Abstract Title Page

Title: The Preschool Rating Instrument for Science and Mathematics (PRISM)

Author(s): Kimberly Brenneman, Judi Stevenson-Garcia, Kwanghee Jung, & Ellen Frede

Abstract Body

Background / Context:

Preschool science and math education are enjoying renewed attention among those concerned with pre-kindergarten education and with improving STEM literacy and achievement among our nation's citizens. Given the critical impact of early learning experiences on long-term educational and societal outcomes (Barnett, 2008; Bowman, Donovan, & Burns, 2001; Committee for Economic Development, 2006; Shonkoff & Phillips, 2000) and the links that have been found between general classroom quality and both school readiness outcomes (e.g., Burchinal, et al., 2008) and later mathematics achievement (Melhuish, et al., 2008 (EPPE); Haahr, Nielsen, Hansen, & Jakobsen, 2005 (PISA)), it is reasonable to hypothesize that provision of high-quality math and science experiences early in development will pay off with increased long-term achievement in these critical domains (NRC, 2005). Direct testing of these hypothesized connections is severely limited by a lack of instrumentation, however.

Until recently, few valid and reliable assessments were available to measure young children's mathematics and science learning in a *comprehensive* way. Now, a number of mathematics assessments have been developed and subjected to testing (Klein, Starkey, & Wakeley, 2000; Ginsburg, 2008; Clements & Sarama, 2008), and progress has been made in science as well (Greenfield, Dominguez, Fuccillo, Maier, & Greenberg, 2008; Greenfield, Dominguez, Greenberg, Fuccillo, & Maier, 2011). To study the relationship between classroom supports and instruction and learning outcomes requires valid and reliable measures of *classroom quality* along with these measures of children’s learning. Recent reviews of assessments for classroom quality in science and math have concluded that few comprehensive measures are yet available for use by researchers or educators and that *science is especially weak* (Brenneman et al., 2011; Snow & Van Hemel, 2009). Among the “under development” instruments identified in the Brenneman et al. review was the Preschool Rating Instrument for Science and Mathematics (PRISM), which is the focus of the present paper.

Purpose / Objective / Research Question / Focus of Study:

The purpose of our work is to develop and validate a classroom observation instrument that objectively measures the presence of classroom materials and teaching interactions that support mathematics and science learning. The overall question that motivates the effort is whether an observational tool can be developed to measure classrooms supports for math and science learning that are correlated with child learning outcomes and school readiness in these domains. Such a tool would make a significant contribution to the field, as it would have both research and professional development applications that would contribute to efforts to improve teaching and learning of key STEM concepts and practices in the critical preschool years.

Setting:

The PRISM is designed to be used in preschool classrooms, both half-day and full-day programs. Initial development, preliminary testing, and continuing psychometric work have taken place in classrooms serving mainly low-income students across multiple states and auspices.

Population / Participants / Subjects:

The PRISM has undergone, and continues to undergo, an iterative development process. Initial versions of the tool were piloted in early childhood classrooms serving under-resourced communities in New Jersey and New York. The current version of the instrument was used to study a total of 199 preschool classrooms in New Mexico, New Jersey, Rhode Island, and Kentucky, the majority of which were state-funded programs serving low-income families. Data from this sample were used to perform preliminary psychometric analyses reported here.

Intervention / Program / Practice:

The vision of "high quality" measured by the PRISM grows from our theoretical belief in the active learner who, nevertheless, will achieve and learn more when exposed to sensitive adult support and enriching materials. This includes educators who provide materials necessary for children to explore mathematics and science concepts and skills in their free play and in guided explorations. Further, it includes knowledgeable teaching staff who notice and extend children's spontaneous mathematical and scientific play and reasoning and provide planned opportunities for children to learn more about a wide range of mathematical and scientific content in meaningful ways. Our goal was the development of an objective, observational measure that would assess, in a comprehensive way, the materials and instructional supports for math and science learning present in preschool classrooms.

The PRISM's items come from analyses of the research literature and recommended instructional practices for early science and mathematics. Items and indicators are informed by the NAEYC/NCTM (2002) standards for early mathematics, the NCTM Focal Points (2006), and the recent National Research Council report on early mathematics education (Cross et al., 2009). Additionally, the items build on our earlier Preschool Classroom Mathematics Instrument (PCMI; Frede, et al., 2005). We broadened the content base to include science because science is an important school readiness domain and because math and science overlap conceptually. This overlap includes reasoning that supports classification, seriation, identifying patterns, problem-solving, prediction, measurement, and data collection and representation (see also Epstein, 2007; NRC, 2005). For science, experts in the field recognize the importance of providing children with the opportunity to explore a range of science concepts, to engage in extended explorations of these, and to incorporate literacy practices into these explorations (NRC, 2005; see also French, 2004; Gelman & Brenneman, 2004; Greenfield et al., 2009; Worth & Grollman, 2003).

