Volume 13 Number 1

Assessment for Preschool Science Learning and Learning Environments

Kimberly Brenneman
Rutgers University
National Institute for Early Education Research

Abstract

Although interest in preschool science is not new in the United States, this area of learning is enjoying renewed attention among those concerned with prekindergarten education and with improving scientific literacy and achievement among the nation's citizens. Despite the increased interest and funding investment in early science education and the expectation that high-quality educational supports will result in improved school readiness and achievement in science and related domains, research and program evaluation efforts are limited by a lack of appropriate assessments for learning and classroom instructional quality in science. This article reports on a number of promising tools and approaches for evaluating children's learning progress in science and the quality of instructional supports for this learning. The article discusses learning and knowledge assessments, which include supporting and assessing science learning during everyday interactions; performance-based assessments for individualized instruction, progress monitoring, and curricular evaluation; direct assessments of science learning; and assessments of science-relevant skills and dispositions. The article also discusses classroom quality measures related to science learning.

Introduction

Although interest in preschool science is not new (see Riechard, 1973, for a review of programs to that date), this area of learning is enjoying renewed attention in the United States among those concerned with prekindergarten education and with improving scientific literacy and achievement among the nation's citizens. In the early education field, makers of widely used and respected comprehensive preschool curricula such as the Creative Curriculum (http://www.teachingstrategies.com/page/CCPS_Studies.cfm) are strengthening their offerings in science, and subject-specific programs have emerged (Brenneman, Stevenson-Boyd, & Frede, 2009). The National Association for the Education of Young Children (NAEYC, n.d.) holds that children should be provided various opportunities and materials to learn key content and principles of science. The Head Start Child Outcomes Framework includes science among its eight readiness domains (U.S. Department of Health and Human Services, 2003). Most U.S. states have articulated learning expectations for preschool science, either as a stand-alone area or as part of expectations for general cognition and language (Scott-Little, Lesko, Martella, & Milburn, 2007; Snow & Van Hemel, 2008). The current presidential administration has pledged to make this domain a priority at the prekindergarten level ("Barack Obama and Joe Biden's Plan," n.d.), and the National Science Foundation's Discovery Research K-12 program has begun to solicit and fund proposals to study science, technology, engineering, and mathematics (STEM) educational programs that support prekindergarten children and those who teach them. Technology and banking industry leaders also support efforts to make STEM accessible to America's preschoolers (e.g., American Honda Foundation, 2010; Motorola Foundation, n.d.). The PNC Foundation (2009) has made a major investment in partnerships between informal science education organizations and preschools through its Grow Up Great with Science! initiative. The Boeing Corporation provides funding support for PBS's *Sid the Science Kid* (http://pbskids.org/sid/), and Northrop Grumman supports *Peep and the Big Wide World* (http://www.peepandthebigwideworld.com), both of which are science-based programs for preschool audiences.

These curricula, policy statements, and funding commitments reflect beliefs that early exposure to STEM concepts will lead to increased comfort with them later in life and that early experiences are critical both for school readiness and as foundations for future learning (see also Beering, 2009). These ideas are attractive given well-established findings of the critical impact of early learning experiences on long-term educational and societal outcomes, especially among underserved populations (Barnett, 2008; Bowman, Donovan, & Burns, 2001; Committee for Economic Development, 2006; Shonkoff & Phillips, 2000), and research that establishes that measures of general classroom quality show moderate relationships to children's learning outcomes [e.g., Burchinal et al., 2008; Melhuish et al., 2008 (EPPE); Haahr, 2005 (PISA)]. Given these findings, it is reasonable to hypothesize that the provision of high-quality *science* learning experiences early in development will pay off with increased long-term achievement in, and student engagement with, science (National Research Council, 2005). Increasing the number of studies that can speak definitively to these issues is imperative, especially if, as a recent large-scale study in Florida suggests, school readiness in science lags behind other domains, at least among at-risk learners (Greenfield, Jirout, et al., 2009).

Despite the increased interest and investment in early science education and the expectation that high-quality educational supports will result in improved school readiness and achievement in science and related domains, evaluation and research efforts have been severely limited by a lack of appropriate instrumentation. The authors of a recent National Research Council (NRC) report on assessment in early childhood (Snow & Van Hemel, 2008) concluded that science assessments could not be included in their discussion because there "simply was not a basis in theory, research, or practice to include… science, despite [its] obvious importance" (p. 59). The early childhood field does not currently possess the tools needed to answer questions that bear directly on the methods by which to support and improve science teaching and learning. These tools and methods will reflect particular visions for early childhood science and education more generally. Thus, a starting point for anyone designing or using an assessment instrument is to clarify goals for children's learning with regard to thinking skills and content. Then one can ask, “Are the informal and formal preschool science education programs that we develop effective for meeting those goals? Are some more effective than others? What are the materials and instructional interactions that typify a science-minded classroom? How do we get the most “bang” for our limited educational buck? How do we ensure that each child has appropriate learning opportunities that build on and extend the excitement, enthusiasm, content knowledge, and reasoning skills that he or she brings to the scientific endeavor?”

