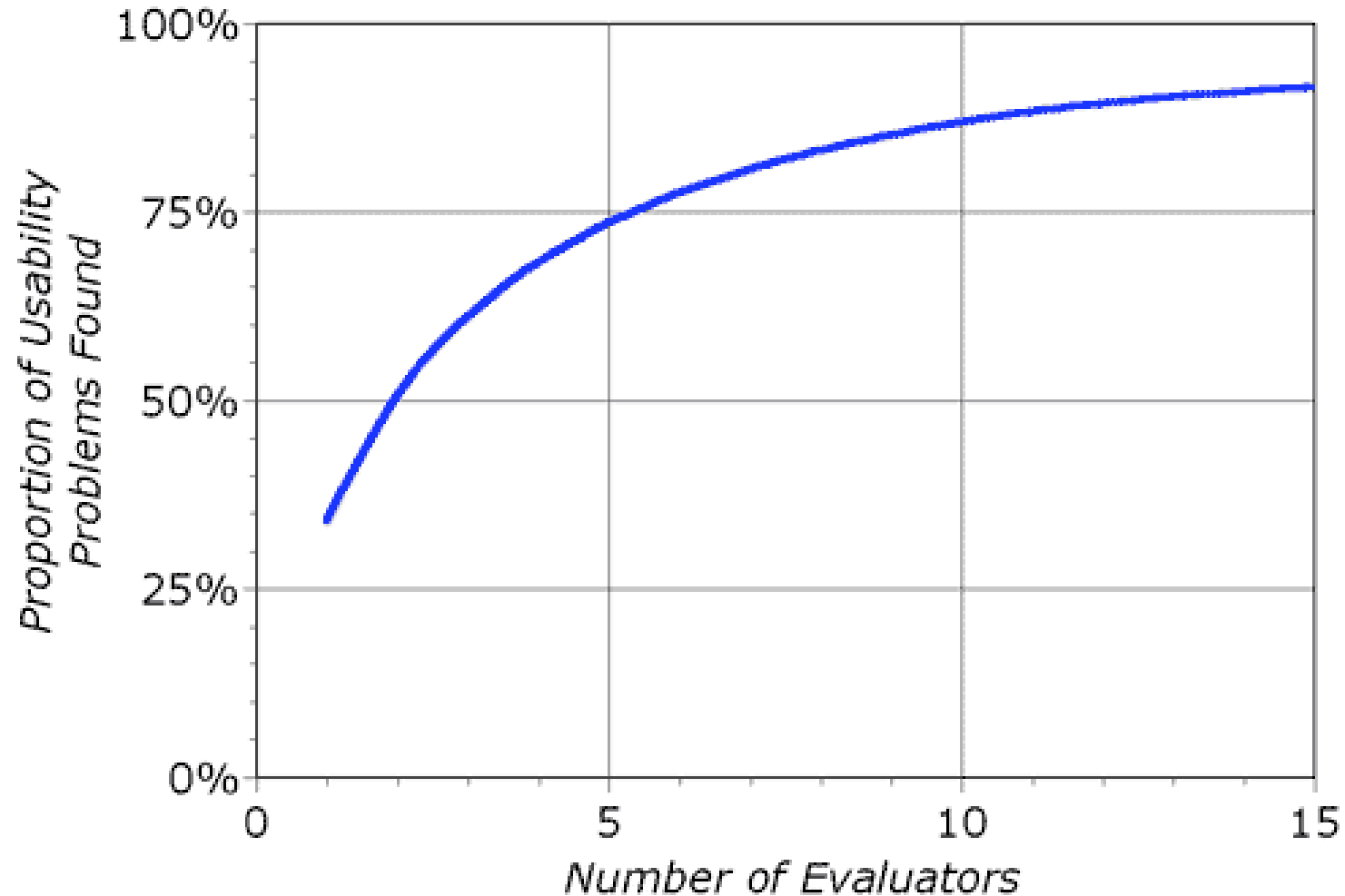
---

## Wer evaluiert?

- \_ Übereinstimmung ist gering
- \_ Zwischen 2 Evaluatoren im Schnitt nur 22% Übereinstimmung (Hertzum & Jacobsen,2003)
  
- \_ Schlimmste Probleme werden nicht von jedem gefunden
  - mehrere Evaluatoren nötig

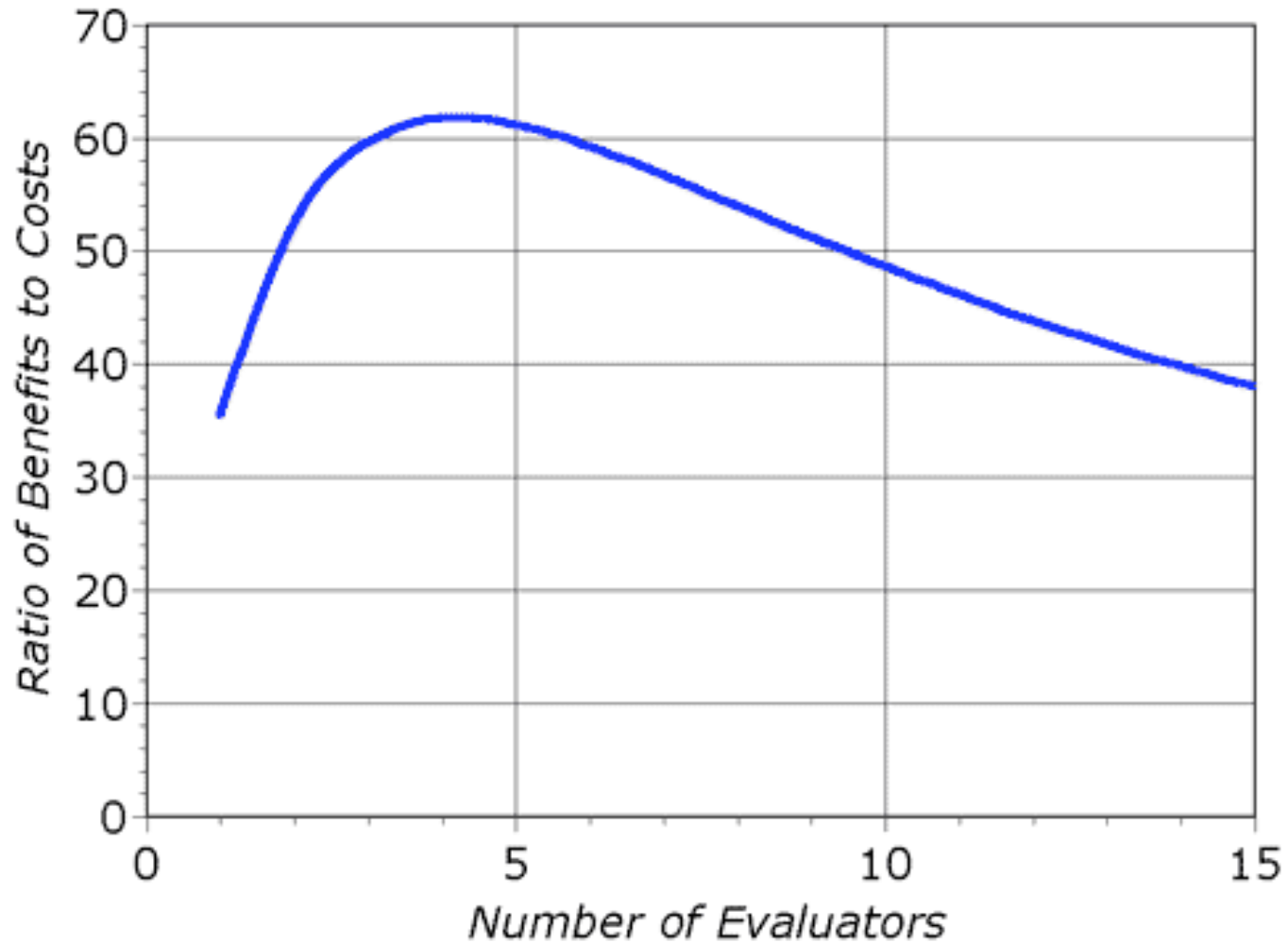
# Heuristic Evaluation

Wie viele Evaluatoren?



# Heuristic Evaluation

Wie viele Evaluatoren?



# Heuristic Evaluation

---

## Ablauf

- \_ Briefing, ggf. Szenario besprechen
- \_ Evaluatoren gehen einzeln das UI anhand der Heuristiken durch und dokumentieren Probleme (1-2h)
- \_ in der Gruppe: Probleme zusammentragen und Schweregrad beurteilen (30 min)
- \_ Debriefing mit Entwicklern



# Heuristic Evaluation

---

## Ablauf

- \_ Jeder Evaluator prüft einzeln, ohne Kommunikation!
  - 1-2 Stunden
  
- \_ 2 Durchgänge:
  - Überblick verschaffen
  - Spezielle Interaktionselemente fokussieren
  
- \_ Produkteigenschaften mit Heuristiken vergleichen und Verstöße protokollieren

# Heuristic Evaluation

---

## Heuristiken

### \_ Faustregeln

- können als Gestaltungsprinzipien oder Evaluationsprinzipien genutzt werden
- decken häufige Probleme ab

# Heuristic Evaluation

---

## Heuristiken

- \_ Sichtbarkeit des Systemstatus
- \_ Übereinstimmung zwischen System und realer Welt
- \_ Benutzerkontrolle und –freiheit
- \_ Konsistenz und Normen
- \_ Fehlervermeidung
- \_ Wiedererkennen vor Erinnern
- \_ Flexibilität und effiziente Nutzung
- \_ Ästhetik und minimalistisches Design
- \_ Unterstützung beim Erkennen, Verstehen und Bearbeiten von Fehlern
- \_ Hilfe und Dokumentation

# Heuristic Evaluation

---

## Heuristiken

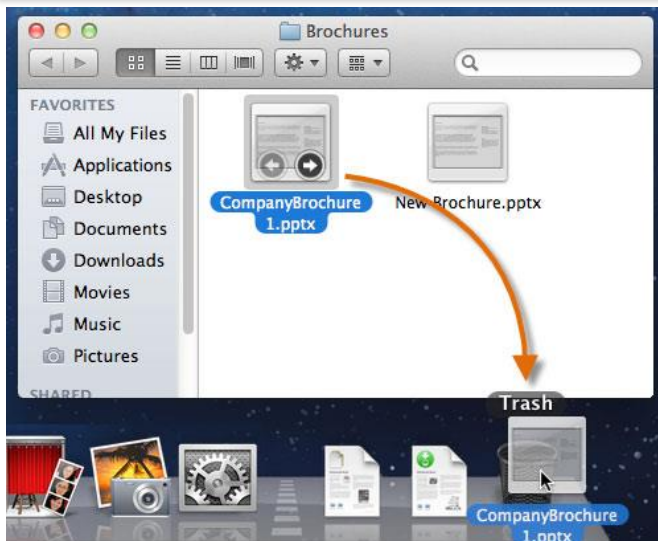
- \_ Übereinstimmung zwischen System und realer Welt
  
- \_ Speak the users' language
- \_ Follow real world conventions
- \_ Pay attention to metaphors

# Heuristic Evaluation

## Heuristiken

- \_ Übereinstimmung zwischen System und realer Welt
- \_ Positivbeispiel

**Datei löschen?**  
**→ in den Papierkorb werfen**



**Datei endgültig löschen?**  
**→ Papierkorb leeren**



[http://content.gcflearnfree.org/topics/202/os\\_delete\\_file\\_mac.jpg](http://content.gcflearnfree.org/topics/202/os_delete_file_mac.jpg)

[http://content.gcflearnfree.org/topics/202/os\\_empty\\_trash.jpg](http://content.gcflearnfree.org/topics/202/os_empty_trash.jpg)

# Heuristic Evaluation

## Heuristiken

- \_ Übereinstimmung zwischen System und realer Welt
- \_ Negativbeispiel

**USB-Laufwerk auswerfen?  
→ in den Papierkorb werfen!?**



[http://www.edb.utexas.edu/education/assets/images/ltc/services/tutorials/tc\\_osx\\_dismount\\_04.jpg](http://www.edb.utexas.edu/education/assets/images/ltc/services/tutorials/tc_osx_dismount_04.jpg)

# Heuristic Evaluation

## Checklist

### Usability Techniques Heuristic Evaluation - A System Checklist

By Denise Pierotti, Xerox Corporation

#### Heuristic Evaluation - A System Checklist

##### 1. Visibility of System Status

The system should always keep user informed about what is going on, through appropriate feedback within reasonable time.