Extensive literature review for content validity yielded an initial draft of the PRISM. We then moved to small pilots in the field, after which modifications were made in response to any confusing indicators, questions, and issues that arose. Indicators were also modified when observers, with strong background knowledge of early mathematics and science, indicated that a particular item was not capturing their qualitative belief about the classroom. This iterative process has yielded a 16-item structured observation tool that assesses the extent to which classroom materials and staff interactions foster a range of mathematical and scientific concepts and reasoning skills for young learners. These items are presented in Table 1 in Appendix B. The separation of the items into two broad areas (materials and staff interaction) reflects our desire to differentiate between classrooms with a lot of "stuff" but few supportive instructional interactions and vice versa.

Ratings for each item are made on a 7-point scale by trained observers. Behavioral descriptors or “anchor points” are present at the 1, 3, 5, and 7 levels. Scores between the anchors (i.e. 2, 4, and 6) are given when all indicators are met at a lower level and at least half but not all are met at the next higher level. Final ratings are made after at least 3 hours of observation in the morning hours of the day in a full-day classroom, with evidence being gathered continuously throughout the observation period. In a half-day classroom the entire session is observed.

Research Design:

The data we would report at the SREE conference were collected as part of larger field studies of the quality and effectiveness of, primarily, state-funded preschool programs. Data from the 199 classrooms provide an adequate sample to assess properties of the PRISM including factor structure and internal consistency. Observers also scored the *Early Childhood Environment Rating Scale - Revised* (ECERS-R; Harms, Clifford, & Cryer, 2005), allowing for an assessment of the concurrent validity of the PRISM with an established measure. Our long-term goal is a dedicated, full-scale study of the PRISM that would allow for final development, field testing, and validation of the instrument. Included in that work will be measures of predictive validity for vocabulary, math, and science outcomes.

Data Collection and Analysis:

Data are collected by trained observers, many of whom have backgrounds in child development and teaching, over a three-hour period. Prior to going into the field, potential data collectors are trained for two days using a training manual, photographs, video and written vignettes, and discussion with a PRISM trainer. After this period, assessors go into classrooms with a gold standard PRISM observer and are shadow scored and must reach 80% agreement with the trainer for three observations before being allowed to collect data independently. Every tenth observation, observers are shadow scored to ensure continued reliability.

Analysis of the PRISM data described here includes descriptive statistics for each item. Confirmatory factor analyses were fit using M-plus, and fit for resulting 3-factor and 4-factor solutions were measured using Chi-square, comparative fit index, and RMSEA. Internal consistency of scales was measured using Cronbach's alpha. Correlations among PRISM factors and ECERS-R factors, as well as overall scores were calculated.

Findings / Results:

Descriptive statistics for the 16 PRISM items are given in Table 2.

Confirmatory factor analyses found solutions that mapped onto theoretical predictions: materials, math interactions, and science interactions for the 3-factor solution, with materials split between math and science for the four-factor solution. Items loaded such that "overlap" items for measurement and categorization loaded with other math items. One item, Item 9: (Staff Interactions to Support) Numerical Operations did not load and is not included in subsequent analyses. Standard measures of fit were good for both models, as shown in Table 3. Correlations among factors were reliable in both the 3- and 4-factor solutions.

The internal consistency for the 3-factor solution was better than that of the 4-factor solution and was as follows: full scale = .78; materials subscale = .75; math interactions subscale = .72, and science interactions = .55.

Moderate correlations with the ECERS-R were found. The overall correlation was .41 ($p<.05$). Correlations between PRISM and ECERS-R factors are given in Table 4.

Conclusions:
The preliminary findings reported here encourage us to continue the development process for the PRISM. Correlations with the ECERS-R are moderate as expected, suggesting that the PRISM is picking up general quality, as well as unique information about observed classrooms. The factor structure of the instrument is strong and matches theoretical predictions, although work remains. Specifically, we will be working to better integrate supports for numerical operational thinking into the instrument, most likely by combining it with another number item. Efforts to improve the internal consistency of the scale, especially for science interactions, will also be made. We hope to undertake this work as part of a dedicated, full-scale study of the PRISM that allows for continued development, field testing, more extensive psychometric explorations, and final validation of the instrument.

