

## **Issues of Assessment in Testing Children Under Age Eight**

**Gwen G. Stevens**  
**Karen DeBord**

### **Abstract**

There are differing opinions about the value of using standardized tests to assess the abilities and knowledge of young children. Research with young children (under 8 years old) indicates that using standardized tests for student grade placement or school retention can prove to be harmful to children's ultimate achievements. This article describes the multiple methods that can appropriately be used to evaluate student progress while shaping assessment policies that represent age-appropriate learning goals without using "high-stakes" accountability testing of individual children prior to third grade.

Using standardized tests to assess the abilities and knowledge of young children is viewed differently among educators, parents, politicians, and members of the community. Tests that are valid and reliable are extremely important tools when they are used to determine or to assure proper treatment for a child's special learning needs. However, research with young children (under 8 years old) indicates that using standardized tests for student grade placement or school retention can prove to be harmful to children's ultimate achievements.

### **Designing assessment systems**

Across the nation, school districts are selecting from among standardized measures, school district designed group-administered measures, and open-ended subjectively scored tests to make decisions about school or teacher accountability and student placement. [\(1\)](#) These tests are part of accountability systems that have been designed differently in each state in an attempt to assure that student learning has occurred. In many cases, test results and accompanying ramifications are neither clearly understood by parents nor policy makers. Often in an effort to design an accountability system, the developmental level of the child seems lost. What are assessment systems used for and what should well-designed assessment systems consider?

An assessment system should include a variety of instruments for various categories or purposes. Clarifying the main purpose of the assessment, determining what should be measured, establishing procedures for data collection, and selecting data sources (child work, standardized tests, teacher report, parent report) are all components in an assessment process. Safeguards, however, should be in place to protect against harmful or questionable assessment-based decisions with the consequences or use of the results spelled out to the parents.

According to Katz (1997) and Kagan (1998), assessment of individual children is currently used

- to determine progression on meaningful developmental achievements,
- to place or promote,
- to detect special needs, learning, and teaching problems,
- to assist with curriculum and instruction decisions,
- to help a child assess his or her own progress,
- to boost learning,
- to evaluate programs,
- to monitor trends, and
- for "high-stakes" accountability.

Common principles can guide assessment policies and practices of young children. First, assessments should benefit children by improving the quality of educational programs or in providing direct services to children. Assessments should link to a specific purpose and be valid, fair, and reliable for that purpose. When designing assessment policies, policy makers should consider that reliability and validity increase with children's age, and that the method of data collection and content through any data collection method is selected should be age appropriate. It is critical to consider language development as well as special learning needs when determining both appropriate methods and in interpreting the meaning of assessment results. Including multiple sources of evidence is key to presenting a clear picture of the child's learning needs. This means collecting information from the child's portfolio of work as well as from reports of parents and teachers (Kagan 1998). The use of more formal assessments (i.e., test results) will increase with age, especially after third grade. These assessments should still be balanced with informal methods such as direct observations which may include looking at

samples of work and drawings, asking other adults about the child, and asking questions of the child either orally or in writing.

As parents and educators plan assessments of young children's learning, Katz (1997) recommends the following:

- Realize that report cards and grades have limitations. There are numerous reasons why achievement scores and report cards with letter grades are not suitable for children at and below third grade. Prior to third grade, differences in developmental timetables are still unstable and are likely to change. Even though teachers need information to determine how well children are progressing, there is little research about how parents use this information.
- Evaluate children's performances so that the findings have real meaning. There are some things that children are tested on that have little significance beyond the test performance itself.
- Children can be encouraged to help set standards to assess their own work. This is a powerful learning tool. To learn how to assume some responsibility of their own progress, children should be encouraged to assess their own progress. Children can be asked questions such as what do you think you need help with? What do you think you are doing good in? What do you need to concentrate on more? Parents and teachers both should be involved with the process.
- Children can be involved in helping the teacher evaluate the classroom and determine the learning climate.