#	Review Checklist	Yes No N/A	Comments
1.1	Does every display begin with a title or header that describes screen contents?	○ ○ ○	
1.2	Is there a consistent icon design scheme and stylistic treatment across the system?	○ ○ ○	
1.3	Is a single, selected icon clearly visible when surrounded by unselected icons?	○ ○ ○	
1.4	Do menu instructions, prompts, and error messages appear in the same place(s) on each menu?	○ ○ ○	
1.5	In multipage data entry screens, is each page labeled to show its relation to others?	○ ○ ○	
1.6	If overwrite and insert mode are both available, is there a visible indication of which one the user is in?	○ ○ ○	
1.7	If pop-up windows are used to display error messages, do they allow the user to see the field in error?	○ ○ ○	
1.8	Is there some form of system feedback for every operator action?	○ ○ ○	
1.9	After the user completes an action (or group of actions), does the feedback indicate that the next group of actions can be started?	○ ○ ○	
1.10	Is there visual feedback in menus or dialog boxes about which choices are selectable?	○ ○ ○	
1.11	Is there visual feedback in menus or dialog boxes about which choice the cursor is on now?	○ ○ ○	
1.12	If multiple options can be selected in a menu or dialog box, is there visual feedback about which options are already selected?	○ ○ ○	
1.13	Is there visual feedback when objects are selected or moved?	○ ○ ○	
1.14	Is the current status of an icon clearly indicated?	○ ○ ○	

# Heuristic Evaluation

---

## Vorteile

- \_ Schnell durchführbar
- \_ Einfach anzuwenden
- \_ Geringe Kosten



# Heuristic Evaluation

---

## Nachteile

- \_ Es werden nur Fehler gefunden, für die den Heuristiken entgegenstehen.
  - Gibt es für spezielle Fehler keine Heuristiken, werden sie nicht entdeckt.
  - Es sind weitere Heuristiken denkbar
- \_ Heuristiken können den Aufmerksamkeitsfokus auf Probleme lenken, die im Produktivgebrauch selten vorkommen.
- \_ Die Selektion des Szenarios hat starke Auswirkung auf die Ergebnisse (bestimmte Fehler sind szenariotypisch)
- \_ Evaluatoren müssen kreativ bei der Fehlersuche sein und eine hohe Expertise besitzen
  - Häufig sind sie nur Usability-Experten, keine Domänenexperten
  - Ergebnisse sind dann trivial und bilden nicht die realen Probleme der Nutzer ab

# Fokusgruppe

	<b>Formativ</b>	<b>Summativ</b>	
<b>Analytisch</b>	Cognitive Walkthrough	Heuristic Evaluation	Experten bewerten
<b>Empirisch</b>	<b>Fokusgruppe</b>		Nutzer bewerten
	Verbesserungspotential identifizieren	Holistische Bewertung	

# Fokusgruppe

---

## Übersicht

- \_ Moderierte Gruppendiskussion mit ausgewählten Teilnehmern zu einem bestimmten Thema, z.B. einer Produktidee
  - Durchführung mehrerer Fokusgruppen zu einem Thema um Verzerrung der Ergebnisse durch spezifische Besonderheiten einer Gruppe zu vermeiden
  - je Gruppe 5-9 Teilnehmer
- \_ Diskussion vorher festgelegter Fragestellungen, die im Rahmen der Produktentwicklung von Interesse sind
  - ggf. Aufgreifen neuer Aspekte, die sich aus Teilnehmerkommentaren ergeben
- \_ Konfrontation mit Sichtweisen der anderen Teilnehmer
  - ausführliche Darstellung, Begründung und Verteidigung der eigenen Position
  - Vorteil gegenüber Einzelinterviews

# Fokusgruppe

---

## Einsatzgebiete und Fragestellung

- \_ Einsatz meist in der frühen Phase der Produktgestaltung
  - Vorstellung von Produktideen mittels Screenshots, Skizzen, Prototypen
- \_ Findet die Produktidee generellen Zuspruch?
- \_ Welche von alternativen Gestaltungsvarianten ist am vielversprechendsten?
- \_ Welche Funktionalitäten sind Nutzern am wichtigsten, und warum?

# Fokusgruppe

---

## Ablauf

### \_ Einleitung

- Vorstellung der Teilnehmer
- Vorstellung des Themas
- Erklärung der Verhaltensregeln

### \_ Hauptteil

- Fokusgruppendifkussion
- Leitung durch Moderator mit Diskussionsleitfaden
- Notieren und Zusammenfassen der wichtigsten Zwischenergebnisse

### \_ Wrap-Up

- Zusammenfassung der wichtigsten Erkenntnisse
- Fazit ziehen

# Fokusgruppe

---

## Stolpersteine

- \_ "Falsche Teilnehmer"
  - Rekrutierung der Teilnehmer beruht auf falschen Vorstellung bzgl. der Nutzergruppe
- \_ Unqualifizierte Teilnehmer
  - Teilnehmer erfüllen Rekrutierungskriterien nicht und können nicht zur Diskussion beitragen
- \_ Groupthink
  - Gruppe bemüht sich (teils unbewusst) um Konsens
  - Dazugehören und sich der Gruppenmeinung anzuschließen wird wichtiger als den eigenen Standpunkt zu vertreten

# Fokusgruppe

---

## Stolpersteine

### \_ Teilnehmer reden zu wenig

- Einsilbige Antworten
- Scheinbar kein Interesse an Diskussionsgegenstand

### \_ (Ein) Teilnehmer redet zu viel

- Teilnehmer, die Probleme haben interessanten Ideen zu formulieren
- Teilnehmer, die scheinbar davon ausgehen, dass die Ideen beim Reden kommen

### \_ Ungünstige Gruppendynamiken

- Ein Teilnehmer dominiert die Gruppe
- Feindselige Bemerkungen unter Teilnehmern
- Diskussion rutscht schnell ins Emotionale ab

# Fokusgruppe

---

## Vorteile

- \_ Sammlung von Meinungen und Anregungen zu Thema oder Prototypen in relativ kurzer Zeit, mit vertretbarem Aufwand
- \_ Flexible Methode
  - Unklare gestellte Fragen können vom Moderator erläutert werden
  - Unklare Antworten können von Teilnehmern erläutert werden
  - Moderator hat Kontrolle, ob Fragen vollständig beantwortet wurden
- \_ Beobachtung erlaubt Entwicklern direkteren Einblick in Reaktionen (potentieller) Nutzer als z.B. Fragebogenstudien



# Fokusgruppe

---

## Nachteile

- \_ Geringe Verallgemeinerbarkeit und Repräsentativität der Ergebnisse
- \_ Schwierigkeit, die "richtigen" Teilnehmer zu gewinnen
- \_ Gefahr der Verfälschung der Ergebnisse der Diskussion durch gruppendynamische Prozesse
- \_ Aufwändige Auswertung der Daten

# Evaluationskategorien

	<b>Formativ</b>	<b>Summativ</b>	
<b>Analytisch</b>	Cognitive Walkthrough	Heuristic Evaluation	Experten bewerten
<b>Empirisch</b>	Fokusgruppe	<b>Experimentalstudie</b>	Nutzer bewerten
	Verbesserungspotential identifizieren	Holistische Bewertung	

# Experiment

---

## Übersicht

- \_ Empirische Untersuchung
- \_ Bedingungen werden gezielt hergestellt
- \_ Auswirkungen auf ausgewählte Variablen werden gemessen
- \_ Methodisch beste Möglichkeit, um Kausalhypothesen zu prüfen

# Experiment

---

## Eigenschaften

### \_ Standardisiertes Setting

- Versuchsanordnung/ -design
- Versuchsteilnehmer
- Nachprüfbar & wiederholbar

### \_ Kausalschlüsse möglich

- Bei Beobachtungen sind nur Zusammenhänge (Korrelationen) messbar
- Wird ein Unterschied künstlich hergestellt und zeigen sich Effekte, sind diese Effekte auf den Unterschied zurückzuführen

# Experiment

---

## Ablauf

H-1: „Das Design von Handys hat einen Einfluss auf die wahrgenommene Attraktivität“

H-2: „Die Verwendung von blauen Handyschalen führt zu positiverer Bewertung der Attraktivität“



# Experiment

## Ablauf

H-1: „Das Design von Handys hat einen Einfluss auf die wahrgenommene Attraktivität“

H-2: „Die Verwendung von blauen Handyschalen führt zu positiverer Bewertung der Attraktivität“



Design  
Farbe der Handyschale  
(rot vs. blau)



Attraktivitätsrating  
AttrakDiff  
(niedrig vs. Hoch)

# Experiment

## Ablauf

H-1: „Das Design von Handys hat einen Einfluss auf die wahrgenommene Attraktivität“

H-2: „Die Verwendung von blauen Handyschalen führt zu positiverer Bewertung der Attraktivität“



Design  
Farbe der Handyschale  
(rot vs. blau)



Attraktivitätsrating  
AttrakDiff  
(niedrig vs. Hoch)

„Die Verwendung von blauen Handyschalen (im Gegensatz zu roten) führt zu höheren Attraktivitätsratings“

# Experiment

## Ablauf

### Hypothesen

H-1: „Das Design von Handys hat einen Einfluss auf die wahrgenommene Attraktivität“

H-2: „Die Verwendung von blauen Handyschalen führt zu positiverer Bewertung der Attraktivität“



### Unabhängige Variablen

Design  
Farbe der Handyschale  
(rot vs. blau)

### Abhängige Variablen

Attraktivitätsrating  
AttrakDiff  
(niedrig vs. Hoch)



