

Unaufmerksamkeit, Faking, Speedster...

Kontrolle der der Datenqualität in User Experience Befragungen

Prof. Dr. Meinald Thielsch
Prof. Dr. Gerrit Hirschfeld



*Workshop „Quantitative Messung von User
Experience“ auf der MuC 2021*

Gründe für fehlerhafte Daten

- Unaufmerksamkeit: Ungenügende Motivation/Konzentration um richtig mitzumachen. Zu schnelle Bearbeitung der Aufgaben (Speedster), oder wissentlich falsche/unzureichende Bearbeitung (Faking) der Aufgabe.
- Urteilsfehler: Selbst Menschen, die top motiviert sind, richtige Antworten zu geben, geben falsche Antworten (Halo-Effekte, Rückschaufehler etc.)
- Lügen: Falsche Antworten, um ein bestimmtes Ergebnis zu erzielen oder einen Eindruck zu erwecken.

Wie kontrolliert ihr die Qualität eurer Daten?

→ Schreibt uns kurz was in den Chat („gar nicht“ ist auch eine mögliche Antwort)

Praxisbeispiel

Ziel: Testung von 30 E-Health Websites mit einer in den Merkmalen Alter, Geschlecht und Bildungsgrad für Deutschland bevölkerungsrepräsentativen Stichprobe > 2000 Personen (= mehr als 50 TN je Website)

- ca. 15.000 Personen zur Studie eingeladen
 - 5.020 beginnen die Studie
 - 1.303 brechen die Studie im Verlauf ab
(529 davon bereits auf den ersten beiden Befragungsseiten)
 - ...79 widersprechen der Datenverwendung
- = 3.638 verwertbare Datensätze. Oder?

Von diesen 3.638 ...

- ... sind 19,6 % unrealistisch schnell, weitere 10,1 % auffällig schnell
- ... absolvieren 28,9 % offensichtlich nicht die Testaufgabe (Website öffnen)
- ... faken 25,9 % bei den Erinnerungsfragen zur Website
- ... beantworten 8,1 % das erste Kontrollitem zur Qualitätssicherung falsch
- ... beantworten 10,2 % das zweite Kontrollitem zur Qualitätssicherung falsch
- ... geben 0,7 % an, das Internet seit 50 Jahren oder länger zu nutzen
- ... geben 0,4 % an, mehr als 20 h pro Tag aktiv online zu sein

Studie zu E-Health: Datenbereinigung

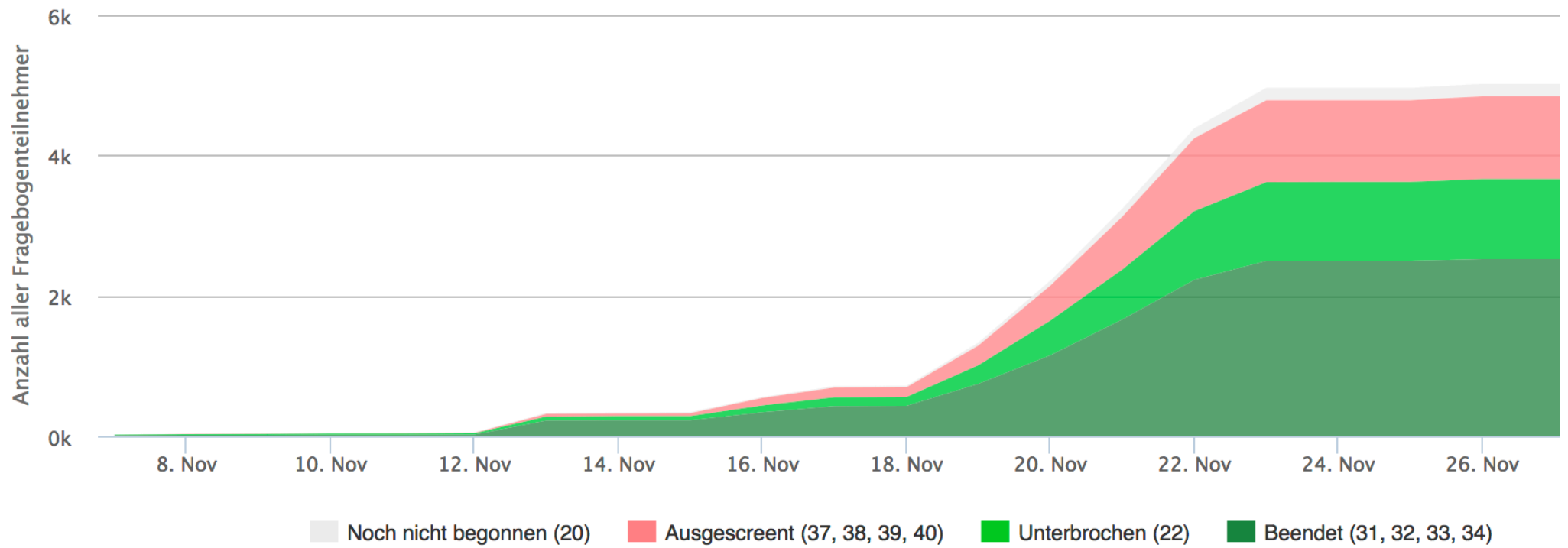
Die verschiedenen Indikatoren werden kombiniert berücksichtigt. Manche sind **K.O.-Kriterien** (z.B. unrealistisch schnelle Antwortzeit, Nichterfüllen der Testaufgabe), bei anderen erfolgt ein Ausschluss nur wenn **mehrere Indikatoren** auffällig sind (z.B. wurde das Falschbeantworten eines Kontrollitems toleriert, wenn das zweite Kontrollitem richtig beantwortet wurde).

