



# Reframing the Assessment Discussion

**Jacqueline Jones**

In the years since the publication of “Framing the Assessment Discussion” (Jones 2004), there has been a significant focus on early childhood education. The field has experienced increases in state-funded preschool programs and in federal support for state systems-building initiatives such as the Race to the Top—Early Learning Challenge program. In addition, the National Academy of Sciences produced two significant consensus reports that speak directly to issues of early childhood assessment (NRC 2008) and to the competencies needed by the early childhood workforce to support young children’s learning and development (IOM & NRC 2015). Yet the early childhood field continues to struggle to reach consensus on appropriate types of assessment for young children and the optimal interpretations and uses of assessment results.

Achieving an understanding of young children’s learning is deeply rooted in teachers’ powers of observation. Up-close, ongoing observation and recording of what children say and do yield valuable information about their interests and emerging understandings. Teachers use this information to create rich learning environments and to implement effective instructional programs for all children (Jablon, Dombro, & Dichtelmiller 2007;

Rous & Hallam 2016). However, across the continuum of pre-K to 12, the national focus on professional accountability and quality programming has evolved into a call for more and more testing of younger and younger children.

It is reasonable to ask for evidence of how young children are developing and learning. It is also reasonable to ask if early childhood programs are providing the most appropriate and effective learning environments. However, as standardized test scores and percentile rankings compete with formative assessments and portfolios of children's work, the nature of the evidence used to answer questions about children's progress and program quality continues to be a matter of considerable debate.

As the accountability and testing debate continues, young children need advocates who are equipped with the knowledge and skills to participate in discourse that is grounded in the basic principles of sound assessment practice. Therefore, responsible early childhood educators need to reach beyond enhancing their skills in observation and documentation and move more deliberately toward developing *assessment literacy*—a deep understanding of the uses and limitations of the full range of assessment options and knowledge of the most appropriate methods to capture young children's learning and development (AERA, APA, & NCME 2014; IOM & NRC 2015; Stiggins 1991). Such methods include teachers' anecdotal notes, samples of children's drawings and constructions, and records of their conversations, as well as a variety of more formal instruments.

## Assessment and Testing

Accurate assessment of young children's learning is a complicated process. The rapid, episodic learning that characterizes early childhood is a significant assessment challenge. Young children may or may not fully engage in a structured assessment task; a 4-year-old may be much more interested in telling the teacher about his family's new pet than in following a set of standardized instructions. Furthermore, young children's understandings may look different from week to week. A child's experiences outside of the classroom, such as a fishing trip with a family member, can reshape her concepts of living and nonliving things and food sources.

Although the number of screening, diagnostic, and achievement instruments has increased over the years, most norm-referenced standardized measures provide a very limited view of early learning. The full picture of learning and development often requires assessors to supplement these measures with formative assessments and observational measures of classroom quality and teacher-child interactions. (For a discussion of assessment-related terms, please see "Common Assessment Terms and How to Use Them: A Glossary for Early Childhood Educators" on pages 15–20 in this volume.)

As the tension continues between providing high-quality, developmentally appropriate instruction and using instructional time to administer and interpret standardized tests that may be disconnected from the curriculum, a fundamental distinction needs to be made between testing and the process of assessment. *Assessment* may be defined as the ongoing process of gathering evidence of learning in order to make informed judgments about instructional practice (NRC 2008). This process occurs continually in almost every early childhood classroom as teachers listen to children's conversations, ask strategic questions to probe their understanding, observe their actions, and make informed judgments about the progress of an individual child or a group of children.

In contrast, a *test* has been defined as a procedure that systematically samples behavior in a specific domain and scores it in a standardized manner (AERA, APA, & NCME 2014). Tests can provide a quick look at specific behaviors at a particular point in time. However, they produce just one type of evidence that might be gathered in the overall assessment process.

Appropriate assessment is an integral part of the teaching and learning process. Sound assessment practices can

- › Highlight children’s knowledge, skills, and interests
- › Document children’s growth over time
- › Describe children’s progress toward specified learning goals
- › Provide constructive feedback to instructional programs and policy makers

When implemented effectively, the assessment process can be a powerful tool for teachers. By collecting a record of children’s growth over time, teachers can use assessment results to advocate for what children know and are able to do when these competencies are not apparent from the results of more standardized measures. In addition, assessment results can be the centerpiece for meaningful conversations between families and educators.