Regardless of its psychometric properties, the PRISM has allowed us to collect a great deal of information about the preschool quality landscape with regard to math and science. It is a highly structured observation tool and data are collected by trained observers. The story these data have to tell is a rather bleak one. Of particular note are the low median scores for interactions, which confirms reports in the literature that very little math and science teaching occurs in preschool classrooms. For science, we found that the majority of classrooms show no interactions at all that support young children's thinking and knowledge-building.

Current recommendations and enthusiasm of education policy makers, governments, industry leaders, curriculum developers, and researchers for pre-K science (and other STEM domains) have not yet been translated so that *they positively impact classroom environments and children's learning*. Nor have we developed the psychometric toolkit that will enable us to confirm that current excitement about STEM is worth translating into practice. To do so, the field requires classroom quality assessments and direct measures of learning in these domains. Our team looks forward to input from the expert SREE audience on the PRISM. We also welcome a discussion of the larger issues surrounding the field's need for valid and reliable instrumentation around early STEM learning.

Appendices

Appendix A. References

Barnett, W. S. (2008). *Preschool education and its lasting effects: Research and policy implications.* Boulder and Tempe: Education and Public Interest Center & Education Policy Research Unit. Retrieved December 11, 2008 from http://epicpolicy.org/publication/preschool-education

Bowman, B.T., Donovan, M.S., & Burns, M.S. (2001). *Eager to learn: educating our preschoolers*. National Research Council Committee on Early Childhood Pedagogy, Commission on Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Brenneman, K., Boller, K., Atkins-Burnett, S., Stipek, D., Forry, N., Ertle, B., French, L., Ginsburg, H., Frede, E., & Schultz, T. (2011). Measuring the quality of early childhood math and science curricula and teaching. In M. Zaslow, I. Martinez-Beck, K. Tout, & T. Halle (Eds.), *Measuring quality in early childhood settings.* Baltimore, MD: Brookes Publishing.

Burchinal, M., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., et al. (2008). Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher-child interactions and instruction. *Applied Developmental Science, 12*, 140-153.

Clements, D.H., & Sarama, J. (2008). Experimental evaluation of the effects of a research-based preschool mathematics curriculum. *American Educational Research Journal, 45,* 443-494.

Cross, C., Woods, T.A., & Schweingruber, H. (Eds.) (2009). *Mathematics learning in early childhood: Paths toward excellence and equity*. National Research Council Committee on Early Childhood Mathematics, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Committee for Economic Development (2006). The economic promise of investing in high-quality preschool: Using early education to improve economic growth and the fiscal sustainability of states and the nation. Washington, DC: Committee for Economic Development.

Epstein, A.S. (2007). *The intentional teacher: Choosing the best strategies for young children's learning.* Washington DC: National Association for the Education of Young Children.

Frede, E., Weber, M., Hornbeck, A., Stevenson-Boyd, J., & Colón, A. (2005). *Prekindergarten Classroom Mathematic Inventory (PCMI)*. New Brunswick, NJ: National Institute for Early Education Research.

French, L. (2004). Science as the center of a coherent, integrated, early childhood curriculum. *Early Childhood Research Quarterly, 19*, 138-149.

Gelman, R., & Brenneman, K. (2004) Science learning pathways for young children. *Early Childhood Research Quarterly, 19*, 150-158.

Ginsburg, H. P. (2008). Early Mathematics Assessment System (EMAS). New York: Author.

Greenfield, D. B., Dominguez, M. X., Fuccillo, J. M., Maier, M. F., & Greenberg, A.C. (2008, June). *The development of an IRT-based direct assessment of preschool science*. Presented at the Ninth National Head Start Research Conference, Washington, DC.

Greenfield, D., Jirout, J., Dominguez., Greenberg, A., Maier, M., & Fuccillo, J (2009). Science in the preschool classroom: A programmatic research agenda to improve science readiness. *Early Education & Development*, *20*, 238-264.

Greenfield, Dominguez, Greenberg, Fuccillo, & Maier (2011). Greenfield, Daryl B., Dominguez, Ximena, Greenberg, Ariela, Fuccillo, Janna, & Maier, Michelle F. (2011, April). *Assessing science readiness in young low-income preschool children.* Paper presented at the biennial meeting of the Society for Research in Child Development, Montreal.

