



Kontrolle von Randbedingungen im Online-Experiment

Timo Gnambs

timo.gnambs@jku.at

Wien, 2010-05-21

8. Wiener Methodenforum



Überblick

1. Merkmale internetbasierter Forschung
2. Varianten internetbasierter Experimente
3. Probleme bei unkontrollierten Web-Experimenten
4. Studie: Kontextfaktor Farbe



1. Merkmale internetbasierter Forschung



Aspekte internetbasierte Forschung

Unkontrollierte Computererhebung an anfallenden Stichproben?

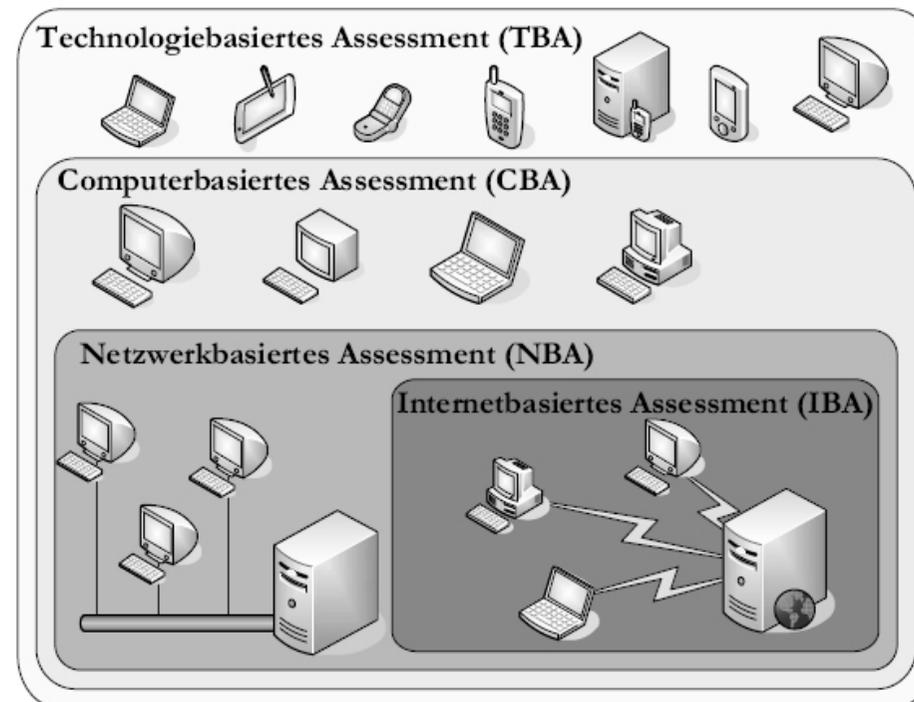
Medium

Modus

Selektion

Gnambs, T., Batinic, B., & Hertel, G. (in Druck). Internetbasierte psychologische Diagnostik. In L. F. Hornke, M. Amelang & M. Kersting (Hrsg.), *Enzyklopädie der Psychologie, Psychologische Diagnostik* (Bd. II/3). Göttingen: Hogrefe.

Technologiebasiertes Assessment



Jurecka, A., & Hartig, J. (2007). Anwendungsszenarien computer- und netzwerkbasierter Assessments. In J. Hartig & E. Klieme (Hrsg.), *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik* (S. 69-79). Berlin: BMBF.



Durchführungsbedingungen

Unkontrollierte Durchführung

(„unproctored testing“)

Kontrollierte Durchführung

Überwachte Durchführung

(„proctored testing“)

Bartram, D. (2006). Testing on the internet: Issues, challenges and opportunities in the field of occupational assessment. In D. Bartram & R. K. Hambleton (Hrsg.), *Computer-Based Testing and the Internet* (S. 13-37). New York: Wiley.



