

ZMATH 2014e.00390

Graf, Tanja; Emmrich, Rico; Harych, Peter; Brunner, Martin

Test administration effects in state-wide mandatory proficiency tests. (Durchführungseffekte bei Vergleichsarbeiten in Jahrgangsstufe 8.)

Empir. Pädagog. 27, No. 4, 459-473 (2013).

Zusammenfassung (DIPF): Die Ergebnisse lehreradministrierter Vergleichsarbeiten (VERA) sollten für Lehrkräfte keine weitreichenden persönlichen Konsequenzen haben (low-stakes testing). Dennoch sehen sich viele Lehrkräfte dem Erwartungsdruck seitens der Eltern, Kolleginnen und Kollegen, Schulleitungen oder Bildungsadministrationen ausgesetzt. Vor diesem Hintergrund sind vermehrt Diskussionen um die Aussagekraft und Vergleichbarkeit der Ergebnisse von VERA aufgekommen. In der in diesem Beitrag vorgestellten Studie wurde deshalb mittels eines experimentellen Untersuchungsdesigns auf Grundlage von vier repräsentativen Stichproben von Schülern ($n_1 = 1306$, $n_2 = 674$, $n_3 = 1173$, $n_4 = 341$) der 8. Klasse in Berlin und Brandenburg Fragen zur Durchführungs- und Auswertungsobjektivität bei VERA für das Fach Mathematik untersucht. Die Ergebnisse zeigen, dass Schülerleistungen im Mittel etwas besser ausfallen können, wenn die Tests von Lehrkräften und nicht von geschulten Testleitern sowie Kodierern durchgeführt bzw. ausgewertet werden. Das Ausmaß solcher Durchführungseffekte ist jedoch nicht einheitlich und variiert in Abhängigkeit von Schulart und Bundesland.

Summary (DIPF): Results derived from teacher administrated statewide assessments of students' achievement should not have any individual consequences for teachers (low-stakes testing). However, teachers are subjected to pressure as a result of expectations on the part of parents, colleagues, principals and school administration. Against this background a debate about the validity and comparability of results obtained from these tests has emerged. Therefore the study presented in this article focuses on effects of test administration during statewide annual assessment of students' achievement in mathematics. In order to examine these effects the authors conducted an experimental research design based on four representative samples of 8th grade students ($n_1 = 1306$, $n_2 = 674$, $n_3 = 1173$, $n_4 = 341$) in the German states Berlin and Brandenburg. Results show that students achievement is on average higher when teachers administrated and marked tests compared to students achievement on tests which are administered by trained test administrators resp. marked by trained raters. The extent of these administration effects is, however, not uniform and varies on school type and state.

Classification: D63

Keywords: empirical study; survey; educational reforms; comparative tests