## Framing the Assessment Conversation

Assessment is more than a single data point. As the National Research Council’s report *Early Childhood Assessment: Why, What, and How* argues,

The selection and use of assessments, in early childhood as elsewhere, should be part of a larger system that specifies the infrastructure for distributing and delivering medical or educational services, maintaining quality, supporting professional development, distributing information, and guiding further planning and decision making. (NRC 2008, 28)

However, these complex discussions on accountability and testing can be framed around a few fundamental, critical assessment-related issues. The following questions do not represent an exhaustive set of the major issues in accountability and the testing of young children. Rather, they attempt to suggest some basic interrelated assessment concerns that teachers, administrators, and families who use assessment information should pose and be able to challenge as they participate in the accountability debate.

### What Is the Purpose of Assessment?

Twenty years after the publication of *Principles and Recommendations for Early Childhood Assessments*, a report to the National Goal Panel (Shepard, Kagan, & Wurtz 1998), it remains the foundation of our understanding of the major purposes of early childhood assessment:

- › Supporting learning
- › Identifying special needs
- › Evaluating programs and monitoring trends
- › Providing high-stakes accountability



Although the primary purpose of early childhood assessment is to improve instruction and thereby support children's learning, identifying special needs and monitoring program quality are also legitimate uses. The high-stakes accountability decisions that result from some assessment data, however, must be carefully considered in the light of appropriate assessment use (AERA, APA, & NCME 2014).

A teacher who needs to gather information about particular children's progress in developing science concepts, for example, might collect the children's drawings, take photographs of their constructions, and record or document their conversations. Classroom- and curriculum-based assessments can help teachers plan more appropriate instructional practices, because these assessments are directly tied to instruction and provide a close look at individual children.

When information about individual children is not needed, evidence about the quality and effectiveness of the educational program might come from assessing a sample of children rather than an entire population. An administrator, for example, may need to collect evidence about whether an instructional program has been effective and how all children, or subgroups of children, seem to be progressing toward a set of learning goals. In this case, it may be helpful to have a sample group of children perform a comparable task that reflects the desired learning goals. The assessment data from such sampling procedures are one part of the evidence that the administrator can use to evaluate program effectiveness and monitor group progress.

Clarifying the purpose of an assessment is a preliminary step in making decisions about the

- › Content and type of knowledge to be assessed
- › Population to be assessed (for example, a subset of children or an entire population)

- › Most appropriate assessment method and instrument
- › Target audience for the assessment information—teachers, parents, policy makers, the children themselves
- › Formats in which assessment results will be reported to the target audience

### **What Content and Type of Knowledge Is Being Assessed?**

The content and type of knowledge to be assessed is critical in determining the assessment method and instrument. For example, if a teacher wants to determine how Sanjay’s gross motor skills are developing, the most appropriate assessment method would be a performance assessment in which the teacher asks Sanjay to do specific gross motor tasks. Moving beyond simply identifying a curriculum domain to be assessed, such as reading or science, to aligning assessment methods with desired learning targets—what children need to know and be able to do in a particular domain—is recommended (Chappuis et al. 2012).

Understanding the type of learning target they want to assess will help educators choose the appropriate assessment method. Chappuis and colleagues (2012) identify the following types of learning targets:

- › **Knowledge:** Does Madison identify a group of objects with a corresponding number?
- › **Reasoning:** How does Elena go about recording her observations of the class gerbil?
- › **Skills:** Can Yuan write his first name using upper- and lowercase letters?
- › **Products:** Can Asia use what she knows about design to build a cardboard model of a house?
- › **Dispositions:** Does Reginald often seek out books on his own and pretend to read them?

### **What Is the Most Appropriate Assessment Method?**

Defining the assessment purpose and the content and type of knowledge to be assessed forms the groundwork for selecting the most suitable assessment method and instrument. Early childhood educators need to understand the range of appropriate assessment options, from classroom-based to norm-referenced measures, and the importance of using multiple methods of assessment.

For example, evidence of children’s ability to perform a task may or may not reveal their dispositions and self-regulation skills, but having this knowledge can help teachers choose effective ways of teaching and interacting with individual children. While a student might read fluently and with comprehension, this does not, by itself, reveal the child’s ability to persist when encountering an unfamiliar word or to focus attention on the text when there is a minor environmental distraction. The teacher might gather evidence of the child’s reading performance by engaging her in an actual reading task and using a checklist, but gauging the child’s persistence, attention, and focus may require further exploration through observation and more structured tasks.

To meet the differing assessment needs of educators, school officials, policy makers, and others, assessment must be part of a comprehensive system. The California Education Code (2014) defines a comprehensive assessment system as “a system of assessments of

pupils that has the primary purposes of assisting teachers, administrators, and pupils and their parents; improving teaching and learning; and promoting high-quality teaching and learning using a variety of assessment approaches and item types” (Section 60602.5[a]). A comprehensive assessment system should include screening measures, formative assessments, and measures of the quality of both the environment and adult–child interactions (ED 2018).