„Die Verwendung von blauen Handyschalen (im Gegensatz zu roten) führt zu höheren Attraktivitätsratings“



# Experiment

## Ablauf

### Hypothesen

H-1: „Das Design von Handys hat einen Einfluss auf die wahrgenommene Attraktivität“

H-2: „Die Verwendung von blauen Handyschalen führt zu positiverer Bewertung der Attraktivität“



Unabhängige Variablen

Abhängige Variablen

### Faktor

Design

Attraktivitätsrating

### Operationalisierung

Farbe der Handyschale

AttrakDiff

### Faktorstufen

(rot vs. blau)

(niedrig vs. Hoch)

### Kausalschluss

„Die Verwendung von blauen Handyschalen (im Gegensatz zu roten) führt zu höheren Attraktivitätsratings“

# Experiment

---

## Variablen

### \_ Unabhängige Variablen (UV)

- werden kontrolliert manipuliert
- Haben Einfluss auf die abhängigen Variablen
- UV werden auch Faktoren genannt
- Jede Ausprägung einer UV (rot, blau) ist eine Faktorstufe
- Beispiele:
  - Instruktionen („hab Spaß mit dem Produkt“ vs. „Erreiche Ziel XY“)
  - Interaktionsmodalitäten (Maussteuerung vs. Touch)

### \_ Abhängige Variablen (AV)

- Werden während des Experiments gemessen
- Können qualitativ (Nutzer-Meinungen) oder quantitativ (Fragebogen-Rating, benötigte Zeit für einen Task) sein

# Querschnitt vs. Längsschnitt

---

## Querschnitt

Zu *einem einzigen Messzeitpunkt* werden viele Personen erhoben

## Längsschnitt

Zu *mehreren Messzeitpunkten* werden viele Personen erhoben

# Querschnitt vs. Längsschnitt

---

## Querschnitt

### Vorteile

- kostengünstig
- schnell durchzuführen
- einfache, effiziente Datengewinnung und –auswertung

### Nachteile

- keine Informationen über zeitliche Veränderungen

# Querschnitt vs. Längsschnitt

---

## Längsschnitt

### Vorteile

- Zeitliche Veränderungen sind messbar
- Querschnitt-Auswertung zusätzlich möglich

### Nachteile

- kosten- und zeitintensiver
- aufwendigere Analysemethoden nötig
- längerer Zeitraum, bis Ergebnisse vorliegen
- geringere Teilnahmebereitschaft (wg. hohem Aufwand der Teilnehmer)

---

# Methodensteckbriefe

---

# ISONORM

## Isonorm 9241-10

- misst die 7 Aspekte der Gebrauchstauglichkeit (Usability)

### Zentrale Frage

„Wie gut wurden die Aspekte der Usability (nach ISO 9241) umgesetzt?“

### Was wird gemessen:

- Aufgabenangemessenheit
- Selbstbeschreibungsfähigkeit
- Steuerbarkeit
- Erwartungskonformität
- Fehlertoleranz
- Individualisierbarkeit
- Lernförderlichkeit

Steuerbarkeit								
Können Sie als Benutzer die Art und Weise, wie Sie mit der Software arbeiten, beeinflussen?								
Die Software ...	---	--	-	-/+	+	++	+++	Die Software ...
bietet keine Möglichkeit, die Arbeit an jedem Punkt zu unterbrechen und dort später ohne Verluste wieder weiterzumachen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	bietet die Möglichkeit, die Arbeit an jedem Punkt zu unterbrechen und dort später ohne Verluste wieder weiterzumachen.
erzwingt eine unnötig starre Einhaltung von Bearbeitungsschritten.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	erzwingt keine unnötig starre Einhaltung von Bearbeitungsschritten.

# Isometrics

## Isometrics

- misst die 7 Aspekte der Gebrauchstauglichkeit (Usability)

## Zentrale Frage

„Wie gut wurden die Aspekte der Usability (nach ISO 9241) umgesetzt?“

## Was wird gemessen:

- Aufgabenangemessenheit
- Selbstbeschreibungsfähigkeit
- Steuerbarkeit
- Erwartungskonformität
- Fehlertoleranz
- Individualisierbarkeit
- Lernförderlichkeit

IsoMetrics<sup>L</sup>

	stimmt nicht	stimmt wenig	stimmt mittelmäßig	stimmt ziemlich	stimmt sehr	Keine Angabe
A.3 Mit der Software kann ich zusammenhängende Arbeitsabläufe vollständig bearbeiten.	1	2	3	4	5	

	nicht wichtig	wenig wichtig	mittelmäßig wichtig	ziemlich wichtig	sehr wichtig	Keine Angabe
Wie wichtig ist dieser Aspekt für Ihren Gesamteindruck von der Software?	1	2	3	4	5	

Können Sie konkrete Beispiele nennen, bei denen Sie dieser Aussage zustimmen können?




# SMEQ

## Subjective Mental Effort Questionnaire

- Der SMEQ misst die kognitive Beanspruchung während der Nutzung

### Zentrale Frage

*„Wie anstrengend wird die Nutzung erlebt?“*

### Was wird gemessen:

- Kognitive Beanspruchung

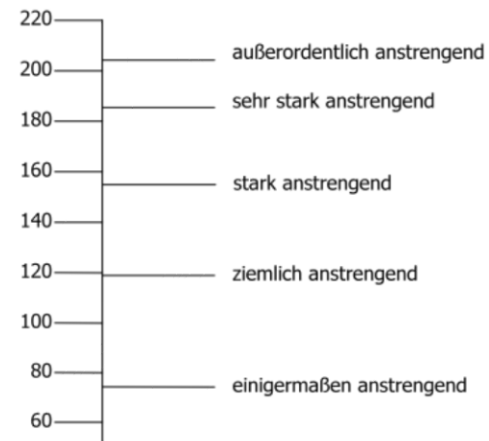
### Leichtgewichtige Alternative zu NASA TLX

#### Ihr Empfinden!

Geben Sie bitte auf der untenstehenden senkrechten Linie an (mit einem Kreuzchen oder Strich), wie anstrengend Sie die gerade bearbeitete Aufgabe empfunden haben.

Bitte machen Sie Ihr Kreuz auf Basis Ihrer **momentanen** Befindlichkeit.

Bitte ausfüllen...



# SAM

## Self Assessment Manikin

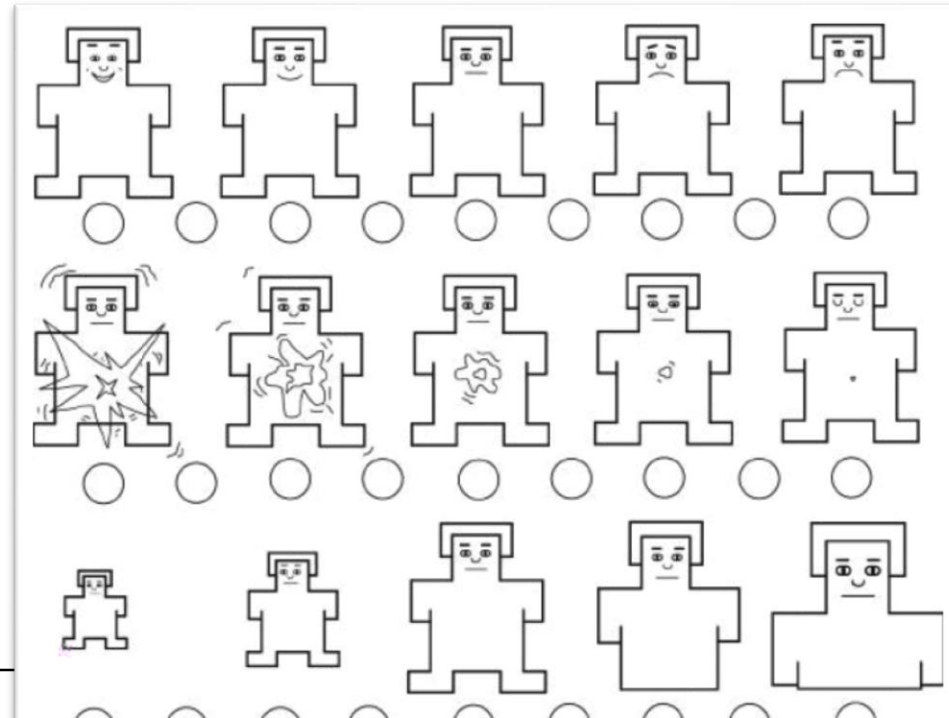
- Sprachfreies Messinstrument zur Messung der Dimensionen Valenz, Arousal und Dominanz.

## Zentrale Frage

*„Welche affektiven Reaktionen entstehen bei der Nutzung?“*

## Was wird gemessen:

- Valenz
- Arousal
- Dominanz



# PANAS

## Positive and Negative Affect Schedule

- misst positiven und negativen Affekt.

## Zentrale Frage

*„Wie fühlt man sich nach der Nutzung?“*

## Was wird gemessen:

- Positiver Affekt
- Negativer Affekt

... wie fühlen Sie sich jetzt?

	gar nicht	ein wenig	einigermaßen	erheblich	äußerst
erschrocken	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
angeregt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
aufmerksam	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
beschämt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
aktiv	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
nervös	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
interessiert	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
feindselig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
gereizt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
freudig erregt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
stark	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
durcheinander	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
stolz	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
entschlossen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## Telekom User Experience

- misst die Erfüllung von Bedürfnissen durch ein Produkt.

## Zentrale Frage

*„Welche Bedürfnisse werden bei der Nutzung angesprochen?“*

## Was wird gemessen:

- Wettkampf
- Sicherheit
- Verbundenheit
- Körperlichkeit
- Einfluss/Popularität
- Bedeutsames bewahren
- Kompetenz / Autonomie
- Stimulation

In der Situation gab mir das Produkt das Gefühl...

Items	gar nicht	ein wenig	einigermaßen	erheblich	äußerst
... etwas Besseres zu besitzen als andere.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... andere zu übertreffen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... im Vergleich zu anderen besser zu sein.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... andere zu übertrumpfen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... ein gut strukturiertes Leben zu führen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... einen routinierten, angenehmen Tagesablauf zu haben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... von Außerordentlichkeit und Regelmäßigkeit.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Routine in mein Leben zu bringen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... mit Menschen, die ich mag, und die mich mögen, verbunden zu sein.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Menschen, die mir wichtig sind, nahe zu sein.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... von Menschen umgeben zu sein.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... in Gesellschaft von Gleichgesinnten zu sein.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... körperlich aktiv zu sein.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... mich körperlich zu betätigen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... meine Gesundheit zu fördern.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... mich pflichtbewusst zu verhalten.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... mich für andere einzusetzen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

# INTUI

## INTUI

- Semantisches Differential zur Messung der Subkomponenten intuitiver Interaktion

### Zentrale Frage

„Wie intuitiv wird die Benutzung erlebt?“

### Was wird gemessen:

- Globale Intuitivität
- Mühelosigkeit
- Bauchgefühl
- Verbalisierungsfähigkeit
- Magisches Erleben

Bitte vergegenwärtigen Sie sich jetzt noch ein Mal die Nutzung des Produkts und **beschreiben Sie Ihr Erleben der Nutzung mit Hilfe der folgenden Aussagenpaare**. Die Paare stellen jeweils extreme Gegensätze dar, zwischen denen eine Abstufung möglich ist.

