

BUILDING EFFECTIVE INSTRUMENT DEVELOPMENT TEAMS: TWO CASE STUDIES

T. May¹ K. Koskey¹ J. Bostic² G. Matney² D. Teuscher³ A. Roth-McDuffie⁴ G. Stone⁵ S. Dingman⁶

¹Drexel ²Bowling Green ³Brigham Young ⁴Washington State ⁵U. Toledo ⁶U. Arkansas

Two case studies are presented to demonstrate the importance of intentionally utilizing subject matter expert and psychometrician expertise throughout the development process for cognitive and affective mathematical instruments. While the cases are unique, both empirically support the need for continued and thoughtful collaborative efforts during mathematical instrument development to produce meaningful measures.

Problem and Purpose Statement

It is a common stance that, when developing new mathematical instruments (cognitive or affective), teams with varying expertise should be convened for the purpose of producing high-quality measures (AERA et al., 2014; Kisker & Boller, 2014). However, research lacks illustrations of how subject matter experts (SMEs) and psychometricians (PSYs) constructing an instrument in tandem influences the quality of resultant measures. A goal of this work was to empirically demonstrate, through two different case studies, the need for purposefully integrating SME and PSY expertise on mathematical instrument development and validation teams.

Case 1: Developing Mathematical Problem-Solving Measures (PSMs) for Students in Grades 3-8

As part of a larger project, PSMs for students ages 8-14 years were developed and vertically equated onto one measure of problem-solving. When placed on a common ruler, fewer items at each grade level were required. Thus, an iterative item reduction process was undertaken first by a PSY who focused on Rasch psychometric indices and item domains, and second by a team of the same PSY and two SMEs who brought mathematics education knowledge to inform item reduction decision making. PSY findings produced strong reliability (Student=0.86, Item=0.99) with equal domain distribution. The combined SME/PSY reliability results slightly improved (Student=0.89, Item=0.99) and also reflected grade level domain relevance in U.S. national standards.

Case 2: Developing a Mathematical Curricular Reasoning Survey (MCRS) for Middle Grades Teachers

Three SME/PSY pairs collaboratively wrote two sections of MCRS items each. Before item writing, SMEs delivered professional development to PSYs on mathematical curricular reasoning. PSYs then presented a survey writing workshop to SMEs. Once all sets of items were co-developed by SME/PSY pairs, team members independently reviewed items and gave written feedback. Two researchers conducted a content analysis of the feedback and identified three themes related to the type of feedback team members provided: *General Item Wording* (47% SMEs, 53% PSYs), *Item - Mathematical CR Construct Alignment* (86% SMEs, 14% PSYs), and *Psychometric Item Writing Principles* (25% SMEs, 75% PSYs). Although both groups had training in the other's domain, experts more commonly offered feedback aligned with their own field.

Cross-Case Common Threads

Intentional SME/PSY collaborative processes thoughtfully established prior to beginning each project resulted in stronger instruments developed in both cases. Further, taking time to reciprocally teach and learn basics of the other's field helped to demystify unfamiliar aspects of the work and led to greater camaraderie across fields.

References

- Kisker, E. E., & Boller, K. (2014). *Forming a team to ensure high-quality measurement in education studies*. U.S. DoE, IES, NCEE, Analytic Technical Assistance and Development.
- AERA, APA, & NCME. (2014). *The standards for educational and psychological testing*. AERA.