This final question is, perhaps, of most interest to the adults who spend their days with young children. Developmental research tells us that long before they attend kindergarten, children possess content knowledge that roughly maps onto the scientific disciplines of physics, chemistry, psychology, and biology and that they have begun to reason in ways that form foundations for later scientific thinking (Duschl, Schweingruber, & Shouse, 2007). Young children also approach the world in ways that remind us of scientists. A powerful illustration of this comes from the following mother-child interaction paraphrased from Callanan and Oakes (1992, pp. 221-222):

> Child: Why does Daddy, James (big brother), and me have blue eyes and you have green eyes?
>
> Her mother tells her she got her eyes from Daddy, says goodnight, and leaves the room.
>
> The child calls her mother back 5 minutes later and says: I like Pee Wee Herman, and I have blue eyes. Daddy likes Pee Wee Herman, and he has blue eyes. James likes Pee Wee Herman, and he has blue eyes. If you liked Pee Wee Herman you could get blue eyes, too.
>
> The mother tells her daughter it would take more than liking Pee Wee to make her own eyes blue. Then she realizes the child doesn’t understand and explains that God gave her green eyes and they can’t be changed.
>
> Child: Could you *try* to like Pee Wee Herman so we could see if your eyes turn blue?

In this short interaction, the child engages multiple inquiry skills including *making and describing observations* about eye colors and TV preferences, *comparing* these, *questioning* the origins of eye color, *reflecting* on what her mother has told her to explain these differences and deciding (perhaps implicitly) that this explanation does not make sense, *generating her own explanation* for the source of the differences she has noted, and *designing a test* to find out whether her causal explanation is correct. While this example may seem extraordinary, those of us who work with children have many similar stories that reveal the capabilities of the young mind, and we feel the responsibility to support, celebrate, and challenge those capabilities. Early childhood as a field awaits strong research evidence that high-quality science learning experiences in preschool lead to long-term benefits for school achievement, scientific literacy, and professional achievements. Until that evidence is available, it remains incumbent upon us to provide children with a full range of enjoyable learning experiences that take advantage of their natural curiosity, desire to know, and deep interest in scientific topics.

Current State of Preschool Science Assessment

The recent National Research Council report *Early Childhood Assessment: Why, What, and How* (Snow & Van Hemel, 2008) defines assessment as “gathering information in order to make informed instructional decisions” (p. 27). Educators and policy makers who would like to make informed decisions for early science instruction are limited in their efforts because science is not among the domains that are well represented in the catalog of reliable and valid assessments available to educators and researchers (see also Brassard & Boehm, 2007). This seems to be true regardless of the purpose one might have for assessing science learning. That is, whether one is a classroom teacher who wishes to assess individual children's learning and skills to guide individualized instruction for her students, a researcher who speaks with a sample of children to assess the effectiveness of a curriculum or curricular program, or a researcher or administrator who observes a classroom to measure the quality of the environment for science learning, few comprehensive tools exist. In what follows, more detail is given on the state of science assessment and on the work of research teams making progress on these fronts.

The discussion begins by briefly addressing the everyday assessment that occurs in preschool classrooms when teachers observe and interact with children, then moves to descriptions of more structured, performance-based assessments used by educators to measure children's progress in scientific knowledge building (and other readiness domains). A discussion of program evaluation follows, with an emphasis on a new standardized measure that can be used in large-scale studies to assess the science readiness landscape for large groups of learners and to provide information about the strengths and weaknesses of particular programs. Finally, instrumentation to measure the quality of supports for science learning in preschool classrooms is reviewed. Given the links between overall classroom quality and children's readiness outcomes, it is assumed that high-quality classrooms for science learning will similarly be associated with positive learning outcomes in the domain. Of course, whether or not this assumption is correct is an empirical question that cannot be answered in the absence of psychometrically valid tools for assessing both learning outcomes and classroom quality.

Learning and Knowledge Assessments

Supporting and Assessing Science Learning during Everyday Interactions

The preschool teacher is charged, every day, with observing children and communicating with them in ways that support their functioning, learning, and thinking in cognitive, social, physical, and emotional areas of development. The adult observes and interacts with children to gain information, then responds with activities, discussions, materials, and questions that encourage children to explore and learn more about the world around them. Meeting this challenge requires that teachers understand child development and the expected sequences of learning across multiple domains. For science, specifically, this means that teachers need to be well versed in the kinds of foundational knowledge that preschoolers already have about science topics, the reasoning skills they possess, and the potential limits of those skills. It also means that teachers need some idea of how learning and development progress in order to support children’s movement along learning pathways or trajectories for science (Duschl et al., 2007; Gelman, Brenneman, Macdonald, & Román, 2009). In short, it requires not just knowing what to teach but how to teach it based both on general understandings of development and on the needs and interests of individual learners with regard to science. Unfortunately, many preschool educators report having concerns about their own knowledge of science and their ability to support children's learning in this domain (Greenfield, Jirout, et al., 2009). These concerns are not surprising given that early education teacher training programs do not emphasize science, either through classroom or practicum training (see Brenneman et al., 2009, for a review). As a result, the teacher who wants to support children's science learning often must spend extra time preparing to teach it by filling in his or her own knowledge gaps (Worth & Grollman, 2003).

The field of early childhood education could better serve young learners and those who teach them by providing more comprehensive and intensive preservice and inservice professional preparation programs in early science. Studies of teacher attitudes and beliefs about science generally, and about teaching it to young children specifically, will enable us to better meet this challenge in a focused manner, as will studies of preschool educators' knowledge of science and pedagogical content knowledge in this domain. Among the key features of an early science assessment system will be tools that allow greater insight into teachers' knowledge and thought processes so that we can respond with programs that better prepare them to assess and support science learning in the preschool classroom.