Vielleicht passen einige Aussagen nicht so gut, kreuzen Sie aber trotzdem bitte immer an, welcher Begriff Ihrer Meinung nach eher zutrifft. Denken Sie daran, dass es keine "richtigen" oder "falschen" Antworten gibt - nur Ihre persönliche Meinung zählt!

	1	2	3	4	5	6	7		
<b>Bei der Nutzung (des Produkts)...</b>									
...handelte ich überlegt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	...handelte ich spontan	G_01
...erreichte ich mein Ziel nur mit Anstrengung	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	...erreichte ich mein Ziel mit Leichtigkeit	M_01
...handelte ich unbewusst, ohne lange über die einzelnen Schritte nachzudenken	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	...führte ich bewusst einen Schritt nach dem anderen aus	P G_02
...ließ ich mich von meinem Verstand leiten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	...ließ ich mich von meinem Gefühl leiten	G_03
...war ich orientierungslos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	...konnte ich mich gut zurechtfinden	M_02
...handelte ich ohne dabei nachzudenken	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	...konnte ich jeden Schritt genau begründen	P G_04
<b>Die Nutzung (des Produkts)...</b>									
...erforderte viel Aufmerksamkeit	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	...ging wie von selbst	M_03
...war begeisternd	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	...war unbedeutend	P X_01
...war einfach	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	...war schwierig	P M_04
...war nichts Besonderes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	...war ein magisches Erlebnis	X_02

# Standard vs. Eigenentwicklung

---

## Standardisierte Methode

Es liegen bereits viele Vergleichswerte vor.

Es wird das *wahre* Konzept gemessen (Akzeptanz in der Forschungsgemeinde).

## Eigenentwicklung

Exakte Anpassung an die eigene Fragestellung / Produkte möglich.

Kosten- und zeitintensiver.

## Aber:

Kombinationen aus standardisierten Methoden und Eigenentwicklungen sind möglich.

# Gütekriterien

---

## Validität

Wie gut misst der Test das, was er messen soll.

## 2 Komponenten der Validität:

### Konvergente Validität

Mehrere Methoden messen das gleiche Konstrukt.

Fragebögen, die Gleiches messen, sollen hoch miteinander korrelieren.

### Diskriminante Validität

Verschiedene Methoden messen unterschiedliche Konstrukte.

Fragebögen, die Unterschiedliches messen, sollen niedrig miteinander korrelieren.

---

# Statistische Methoden

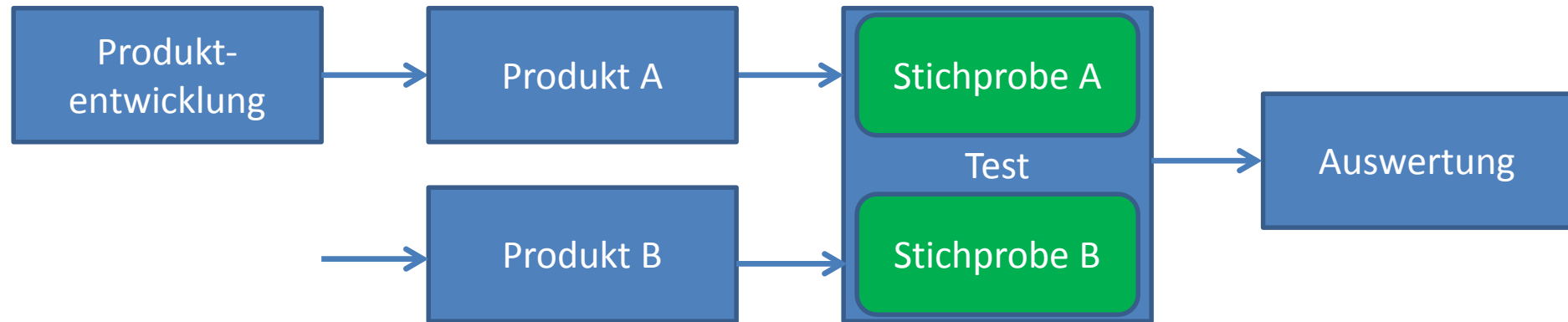
---



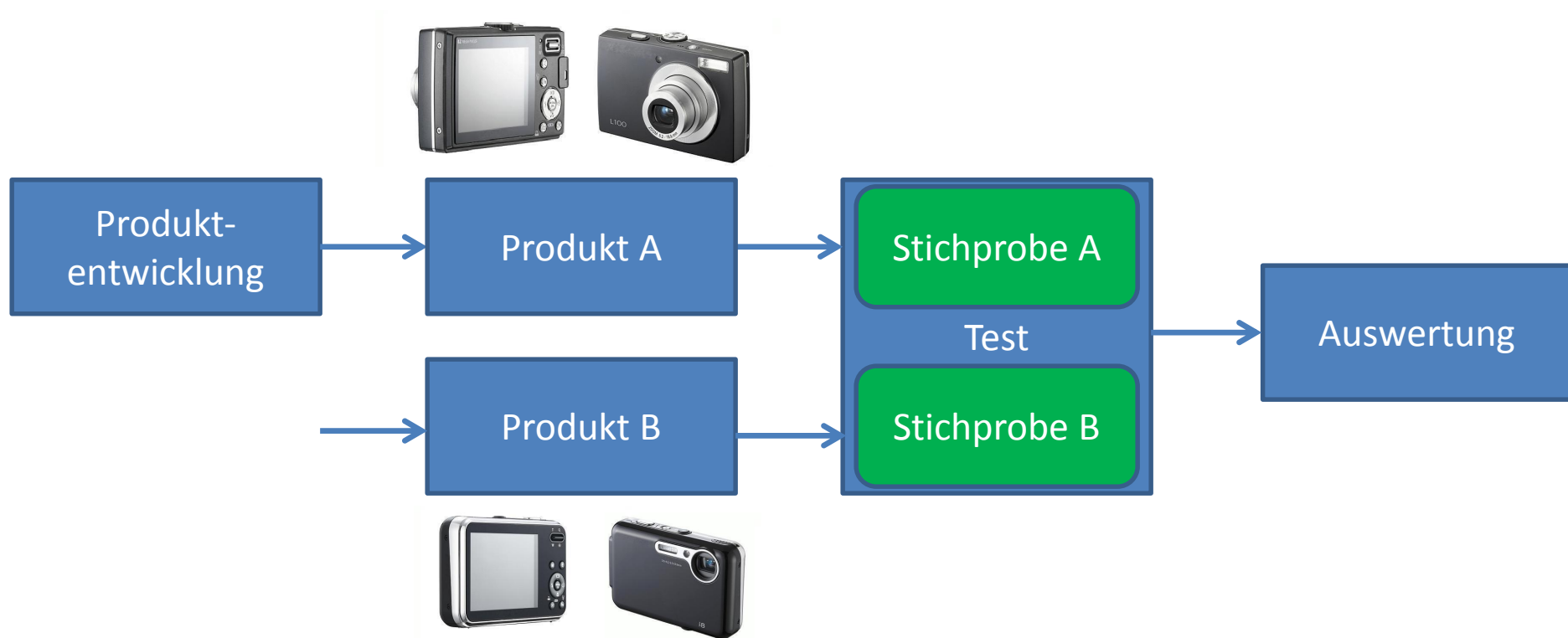
# Statistische Methoden

---

- \_ Testen und Messen ist nur der erste Schritt
  
- \_ Schlüsse und Ableitungen können nur gezogen werden, wenn die Messwerte systematisch und nicht zufällig entstanden sind



