



**Universität  
Zürich<sup>UZH</sup>**

Psychologisches Institut

---

# Studie 2

**Forschungsdesign**

Alain Stocker

2021

# INHALT

---

1	Studie 2.....	3
1.1	Methoden.....	3
1.1.1	Datengenerierung.....	3
1.1.2	Versuchsanordnung.....	4
2	Literaturverzeichnis .....	6

# 1 STUDIE 2

---

Besonders mit Hinblick auf die praktische Anwendung ist die gewählte Methode der Datengenerierung in Studie 1 suboptimal. Für die Datengenerierung wurden sämtliche Faktorladungen auf 1 fixiert, was zwar eine komfortable Datengenerierung in Monte-Carlo-Studien erlaubt, jedoch weit entfernt von der Realität des Forschungsalltags ist. Messinstrumente wie beispielsweise der PHQ-9 Fragebogen (Kroenke & Spitzer, 2002) zur Erfassung von Depressivität weisen für jedes Item unterschiedliche Faktorladungen auf (siehe Titov et al., 2011; Kocalevent et al., 2013). Eine zweite Simulationsstudie soll diesem Umstand Rechnung tragen und eine Brücke von der Theorie zur Praxis schlagen. In explorativer Manier soll die Auswirkung unterschiedlicher Faktorladungen auf die Schätzung eines Regressionskoeffizienten durch das SEM im Vergleich zum LM untersucht werden.

## 1.1 METHODEN

### 1.1.1 *Datengenerierung*

Die Datengenerierung orientierte sich am Schema in Studie 1, allerdings mit drei Unterschieden. Erstens wurden nun ausschliesslich 5 Indikatoren pro latenter Variable verwendet. Da diverse Screening-Verfahren zur Erfassung psychischer Gesundheit oder Störungen über 5 Items verfügen – beispielsweise der SCARED (Birmaher et al., 1999) oder der WHO-5 (World Health Organization, 1998) – scheinen 5 Indikatoren praxisnaher. Zweitens wurde die Stichprobengrösse auf 500 fixiert, da die Auswirkungen unterschiedlicher Stichprobengrössen auf diverse quantitative Grössen bereits in Studie 1 ausgiebig untersucht wurden. Drittens wurden die Mittelwerte der Reliabilitäten  $\omega_x$  und  $\omega_y$  nach McDonald (1999) für jede Versuchsbedingung auf den Wert  $\bar{\omega}_x = \bar{\omega}_y = 0.7$  fixiert, was den von Moosbrugger und Kelava (2012) empfohlenen Mindestwert des Reliabilitätskoeffizienten für einen guten Test darstellt.

Die Berechnung der Reliabilitäten kann Formel Studie 1 entnommen werden. Aus der Fixierung von  $\omega_x$  und  $\omega_y$  auf einen Mittelwert von 0.7 pro Versuchsbedingung erge-

ben sich bestimmte Werte für die Faktorladungen und Fehlervarianzen. Die Reliabilitäten  $\omega_x$  und  $\omega_y$  wurden über spezifische Ausprägungen der Faktorladungen  $\lambda_{x_m}$  und  $\lambda_{y_m}$  sowie der Fehlervarianzen  $\sigma_{\delta_m}^2$  und  $\sigma_{\varepsilon_m}^2$  mit  $m = \{1, \dots, M\}$  und  $M = 5$  operationalisiert. Im Gegensatz zu Studie 1 waren die Faktorladungen und Fehlervarianzen jedoch nicht auf einen festen, von der Versuchsbedingung vorgegebenen Wert fixiert, sondern wurden für jede Replikation neu aus Uniformverteilungen gezogen, deren Ober- und Untergrenze je nach Versuchsbedingung variierten.