One such tool is the Preschool Teachers’ Attitudes and Behaviors towards Science (P-TABS), a newly validated measure of preschool educators’ attitudes and beliefs about science developed by researchers at the University of Miami. P-TABS can be used to gain a clearer picture of the ideas that teachers have about science and to assess the effects of professional development on these ideas (Maier,Greenfield, & Bulotsky-Shearer, 2011). The Education Development Center has developed and validated a measure of teacher pedagogical content knowledge as part of the Science Teaching and Environment Rating Scale (STERS; described further below). These Science Teacher Performance Tasks have been used to measure positive changes in teachers’ science knowledge as a result of participating in an intensive professional development program (Clark-Chiarelli, Gropen, Chalufour, Hoisington, Fuccillo, & Thieu, 2011).

Performance-based Assessments for Individualized Instruction, Progress Monitoring, and Curricular Evaluation

A particular kind of professional support could come in the form of child observation and assessment frameworks *and training of teachers to use them* in the classroom. Educational Testing Service's (ETS, n.d.) PATHWISE Understanding Early Science Learning provides early educators with an assessment framework and strategies to systematically collect and use children's behavior, language, and work products to guide instruction. The authors of PATHWISE suggest that "a first purpose of assessment in early science education is to help teachers observe, record, and reflect upon children's investigations of the natural world” (p. 1). In this view, assessment is less about identifying children's strengths and weaknesses than about supporting teachers as observers and interpreters of children's knowledge-building processes so that they can better support these processes (Chittenden & Jones, 1999). A similar approach to early science assessment is a key part of the constructivist classroom (Edmiaston, 2002). Under this theoretical orientation, assessment serves dual purposes, to document and interpret children's knowledge and reasoning while simultaneously evaluating how classroom activities and instruction encourage or hinder learning.

Under both approaches described (Chittenden & Jones, 1999; ETS, n.d., Edmiaston, 2002), the evaluation process involves identifying evidence of children's science learning during everyday classroom activities by collecting data over time from multiple sources. These sources include actions, talk, and artifacts that children create individually and in collaborative groups. Individual student portfolios composed of teachers' descriptions of ongoing behavior and conversations as well as children's work products (drawings, concept webs, science journals, sculptures, models, and so on) provide evidence used to assess children's understandings (see also Gelman et al., 2009; Worth & Grollman, 2003). This information is interpreted and applied to inform instruction and support new learning. As teachers practice these assessment procedures, they become more skilled as observers of children's scientific thinking and behavior and are in an increasingly better position to support preschoolers' learning and development in science and other related domains.

The focus on collecting and interpreting anecdotes and documentation of children's science learning can also feed into *comprehensive* progress-monitoring tools that span critical learning and development domains that include, but are not limited to, science. Assessments such as the Work Sampling System (Dichtelmiller, Jablon, Marsden, & Meisels, 2001), the Child Observation Record (HighScope Educational Research Foundation, 2003), and the Early Learning System (Riley-Ayers, Stevenson-Garcia, Frede, & Brenneman, in press) provide structures for tracking student progress in science learning, and other learning areas, using portfolios to inform teacher report. Teachers who use the Galileo System (Bergan, Burnham, Feld, & Bergan, 2009), in which they judge whether particular readiness skills are learned based on having observed a child demonstrating the skill or knowledge under three different circumstances, similarly would benefit from gathering evidence of children's science learning as they complete their ratings. Assessments of this type do introduce data collection burdens on teachers. However, such data collection is done with the goal of providing information about individual students as learners of science, math, language, literacy, socioemotional skills, motor skills, and so on, in order to help the teacher better tailor instruction to children who require further support, or challenge, in a particular area. Results from these assessments can be used to provide local information for teachers and schools to assess individual learning profiles at particular time points, to track growth over time, and, when aggregated, to assess whether curricular programmatic goals are being met.

Direct Assessments of Science Learning

Direct assessments of learning for purposes of program evaluation sometimes take advantage of established tasks from the developmental psychology and educational literatures or have been adapted from them. Van Egeren and colleagues (Van Egeren, Watson, & Morris, 2008) developed a child outcomes assessment battery to evaluate the Head Start on Science program. Measures included evidence evaluation, biology knowledge, hypothesis evaluation, and theory of mind tasks drawn from the developmental literature (Sodian, Zaitchik, & Carey, 1991, for evidence evaluation; Hatano & Inagaki, 1994, for biology; and Ruffman, Perner, Olson, & Doherty, 1993, for hypothesis evaluation and theory of mind tasks).

Measures used to assess effects of the Preschool Pathways to Science Program have included tasks similar to those used in developmental work, such as tests of children's understanding of the sources of their knowledge or their knowledge about setting up an informative experimental test (see Gelman et al., 2009).

Evaluation of the Marvelous Explorations through Science and Stories (MESS) program implemented in Head Start classrooms also used a combination of home-grown measures and those drawn from the developmental literature (such as theory of mind tasks) to evaluate program effectiveness in bringing about growth in children’s science skills and knowledge of conceptual content, such as animal life cycles and defense mechanisms (S. Ellis, personal communication, August 31, 2010). Language skills were also assessed using the Expressive and Receptive One Word Vocabulary Tests (EOWVT and ROWVT).