# Statistische Methoden



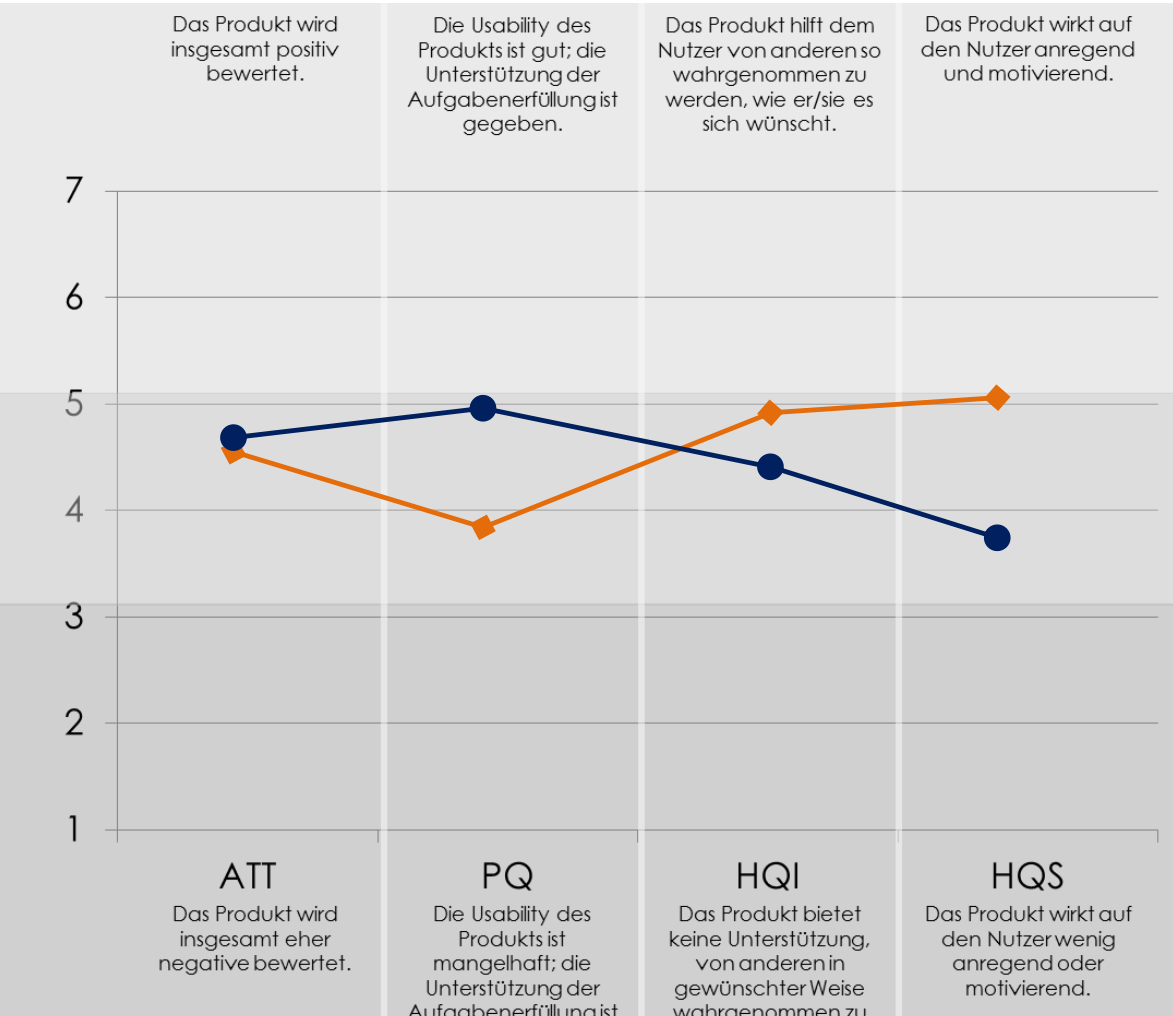
# Statistische Methoden



**Samsung L100**



**Samsung i8**



# Statistische Methoden

## Pragmatische Qualität



**Samsung L100**

Mean: 5,0

vs



**Samsung i8**

Mean: 3,8

# Statistische Methoden

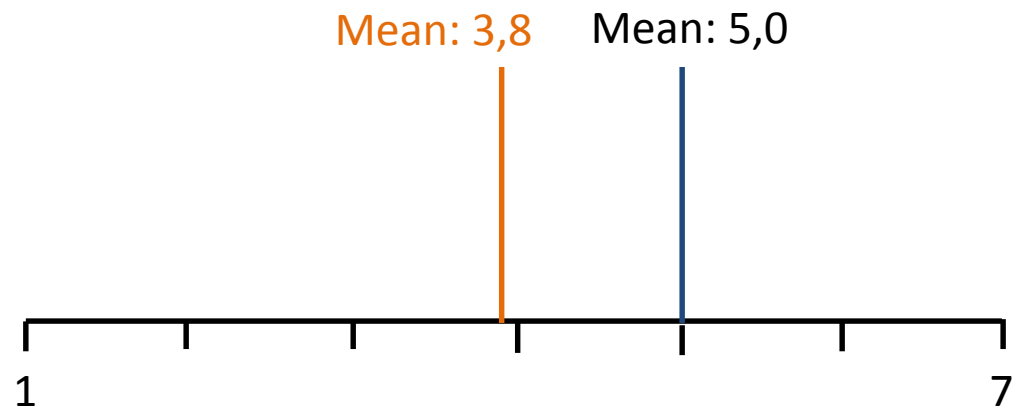
## Pragmatische Qualität



**Samsung L100**



**Samsung i8**



# Statistische Methoden

## Pragmatische Qualität



**Samsung L100**

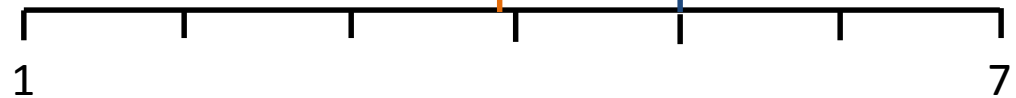


**Samsung i8**

**Bedeutsamer Unterschied?!**

Mean: 3,8

Mean: 5,0



# Statistische Methoden

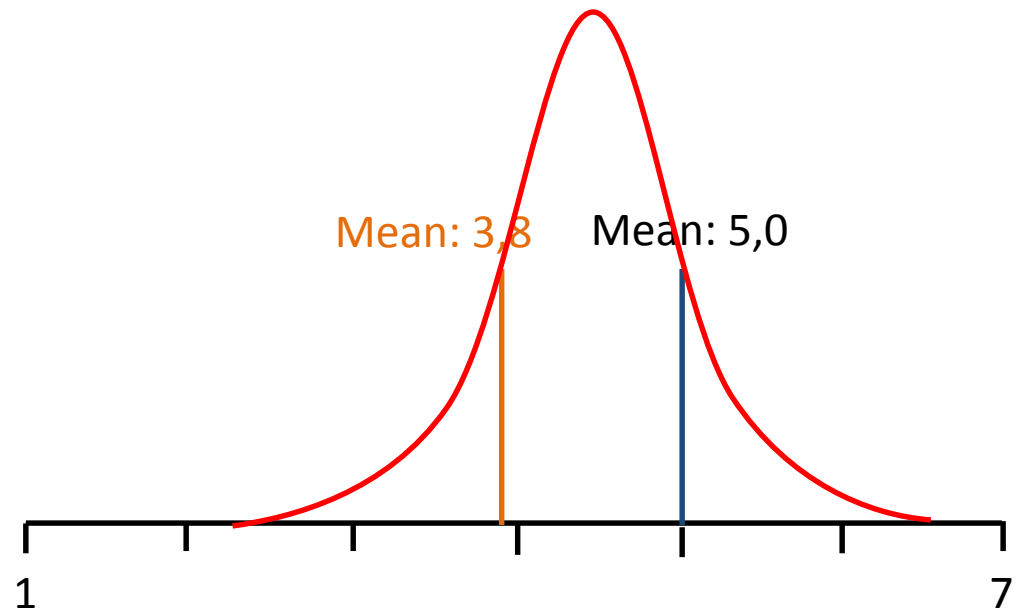
## Pragmatische Qualität



**Samsung L100**



**Samsung i8**





# Statistische Methoden

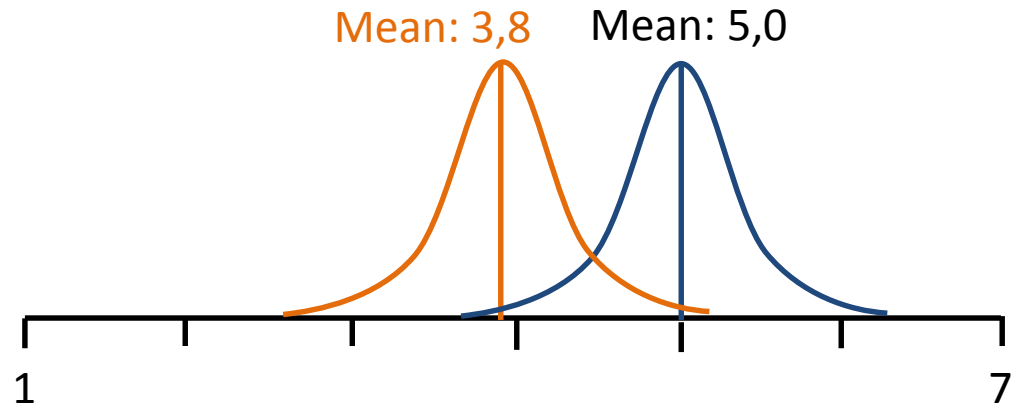
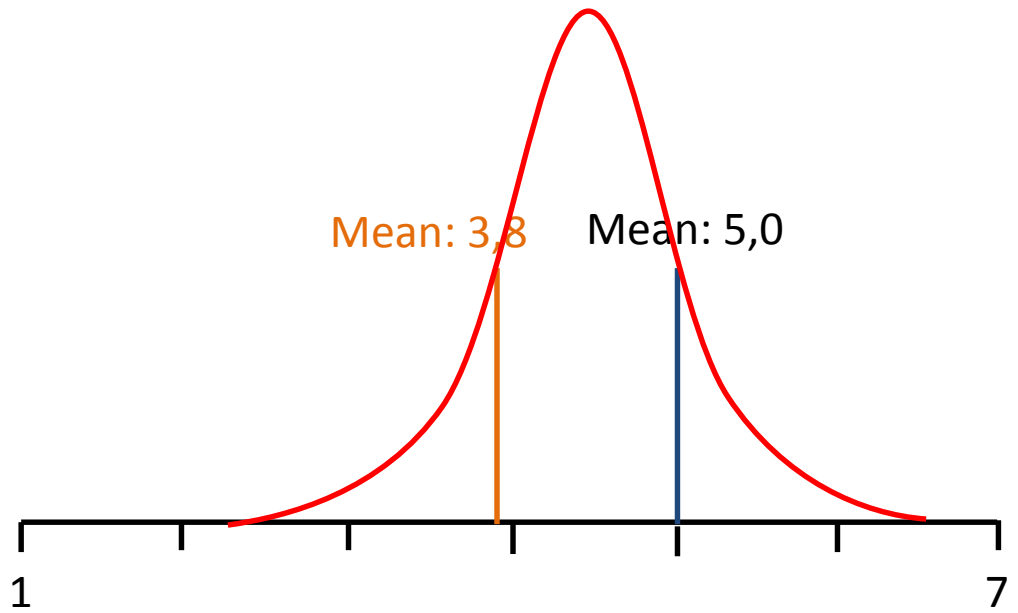
Pragmatische Qualität



Samsung L100



Samsung i8



## T-Test für unabhängige Stichproben

- Prüft ob zwei Stichproben aus der gleichen Grundgesamtheit stammen oder aus unterschiedlichen Teilpopulationen
- $H_0$ : Es gibt keine Unterschiede zwischen den Gruppen  
(→ gleiche Grundgesamtheit; Unterschiede entstanden zufällig)
- $H_1$ : Es gibt Unterschiede zwischen den Gruppen  
(→ unterschiedliche Teilpopulationen; Unterschiede entstanden systematisch)

## T-Test für unabhängige Stichproben

- Prüft ob zwei Stichproben aus der gleichen Grundgesamtheit stammen oder aus unterschiedlichen Teilpopulationen
- $H_0$ : Es gibt keine Unterschiede zwischen den Gruppen  
(→ gleiche Grundgesamtheit; Unterschiede entstanden zufällig)
- $H_1$ : Es gibt Unterschiede zwischen den Gruppen  
(→ unterschiedliche Teilpopulationen; Unterschiede entstanden systematisch)

## — Ergebnis des T-Tests:

- T-Wert
- p-Wert („Irrtumswahrscheinlichkeit“)

## p-Wert

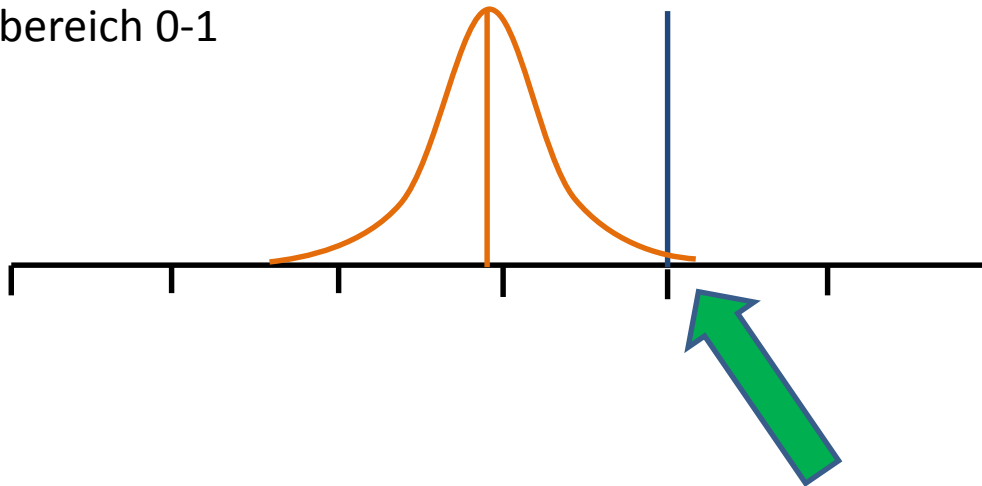
- „Irrtumswahrscheinlichkeit“
- „Irrtum“ wäre:  $H_0$  wird verworfen, obwohl sie eigentlich wahr ist (man nimmt Unterschiede an, obwohl es keine gibt).
- Wunschergebnis: möglichst kleiner p-Wert