### **How Will the Assessment Results Be Evaluated?**

Samples of children’s work, teachers’ anecdotal notes, and performances on standardized measures are eventually evaluated against appropriate learning goals and standards, the performance of a similar group of children, or a scoring rubric. At some point an informed judgment or evaluation must be made that will modify an instructional program, generate further assessment, or provide feedback on program quality. Carefully gathered evidence is of little use unless it begins to answer questions about how young children are developing and learning and whether programs are providing the most appropriate, effective learning environments. Therefore, learning goals and standards must be appropriate, and any comparison groups must be as similar as possible to the child or group being assessed.

*Mrs. Jacobs knows that by the end of third grade, her current kindergarteners are expected to have mastered this state science standard: Keep records that describe observations, carefully distinguish actual observations from ideas and speculation, and are understandable weeks and months later. To support her students’ development of these skills, she designs a unit on change and decay. As part of the unit, the children conduct regular, ongoing observations of two pumpkins, one of which has been split into two pieces and one of which remains intact. The children engage in group discussions about their observations of the changes occurring in the sliced pumpkin and on the outside of the intact pumpkin. They keep science journals in which they draw and write their observations and impressions. All classroom-based and formal assessment information is weighed against the children’s progress toward the science standard.*

### **Are Assessments Results Reported Clearly and Accurately?**

Understanding assessment results can be daunting. The most well-constructed, appropriate assessment is useless if the intended audience—teachers, families, center directors, principals, and/or policy makers—cannot understand the results. Young children and educators are not well served when instructional and policy decisions are made on the basis of assessment results that the intended audience does not understand.

According to Standard 6.10 in *Standards for Educational and Psychological Testing*, “When test score information is released, those responsible for testing programs should provide interpretations appropriate to the audience. The interpretations should describe in simple language what the test covers, what the scores represent, and how the scores are intended to be used” (AERA, APA, & NCME 2014, 119). In addition, Standard 6.11 states, “When automatically generated interpretations of test response protocols or test performance are reported, the sources, rationale, and empirical basis for these interpretations should be available, and their limitations should be described” (119).

When Mr. Scott meets with the families of the preschoolers in his classroom to talk about the children’s development, he shows them samples of their children’s drawings, their constructions, and transcripts of conversations he has collected over time. He discusses the children’s work in the context of trajectories for typical development while acknowledging each child’s specific patterns of learning. The children’s work samples frame the conversations with families about how their children’s learning is progressing and what the next stages of development might look like.

### **How Are the Assessment Results to Be Used?**

Perhaps the most important element of any assessment is the ultimate use of the assessment information. Use of the data to make decisions—such as extending a lesson for a few additional days, identifying a child as needing special services, or providing additional resources to a program—should be linked to the stated purpose of the assessment process, aligned to intended use of the assessment method or instrument, and based on a thorough understanding of the assessment results. Applying assessment results should cause no harm.

### **Conclusion**

No assessment, by itself, can improve the quality of instruction or enhance children’s outcomes. Rather, assessment data can and should serve as a catalyst for continuous quality improvement. The assessment issues described in this article are basic to a reasoned discourse on accountability and testing of young children. If the conversation is based on principles of sound measurement practice, the fields of early childhood education and educational measurement will be challenged in new ways to act as responsible advocates for children.

The Institute of Medicine and National Research Council report (IOM & NRC 2015) outlines a set of competencies that are important for all early childhood educators who work with children from birth through age 8. These include the following understandings related to assessment:

- **Core knowledge base:** Knowledge of principles for assessing children that are developmentally appropriate; culturally sensitive; and relevant, reliable, and valid across a variety of populations, domains, and assessment purposes
- **Practices to help children learn:** Ability to select, employ, and interpret a portfolio of both informal and formal assessment tools and strategies; to use the results to understand individual children’s developmental progression and determine whether needs are being met; and to use this information to individualize, adapt, and improve instructional practices (328–29)

Those in leadership and administrative roles also need the following competencies related to assessing children (IOM & NRC 2015):

- Knowledge of assessment principles and methods to monitor children’s progress and ability to adjust practice accordingly
- Ability to select assessment tools for use by the professionals in their setting (344)

## Culturally and Linguistically Appropriate Assessment

The assessment of young children is, by its very nature, a culturally and linguistically loaded interaction. This must be acknowledged and understood at a deep level. Consider, for example, the following factors surrounding any situation in which a child's learning is being assessed.

**Teacher characteristics.** Teachers observe and document young children's learning through the lens of their own personal and professional experiences. Because no one is naturally blind to differences in color, race, gender, socioeconomic status, and other personal characteristics, educators must combine a deep understanding of child development with personal introspection to mitigate the implicit biases that can skew their interpretation of children's behavior.