Assessments used to measure the benefits of the ScienceStart! Curriculum on children's language development include the well-established Peabody Picture Vocabulary Test III (PPVT; Dunn & Dunn, 1997), which has revealed benefits of the program (French, 2004).

In each of these cases, researchers (rather than classroom teachers) assessed learning as a way to evaluate the effectiveness of curricular programs and interventions for science. Until recently, however, the field has had no comprehensive assessment to directly test children's knowledge of science content and processes in a valid, reliable way. This gap in instrumentation has hindered efforts to research and evaluate preschool science programs and curricula.

A number of years ago, Daryl Greenfield and colleagues began development of such a tool. They began by reviewing state learning expectations for early science and those preschool curricula that included or focused on science, with the goal of creating a blueprint of content and process skills emphasized by the states and by those curricula. An initial item pool that reflected these content and process skills was created. Expert review and pilot testing were used to choose the final item pool and to further ensure the construct validity of the instrument. Results of testing in Head Start classrooms showed that the assessment was sensitive to a range of knowledge and skills, captured growth over the school year in science skills, and showed moderate, positive correlations with vocabulary and learning behaviors scores (Greenfield, Dominguez, et al., 2009).

The team's ongoing work involves the development and use of an 80-item version of the test to use to evaluate the impact of the Early Childhood Hands-On Science (ECHOS) professional development and curriculum program on children's learning in science and other domains. Additionally, the original flipbook version of the science assessment will serve as the basis for the development of a computerized version, *Lens on Science*. Extensive psychometric evaluation will be completed with the ultimate goal of delivering an assessment that can be used in research and program evaluation nationally (Greenfield, Dominguez, Greenberg, Fuccillo, & Maier, 2011). Such an assessment will allow states, school districts, or other educational entities to know generally where their students are with respect to science learning upon kindergarten entry, which, in turn, can inform educational decision making with regard to programmatic, curricular, or instructional changes to improve learning.

Assessments of Science-Relevant Skills and Dispositions

Other areas of child development certainly influence, and are influenced by, science learning. Thus, one might reasonably look to other important areas of child learning and development for evidence of related skills and knowledge. For example, social skills have an impact on scientific inquiry, because children engaging in such inquiry in school must learn to share and present evidence for their opinions during scientific discussion, to respect others' opinions during discussions, and to cooperate with peers and adults during group experiments or inquiry experiences. In fact, in their review of state learning standards and curricula for preschool, Greenfield and colleagues (Greenfield, Jirout, et al., 2009) identify cooperation as one of eight critical inquiry skills.

Similarly, an individual child's *approaches to learning*—including initiative, motivation, persistence, and curiosity—should influence the nature of spontaneous explorations. Identified as a critical domain of child learning and development by the National Education Goals Panel (1995), *approaches to learning* is among the domains of assessment described in detail in the recent NRC volume on assessment in early childhood (Snow & Van Hemel, 2008). While the reader is referred to that volume for a comprehensive discussion of this domain and assessments, one goal of this paper is to describe relevant developments that have not made their way into the larger literature. One such effort is being undertaken by Jamie Jirout and David Klahr (2010a, 2010b) who are developing and validating a measure of children's scientific curiosity.

Jirout has developed a game-like measure that manipulates uncertainty or ambiguity within an information-gathering situation as a way to assess individual learners' levels of curiosity. The present computerized version of the game Underwater Exploration! presents situations in which children can re-confirm known information (that is, at a level of no or low uncertainty), explore under conditions of moderate uncertainty (i.e., one of a few fish could appear behind a window), or explore under conditions of high uncertainty (i.e., any fish could appear). The game is adaptive in ways that provide detailed information about an individual child's preferred levels of uncertainty. That is, a child's choices allow the researcher to assess his or her comfort with situations in which correct answers are more or less certain. The behavioral assessment correlates positively with different scales of the Preschool Learning Behavior Scale, including competence motivation, attention/persistence, attitudes toward learning, and the total score of the scale (Jirout & Klahr, 2010a, 2010b). Jirout's motivation for development of a curiosity measure for preschoolers and kindergartners comes from the fact that "curiosity" is mentioned so often as a dispositional aspect of school readiness, yet the field has neither an accepted definition of curiosity nor a psychometrically validated measure of it. The instrument will allow for assessment of the extent to which educational programs support and increase children's curiosity, which should motivate increased exploratory behaviors by children and lead to greater learning (Jirout & Klahr, 2010b).

In sum, measuring individual children's science learning can take a variety of forms, and the choice of forms should be motivated by the purpose for which information is being gathered. Teachers observe, listen, and question in order to assess children's ideas and understandings in the moment, during everyday classroom activities. Performance-based assessment tools, informed by children's ongoing behaviors and work products under unstructured and semi-structured circumstances, provide formative assessment of children's learning and can be used by teachers to design new learning experiences to better support and challenge learners in science and other areas. In addition to progress monitoring for individual learners, information from assessments can be used to assess the degree to which a particular curricular program is related to growth in children's science learning. (Note that the validity of this information, or of comparisons among programs, is warranted only to the extent that teachers using the tools have been trained to adequate levels of reliability and are checked regularly to ensure fidelity to assessment procedures and, thus, the comparability of information across classrooms or programs.)

The University of Miami direct assessment in flipbook form and the forthcoming computerized *Lens on Science* version are standardized measures appropriate to assess the strengths and weaknesses in programs with regard to the extent to which they prepare young learners for kindergarten.

