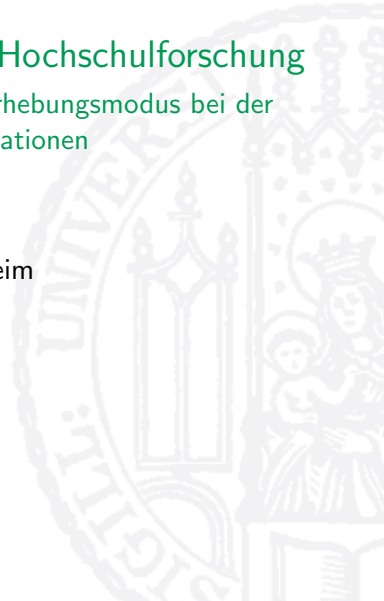


Experimente in der empirischen Hochschulforschung

Drei Teilstudien zu den Effekten des Erhebungsmodus bei der studentischen Lehrevaluationen

Tobias Wolbring
Universität Mannheim
Edgar Treischl
LMU München



Experimente in der Hochschulforschung

Causa finita?

- Experimentaldesign als Goldstandard zum Nachweis kausaler Effekte
- Experimente in der Hochschulforschung häufig abgelehnt
- Problem: Potentielle Verzerrung von Studien mit Vorher-Nachher-Vergleich
- Evidenzbasierte Entscheidung bedarf des Investment in *shoe leather* (Freedman 1991)
- Beispiel: Umstellung von papierbasierten auf onlinebasierte Evaluationsverfahren
- Forschungsfrage: Welchen Effekt hat der Erhebungsmodus auf die Lehrveranstaltungsevaluation?

- *Unterschiedliche Forschungsergebnisse*
 - ▶ Keine Reduktion (Simons & Pötschke 2006; Perrett 2013)
 - ▶ Reduktion des Rücklaufs (Handwerk et al. 2000; Tomsic et al 2000; Watt et al. 2002; Nulty 2008; Dommeyer et al. 2004; Meinefeld 2010)
 - ▶ Keine Verzerrung (Dommeyer et al 2004; Liegle & Mc Donald 2005)
 - ▶ Verzerrung der Evaluationsergebnisse (Meinefeld 2010; Kherif 2011)
- *Unterschiedliche Forschungsdesigns*
 - ▶ 11 von 15 nutzen quasi-experimentelle Forschungsdesigns: Messwiederholungsdesigns (Meinefeld 2010, Dresel & Tinsner 2008), Sekundärdatenanalyse (Simon et al 2013, Adams & Umbach 2012)
 - ▶ Experimentaldesigns in der Minderheit (Layne et al. 1999; Sax et al 2003; Dommeyer et al. 2004, Perrett 2013)
- *Unterschiedliche Online Erhebungsmodi: TAN, E-Mail*

Design I: Split-Half

Randomisierung der Studierenden in der Lehrveranstaltung

$R_{student}$	X_1	0	<i>Online Survey</i>
$R_{student}$	X_2	0	<i>Paper Survey</i>

- Randomisierung auf Individualebene
- *Übliche* papierbasierte Lehrevaluation
- Bearbeitungszeit ca. 10 Minuten während der Veranstaltung
- Online-Lehrevaluation mit TAN-Verfahren (Studierende mit mobilem Endgerät konnten sofort in der VL evaluieren)
- Split-Half in 11 Vorlesungen mit insgesamt 965 Fälle

$$\begin{array}{l|ll}
 R_{\text{twincourse}} & X_1 & 0 & \text{Online Survey} \\
 R_{\text{twincourse}} & X_2 & 0 & \text{Paper Survey}
 \end{array}$$

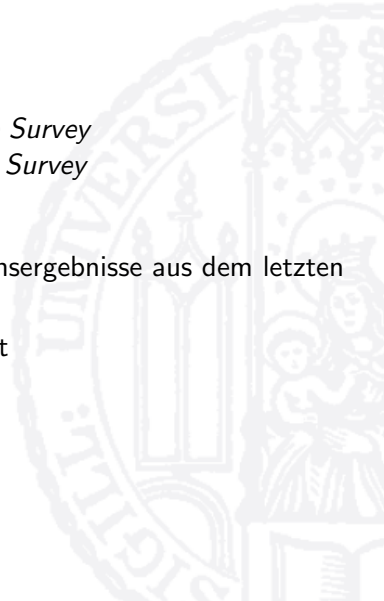
- Evaluation von Lehrveranstaltungen von Dozierenden mit dem gleichen Inhalt, aber an verschiedenen Zeitpunkten
- Randomisierung der Zwillingslehrveranstaltungen
- Keine Kontrolle der individuellen Aufteilung der Studierenden
- Kontaktaufnahme per E-Mail
- 44 Lehrveranstaltungen mit 485 Fälle