Internetbasierte Forschung

- Betont vor allem das Durchführungsmedium als zentrales Element der Datenerhebung (**Medium**).
- Kann sowohl unkontrolliert wie auch überwacht (**Modus**) erfolgen.
- Kann sowohl an anfallenden Stichproben wie auch echten Zufallsstichproben (**Selektion**) erfolgen.

Gnambs, T., Batinic, B., & Hertel, G. (in Druck). Internetbasierte psychologische Diagnostik. In L. F. Hornke, M. Amelang & M. Kersting (Hrsg.), *Enzyklopädie der Psychologie, Psychologische Diagnostik* (Bd. II/3). Göttingen: Hogrefe.



2. Varianten internetbasierter Experimente

Web-Experimente mit standardisierter Befragung

- Online-Fragebogen mit experimentell variierten Elementen
- „Klassische“ Form des Web-Experiments

Welche Art von Mobilfunkplan nutzen Sie derzeit? Haben Sie einen Vertrag (Abonnement, Sie zahlen eine Rechnung am Ende eines jeden Monats) oder nutzen Sie eine Prepaid-Karte?

Vertrag (Abonnement)

Prepaid-Karte

Wer entscheidet über die Wahl Ihres Mobilfunkanbieters?

Ich

Firma / Arbeitgeber

Andere

Keine Angabe

Welcher ist Ihr hauptsächlichster Mobilfunkanbieter?

Swisscom

Orange

Sunrise

Tele 2

Coop-Mobile

M-Budget

Cablecom

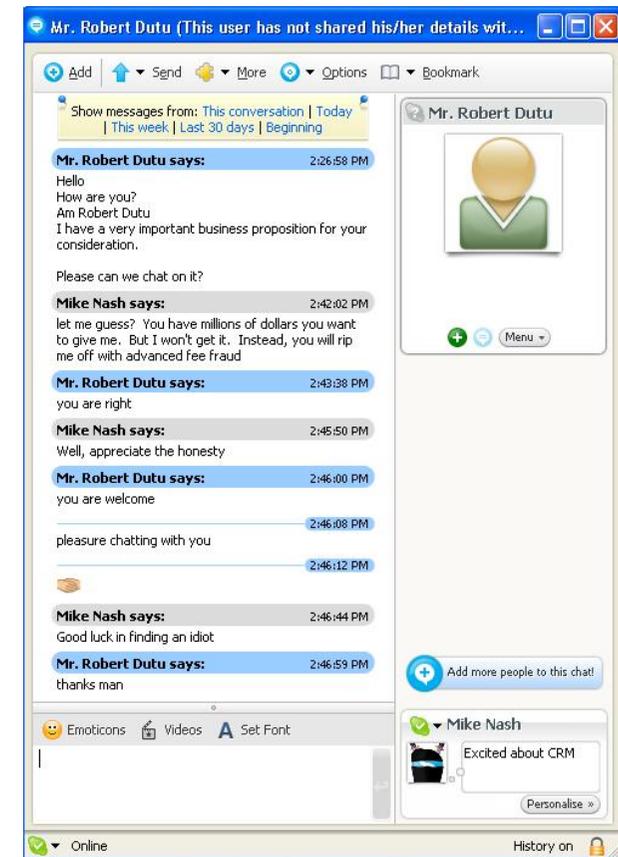
Anderer

[»»»](#)

Qualitative Web-Experimente

- Chat-Interviews mit experimentell variierten Fragen
- Häufig als Online-Feldexperiment
- Z.B. zur Untersuchung von sozial stark negativ konnotierten Einstellungen und Verhalten (vgl. Glaser, Dixit & Green, 2002).

Glaser, J., Dixit, J., & Green, D. P. (2002). Studying hate crime with the Internet: What makes racists advocate racial violence? *Journal of Social Issues*, 58 , 177-193.



Virtueller Supermarkt

- 3D-Simulationen traditioneller Einkaufsmärkte für virtuelle Verkaufstests
- Experimentelle Variation verschiedener Produktgestaltungen oder -platzierungen



Virtueller Supermarkt des Lehrstuhls für Wirtschaftsinformatik der RWTH Aachen.