Classroom Quality Measures Related to Science Learning

If young children's science readiness is to improve so that it no longer shows the flattest growth curves and lowest overall achievement among the Head Start readiness domains (Greenfield, Jirout, et al., 2009), then assessments for learners are important. So, too, are assessments of the environments in which children grow and develop as science learners. To improve outcomes, educators and policy makers need to know what kinds of materials and classroom interactions are linked to better learning. Environmental quality measures can contribute to this endeavor in a variety of ways. A structured observation tool describes the features of a high-quality learning environment and can be used by educators and administrators to evaluate their programs relative to the benchmarks described by the tool or to other programs that have been assessed using the same tool. These evaluations can be used to identify areas in need of improvement and to guide professional development for educators.

Classroom quality measures can be used at multiple time points to monitor efforts to indicate ways that program and environmental quality might be improved. In these cases, the structured observations are completed by an external observer, not the classroom teacher. Another kind of classroom quality measure would involve a self-evaluation for teachers (and perhaps an evaluation by coaches or mentors) to inform and improve their instructional interactions with children (e.g., Frede, Stevenson-Garcia, & Brenneman, 2010). Finally, classroom quality measures could be used for program accountability purposes, if they were psychometrically validated and reliably administered (Snow & Van Hemel, 2008).

Mirroring the situation for child outcome assessment instruments, measures of classroom quality with regard to supports for science are not widely available. A working group that reviewed the available tools for assessing instructional supports for mathematics and science in preschool-grade 3 care settings concluded that the early childhood field's assessment tools are limited in both areas but that *science is particularly weak* (Brenneman et al., in press). This sentiment is echoed by other authors (Greenfield, Jirout, et al., 2009; Snow & Van Hemel, 2008), with the recent Snow and Van Hemel (2008) report concluding that most existing classroom environment observation measures assess the learning environment at a very general level, and only a few adequately assess practices related to cognition or academic skill domains such as science. The following sections outlines some measures that do exist, in varying states of development.

ECERS-R

An extension of the Early Childhood Environment Rating Scale–Revised (ECERS-R; Harms, Clifford, & Cryer, 2005), the Early Childhood Environment Rating Scale-Extension (ECERS-E) was developed in response to the overall lack of attention to literacy, mathematics, science, and diversity in the ECERS-R (Sylva, Siraj-Blatchford, & Taggart, 2003). The ECERS-E measures classroom science supports more extensively than any other published, widely available classroom observation instrument. Observers are required to evaluate two items that involve the presence of natural materials and the presence of classrooms area(s) dedicated to science and science resources. Observers also choose to score one item among the three remaining science activity/science processes items (nonliving, living processes and the world around us, and food preparation) after determining which kind of science learning experience is most apparent during the observation. This approach might represent a solution to the issue that observers spend a limited amount of time in a classroom and cannot be expected to observe the full range of science activities; however, important areas of science learning are either not represented in the instrument or remain unevaluated if another area is more apparent during the observation period. The psychometric properties of the ECERS-E include inter-rater reliability correlations above .88, weighted kappa coefficients that range from .83 to .97, and a high degree of concurrent validity with the ECERS-R (.78). The average total ECERS-E score shows significant, positive associations with children's scores for prereading, nonverbal reasoning, and early number concepts. The science scale alone did not show a significant relationship with child outcomes (see Halle & Vick, 2007; Sylva et al., 2003, for reviews).

STERS and PRISM

Two instruments that assess a more comprehensive range of science materials, concepts, and reasoning skills have been developed by teams from the Education Development Center (EDC) (Chalufour, Worth, & Clark-Chiarelli, 2006) and the National Institute for Early Education Research (NIEER) (Stevenson-Garcia, Brenneman, Frede, & Weber, 2010). EDC's measure, the Science Teaching and Environment Rating Scale (STERS), was created in response to the need to measure changes in the quality of classroom science instruction to evaluate the effectiveness of a professional development intervention. The STERS uses classroom observation and a teacher interview to rate the extent to which the teaching staff (1) creates a physical environment for inquiry and learning, (2) facilitates direct experiences to promote conceptual learning, (3) promotes use of scientific inquiry, (4) creates a collaborative climate that promotes exploration and understanding, (5) provides opportunities for extended conversations, (6) builds children's vocabulary, (7) plans in-depth investigations, and (8) assesses children's learning. Each of these components is rated using a 4-point rubric (1 = deficient through 4 = exemplary) that describes the sorts of materials and interactions one would find in a classroom that meets each numerical level. The authors report high internal consistency for the STERS (Cronbach's alpha = .96), and further investigation of the psychometric properties of the instrument are ongoing (Clark-Chiarelli, Gropen, Chalufour, Hoisington, Fuccillo, & Thieu, 2011).

NIEER's Preschool Rating Instrument for Science and Mathematics (PRISM) is a comprehensive, 16-item instrument designed to measure the presence of classroom materials and teaching interactions that support both mathematics and science learning. The science items focus on materials and teaching interactions that support explorations of biological and nonbiological science; encourage reading about, writing about, and representing science; encourage investigations and discussions of scientific concepts; support observing, predicting, comparing, and contrasting; and encourage recording of scientific information in journals, graphs, and other representational formats. In addition, items on measurement and classification cross the math and science domains. A full validation study and continued exploration of factor structure is planned for the PRISM. Preliminary analyses indicate acceptable internal consistency (Cronbach's alpha = .78) and moderate concurrent validity with the ECERS-R (R = .41) (Brenneman, Stevenson-Garcia, Frede, & Jung, 2011).

