

Zur psychometrischen Qualität von neu entwickelten internetadministrierten Leistungstests

Wilhelm, O.¹, Witthöft, A., McKnight, P. & Größler, A.

Internetvermittelte Testungen der Leistungsfähigkeit von Personen unterscheiden sich von herkömmlichen Untersuchungen in erster Linie durch die Beschaffenheit der Testsituation (mit bzw. ohne durch eine Person vermittelte Instruktion, mit bzw. ohne Kontrolle des Verhaltens der Versuchsperson uvm.). Betrachtet man Tests als spezifische psychologische Experimente, sollte unter sonst gleichen Bedingungen der herkömmlichen Testung der Vorzug gegeben werden. Andererseits ist es legitim und sachgerecht, daß Kriterien wie etwa "erforderliche Ressourcen" "Verfügbarkeit von Experten" "Überbrückung geographischer Distanz" bei der Wahl der Untersuchungsmethode berücksichtigt werden. Aus angewandter Perspektive, ergibt sich häufig das Problem, unter sehr fähigen Personen die fähigste zu identifizieren, etwa im Rahmen der Personalauswahl. In herkömmlichen Untersuchungen stehen jedoch typischerweise nicht genug Experten freiwillig und kostenlos zur Verfügung. Daher wurden neue Wissenstests für Domänen entwickelt, in denen unserer Kenntnis nach keine geeigneten Verfahren verfügbar sind, solche Verfahren aber für praktische Belange durchaus wünschenswert wären: Naturwissenschaft und Betriebswirtschaft. Für beide Domänen wurden je zwei Paralleltests in zwei Sprachen entwickelt, ggf. angepaßt und via Internet bei selbstselektierten Stichproben eingesetzt.

Die Qualität der Daten wurde auf verschiedenen Ebenen beleuchtet und kritisch geprüft. Dazu zählen neben Prozeduren der klassischen Testtheorie auch probabilistische Verfahren der Prüfung von Skaleneigenschaften. Für beide Verfahren konnte eine Prüfung einer theoretischen Binnenstruktur vorgenommen werden. Beim betriebswirtschaftlichen Test konnten die empirisch gefundenen Ergebnisse mit Expertenratings der Schwierigkeit und Subskalenzugehörigkeit verglichen werden. Darüber hinaus war für den betriebswirtschaftlichen Test ein Vergleich mit Testleistungen von Experten möglich, die auf herkömmlichem Wege gewonnen wurden. Im naturwissenschaftlichen Test wurde die Prüfung der Binnenstruktur der Leistung einer experimentellen Manipulation unterzogen. Schließlich wurde die Äquivalenz beider Paralleltests beider Verfahren und die Übereinstimmung relevanter

¹ Oliver Wilhelm, Lehrstuhl Psychologie III, Universität Mannheim, D-68131 Mannheim, 0621 292 1482, wilhelm@tnt.psychologie.uni-mannheim.de, [http://www.psychologie.uni-mannheim.de/psycho2/forsch/wmc/tests/online_tests .htm](http://www.psychologie.uni-mannheim.de/psycho2/forsch/wmc/tests/online_tests.htm)

Statistiken in beiden Verfahren bestimmt. Zur Validierung der neuen Meßinstrumente wurden verschiedene biographische Fragen, wie etwa höchster erreichter Schulabschluß, relevante Vorbildung und Proxyvariablen für das Interesse an der jeweiligen Domäne herangezogen (etwa regelmäßig rezipierte Zeitungen, Zeitschriften und Fernsehsendungen). Die Ergebnisse der Analysen sprechen insgesamt gesehen für eine Qualität der Daten, die es gerechtfertigt erscheinen läßt, internetadministrierte Tests für die hier vorliegenden Zwecke einzusetzen. In Zukunft soll insbesondere die Sicherstellung der Verhaltensverankerung der Testverfahren als entscheidender Aspekt der Validität der Testverfahren beachtet werden. Daneben sollen neuartige Verfahren zur Bestimmung von Personenparametern erprobt werden, die bei gleicher Testlänge aussagekräftiger sind als übliche Richtig- / Falsch und Mehrfachwahl Aufgabenformate.

Stichwörter: Kalibrierung, Äquivalenz, Validität, Wissenstest