Verfügbar unter <http://www.winfor.rwth-aachen.de/inhalte/forschung/vrmafo/supermarkt/eingabe.html>

Virtuelle Spielumgebungen

- Gestaltung virtueller Spielumgebungen
- Z.B. Visualisierung von Produktionsabläufen



Frey, A., Hartig, J., Ketzler, A., Zinkernagel, A., & Moosbrugger, H. (2007). The use of virtual environments based on a modification of the computer game Quake III Arena® in psychological experimenting. *Computers in Human Behavior*, 23, 2026-2039.



Experimentelle Logfileanalyse und Verhaltensbeobachtung

```
14.45.137 - - [23/Jun/2008:04:19:21 +0200] "GET /kontakt/email/ HTTP/1.0" 200 17823 "-" "Mozilla/5.0 (compatible; Ask Jeeves/Teoma; +ht  
49.65.40 - - [23/Jun/2008:04:19:21 +0200] "GET /rezensionen/liste HTTP/1.1" 200 42329 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http;  
.8.105 - - [23/Jun/2008:04:26:56 +0200] "GET /robots.txt HTTP/1.0" 404 14494 "-" "Mozilla/5.0 (compatible; Yahoo! Slurp; http://help.yahoo  
95.51.233 - - [23/Jun/2008:04:26:58 +0200] "GET /impressum/haftungsausschluss HTTP/1.0" 200 19071 "-" "Mozilla/5.0 (compatible; Yahoo!  
49.65.40 - - [23/Jun/2008:04:48:26 +0200] "GET /links/netzwerkpartner HTTP/1.1" 200 17935 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +h  
208.132.53 - - [23/Jun/2008:04:58:17 +0200] "POST /sites/default/reviews/xml.php HTTP/1.0" 200 5032 "-" "-"  
.22.89 - - [23/Jun/2008:04:58:17 +0200] "GET /rezensionen/rezension/65/ HTTP/1.0" 200 17466 "-" "Mozilla/5.0 (compatible; Yahoo! Slurp; h  
22.182 - - [23/Jun/2008:05:04:18 +0200] "GET /robots.txt HTTP/1.0" 404 14494 "-" "Mozilla/5.0 (compatible; Yahoo! Slurp; http://help.yaho
```

- Beobachtung von Nutzerverhalten
- Experimentelle Variation von Gestaltungselementen bei Webauftritten in Usabilitystudien
- Sozialpsychologische Online-Feldexperimente zu ethnische Vorurteilen in online Auktionshäusern (vgl. Shohat & Musch, 2003)

Shohat, M., & Musch, J. (2003). Online auctions as a research tool: A field experiment on ethnic discrimination. *Swiss Journal of Social Psychology*, 62, 193-1945.



Internetbasierte Experimente

▪ **Web-Experimente**

- Teilnehmer greifen auf Stimulusmaterial auf Experimentalserver zu (z.B. Fragebogen mit experimenteller Manipulation)
- Im Vergleich zu traditionellen Laborexperimenten geringere Standardisierung

▪ **Online-Feldexperimente**

- Erhebt i.R. „natürliches“ Verhalten
- Häufig keine a priori Aufklärung über Experimentalteilnahme



Arten internetbasierter Experimente

	Befragen	Beobachten
Web-Experiment	Fragebogen Interview	Simulationen
Online-Feldexperiment	Interview	Logfileanalyse Verhaltensanalyse



3. Probleme von Web-Experimenten



Probleme von unkontrollierten Web-Experimenten

- **Systematischer Dropout**

Teilnahmeabbrüche sind mit der experimentellen Manipulation konfundiert und treten in einer oder mehreren Versuchsbedingungen gehäuft auf.

- **Geringe Situationskontrolle**

Eine größere Varianz der Umgebungsvariablen (Ausstattung, auftretende technische Störungen etc.) kann die interne Validität des Experiments beeinträchtigen.