Directions for Further Research

As described in the introduction to this article, there is currently a great deal of enthusiasm for preschool STEM learning among education policy makers, U.S. federal and state governments, industry leaders, curriculum developers, and researchers. To capitalize on this interest and to translate it into clear educational policy and practice recommendations require strong research-based evidence about the instructional environments and interactions that provide positive learning experiences for young children.

Such evidence should come from tools that must be both based on research and empirically tested to ensure that they are valid, reliable, and linked to children's learning outcomes. Based on the inquiries that colleagues and I receive about the existence (or lack thereof) of such evaluation tools, it is clear that there is a very real demand in the field. Instruments that measure classroom supports for science learning in a comprehensive way will be of use as objective measures that can be used to compare classrooms, curricula, and programs using a common ruler, allowing us to evaluate these in a rigorous way and to answer the questions first posed in the introduction: Are the informal and formal preschool science education programs that we develop effective for meeting our goals for children's learning? Are some more effective than others? What are the materials and instructional interactions that typify a science-minded classroom? How do we get the most "bang" for our limited educational buck? How do we ensure that each child has appropriate learning opportunities that build on, and extend, the excitement, enthusiasm, content knowledge, and reasoning skills that he or she brings to the scientific endeavor?

Conclusion

Assessment in preschool justifiably concerns many people; they worry about the negative effects of certain kinds of assessments on young children. They fear that students experience feelings of inadequacy, confusion, pressure, or boredom if they are tested. The assessments described here include some that take advantage of the work products, conversation, and activities that are naturally part of children's experiences during the course of a typical preschool day. Other assessments may require that children take time out of their day, but these are often designed to be game-like and interesting for children. Assessment of young children also raises concerns if data from preschoolers, whose performances are more variable than those of older learners and who do not know the "importance" of performing well, are used to inform high stakes decisions about program and school effectiveness. As with assessment more generally, it is critical that the instruments be used for the purposes for which they were designed. Reviews of these issues can be found in recent review volumes (Brassard & Boehm, 2007; Snow & Van Hemel, 2008).

While this article in no way dismisses these important concerns, assessment is not optional for preschool science education. The question is not *if* we will assess science learning but *how* we can do so in ways that are appropriate for the questions being asked by teachers, administrators, researchers, and policy makers; that are viewed as useful by those who work in classrooms, administration, research, and policy; and that fit as seamlessly as possible into the lives of the learners being assessed.

The field currently lacks adequate instrumentation in early science, but progress is being made, both in the assessments available and in the ways that early childhood professionals can support young science learners. Much of this work, however, resides outside of the published literature; thus, one goal of this article is to start a conversation about the current state of early science assessment instruments, with the expectation that others will add to the inventory begun here. Together, those of us who study science learning and those who teach young science learners can generate a blueprint for the assessment toolkit that must be developed if we are to fully support the preschoolers of today as they learn and grow into the students, citizens, and STEM professionals of tomorrow.

Acknowledgments

This paper was originally prepared for the STEM in Early Education and Development (SEED) Conference held in May 2010 at the University of Northern Iowa, organized by Betty Zan. Many thanks to Betty and to Ingrid Chalufour, Nancy Clark-Chiarelli, Daryl Greenfield, Michele Maier, and Jamie Jirout for sharing details about their work. I thank Ellen Frede and an anonymous reviewer for their comments on earlier drafts. Their generosity and suggestions improved the paper; any errors are mine alone. [**Editors' note:** Other papers from the SEED Conference are available in *ECRP*'s Beyond This Issue section.]

References

American Honda Foundation. (2010). Retrieved April 3, 2010, from http://corporate.honda.com/america/philanthropy.aspx?id=ahf

Barack Obama and Joe Biden's plan for lifetime success through education. (n.d.). Retrieved March 22, 2011, from http://www.barackobama.com/pdf/issues/PreK-12EducationFactSheet.pdf

Barnett, W. Steven. (2008). *Preschool education and its lasting effects: Research and policy implications.* Boulder, CO, & Tempe, AZ: Education and Public Interest Center & Education Policy Research Unit. Retrieved December 11, 2008, from http://epicpolicy.org/publication/preschool-education
Editor's note: This url has changed:
http://nepc.colorado.edu/publication/preschool-education

Beering, Steven C. (2009). *Actions to improve science, technology, engineering, and mathematics (STEM) education for all American students* [Letter from the National Science Board STEM education outlining recommendations for the President-Elect Obama administration]. Retrieved January 1, 2009, from http://www.nsf.gov/nsb/publications/2009/01_10_stem_rec_obama.pdf

Bergan, John Richard; Burnham, Christine Guerrera; Feld, Jason K.; & Bergan, John Robert. (2009). *The Galileo Pre-K online system for the electronic management of learning.* Tucson, AZ: Assessment Technology. Retrieved March 22, 2011, from http://www.ati-online.com/galileopreschool/PreWelcomeTechManual.html

Bowman, Barbara T.; Donovan, M. Suzanne; & Burns, M. Susan (Eds.). (2001). *Eager to learn: Educating our preschoolers*. Washington, DC: National Academy Press.