### **Use of assessment results to shape policy**

Assessments, when used as an ongoing process can shape teaching, help teachers and parents make decisions about child learning challenges, and help children master particular skills. High quality assessment can reveal important information that will support not only individual child learning, but will provide educational accountability data at the district, state, and national level. The data can then be used to improve services and educational programs.

High quality assessment results can help shape school policy, lead to improvements in teacher performance, and determine factors that may affect learning such as health care, child care, and preschool services. To assure children receive needed services, direct measures are needed to track children's early learning. However, as accountability pressures have mounted, many

teachers have turned their focus from student learning to preparing curriculum focusing on what is to be tested.

### **Difficulty in assessment design**

Assessments must be designed with attention given to the vast diversity of children. During young childhood (birth to age 8) children's rates of motor, physical, and linguistic development outpace growth rates at all other stages. Variance in growth coupled with environmental diversity influences the child's growing intellect in all domains. Schweinhart (1993) notes that assessment tools must be developmentally appropriate, reliable, valid, and user-friendly. These principles do not hold true when we review actual assessment systems.

The very skills best learned and assessed in context -- reading, math, and language -- are the skills assessed generally by school districts with standardized, group-administered, paper-pencil tests which are inappropriate for young children. Young children do not understand the purpose of formal testing. Moreover, tests that are given at one point in time may not give a total picture of a child's abilities. On a given test day, children may experience externally imposed anxiety or apathy contributing to the questioned reliability of such measures. There are several studies that have documented the associated stress young children experience during testing (Wodtke et al. 1989; Fleege et al. 1992). Still, test designers profess that group-administered tests are more objective than the documented work of children with supportive intentional observations in the context of daily activities (Meisels 1993).

Tests, which generally fail on the criteria of developmentally appropriate, have the power to change perceptions of children about themselves and about their educational experiences. Test scores are more likely to disclose more about a child's test-taking skills than his or her knowledge. Group administered tests can only focus on the "acquisition of simple facts, low-level skills, superficial memorization, and isolated evidence of achievement" (Meisels 1993, 35). With the trend to use group administered, objectively scored, paper-pencil tests for grade retention, the negative effects of retention overtime are likely to outweigh the positive benefits of using an appropriate accountability system to enhance student learning.

The stringent accountability standards of today are linked strongly to grade retention. Grade retention, however, as reported by The National Association of School Psychologists (NASP), is not attributed with subsequent school success. NASP indicates that most children do not "catch up" when held back and that although some retained students do better at first, these children often fall behind again in later grades (1996). Students who are held back tend to get into trouble, dislike school, and feel badly about themselves more often than children who go on to the next grade. In addition to the conclusions that NASP has drawn from the research, the weakened self-esteem that accompanies retention plays a role in how well the child may cope in the future.

Research has shown that children view the thought of failing a grade to be almost as stressful as the death of a parent or blindness (Sevener 1990). "Even more staggering is the fact that being held back twice makes dropping out of school a virtual certainty" (Setencich 1994, 7).

### **Alternatives to testing for grade placement**

Children have the right to have schools and teachers held accountable for providing high-quality instruction. Furthermore, documentation of children's progress is needed to assure needed intervention with special services at a time when, perhaps, these services will help them the most. And lastly, decision makers need accountability data. But the unintended negative outcomes for the teaching and learning of young children should be minimized. Group-administered and standardized performance-based tests should be reconsidered if they are used as the basic indicator of learning.

Educators cannot be held accountable for assuring that all children are above average. However, they are accountable for applying teaching strategies that are effective and appropriate for each child's learning needs. A child-centered, holistic approach to the accountability process is needed. This means early childhood educators must define what young children should be able to do and what they should know in order to create accountability systems that honor and safeguard high quality learning environments.

Tutoring, mixed-age classrooms, and smaller class size are a few of the strategies to building stronger educational systems. Tutoring supports student learning through individual attention while working throughout the year with children in academic areas that are difficult for them. With mixed-age classes, students learn at their own rate and advance to the next stage when they have mastered the required skills without the restriction of grade-level labeling. Students who receive individualized instruction where the teaching method is tailored to the individual student's style of learning provides a solid learning experience (Robertson 1997). This may require a lower student-teacher ratio and smaller classes, however, the benefits of such would replace the efforts to help students "catch-up" with the standards that tests are designed to assess.