Ansätze zur Kontrolle von Dropout (Reips, 2002)

- **Warm-Up Technik**

Experimentelle Variation wird so weit wie möglich verzögert, da Teilnahmeabbrüche zumeist am Anfang vorkommen.

Kann Dropout während des Experiments auf bis zu 2% senken (Reips, 2009)

- **Incentives**

Belohnung steigern Bereitschaft Experimente zu beginnen und bis zum Ende fertig zu machen (Göritz, 2006)

Aber: wenig über Datenqualität bekannt.

- **High-Hurdle Technik:**

Dropout wird am Anfang durch hohe Einstiegsbarrieren provoziert (z.B. lange Ladezeiten, lange Texte, schlechtes Design). Dadurch brechen wenig motivierte Personen frühzeitig ab, während hoch motivierte Personen die Einstiegshürde überwinden und bis zum Ende teilnehmen

Aber: keine empirische Bestätigung (Göritz & Stieger, 2008)!

Reips, U.-D. (2002). Standards for Internet-based experimenting. *Experimental Psychology*, 49, 243-256.

Göritz, A. (2006). Incentives in web studies: Methodological issues and a review. *Journal of Internet Studies*, 1, o. S. Verfügbar unter <http://www.ijis.net> [2010-05-20].

Göritz, A. & Stieger, S. (2008). The high-hurdle technique put to the test: Failure to find evidence that increasing loading times enhances data quality in Web-based studies. *Behavior Research Methods*, 40, 322-327.



Seriousness-Check

How do you intend to browse the Web pages of this study?

- I would like to seriously participate now.
- I would like to look at the pages only.

- Um Personen zu identifizieren, die lediglich aus Interesse das Studienmaterial ansehen, aber nicht ernsthaft teilnehmen.
- Reduziert die Dropout-Rate von ca. 30-50% auf 10-15% der Teilnehmer (Reips, 2009)

Reips, U.-D. (2009). Internet experiments: methods, guidelines, metadata. *Human Vision and Electronic Imaging XIV* 7240 (1), 724008.



Sports Participation

Most modern theories of decision making recognize the fact that decisions do not take place in a vacuum. Individual preferences and knowledge, along with situational variables can greatly impact the decision process. In order to facilitate our research on decision making we are interested in knowing certain factors about you, the decision maker. Specifically, we are interested in whether you actually take the time to read the directions; if not, then some of our manipulations that rely on changes in the instructions will be ineffective. So, in order to demonstrate that you have read the instructions, please ignore the sports items below, as well as the continue button. Instead, simply click on the title at the top of this screen (i.e., "sports participation") to proceed to the next screen. Thank you very much.

Which of these activities do you engage in regularly?
(click on all that apply)

skiing

soccer

snowboarding

running

hockey

football

swimming

tennis

basketball

cycling

Continue

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, 45, 867-872.



Potentielle Kontextfaktoren bei unkontrollierten Web-Experimenten

Technologiefaktoren

- Bildschirm
- Geschwindigkeit
- Mousetyp
- Prozessorleistung
- ...

Interfacefaktoren

- Scrollen
- Item Navigation
- Itemdarstellung
- Farbgestaltung
- ...

Personenfaktoren

- Geschlecht
- Allgemeine kognitive Fähigkeit
- Erfahrung mit Computern
- Computerängstlichkeit
- ...

Leeson, H. V. (2006). The mode effect: A literature review of human and technological issues in computerized testing. *International Journal of Testing*, 6, 1-24.

Waters, S., & Pommerich, M. (2007, April). *Context effects in internet testing: A literature review*. Vortrag auf der XXII. Konferenz der Society of Industrial and Organizational Psychology in New York (USA).



