

Zur Quantifizierung von Positionseffekten

Modellierungsstrategien im Vergleich

Marco J. Maier

Marco.Maier@edu.uni-klu.ac.at

Rainer W. A. Alexandrowicz

Rainer.Alexandrowicz@uni-klu.ac.at

Alpen-Adria-Universität Klagenfurt

Positionseffekte bei psychologischen Tests sind keineswegs ein neues Phänomen und werden schon seit den 50er-Jahren untersucht. Kurz gesagt: In Abhängigkeit der Position eines Items in einem Test muss nicht nur Itemschwierigkeit an sich (unbedingte Itemschwierigkeit), sondern auch der Einfluss der Position berücksichtigt werden.

Im Laufe der Zeit wurden verschiedene Modellierungsansätze, von einfachen 2-Gruppen Designs bis zu ausgefeilten IRT Methoden, verwendet. Ein klassischer Ansatz, um sowohl unbedingte als auch bedingte Itemschwierigkeiten zu schätzen, ist die Verwendung des LLTMs.

Ein Rasch-homogener, computergestützter Reasoning-Test wurde in 12 Versionen vorgegeben, die dieselben Items in unterschiedlicher Reihenfolge beinhalteten (Lateinisches Quadrat). Die Datenanalyse mit einem LLTM, zeigt einen unklaren Effekt der keinen ausgeprägten Positionseffekt erkennen lässt. Die Modellierung von Positionseffekten auf diese Art ist zweifelhaft wenn Itembearbeitungszeiten unlimitiert sind und daher stark variieren.

Aus diesem Grund haben wir einen anderen Modellierungsansatz gewählt, der auf der Reformulierung des LLTMs als generalisiertes lineares Mehrebenenmodell (random-intercept logistic regression) fußt. In diesem werden Itemantworten (Ebene 1) in Personen (Ebene 2) geschachtelt. Der Vorteil dieses Modells ist, dass es kontinuierliche

Prädiktoren auf Itemebene erlaubt und somit die Itemstartzeit – statt diskreten Positionseffekten – mit einer Art „Wachstumskurvenmodell“ (technisch ein Polynom) modellierbar ist. Dieses Modell weist eine deutlich bessere Passung auf und offenbart einen klaren U-förmigen Effekt der Startzeit auf die subjektiv wahrgenommene Itemschwierigkeit.

Beide vorgestellten Ansätze haben ihre Vor- und Nachteile: Die LLTM-Variante ermöglicht CML-Parameterschätzung, die einerseits verteilungsunabhängig hinsichtlich der Personenparameter ist und andererseits Conditional Likelihood-Ratio-Tests zulässt. Die Modellierung von Positionseffekten als „diskrete“ Parameter liefert jedoch keine überzeugenden Ergebnisse. Der Regressionsansatz andererseits bietet mehr Flexibilität bei den Prädiktoren auf Item- und Personenebene, wird jedoch per MML geschätzt wodurch man die oben genannten Vorteil der CML-Methode einbüßt.

In unserem Fall hat sich letztere Strategie als aufschlussreicher bewährt: Wir konnten Abweichungen der Itemschwierigkeiten durch die spezifische Itemstartzeiten feststellen, was einen psychologischen Test und seine Ergebnisse bei CAT-Vorgabe unzulässig machen würde. Ein Hauptbestandteil der Diskussion ist es, ob der methodologische Zweck (neue Erkenntnisse über Positionseffekte) in diesem Fall die Mittel (Aufgabe der CML-Schätzung) heiligt.