### **Summary and recommendations**

Assessments are a natural part of instructional activities. Policy makers who are informed of the research associated with performance for young children and strategies for appropriate assessment can encourage the development or provision of assessment policies that represent age-appropriate learning goals. Given the possibility for misuse of cognitive screening measures, states that make screening tests mandatory should monitor their use and take precautions to avoid inappropriate uses. Consideration of investing in smaller class sizes, and teacher training about ongoing assessment can replace stressful and unreliable testing procedures.

Before the end of third grade there should be no "high-stakes" accountability testing of individual children (Kagan 1999). Assessment systems designed to gather data at the state or national level with children before age 8 should focus on social indicators that describe the conditions of learning. Student accountability systems should involve observation and documentation of development that tracks how well students are learning and developing.

In the future, the focus should be on children, their individual developmental stages, and how they learn. Teachers are accountable for student learning but not to the point that they feel pressured to teach to the test instead of instilling learning skills that will last a lifetime.

### Endnote

1. The tests that are emerging are not those associated with individualized education plans (IEPs) designated for students with special learning needs. [Return to text.](#)

### References

Fleege, P. O., R. Charlesworth, D. C. Burts, and C. H. Hart. 1992. Stress begins in Kindergarten: A look at behavior during standardized testing. *Journal of Research in Childhood Education*, 7: 20-26.

Gerber, A. 2000. An alarming trend, teachers are the cheaters. *USA Today*. July 13.

Henry, T. 2000. Test pressure blamed for cheating. *USA Today*. July 13.

Henry, T. 2000. It's about helping all children success. *USA Today*. July 13.

Kagan, S. L., L. Shepard, and E. Wurtz. 1998. *Principles and Recommendations For Early Childhood Assessments*. Washington D.C.: U.S. Government Printing Office.

Kagan, S. L. 1999. A5 Redefining 21st-Century Early Care and Education. *Young Children*. 54(6):2-3.

Katz, L. G. 1997. A Development Approach to Assessment of Young Children. *ERIC Digest*.

Meisels, S.J. 1993. Remaking the classroom Assessment with the work sampling system. *Young Children*. 48(5):34-40.

National Association of School Psychologists. 1996. Should my child repeat a grade? Bethesda, Md.: NASP. Available at <http://www.npin.org/library/1996/n00334/n00334.html>

Robertson, A. S. 1997. When retention is recommended, What should parents do? ERIC Clearinghouse on Elementary and Early Childhood Education. Champaign Ill.: *ERIC Digest*. ED408102

Schweinhart, L. 1993. Observing young children in action: The key to early childhood assessment. *Young Children*. 48(5):29-34.

Setencich, J. 1994. "Impact of early grade retention on the academic achievement and self-esteem of seventh and eighth grade." Paper presented at the Annual Convention of the National Association of School Psychologists, Seattle, Wash.

Sevener, Donald. 1990. Retention: More malady than therapy. *Synthesis*. 1(1):1-4.

Testing, Tracking, & Timing: Challenging policies that deny equity and access to young children. 1989 *The Early Childhood Advocate* 1(1):1,4.

Wodke, K.H. , F. Harper, M. Schommer, and P. Brunelli. 1989. How standardized is school testing? An exploratory observational study is standardized groups testing in kindergarten. *Educational Evaluation and Policy Analysis*. 11:223-235.

### **Author**

Gwen G. Stevens, Family and Consumer Extension Agent, North Carolina Cooperative Extension Service; and Karen B. DeBord, Ph.D., Associate Professor and Child Development Specialist, NC State University.

### **Cite this article:**

Stevens, Gwen and Karen DeBord. 2001 Issues of assessment in testing children under age eight. *The Forum for Family and Consumer Issues* 6(2).