Technologiefaktoren

- **Bildschirmauflösung:** Beeinflusst sowohl Größe wie auch Umfang der dargestellten Informationen.
- **Bildschirmgröße:** Beeinflusst die Größe der dargestellten Informationen aber nicht den Umfang der Informationen.
- **Bildschirmtyp** (Röhrenbildschirm vs. LCD): Beeinflusst das Ausmaß der Bildschirm Spiegelung.
- **Prozessorleistung:** Beeinflusst die Genauigkeit von Zeitmessungen
- **Maustyp:** Beeinflusst die Genauigkeit von Reaktionszeitmessungen
- **Internetverbindung:** Erzeugt eine Verzögerung beim Laden neuer Informationen und kann zu Teilnahmeabbrüchen oder erhöhter Testangst bei Speed-Tests führen.



Interfacefaktoren

- **Schriftgröße:** Beeinflusst den Textumfang, der auf dem Bildschirm angezeigt wird.
- **Schriftart:** Beeinflusst Lesegeschwindigkeit.
- **Zeilenlänge:** Beeinflusst Lesegeschwindigkeit und Behaltensleistung.
- **Zeilenabstand:** Beeinflusst Textumfang pro Seite und Lesegeschwindigkeit.
- **Scrollen:** Beeinflusst das Ausmaß an Informationen, die auf dem Bildschirm dargestellt werden.
- **Abstände:** Beeinflusst Behaltensleistung



Umgang mit Kontexteffekten

- **Instruktion der Testnehmer** (Tippins et al., 2006)
 - Teilnehmer werden darüber informiert, welche Computerkonfigurationen für die Teilnahme notwendig ist
 - Problem: Nur wenig sind bereit bzw. in der Lage ihren Computer entsprechend zu adaptieren
- **Automatische Testkalibrierung** (Bennett, 2003)
 - Das System kalibriert die Darstellung in Abhängigkeit des jeweiligen Computersystems, um eine optimale Darstellung zu gewährleisten
 - Problem: In der Praxis häufig technisch sehr aufwändig.
- **Gestaltung für den kleinsten gemeinsamen Nenner:**
 - Der Fragebogen wird so einfach wie möglich gehalten.

Bennett, R. E. (2003). *Online assessment and the comparability of score meaning*. ETS.

Tippins, N., Beaty, J., Drasgow, F., Gibson, W., Pearlman, K., Segall, D., & Shepherd, W. (2006). Unproctored internet testing in employment settings. *Personnel Psychology, 59*, 189-225.



Personenfaktoren

- Computerängstlichkeit
- Mangelnde Computer- und Testerfahrung
- Sozioökonomischer Status, Alter
- Lesefähigkeit
- Tippgeschwindigkeit



4. Studie: Kontextfaktor Farbe



Kontexteffekt Farbe?

Farbliche Hinweisreize in Rot können sich in Leistungsbeeinträchtigungen manifestieren.

- **Barton & Hill (2005):** Sportmannschaften, die gegen Gegner in roten Trikots antreten verlieren überdurchschnittlich häufig.
- **Elliot et al. (2007):** Rot beeinträchtigt Leistungsmotivation „Furcht vor Misserfolg“ und führt zu schlechteren Leistungen
- **Ioan et al. (2007):** Schlechtere Leistungen in Stroop-Aufgaben bei roten Farbwörtern – allerdings nur bei Männern

Barton, R. A., & Hill, R. A. (2005). Sporting contests: Seeing red? Putting sportswear in context. *Nature*, 437, E10-E11.

Elliot, A. J., et al. (2007). Color and psychological functioning: The effect of red on performance attainment. *Journal of Experimental Psychology: General*, 136, 154-168.

Ioan, S. et al. (2007). Red is a distractor for men in competition. *Evolution and Human Behavior*, 28, 285-293.



Empirische Untersuchungen

Fragestellung:

Beeinträchtigen minimale farbliche Hinweisreize in Rot den Abruf deklarativen Wissens im Rahmen von online Befragungen?

