

ZMATH 2016d.00284

Böhme, Katrin; Richter, Dirk; Weirich, Sebastian; Haag, Nicole; Wendt, Heike; Bos, Wilfried; Pant, Hans Anand; Stanat, Petra

Do we measure the same? On the comparability of the IQB state comparison 2011 with the international studies IGLU and TIMSS. (Messen wir dasselbe? Zur Vergleichbarkeit des IQB-Ländervergleichs 2011 mit den internationalen Studien IGLU und TIMSS 2011.)

Unterrichtswissenschaft 42, No. 4, 342-365 (2014).

Zusammenfassung (DIPF): In regelmäßigen Abständen werden in Deutschland nationale und internationale Schulleistungsstudien durchgeführt, um den Leistungsstand von Schülerinnen und Schülern in ausgewählten Kompetenzbereichen zu überprüfen. Es wird der Frage nachgegangen, inwiefern die zur Messung eines Kompetenzbereichs eingesetzten Instrumente zwischen den Studien vergleichbar sind. In der vorliegenden Studie wird untersucht, inwieweit der nationale Institut zur Qualitätsentwicklung im Bildungswesen (IQB) Ländervergleich 2011 und die internationalen Studien Internationale Grundschul-Lese-Untersuchung (IGLU) und Trends in International Mathematics and Science Study (TIMSS) 2011 miteinander vergleichbar sind und welcher Zusammenhang zwischen den gemessenen Konstrukten besteht. Die Ergebnisse weisen darauf hin, dass die nationalen und internationalen Studien weitgehende Übereinstimmungen in der Zielpopulation und in den Durchführungsbedingungen aufweisen. (Kleinere) Unterschiede bestehen in der Rahmenkonzeption, den Testinstrumenten, den statistischen Analysemethoden und den aus den Ergebnissen ableitbaren Schlussfolgerungen. Empirisch zeigen sich für die mathematische Kompetenz hohe latente Zusammenhänge. Für die Lesekompetenz sind diese Zusammenhänge geringer, erhöhen sich jedoch nach Berücksichtigung der Testletstruktur. Konsequenzen für die Überführbarkeit von Testergebnissen von der nationalen auf die internationale Metrik werden diskutiert.

Summary: In Germany, both national and international large-scale assessment studies are conducted at regular intervals in order to determine students' achievement in selected domains. However, research investigating the comparability of instruments within given domains between national and international assessments is scarce. Therefore, we compared the construct definitions and test instruments of the German National Assessment Study (IQB-Ländervergleich) 2011 and the Progress in International Reading Literacy Study (PIRLS) and TIMSS 2011 assessments. The results indicate that the national and international studies are for the most part comparable in their target population, and testing procedures. There are, however, some differences in the construct definitions, in the design of testing materials, the applied IRT-models, and the inferences drawn from the results. Empirically, we found a high correlation between the national and international measures of student achievement in mathematics and a somewhat lower correlation for reading. The latter correlation increases when the testlet structure of the tasks is taken into account. Conclusions for converting test results from the national to the international scale are discussed.

Classification: D20 D60

Keywords: elementary school; construct equivalence; linking; reading; large-scale assessment

http://www.beltz.de/fachmedien/erziehungs_und_sozialwissenschaften/zeitschriften/unterrichtswissenschaft/article/Jour