**Child characteristics.** Children grow and develop in the context of their families and communities, and they, like teachers, bring their own unique set of experiences to any assessment setting. For example, the language used to communicate with a child as well as that used by the child to demonstrate his knowledge and skills are critical variables in accurately discerning the child's abilities. When a child does not fully understand the language used by an adult in an assessment situation, it should not be surprising that the child's performance appears to be poor. Further, when the adult does not fully understand the language of a child she is assessing, accurately documenting and evaluating the child's conversation and writing may not be possible.

**Test characteristics.** Standardized, norm-referenced measures are often viewed as benefiting some groups while disadvantaging others because of the focus of test items or the language of the assessment. Nuanced distinctions that can favor one subgroup over another must be explored and addressed. The purpose of assessment instruments is to differentiate among those with varying degrees of knowledge, skills, and dispositions around a specific topic. However, achieving fairness in any assessment context requires an assurance that the results differentiate by competence in the domain that is assessed rather than by racial or ethnic subgroup, gender, or socioeconomic status. Without such assurance, the fairness of the assessment is in doubt, and there can be little confidence in the interpretation of the results. Achieving fairness is not easy. The individual test items presented on a standardized assessment cannot have precisely the same familiarity to all students. This is why it is so important that an assessment be used with children who are similar to the population on whom the test was normed.

A completely culture-free test would be very difficult to construct; even a nonverbal test would have some cultural components. However, it is critical for educators and officials working in districts and state educational offices to understand how factors unrelated to the purpose of a test may affect the way it is administered, how children perform, and how the results are interpreted and used.

And early childhood teacher preparation programs have the opportunity and the challenge to enhance their assessment-related content so that candidates have a repertoire of assessment strategies and the knowledge and skills needed for accurate use and interpretation of assessment data.

All educators can become better advocates for young children when they are able to demystify assessment and testing and understand the strengths and limitations of the range of assessment options. As educators build their assessment literacy, they can inform families and hold policy makers responsible for supporting sound assessment practices for young children and the programs that serve them.



## Reflection Questions

1. Why does assessment matter?
2. How might you explain the difference between assessment and tests to a colleague or child's family?
3. This article mentions a number of benefits of assessment. What other benefits can you identify or have you seen for children in your own program?
4. Consider the ways you currently assess the different types of knowledge and skills children are expected to have and do. What barriers have you encountered? What solutions could you implement to overcome these barriers?
5. How do you use the information you gather about children's learning and development to inform your practice? In what ways might you make even more meaningful connections between assessment and instruction?

## References

- AERA (American Educational Research Association), APA (American Psychological Association), & NCME (National Council on Measurement in Education). 2014. *Standards for Educational and Psychological Testing*. Washington, DC: AERA.
- Chappuis, J., R.J. Stiggins, S. Chappuis, & J.A. Arter. 2012. *Classroom Assessment for Student Learning: Doing It Right—Using It Well*. 2nd ed. Upper Saddle River, NJ: Pearson.
- ED (US Department of Education). 2018. "Definitions." Accessed January 14. [www.ed.gov/early-learning/elc-draft-summary/definitions](http://www.ed.gov/early-learning/elc-draft-summary/definitions).
- IOM (Institute of Medicine) & (NRC) National Research Council. 2015. *Transforming the Workforce for Children Birth Through 8: A Unifying Foundation*. Washington, DC: National Academies Press.
- Jablon, J.R., A.L. Dombro, & M.L. Dichtelmiller. 2007. *The Power of Observation: Birth to Age Eight*. 2nd ed. Washington, DC: Teaching Strategies.
- Jones, J. 2004. "Framing the Assessment Discussion." *Young Children* 59 (1): 14–18.
- NRC (National Research Council). 2008. *Early Childhood Assessment: Why, What, and How*. Washington, DC: National Academies Press. doi:10.17226/12446.
- Rous, B., & R. Hallam. 2016. "Screening and Supporting Children at Risk for Developmental Delay or Disability." In *The Leading Edge of Early Childhood Education: Linking Science to Policy for a New Generation*, eds. N.K. Lesaux & S.M. Jones, 117–34. Cambridge, MA: Harvard Education Press.
- Shepard, L., S.L. Kagan, & E. Wurtz, eds. 1998. *Principles and Recommendations for Early Childhood Assessments*. Washington, DC: National Education Goals Panel.
- Stiggins, R.J. 1991. "Assessment Literacy." *Phi Delta Kappan* 72 (7): 534–39.

## About the Author

**Jacqueline Jones**, PhD, is president and CEO of the Foundation for Child Development. She served as the country's first deputy assistant secretary for Policy and Early Learning in the US Department of Education and was the assistant commissioner for the Division of Early Childhood Education in the New Jersey State Department of Education.

Photographs: pp. 6, 9, © Getty Images