# Statistische Methoden

## p-Wert

- „Irrtumswahrscheinlichkeit“
- „Irrtum“ wäre:  $H_0$  wird verworfen, obwohl sie eigentlich wahr ist (man nimmt Unterschiede an, obwohl es keine gibt).
- Wunschergebnis: möglichst kleiner p-Wert

- Wertebereich 0-1



## p-Wert

- „Irrtumswahrscheinlichkeit“
- „Irrtum“ wäre:  $H_0$  wird verworfen, obwohl sie eigentlich wahr ist (man nimmt Unterschiede an, obwohl es keine gibt).
- Wunschergebnis: möglichst kleiner p-Wert
- Wertebereich 0-1
  - $p > .1$  → nicht signifikant
  - $p < .1$  → „tendenziell signifikant“
  - $p < .05$  → signifikant
  - $p < .01$  → sehr signifikant

## „Signifikanz“

- Signifikanz = Bedeutsamkeit
- In Zusammenhang mit statistischen Methoden wird mit Signifikanz meist die **statistische Signifikanz** gemeint

## „Signifikanz“

- Signifikanz = Bedeutsamkeit
- In Zusammenhang mit statistischen Methoden wird mit Signifikanz meist die **statistische Signifikanz** gemeint
- Statistische Signifikanz = Unterschiede sind nicht durch **Zufall** entstanden (das sagt noch nichts über die Größe der Unterschiede aus)
- Praktische Signifikanz = Bedeutsame **Größe** der Unterschiede (wird quantifiziert durch die *Effektgröße*)
- Statistische und praktische Signifikanz sind völlig unabhängig voneinander. Es existieren alle Kombinationen (zudem können beide Größen bewusst „gepusht“ werden)



## T-Test für unabhängige Stichproben

- Ergebnis wird durch viele Faktoren beeinflusst:
  - Mittelwertsdifferenz  
(wie weit liegen die Mittelwerte auseinander)
  - Varianzen  
(wie homogen waren die Antworten im Fragebogen)
  - Freiheitsgrade  
(wieviele Personen haben an der Studie teilgenommen)

# Statistische Methoden

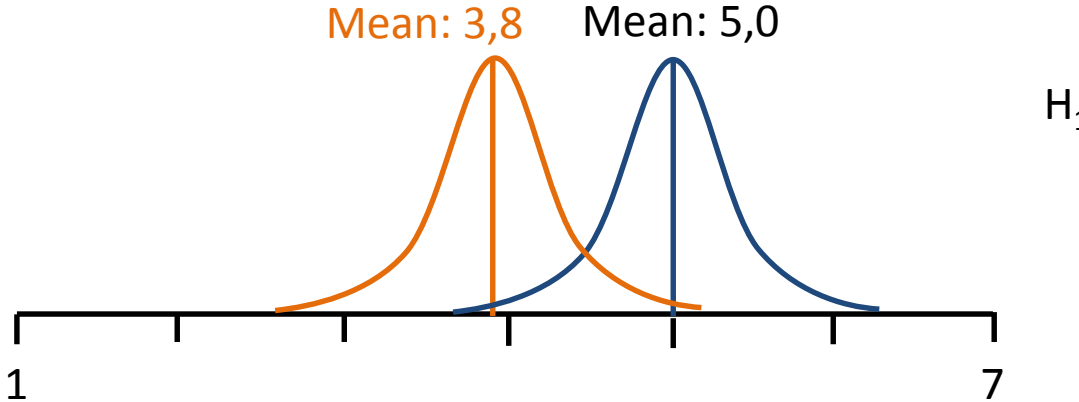
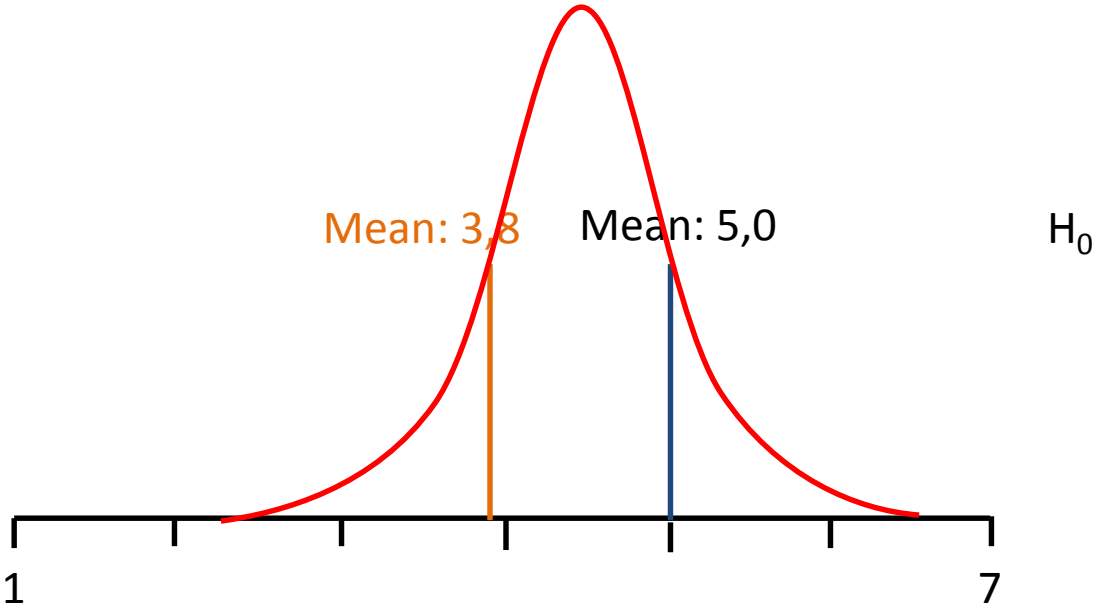
## Pragmatische Qualität



Samsung L100



Samsung i8



# Statistische Methoden

## Pragmatische Qualität



**Samsung L100**



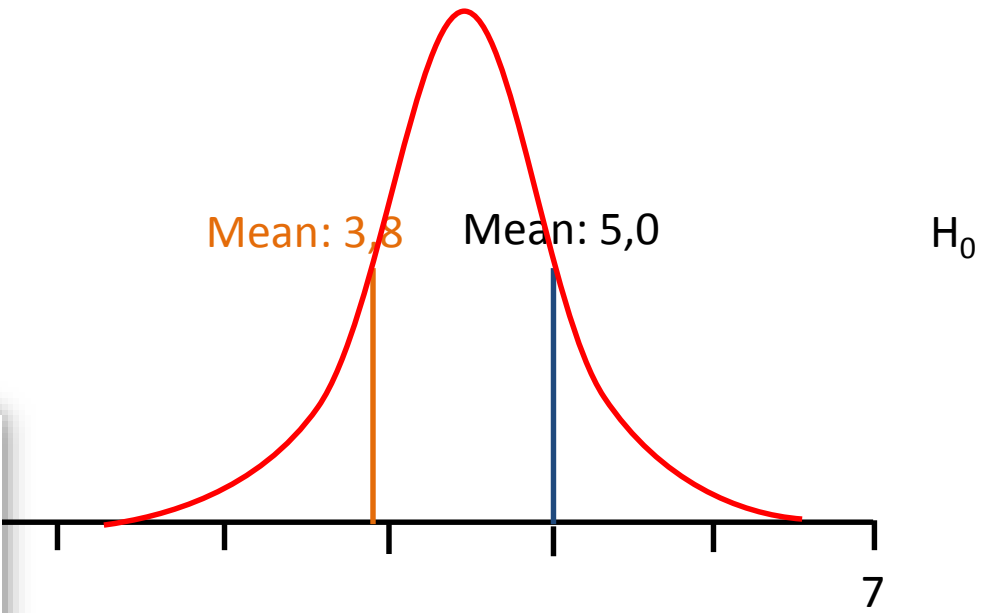
**Samsung i8**

**Gruppenstatistiken**

Handy	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
PQ Samsung i8	10	3,7900	1,29996	,41108
Samsung L100	10	5,0000	1,05409	,33333

**Test bei unabhängigen Stichproben**

	Levene-Test der Varianzgleichheit		T-T		
	F	Signifikanz	T	df	Sig. (2-seitig)
PQ Varianzen sind gleich	,000	,993	-2,286	18	,035
Varianzen sind nicht gleich			-2,286	17,263	,035



# Statistische Methoden

## Pragmatische Qualität



**Samsung L100**



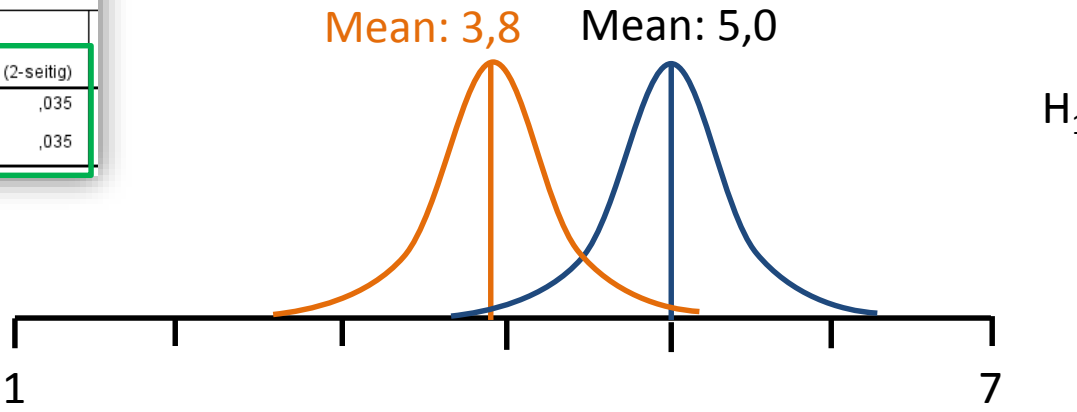
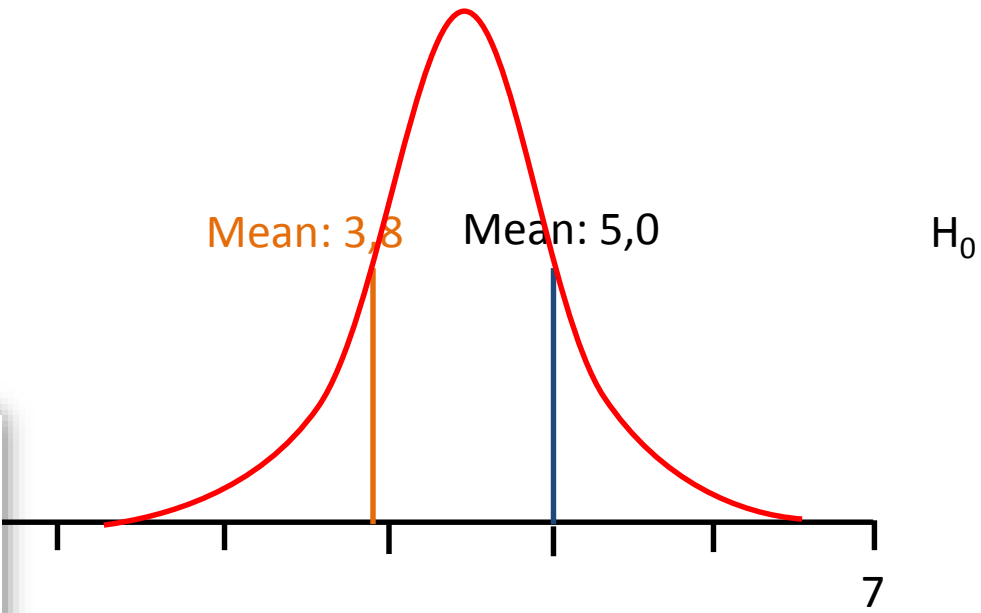
**Samsung i8**