Design III: Vergleich über die Zeit

Randomisierung von aktuellen Lehrveranstaltungen

R_{course}	O_{t-1}	X_1	O_t	<i>Online Survey</i>
R_{course}	O_{t-1}	X_2	O_t	<i>Paper Survey</i>

- Als Referenzpunkt dienten Lehrevaluationsergebnisse aus dem letzten Sommer- bzw. Wintersemester
- Studierende wurde per E-Mail kontaktiert
- Randomisierung auf Kursebene
- 33 Lehrveranstaltungen mit 587 Fällen



Ergebnisse I: Split-Half

Gesamtbeurteilung

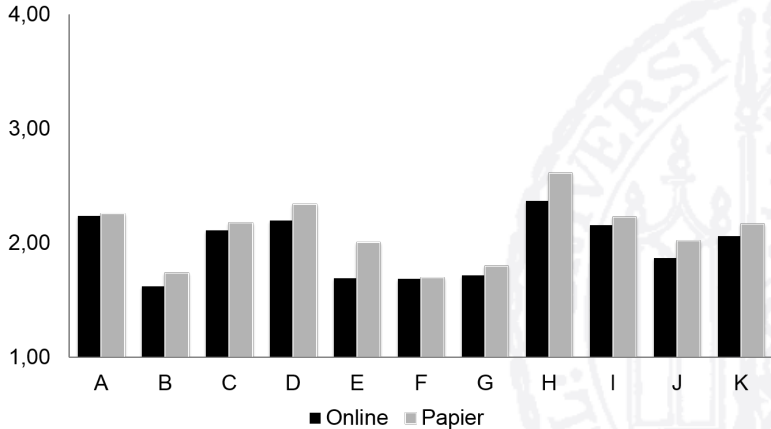


Abbildung: Durchschnittliche Veranstaltungsbewertung auf Veranstaltungsebene

Ergebnisse II: Zwillingslehrveranstaltungen

Gesamtbeurteilung

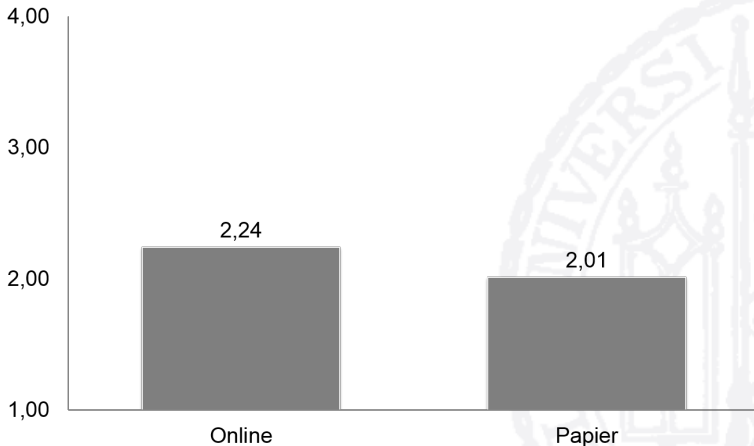


Abbildung: Gesamtbeurteilung der Zwillingslehrveranstaltungen

Ergebnisse II: Zwillingsslehrveranstaltungen

Gesamtbeurteilung

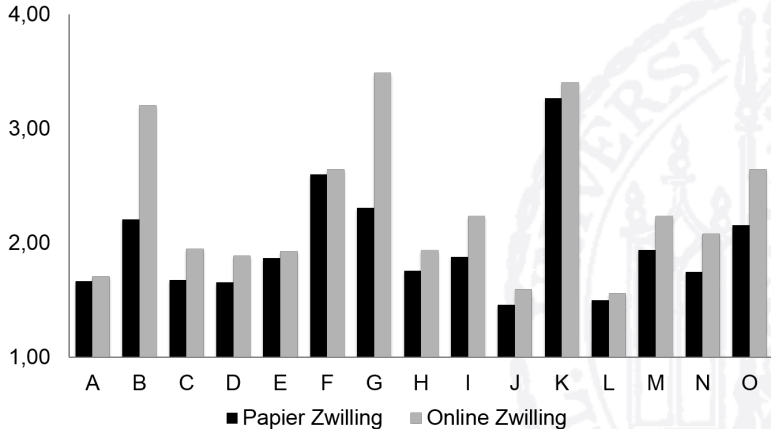
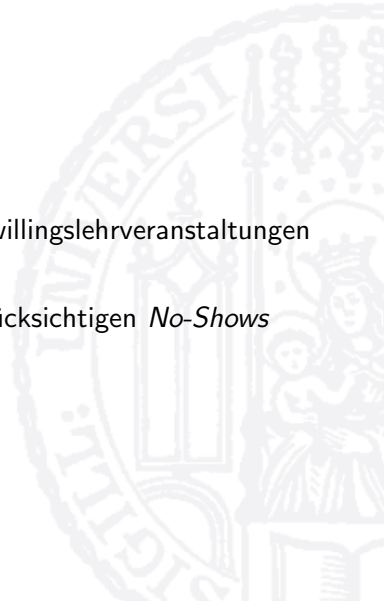


