

CONTENT VALIDITY EVIDENCE FOR NEW PROBLEM-SOLVING MEASURES (PSM3, PSM4, AND PSM5)

Jonathan Bostic
Bowling Green State University
bosticj@bgsu.edu

Gabriel Matney
Bowling Green State University
gmatney@bgsu.edu

Toni Sondergeld
Drexel University
tas365@drexel.edu

Gregory Stone
Metriks Amerique
gregorystone@metriks.com

Keywords: assessment and evaluation, elementary school, measurement, problem solving

Instrument development should adhere to the *Standards for Educational and Psychological Testing* (AERA et al., 2014). “Content-oriented evidence of validation is at the heart of the [validation] process” (AERA et al., 2014, p. 15) and is one of the five sources of validity evidence discussed. It usually involves exploring the connections between standards and test content. The research question for this study is: What is the evidence related to test content for the three instruments called the PSM3, PSM4, and PSM5? The study’s purpose is to describe content validity evidence related to new problem-solving measures currently under development. In previously published work, we have explored validity evidence related to test content for problem-solving measures (PSM6, PSM7, and PSM8) that address middle grades mathematics standards (see Bostic & Sondergeld, 2015; Bostic, Sondergeld, Folger, & Kruse, 2017). We chose a design-science based methodology to develop the PSM series. A design science-based methodology is useful for measure development, gathering data from the measure, drawing reasonable conclusions from the data, revising the measure, and repeating the cycle. Three forms of data were collected sequentially to explore test content validity evidence. Our data analysis approach used traditional methods (Sireci & Faulkner-Bond, 2014). Broadly speaking, all reviewers on the expert panel agreed the items were open, complex, and realistic. Mathematicians confirmed that each item could be solved in two or more ways. Students expressed that the items were complex, solvable, and realistic.

Results indicated that items were both representative and relevant of the construct. Supplementing this conclusion with the knowledge that the definition is sufficiently bounded and test construction followed the *Standards* (AERA et al., 2014), leads to a conclusion that the PSMs have adequate test content validity evidence. This is the initial step in building a validity argument; next steps are to gather evidence related to response processes, relations to other variables, internal structure, and consequences from testing.

Acknowledgments

This material is based upon work graciously supported by the National Science Foundation (DRK-12 1720646& 1720661). Any opinions, ideas, or findings expressed in this materials are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

- AERA, APA, & NCME. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Bostic, J., Sondergeld, T., Folger, T. & Kruse, L. (2017). PSM7 and PSM8: Validating two problem-solving measures. *Journal of Applied Measurement*, 18(2), 151-162.
- Bostic, J., & Sondergeld, T. (2015). Measuring sixth-grade students’ problem solving: Validating an instrument addressing the mathematics Common Core. *School Science and Mathematics Journal*, 115, 281-291.
- Sireci, S. & Faulkner-Bond, M. (2014). Validity evidence based on test content. *Psichothema*, 26(1), 100-107.