Brassard, Marla R., & Boehm, Ann E. (2007). *Preschool assessment: Principles and practices.* New York: Guilford Press.

Brenneman, Kimberly; Boller, Kimberly; Atkins-Burnett, Sally; Stipek, Deborah; Forry, Nicki; Ertle, Barbina; et al. (in press). Measuring the quality of early childhood math and science curricula and teaching. In Martha Zaslow, Ivelisse Martinez-Beck, Kathryn Tout, & Tamara Halle (Eds.), *Quality measurement in early childhood settings.* Baltimore, MD: Paul H. Brookes.

Brenneman, Kimberly; Stevenson-Boyd, Judi; & Frede, Ellen C. (2009). Math and science in preschool: Policies and practice. *Preschool Policy Brief, 19*. New Brunswick, NJ: National Institute for Early Education Research.

Brenneman, Kimberly; Stevenson-Garcia, Judi; Frede, Ellen;& Jung, Kwanghee. (2011, April). *Assessing instructional quality for science and math: Preschool Rating Instrument for Science and Mathematics (PRISM)*. Paper presented at the biennial meeting of the Society for Research in Child Development, Montreal.

Burchinal, Margaret; Howes, Carollee; Pianta, Robert; Bryant, Donna; Early, Diane; Clifford, Richard; 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*(3), 140-153.

Callanan, Maureen A., & Oakes, Lisa M. (1992). Preschoolers' questions and parents' explanations: Causal thinking in everyday activity. *Cognitive Development, 7*(2), 213-233.

Chalufour, Ingrid; Worth, Karen; & Clark-Chiarelli, Nancy. (2006). Science teaching environment rating scale (STERS). Newton, MA: Education Development Center.

Chittenden, Edward, & Jones, Jacqueline. (1999). Science assessment in early childhood programs. In *Dialogue on early childhood science, mathematics, and technology education*. Washington, DC: Project 2061, American Association for the Advancement of Science.

Clark-Chiarelli, Nancy; Gropen, Jess; Chalufour, Ingrid; Hoisington, Cynthia; Fuccillo, Janna; & Thieu, Yen. (2011, April). *Assessing teacher and classroom outcomes form a professional development program in science for preschool teachers.* Paper presented at the biennial meeting of the Society for Research in Child Development, Montreal.

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: Author.

Dichtelmiller, Margo L.; Jablon, Judy R.; Marsden, Dorothea B.; & Meisels, Samuel J. (2001). *The work sampling system omnibus guidelines: Preschool through third grade* (4th ed.). New York: Rebus.

Dunn, Lloyd M., & Dunn, Leota M. (1997). *Peabody picture vocabulary test, third edition (PPVT–1II).* Circle Pines, MN: American Guidance Service.

Duschl, Richard A.; Schweingruber, Heidi A.; & Shouse, Andrew W. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academies Press; Board on Science Education, Center for Education, Division of Behavioral and Social Sciences and Education.

Edmiaston, Rebecca. (2002). Assessing and documenting learning in constructivist classrooms. In Rheta DeVries, Betty Zan, Carolyn Hildebrandt, Rebecca Edmiaston, & Christina Sales (Eds.), *Developing constructivist early childhood curriculum: Practical principles and activities* (pp. 53-67). New York: Teachers College Press.

Educational Testing Service. (n.d.). *PATHWISE understanding early science learning.* Princeton, NJ: Author.

Frede, Ellen; Stevenson-Garcia, J., & Brenneman, Kimberly. (2010). *Self-evaluation for science and math education (SESAME).* New Brunswick, NJ: Author.

French, Lucia. (2004). Science as the center of a coherent, integrated early childhood curriculum. *Early Childhood Research Quarterly, 19*(1), 138-149.

Gelman, Rochel; Brenneman, Kimberly; Macdonald, Gay; & Román, Moisés. (2009). *Preschool pathways to science (PrePS): Facilitating scientific ways of thinking, talking, doing, and understanding*. Baltimore, MD: Paul H. Brookes.

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.

Greenfield, Daryl B.; Dominguez, Ximena; Fuccillo, Janna M.; Maier, Michelle F.; Greenberg, Ariela C.; & Penfield, Randall. (2009, April). *Development of an IRT-based direct assessment of preschool science*. Paper presented at the biennial meeting of the Society for Research in Child Development, Denver, CO.

Greenfield, Daryl B.; Jirout, Jamie; Dominguez, Ximena; Greenberg, Ariela; Maier, Michelle; & Fuccillo, Janna. (2009). Science in the preschool classroom: A programmatic research agenda to improve science readiness. *Early Education and Development, 20*(2), 238-264.

Haahr, Jens Henrik (with Nielsen, Thomas Kibak; Hansen, Martin Eggert; & Jakobsen, Søren Teglegaard). (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

Halle, Tamara, & Vick, Jessica. (2007). *Quality in early childhood care and education settings: A compendium of measures.* Washington, DC: Office of Planning, Research, and Evaluation. Administration for Children and Families, U.S. Department of Health and Human Services.

Harms, Thelma; Clifford, Richard M.; & Cryer, Debby. (2005). *Early childhood environment rating scale* (Rev. ed.). New York: Teachers College Press.

Hatano, Giyoo, & Inagaki, Kayoko. (1994). Young children's naive theory of biology. *Cognition, 50,* 171-188.