Abbildung: Unterschiede in den einzelnen Zwillingsslehrveranstaltungen

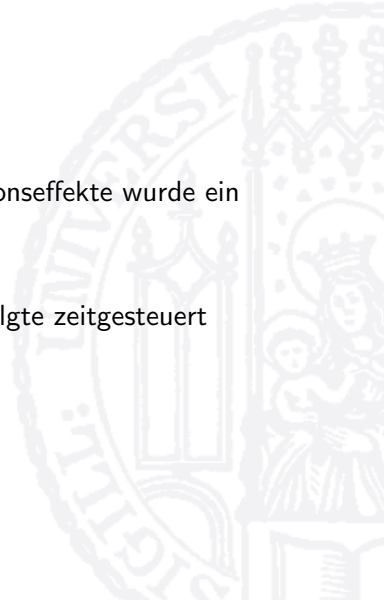
- Keine Unterschiede im Split-Half-Design
- Geringe, signifikante Unterschiede bei Zwillingslehrveranstaltungen
- Analoges Ergebnis im Design III
- Verdacht: E-Mail basierte Verfahren berücksichtigen *No-Shows*



Design III: Erweiterung des Fragebogens

Erweiterung beim Vergleich über die Zeit

- Zur Überprüfung möglicher Selbstselektionseffekte wurde ein zusätzliches Item aufgenommen
- Unterscheidung durch „Codewort“
- Versand der Einladungen per E-Mail erfolgte zeitgesteuert



Ergebnisse III: Codewort

Unterschiede in den Bewertung nach Anwesenheit der Teilnehmer

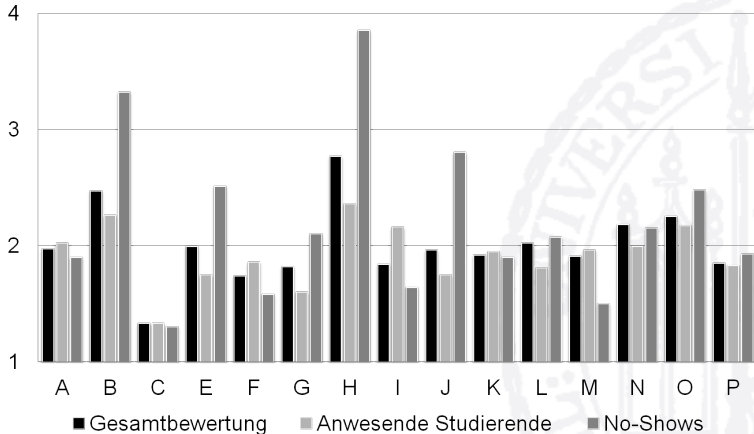


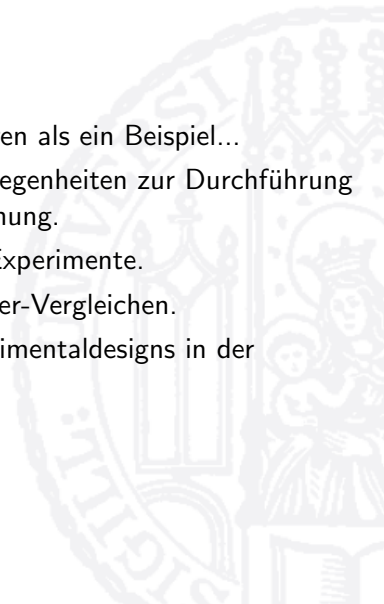
Abbildung: Bewertung der online durchgeführten Lehrveranstaltungen (Design III)

- 1 Nach Möglichkeit ein einheitlicher Erhebungsmodus
- 2 Inklusion der No-Shows
- 3 Filter nach Fehlhäufigkeit
- 4 Ausschluss der Drop-Outs
- 5 Gründe für das Fernbleiben wurden offen erfragt

Grund	in %
Lehrqualität	30,22
Zeit	26,87
Selbststudium	16,42
Bereits besucht	10,07
Job	6,34
Sonstiges	4,54

Einführung von Online Lehrevaluationsverfahren als ein Beispiel...

- 1 für das Vorhandensein von günstigen Gelegenheiten zur Durchführung von Experimenten in der Hochschulforschung.
- 2 für die Existenz ethisch unbedenklicher Experimente.
- 3 zur Problematisierung von Vorher-Nachher-Vergleichen.
- 4 als Illustration für unterschiedliche Experimentaldesigns in der Hochschulforschung.



Vielen Dank für die Aufmerksamkeit!

