

Use of Multiple Measures

This study has confirmed the repeated finding of many investigations going back at least to the NLSMA reports prepared by the School Mathematics Study Group in the 1960s: Mathematics achievement is a multidimensional phenomenon. More recent research comparing student learning under traditional mathematics instruction to student learning under Standards-based mathematics instruction has found that the results can vary, depending on what type of test is used to measure achievement (Wood & Sellers, 1996; Huntley, et al., 2000). Tests emphasizing procedures and symbol manipulation tend to favor traditional instruction, whereas tests emphasizing problem solving in context tend to favor reform instruction. This means that studies reporting changes in “mathematics achievement” can be highly misleading, unless they provide a detailed description of precisely what was tested in their measure of “achievement”. It is possible to construct a test that will favor almost any program, so long as the test covers only the specific skills at which that program excels.

The current study used an assessment containing three very different measures of mathematics achievement and found that results differed depending on the measure. Further, considerable insight into what students did and did not learn was provided by a detailed examination of student responses to specific questions within each of the measures. Given the controversy that has surrounded reform mathematics programs like IMP, it is particularly important for researchers reporting results to explain just what they measured, and whenever possible to use multiple measures that will assess a range of student abilities.