### 1.1.2 Versuchsanordnung

Sowohl für die Faktorladungen als auch die Fehlervarianzen wurde die Range der jeweiligen Uniformfunktion – also deren Ober- und Untergrenze – systematisch über zwei Stufen variiert. Für sämtliche Faktorladungen einer latenten Variable und für alle Fehlervarianzen der dazugehörigen Indikatoren konnte die Range  $R$  je nach Versuchsbedingung entweder klein oder gross sein (siehe Tabelle 1). Innerhalb einer Versuchsbedingung galt somit  $R(\lambda_{x_m}) = R(\lambda_x)$  sowie  $R(\sigma_{\delta_m}^2) = R(\sigma_{\delta}^2)$  für alle manifesten Variablen  $x_m$ . Selbiges galt bei  $y_m$  für die Ausprägung der Ranges von  $\lambda_{y_m}$  sowie  $\sigma_{\varepsilon_m}^2$ , weshalb der Index  $m$  weggelassen werden kann. Während für Versuchsbedingungen mit kleiner Range die Faktorladungen auf 1 fixiert waren, konnten sie in Versuchsbedingungen mit grosser Range Werte zwischen 0.5 und 1.5 annehmen. Die Fehlervarianzen hingegen nahmen bei Versuchsbedingungen mit kleiner Range Werte zwischen 1.14 und 3.14 an und bei solchen mit grosser Range Werte von 0.14 bis 4.14.

**Tabelle 1**

*Ober- und Untergrenzen der Uniformverteilung für Faktorladungen und Fehlervarianzen*

Parameter	Klein [Minimum, Maximum]	Gross [Minimum, Maximum]
$R(\lambda_x)$	[1, 1]	[0.5, 1.5]
$R(\lambda_y)$	[1, 1]	[0.5, 1.5]
$R(\sigma_{\delta}^2)$	[1.14, 3.14]	[0.14, 4.14]
$R(\sigma_{\varepsilon}^2)$	[1.14, 3.14]	[0.14, 4.14]

Die Stufen für  $R(\lambda_x)$ ,  $R(\lambda_y)$ ,  $R(\sigma_{\delta}^2)$  und  $R(\sigma_{\varepsilon}^2)$  wurden miteinander gekreuzt. Für jede

mögliche Kombination dieser vier Faktoren wurden über 1'000 Replikationen Messdaten generiert. Im Anschluss wurden an jeden Datensatz einer Replikation ein SEM sowie ein LM angepasst. Somit handelt es sich bei Studie 2 um ein  $2 \times 2 \times 2 \times 2 \times 2$ -Design mit den beiden Methoden als Within-Faktoren und total 32 Versuchsbedingungen, welche in 32'000 Datensätzen resultierten.

## 2 LITERATURVERZEICHNIS

---

- Birmaher, B., Brent, D. A., Chiappetta, L., Bridge, J., Monga, S. & Baugher, M. (1999). Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED): A Replication Study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38(10), 1230–1236. <https://doi.org/10.1097/00004583-199910000-00011>
- Kocalevent, R.-D., Hinz, A. & Brähler, E. (2013). Standardization of the depression screener patient health questionnaire (PHQ-9) in the general population. *General Hospital Psychiatry*, 35(5), 551–555. <https://doi.org/10.1016/j.genhosppsych.2013.04.006>
- Kroenke, K. & Spitzer, R. L. (2002). The PHQ-9: A New Depression Diagnostic and Severity Measure. *Psychiatric Annals*, 32(9), 509–515. <https://doi.org/10.3928/0048-5713-20020901-06>
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum Associates.
- Moosbrugger, H. & Kelava, A. (2012). Qualitätsanforderungen an einen psychologischen Test (Testgütekriterien). In H. Moosbrugger & A. Kelava (Hg.), *Springer-Lehrbuch. Testtheorie und Fragebogenkonstruktion* (2. Aufl., S. 8–25). Springer.
- Titov, N., Dear, B. F., McMillan, D., Anderson, T., Zou, J. & Sunderland, M. (2011). Psychometric comparison of the PHQ-9 and BDI-II for measuring response during treatment of depression. *Cognitive Behaviour Therapy*, 40(2), 126–136. <https://doi.org/10.1080/16506073.2010.550059>
- World Health Organization. (1998). *Mastering Depression in Primary Care*. World Health Organization, Regional Office for Europe, Psychiatric Research Unit, Frederiksborg General Hospital.