Methode:

Unkontrolliertes Web-Experiment an anfallender Stichprobe

Experimentelle Manipulation

Studie I



Fortschrittsanzeige (ca. 100px):
grün vs. rot

Studie II

	Blau	Rot/Blau	Rot
Seite 1	Weiter	Weiter	Weiter
Seite 2	Weiter	Weiter	Weiter
Seite 3	Weiter	Weiter	Weiter

Button zum Weiterblättern:
blau vs. rot/blau vs. rot

Gesamter Fragebogen in schwarz/weiß



26%

Literatur

Wer schrieb den Roman "Anna Karenina"?

Wer schrieb den Roman "Der Zauberberg"?

Welcher Titelheld taucht bei Homer und James Joyce auf?

Wer schrieb "Utopia"?

Wessen Hauptwerk ist der "Don Quijote de la Mancha"?

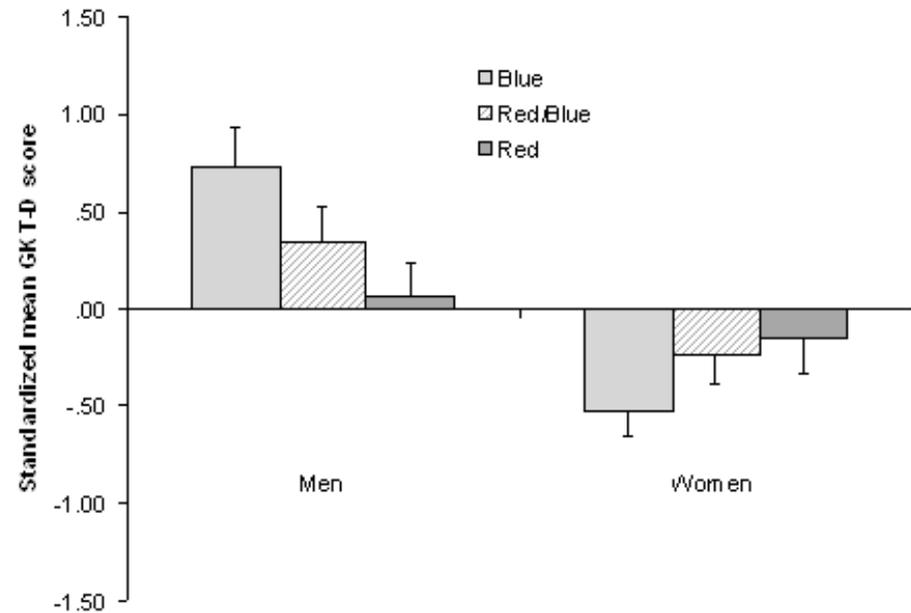
Weiter



Stichproben

	Studie I	Studie II
	Studierende versch. Universitäten	Studierende Universität Linz
<i>N</i>	126	183
<i>N</i> / Bedingung	~60 Personen	~60 Personen
Anteil weiblich	67%	51%
Alter	$M = 25 (SD = 7.3)$	$M = 24 (SD = 4.9)$
Bildung	80% Studierende	77% Studierende
Randomisierung	Keine Unterschiede bez. Alter, Geschlecht, Bildung	Keine Unterschiede bez. Alter und Geschlecht; etwas mehr Absolventen in Bedingung rot/blau

Ergebnisse



$$F(df=1;N=126) = 4.16 (p = .04, \eta_p^2 = .03)$$

$$F(df=2;N=183) = 5.24 (p = .01, \eta_p^2 = .05)$$

Gnambs, T., Appel, M. & Batinic, B. (submitted). *Color red in web-based knowledge testing.*



Zusammenfassung

- Kleine Farbvariationen können Leistungen beeinträchtigen
- Selbst in stark unkontrollierten Settings unkontrollierter Web-Tests
- Effekt tritt allerdings nur bei Männern auf (vgl. auch Ioannidis et al., 2007)



Herzlichen Dank für Ihre Aufmerksamkeit!

Timo Gnambs

Institut für Pädagogik und Psychologie

Universität Linz

timo.gnambs@jku.at

<http://timo.gnambs.at>