Gruppenstatistiken

	Handy	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
PQ	Samsung i8	10	3,7900	1,29996	,41108
	Samsung L100	10	5,0000	1,05409	,33333

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-T		
		F	Signifikanz	T	df	Sig. (2-seitig)
PQ	Varianzen sind gleich	,000	,993	-2,286	18	,035
	Varianzen sind nicht gleich			-2,286	17,263	,035



# Statistische Methoden

## Pragmatische Qualität



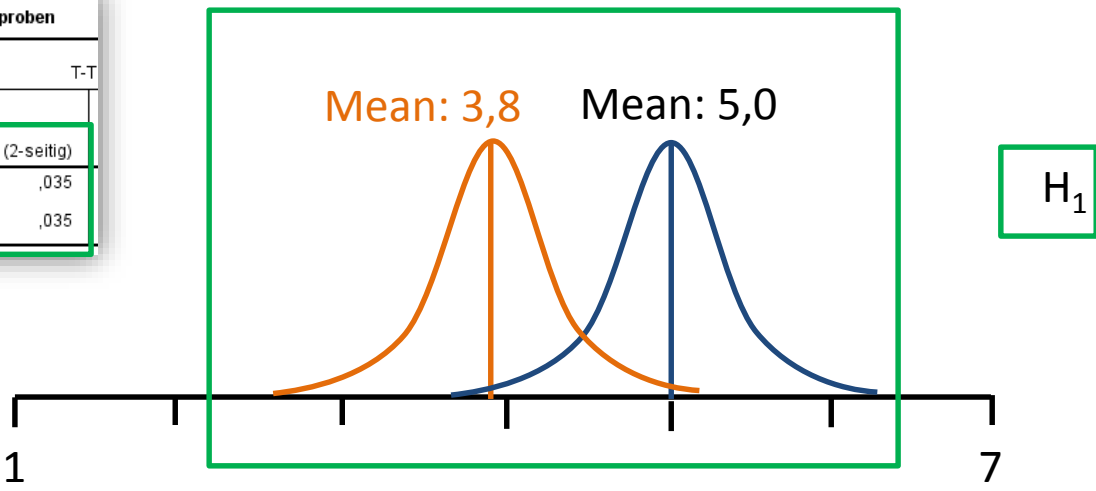
**Samsung L100**



**Samsung i8**

	Handy	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
PQ	Samsung i8	10	3,7900	1,29996	,41108
	Samsung L100	10	5,0000	1,05409	,33333

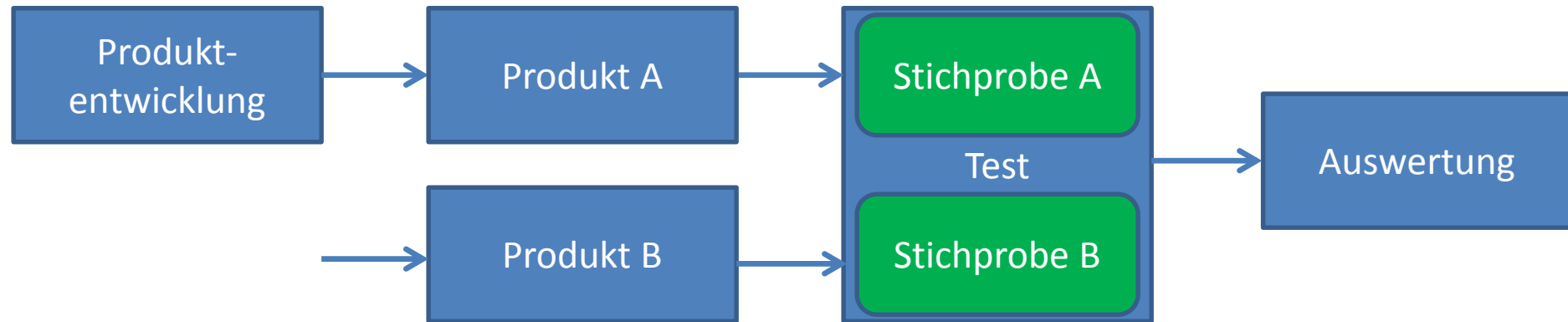
		Levene-Test der Varianzgleichheit		T-T		
		F	Signifikanz	T	df	Sig. (2-seitig)
PQ	Varianzen sind gleich	,000	,993	-2,286	18	,035
	Varianzen sind nicht gleich			-2,286	17,263	,035