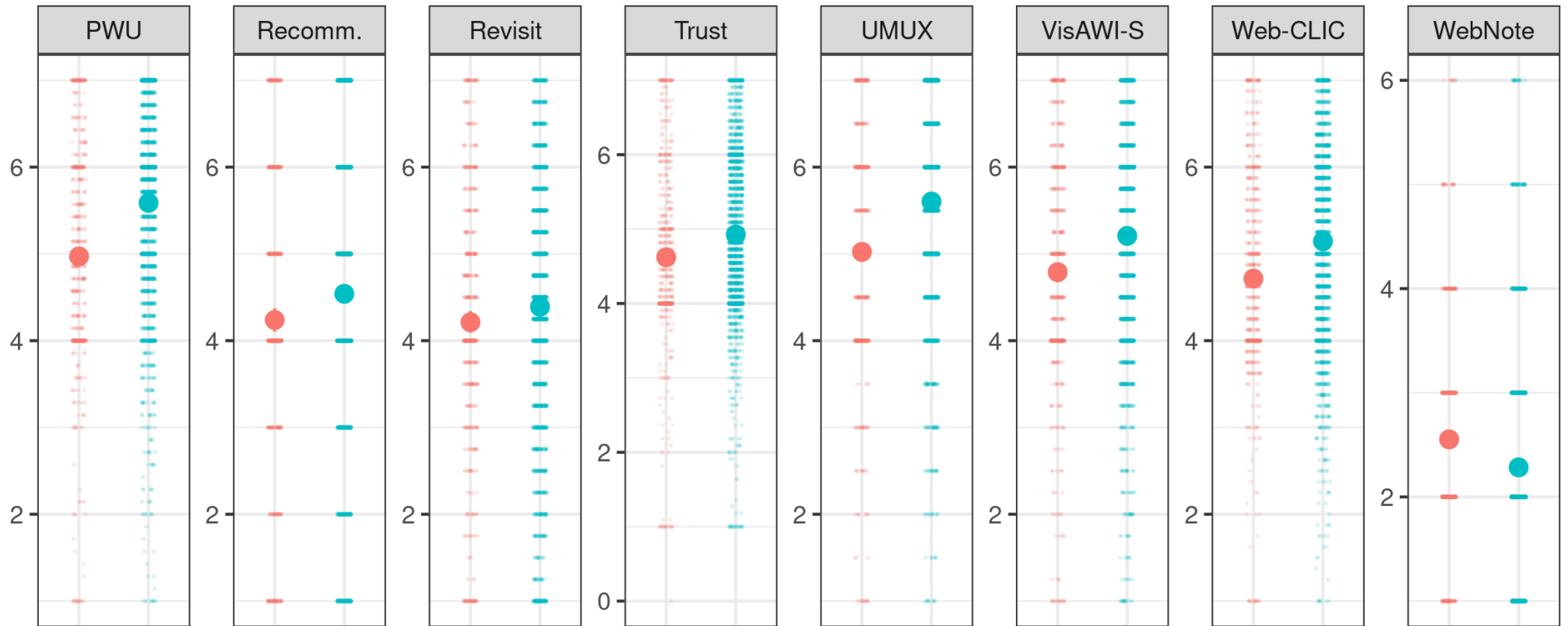
Ergebnis: Von den ursprünglichen $n = 3.638$...

