

Technical Report – Technische Informationen zur Erhebung des bildungswissenschaftlichen Wissens im Forschungsprogramm BilWiss

Erhebung 1. Messzeitpunkt (t1)

Der finale Itempool zur Erfassung des bildungswissenschaftlichen Wissens enthielt 289 Items, daher wurde ich Multi-Matrix-Design verwendet, im Zuge dessen die Items auf 12 Testhefte verteilt wurden. Dabei beantwortete jede Person Items zu drei Inhaltsbereichen. Innerhalb dieses Rotationsdesigns (*pairwise-link rotation*) konnten vollständige Kovarianzmatrizen innerhalb jedes Wissensbereichs gebildet werden, die mehrdimensionale IRT-Analysen ermöglichten. Eine vollständige Kovarianzmatrix zwischen allen Wissensbereichen konnte allerdings nicht gebildet werden. Sechs eindimensionaler 2-PL-Partial-Credit-Modelle (Muraki, 1992) wurden für die folgenden sechs Wissensdimensionen berechnet:

- Unterrichtsgestaltung,
- Lernen und Entwicklung,
- Diagnostik und Evaluation,
- Bildungstheorie,
- Schule als Bildungsinstitution,
- Lehrerberuf als Profession.

Personenparameter wurden auf der Grundlage der Itemparameter geschätzt (EAP-Werte mit Standardfehlern), mithilfe derer die EAP-Reliabilität berechnet wurde (Rost, 2004). Die IRT-Analysen wurden mit der Statistiksoftware Mplus 6 berechnet (Muthén & Muthén, 1998–2013). Da der Wissenstest nicht veröffentlicht wurde und sich weiterhin im Optimierungsprozess befindet, sind im Datensatz lediglich die Personenparameter mit dazugehörigen Standardfehlern inkludiert. Veränderungen der Werte aufgrund von Skalenoptimierungen sind also nicht ausgeschlossen. Die Personenparameter sind auf einer Logit-Skala angeordnet. Der Nullpunkt (Skalenmittelpunkt) spiegelt das mittlere Ausmaß an Wissen der Stichprobe dieser Erhebung (t1) auf der entsprechenden Dimension wider.

Erhebung 2. Messzeitpunkt (t2)

Zu t2 wurde das bildungswissenschaftliche Wissen nicht erfasst.

Erhebung 3. Messzeitpunkt (t3)

Zu t3 wurde eine neu entwickelte Kurzversion des Tests zum bildungswissenschaftlichen Wissen eingesetzt, die insgesamt 57 Items aus allen sechs Testdimensionen enthielt. Die Personenparameter wurden aufgrund fixierter Itemkennwerte aus den zu t1 berechneten Modellen geschätzt.

Erhebung 4. Messzeitpunkt (t4)

Zu t4 wurde eine optimierte Kurzversion des Tests zum bildungswissenschaftlichen Wissen eingesetzt, die insgesamt 56, teilweise überarbeitete Items aus allen sechs Testdimensionen enthielt. Die Personenparameter wurden erneut aufgrund fixierter Itemkennwerte aus den zu t1 berechneten Modellen geschätzt, für neue bzw. veränderte Items wurden die Itemparameter frei geschätzt.

Literatur

- Muraki, E. (1992). A Generalized Partial Credit Model: Application of an EM Algorithm. *Applied Psychological Measurement*, 16(2), 159–176. doi: 10.1177/014662169201600206
- Muthén, L. K., & Muthén, B. (1998–2013). *Mplus (Version 6)*. Los Angeles, CA: Muthén & Muthén.
- Rost, J. (2004). *Testtheorie - Testkonstruktion* (2 ed.). Bern: Verlag Hans Huber.