HighScope Educational Research Foundation. (2003). *The preschool child observation record (COR).* Ypsilanti, MI: HighScope Press.

Jirout, Jamie, & Klahr, David. (2010a, June). *Assessing and validating measures of curiosity in preschool children.* Poster presented at the Fifth Annual Institute of Education Sciences Research Conference, National Harbor, MD.

Jirout, Jamie, & Klahr, David. (2010b). *The measurement of curiosity: A review of the literature.* Manuscript in preparation.

Maier, Michelle F.; Greenfield, Daryl B.; & Bulotsky-Shearer, Rebecca J. (2011, April). *Development and initial validation of a preschool teachers' attitudes and beliefs toward science questionnaire.* Paper presented at the biennial meeting of the Society for Research in Child Development, Montreal

Melhuish, Edward C.; Sylva, Kathy; Sammons, Pam; Siraj-Blatchford, Iram; Taggart, Brenda; Phan, Mai B.; et al. (2008). Preschool influences on mathematics achievement. *Science, 321(5893), 1161-1162. Retrieved March 24, 2011, from http://www.sciencemag.org/cgi/content/full/321/5893/1161*

Editor's note: This url has changed: http://www.sciencemag.org/content/321/5893/1161.full

Motorola Foundation. (n.d.). *Innovation generation grants*. Retrieved March 22, 2011, from http://responsibility.motorola.com/index.php/society/comminvest/education/igg/

National Association for the Education of Young Children. (n.d.). *Standard 2: NAEYC accreditation criteria for curriculum*.

Retrieved March 22, 2011, from https://oldweb.naeyc.org/academy/standards/standard2/

National Education Goals Panel. (1995). *Reconsidering children's early development and learning: Toward common views and vocabulary*. Washington, DC: Author.

National Research Council. (2005). *Mathematical and scientific development in early childhood: A workshop summary*. Washington, DC: National Academies Press.

PNC Foundation. (2009). *PNC Grow Up Great collaborates with 14 science centers to boost science skills of preschoolers*. Downloaded April 3, 2010, from http://www.pncgrowupgreat.com/media/downloads/PNC_GrowUpGreat_SciencePressRelease.pdf

Riechard, Donald E. (1973). A decade of preschool science: Promises, problems, and perspectives. *Science Education, 57*(4), 437-451.

Riley-Ayers, Shannon; Stevenson-Garcia, Judi; Frede, Ellen; & Brenneman, Kimberly. (in press). *Early learning scale*. Anaheim, CA: Lakeshore Learning.

Ruffman, Ted; Perner, Josef; Olson, David R.; & Doherty, Martin. (1993). Reflecting on scientific thinking: Children's understanding of the hypothesis-evidence relation. *Child Development, 64*(6), 1617-1636.

Scott-Little, Catherine; Lesko, Jim; Martella, Jana; & Milburn, Penny. (2007). Early learning standards: Results from a national survey to document trends in state-level policies and practices. *Early Childhood Research & Practice, 9*(1). Retrieved March 22, 2011, from http://ecrp.uiuc.edu/v9n1/little.html

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

Snow, Catherine E., & Van Hemel, Susan B. (Eds.). (2008). *Early childhood assessment: Why, what, and how.* Washington, DC: National Academies Press.

Sodian, Beate; Zaitchik, Deborah; & Carey, Susan. (1991). Young children's differentiation of hypothetical beliefs from evidence. *Child Development, 62*(4), 753-766.

Stevenson-Garcia, Judi; Brenneman, Kimberly; Frede, Ellen; & Weber, Marcie. (2010). *Preschool rating instrument for science and mathematics (PRISM)*. New Brunswick, NJ: National Institute for Early Education Research.

Sylva, Kathy; Siraj-Blatchford, Iram; & Taggart, Brenda. (2003). *Assessing quality in the early years: Early childhood environmental rating scale extension (ECERS-E).* Stoke-on-Trent, UK: Trentham Books.

U.S. Department of Health and Human Services, Administration on Children, Youth and Families/Head Start Bureau. (2003). *The Head Start path to positive child outcomes.* Washington, DC: Author. Retrieved December 12, 2008, from http://www.ecechicago.org/resources/pdfs/headstart01.pdf

Van Egeren, Laurie A.; Watson, Dyane P.; & Morris, Bradley J. (2008, June). *Head start on science: The impact of an early childhood science curriculum.* Paper presented at the Ninth National Head Start Research Conference, Washington, DC.

Worth, Karen, & Grollman, Sharon. (2003). *Worms, shadows, and whirlpools: Science in the early childhood classroom*. Portsmouth, NH: Heinemann.

Author Information

Kimberly Brenneman conducts research on early science and mathematics learning and supports for these in preschool classrooms. Her work at NIEER involves the development and validation of assessments of instructional quality and learning and the design of professional development resources to improve teaching in science and mathematics. Dr. Brenneman is an author of *Preschool Pathways to Science (PrePS): Facilitating Scientific Ways of Thinking, Talking, Doing, and Understanding* (Brookes Publishing). She also serves as an education advisor for *Sid the Science Kid*, a PBS television series and Web site that promote exploration, discovery, and science readiness among young children.

Kimberly Brenneman, Ph.D.
National Institute for Early Education Research
120 Albany St.
Suite 500
New Brunswick, NJ 08901
Telephone: 732-932-4350, ext. 239
Fax: 732-932-4360
Email: kbrenneman@nieer.org