→ 37,8 % Ausschluss ($n = 1.374$)

→ 26,9 % zumindest eine Auffälligkeit ($n = 981$)

→ 35,3 % unauffällig ($n = 1.283$)





status ● ausgeschlossene Fälle ● Fälle nach Qualitätscheck

Maßnahmen zur Sicherung der Datenqualität

Sicherung der Datenqualität

1. **Freiwilliger Selbstausschluss** am Ende der Umfrage (kann auch mit einem offenen Anmerkungsfeld kombiniert werden)
2. **Bearbeitungsdauer** → Identifikation von Speedster
 - a) auf Basis einer realistischen Schätzung, wie viel Zeit die valide Bearbeitung einer Umfrage mindestens in Anspruch nimmt.
 - b) Häufigkeitsdiagramm der Bearbeitungszeiten: Speedster manchmal direkt visuell in der Darstellung zu erkennen
 - c) Analyse der Antwortdauer pro einzelner Befragungsseite (Check, wer z.B. Instruktionen einer Testaufgabe nicht oder nur sehr kurz gelesen hat)
3. **Explizite Kontrollfragen**: Meade und Craig (2012) empfehlen ungefähr alle 50 bis 100 Fragen ein instruiertes Testitem (z. B. „Bitte antworten Sie bei dieser Frage zur Qualitätssicherung mit stimme zu.“) einzufügen.

4. Prüfung auf *unrealistische Angaben*:
 - a) auffällige Antwortmuster und eingeschränkte Varianz (z.B. eine Person klickt immer die gleiche Antwortoption unabhängig von Art und Richtung der Frage)
 - b) unmögliche Antworten (z. B. eine Internetnutzung seit mehr Jahren als dieses existiert oder eine tägliche Onlinenutzung von > 24 Stunden) geprüft werden.
 - c) Einträge in Freitextfeldern oder in ein Anmerkungsfeld zur Umfrage selbst (oftmals am Ende der Studie gegeben)
5. Gezielte *technische Kontrolle*: Werden in einer Online-Umfrage bestimmte Aufgaben gestellt, so kann beispielsweise mittels eines JavaScripts die Verwendung unerlaubter Hilfsmittel (siehe Diedenhofen & Musch, 2017) oder auch die Ausführung der Aufgabe (z. B. das Öffnen einer Testwebsite, vgl. Thielsch & Hirschfeld, 2021) kontrolliert werden.

Fazit

Wir müssen uns der Tatsache stellen, dass nicht jede Angabe in unseren Studien immer korrekt und sinnvoll auswertbar ist.

Es gibt sehr gute Methoden, um Probleme die sich in Unaufmerksamkeit begründen zu entdecken.

Echte Urteilsfehler und Lügen sind schwieriger zu erkennen, aber auch hierzu gibt es verschiedene Möglichkeiten die Datenqualität zu sichern und zu erhöhen.

Wer mehr wissen möchte...

...liest unseren Beitrag im Tagungsband – und dann noch:

P. G. Curran (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology*, 66, 4–19.
<https://doi.org/10.1016/j.jesp.2015.07.006>.

A. W. Meade, & S. B. Craig (2012). Identifying careless responses in survey data. *Psychological Methods*, 17(3), 437–455. <https://doi.org/10.1037/a0028085>

Vielen Dank!

Prof. Dr. Meinald T. Thielsch



www.meinald.de



@MeinaldThielsch

Prof. Dr. Gerrit Hirschfeld



www.gerrithirschfeld.de

Gefördert von der Bundeszentrale für
gesundheitliche Aufklärung im Auftrag
des Bundesministeriums für Gesundheit

The logo for the Bundeszentrale für gesundheitliche Aufklärung (BZgA) features the letters 'BZgA' in a bold, white, sans-serif font, set against a solid grey rectangular background.

**Bundeszentrale
für
gesundheitliche
Aufklärung**