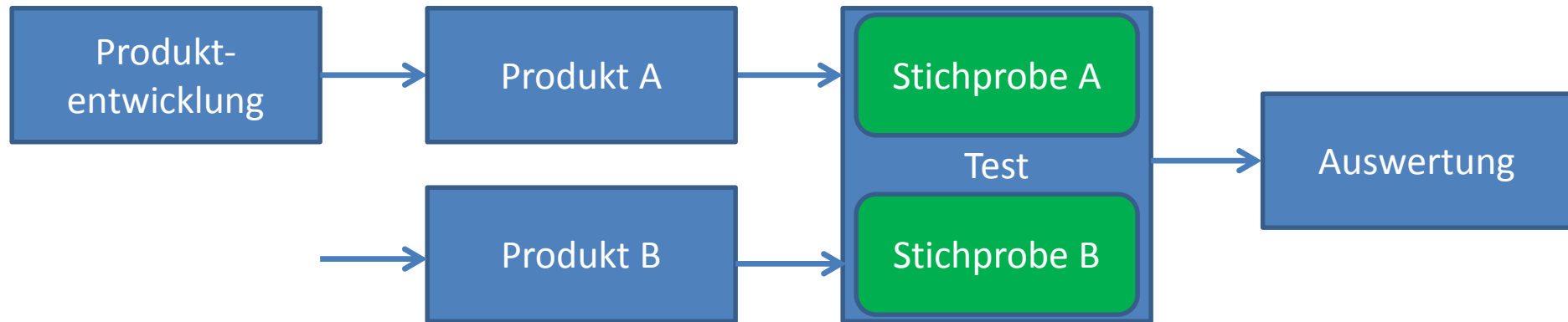


---

**Welcher Test ist der Richtige?**

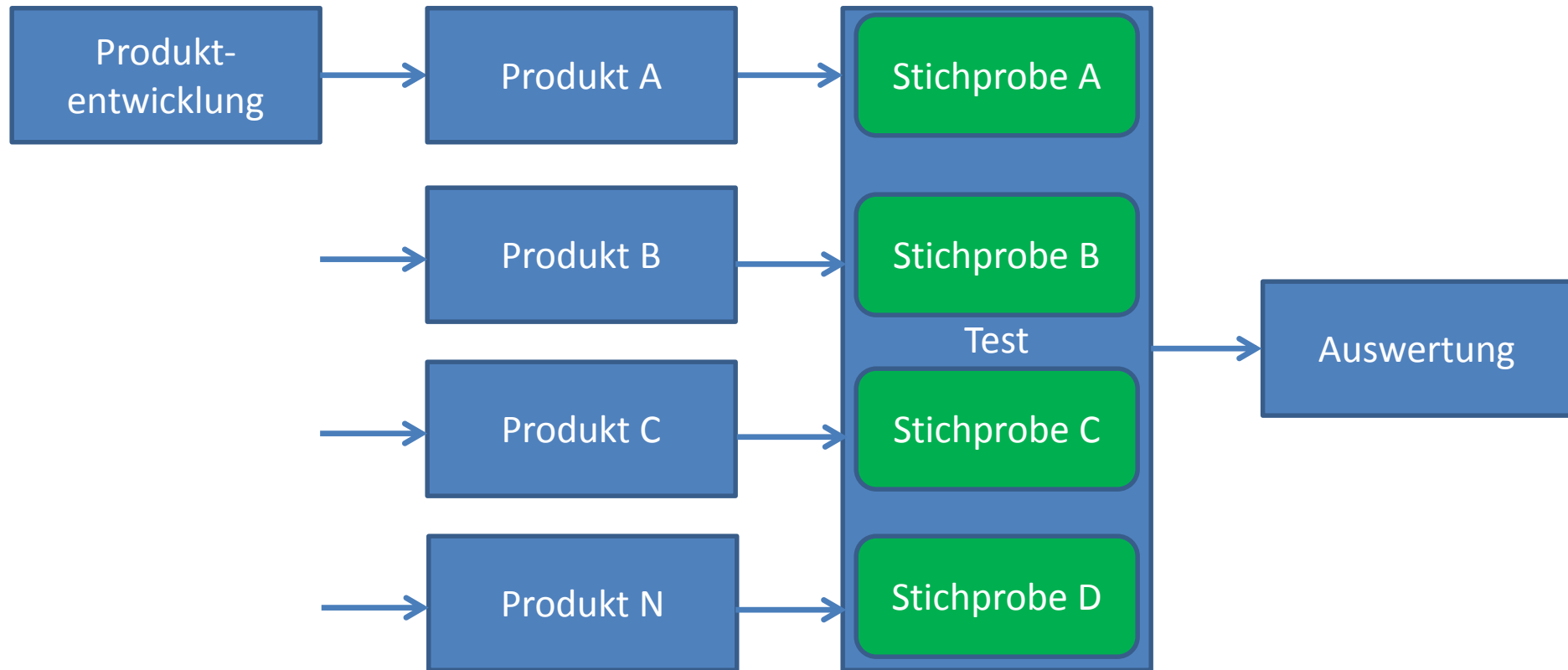
---



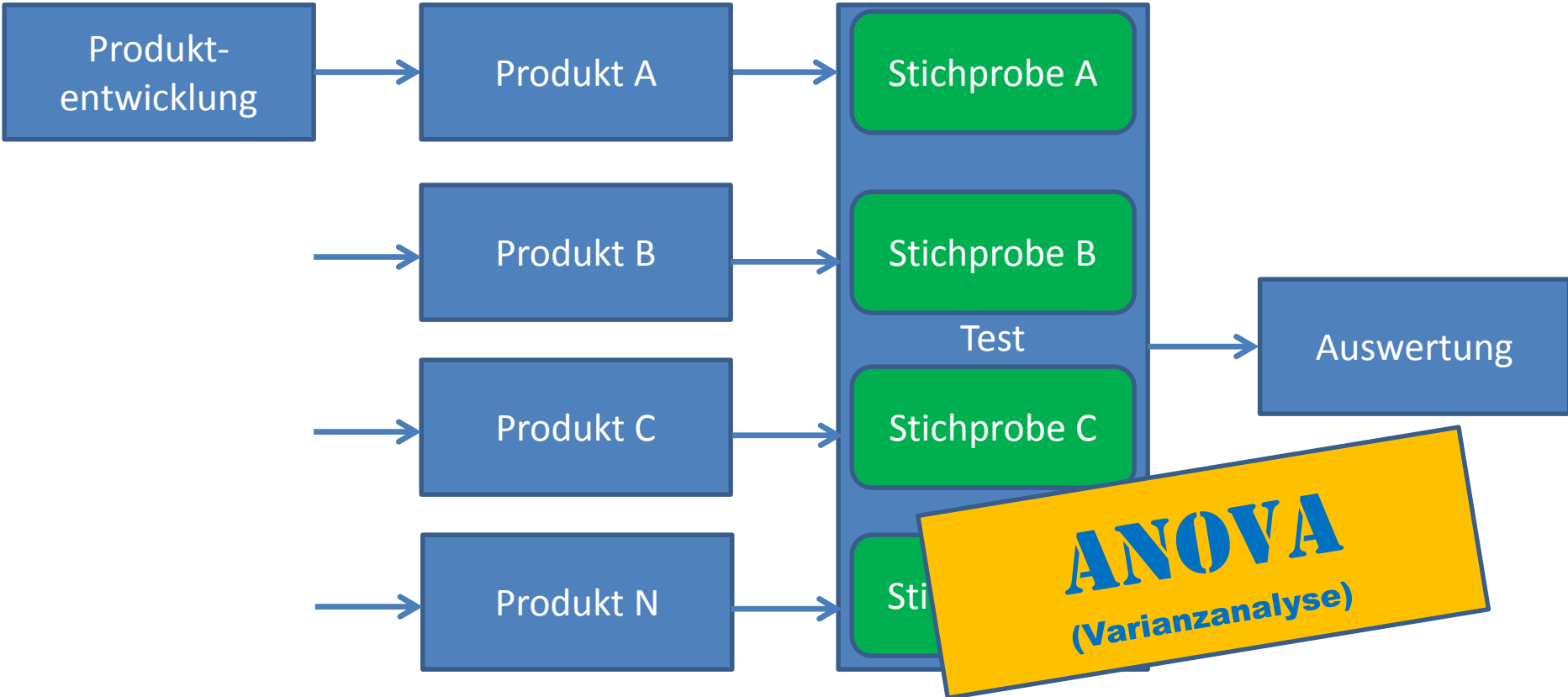


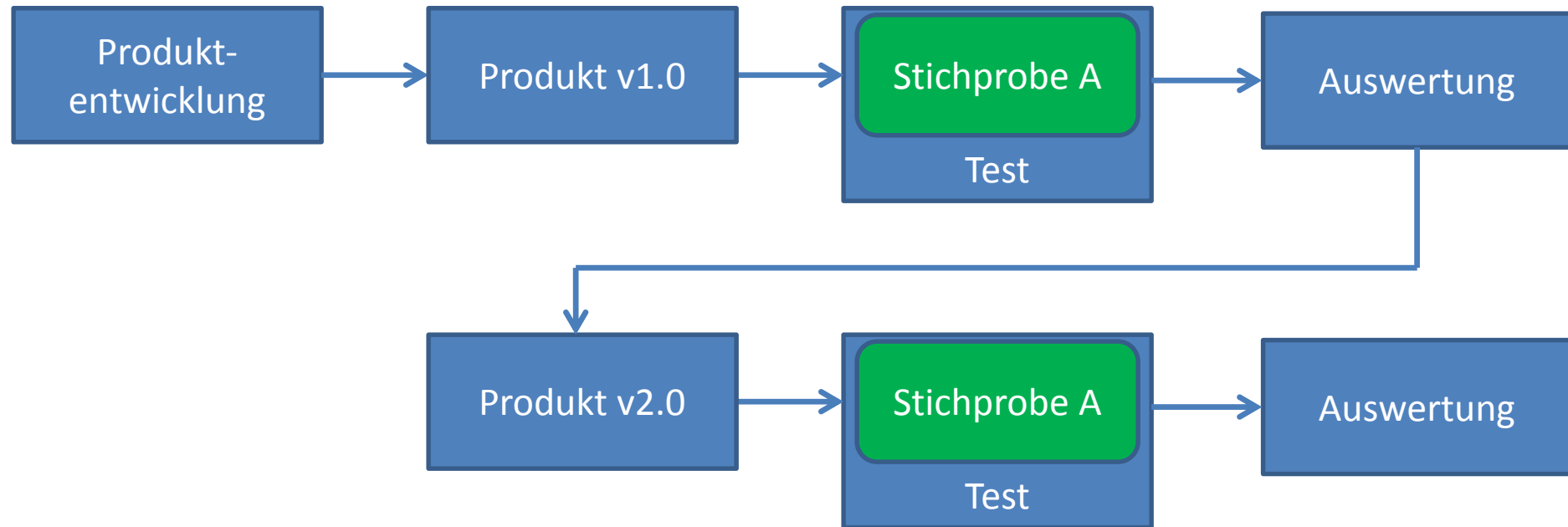
**T-TEST**  
(unabhängige Stichproben)



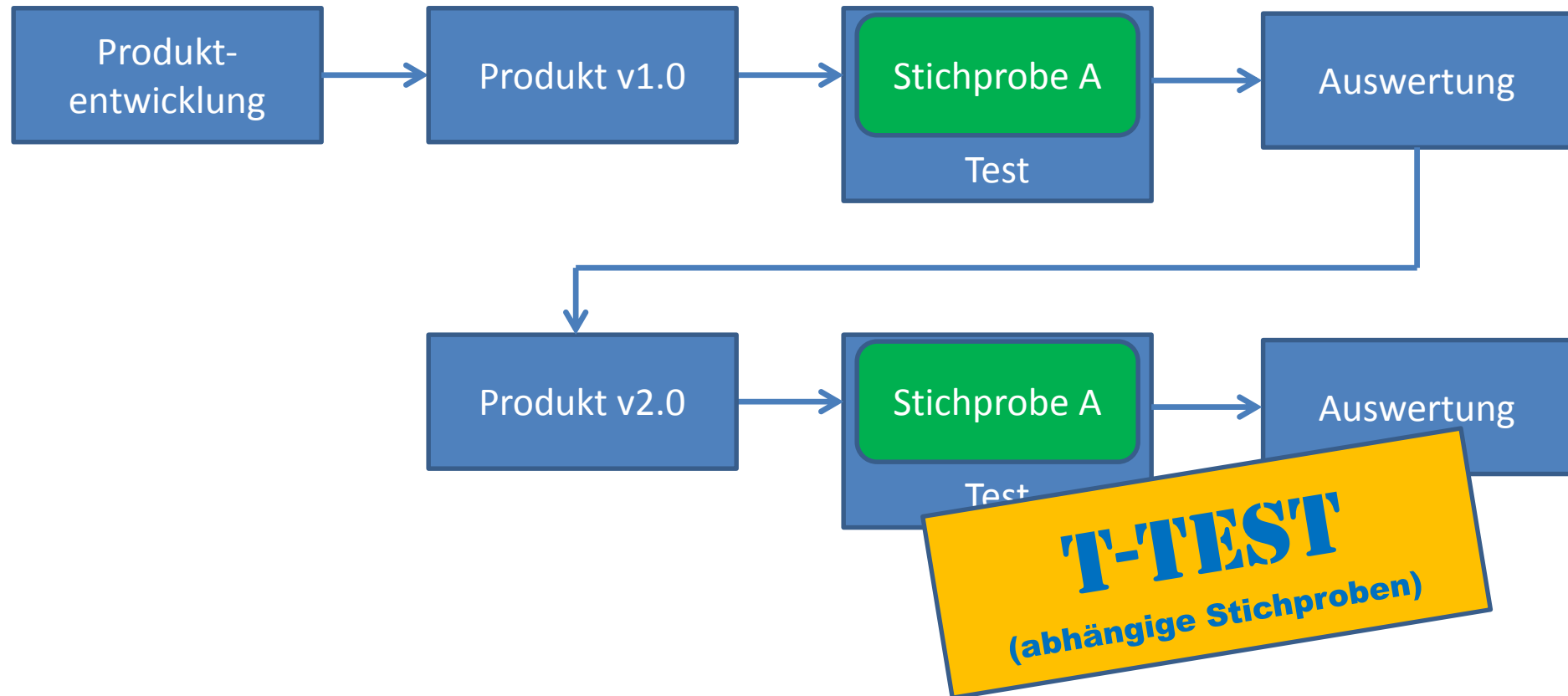


# Statistische Methoden





# Statistische Methoden



- \_ Was soll geprüft werden?
  - Unterschiede vs. Zusammenhänge
    - T-Test, Varianzanalyse, U-Test, Wilcoxon...
    - Korrelation, Regression

# Statistische Methoden

---

## \_ Was soll geprüft werden?

- Unterschiede vs. Zusammenhänge
  - T-Test, Varianzanalyse, U-Test, Wilcoxon...
  - Korrelation, Regression

## \_ Welche Datenstruktur liegt vor?

- Intervalldaten vs. Ordinaldaten vs. Nominaldaten
- Anzahl Gruppen
  - Abhängig vs. unabhängig
- Anzahl abhängiger Variablen

- \_ Sind die Voraussetzungen erfüllt?
  - Z.b. T-Test:
    - Normalverteilte Daten
    - Homogene Varianzen
    - Intervallskalenniveau
  - Was, wenn die Voraussetzungen nicht erfüllt sind?
    - Fehlerkorrekturen (z.b. Anpassung der Freiheitsgrade)
    - Auf Test mit geringeren Anforderungen ausweichen

---

# **Stolpersteine der Evaluation**

---



# Stolpersteine der Evaluation.

---

Die Methode bestimmt das Problem

Bestimmte Methoden sind akzeptierter als andere

Also wird zuerst die Methode gewählt und dann ein entsprechendes Problem, "Was könnte ich denn mal mit dem Eyetracker untersuchen?"

Evaluation ist oft bewusst konfirmatorisch

Es werden (implizit) Situationen kreiert, in denen das System nicht so schlecht abschneidet

Statt den expliziten Versuch der Falsifizierung zu unternehmen

# Stolpersteine der Evaluation.

---

Mehr ist nicht unbedingt besser

Gefahr durch Überforderung des Nutzers (Ermüdungseffekte und sinkende Akzeptanz).

Oft ist schon ein Urteil aussagekräftig (viele Messinstrumente messen ähnliches).

Vielen Dank für die Aufmerksamkeit!

Daniel Ullrich

[daniel.ullrich@ifi.lmu.de](mailto:daniel.ullrich@ifi.lmu.de)