Haahr, J. H., Nielsen, T. K., Hansen, M. E., & Jakobsen, S. T. (2005). *Explaining student performance: Evidence from the international PISA, TIMMS, and PIRLS surveys.* Danish Technological Institute. Retrieved December 12, 2008 from http://www.pisa.oecd.org/dataoecd/5/45/35920726.pdf

Harms, T., Clifford, R., & Cryer, D. (2005). *Early Childhood Environment Rating Scale (ECERS-R)*, revised edition. New York, NY: Teacher College Press.

Klein, A., Starkey, P., & Wakeley, A. (2000). *Child Math Assessment: Preschool Battery (CMA)*. Berkeley: University of California, Berkeley.

Melhuish, E. C., Sylva, K., Sammons, P., Siraj-Blatchford, I., Taggart, B., Phan, M. B., et al. (2008). Preschool influences on mathematics achievement. *Science, 29*, 1162-1162. http://www.sciencemag.org/cgi/content/full/321/5893/1161

NAEYC/NCTM (2002). Early childhood mathematics: Promoting good beginnings. Washington, DC: National Association for the Education of Young Children.

National Council of Teachers of Mathematics (2006). *Curriculum focal points for prekindergarten through grade 8 mathematics: A quest for coherence.* Reston, VA: NCTM.

National Research Council (2005). *Mathematical and scientific development in early childhood*. Summary of a 2005 National Academy of Sciences workshop.

Shonkoff, J. P., & Phillips, D. A. (Eds.). (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academy Press.

Snow, C.E., & Van Hemel, S.B. (Eds.) (2009). *Early childhood assessment: Why, what, and how.*   Board on Children, Youth, and Families, Board on Testing and Assessment, Division of Behavioral and Social Sciences and Education.

Worth, K., & Grollman, S. (2003). Worms, shadows, and whirlpools: Science in the early childhood classroom. Heinemann.

Appendix B. Tables and Figures

Table 1: The 16 items on the Preschool Rating Instrument for Science and Mathematics (PRISM)

Materials
1. Materials for counting, comparing, estimating, and recognizing number symbols
2. Materials for measuring and comparing amount: Volume, weight, length, height, distance, time, and area
3. Materials for classifying and seriating
4. Materials for geometry and spatial positions/relationships
5. Materials for biological and non-biological science explorations
6. Materials to support reading about and representing science

Staff Interactions
7. Counting for a purpose
8. Identifying and writing numerals and numerical symbols
9. Numerical operations
10. Identifying and using geometric shapes
11. Spatial positions/relationships
12. Measuring and comparing amounts
13. Classification and seriation
14. Science explorations, experiments, and discussions
15. Observing and predicting
16. Recording science information

Table 2: PRISM Items: Descriptive Statistics

Item (See Table 1 for item content)	Mean (SD)	Median	Minimum	Maximum
1	4.90 (1.82)	5	1	7
2	4.53 (1.27)	4	2	7
3	4.31 (1.83)	5	1	7
4	4.31 (1.31)	4	1	7
5	4.07 (1.56)	4	1	7
6	4.04 (1.50)	4	1	7
7	4.08 (1.82)	4	1	7
8	1.75 (0.95)	1	1	7
9	2.13 (1.31)	1	1	7
10	1.82 (1.13)	1	1	7
11	2.49 (1.52)	2	1	7
12	1.97 (1.35)	1	1	7
13	1.99 (1.29)	1	1	6
14	1.83 (1.21)	1	1	7
15	1.83 (1.06)	2	1	7
16	1.54 (1.07)	1	1	7

Table 3: Measures of Fit for 3- and 4-factor models

Measures of Fit	3-Factor Model (15 Items)	4-Factor Model (15 Items)
Chi-Square Value (df)	96.41 (87)	89.51 (84)
Comparative Fit Index (CFI)	.98	.99
RMSEA	.02	.02

Table 4: Correlations Between PRISM and ECERS-R Factors

	PRISM: Materials	PRISM: Math Interactions	PRISM: Science Interactions	Overall PRISM Score
ECERS-R: Teaching and Interactions	.37	.32	.31	.44
ECERS-R: Provisions for Learning	.41	.13*	.26	.35
Overall ECERS-R Score	.48	.16	.27	.41

* $p<.10$, all other correlations, $p<.05$