

Research Brief

Effect of Prior Grade-Level Content on Pretest Reliability

by Sarah Callahan, Ph.D.
Assessment Technology Incorporated



Overview: Many districts and schools administer a comprehensive pretest at the beginning of the school year and a comprehensive posttest at the end of the year with the goal of assessing student growth over the entire year. Increasingly, districts and schools are also administering comprehensive pretests and posttests with the goal of providing data about student growth for use in the evaluation of instructional effectiveness. Given these goals, it is important to maximize the reliability of these assessments so that they provide the most precise estimates of student ability and growth. In some cases, the reliability of traditional pretests may be lower than is desirable since students may perform poorly on an assessment composed of items assessing a set of current grade level standards on which they have not yet received any instruction. In the 2011-12 school year, ATI implemented a new approach to the design of pretests that was intended to maximize reliability. Under this new approach, each pretest includes items that assess current grade-level standards as well as items that assess prior grade level standards on which students previously received instruction. This document presents the results of research conducted by Assessment Technology Incorporated (ATI) to investigate the effects of the inclusion of items that assess prior grade level standards on the reliability of the assessments.

Design: In the 2010-11 school year, all the items on each ATI pretest assessed current grade-level standards. In the 2011-12 school year, half of the items on each ATI pretest assessed current grade-level standards and half of the items assessed prior grade-level standards. For this research, a sample of 164 pretests was identified from six districts that administered corresponding sets of ATI pretests in the same grades and content areas in the 2010-11 school year (82 pretests) and in the 2011-12 school year (82 pretests). Collectively, the sample included pretests from a wide variety of grade levels and content areas including math (grades two through eight and high school algebra and geometry), reading/English language arts (grades one through 10), and science (grades four, six through eight, and high school biology). All pretests selected for the sample included scores from at least 150 students.

Analysis: ATI analyzed each pretest using procedures based in Item Response Theory (IRT). This analysis produced an ability estimate for each student, the Developmental Level (DL) score as well as an estimate of the reliability of the assessment. The IRT procedures employed by ATI place scores from each assessment within a grade level and content area on a common scale so that DL scores can be directly compared across assessments within a grade level and content area. These procedures also enable ATI to place scores from pretests and posttests on a common scale even when a subset of the items on the pretest assess prior grade level content.

Results: On average, pretests displayed adequate levels of reliability in both the 2010-11 school year ($M=0.84$; $SD=0.06$) and the 2011-12 school year ($M=0.87$; $SD=0.05$); however, higher levels of reliability were observed for pretests in the 2011-12 school year than in the 2010-11 school year ($t(156)=3.73$, $p=0.0003$), suggesting that the inclusion of prior grade-level content had the intended positive effect on reliability.

Conclusion: This research suggests that one way to maximize the reliability of pretests is to include items that assess prior grade level standards. An added benefit of this approach is that prior grade-level items can provide useful information about student mastery of prior grade-level standards that can be used to guide reteaching and enrichment. Based on the research described here, districts and schools that desire to use pretests and posttests to assess student growth from the beginning to the end of the year in the context of instructional effectiveness or other types of initiatives may wish to consider including prior grade-level content on pretests.